JMIR mHealth and uHealth

Impact Factor (2024): 5.4
Volume 9 (2021), Issue 9 ISSN 2291-5222 Editor in Chief: Lorraine Buis, PhD, MSI

Contents

Reviews

Original Papers

Tutorial

Corrigenda and Addendas

Review

Using Acoustic Speech Patterns From Smartphones to Investigate Mood Disorders: Scoping Review

Olivia Flanagan[1], BA, MSc; Amy Chan[2], BPharm (Hons), PhD; Partha Roop[3], BE, MTech, PhD; Frederick Sundram[1], BMedSc, MB BCh BAO, FRCPsych, MA, MSc, PhD

[1]Department of Psychological Medicine, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand

[2]School of Pharmacy, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand

[3]Faculty of Engineering, University of Auckland, Auckland, New Zealand

Corresponding Author:
Frederick Sundram, BMedSc, MB BCh BAO, FRCPsych, MA, MSc, PhD
Department of Psychological Medicine
Faculty of Medical and Health Sciences
University of Auckland
Building 507, Level 3
28 Park Avenue, Grafton
Auckland, 1023
New Zealand
Phone: 64 9 923 7521
Email: f.sundram@auckland.ac.nz

Abstract

Background: Mood disorders are commonly underrecognized and undertreated, as diagnosis is reliant on self-reporting and clinical assessments that are often not timely. Speech characteristics of those with mood disorders differs from healthy individuals. With the wide use of smartphones, and the emergence of machine learning approaches, smartphones can be used to monitor speech patterns to help the diagnosis and monitoring of mood disorders.

Objective: The aim of this review is to synthesize research on using speech patterns from smartphones to diagnose and monitor mood disorders.

Methods: Literature searches of major databases, Medline, PsycInfo, EMBASE, and CINAHL, initially identified 832 relevant articles using the search terms “mood disorders”, “smartphone”, “voice analysis”, and their variants. Only 13 studies met inclusion criteria: use of a smartphone for capturing voice data, focus on diagnosing or monitoring a mood disorder(s), clinical populations recruited prospectively, and in the English language only. Articles were assessed by 2 reviewers, and data extracted included data type, classifiers used, methods of capture, and study results. Studies were analyzed using a narrative synthesis approach.

Results: Studies showed that voice data alone had reasonable accuracy in predicting mood states and mood fluctuations based on objectively monitored speech patterns. While a fusion of different sensor modalities revealed the highest accuracy (97.4%), nearly 80% of included studies were pilot trials or feasibility studies without control groups and had small sample sizes ranging from 1 to 73 participants. Studies were also carried out over short or varying timeframes and had significant heterogeneity of methods in terms of the types of audio data captured, environmental contexts, classifiers, and measures to control for privacy and ambient noise.

Conclusions: Approaches that allow smartphone-based monitoring of speech patterns in mood disorders are rapidly growing. The current body of evidence supports the value of speech patterns to monitor, classify, and predict mood states in real time. However, many challenges remain around the robustness, cost-effectiveness, and acceptability of such an approach and further work is required to build on current research and reduce heterogeneity of methodologies as well as clinical evaluation of the benefits and risks of such approaches.

(JMIR Mhealth Uhealth 2021;9(9):e24352) doi:10.2196/24352

KEYWORDS

smartphone; data science; speech patterns; mood disorders; diagnosis; monitoring

Introduction

Background

Mood disorders are common with 12-month prevalence rates ranging between 6.6% and 11.9% [1] and represent significant personal, social, and economic burden [2,3]. However, these disorders remain underrecognized and undertreated [4]. Early objective identification of warning signs that are associated with such disorders can facilitate time-sensitive interventions and early responses by the health care provider [5]. However, current methods of mental health assessment are limited in their capacity to accomplish this due to the following reasons. First, mental health assessments rely on self-reporting and clinical interviews, which depend on individuals' memories and consequently are susceptible to recall and other biases [6]. Second, assessments often take place in clinical contexts by health care professionals, which may be substantially different from one's usual environment and thus limits ecological validity [7]. Finally, individuals may not recognize the need to seek treatment until symptoms reach a level of severity that warrants clinical attention, making treatment more difficult than if the mood changes had been detected earlier [8]. Moreover, the COVID-19 pandemic is having a profound impact on our way of life and mental well-being [9-11]. Increased fear, uncertainty, and anxiety as well as the public health measures taken to manage the pandemic and social/economic crisis mean that people are more vulnerable to developing mood disorders and engagement with health care providers is even more difficult [12]. There is thus a need for better tools, which can provide objective mental health assessments on an ongoing basis and within a home setting, to enable earlier and accurate diagnosis of mood disorders and detection of changes in mental state.

There has been increasing interest in the use of data-driven approaches in the detection and monitoring of health and disease [13,14]. The rapid growth of smart-sensor integration in smartphones allows the collection of objective quantitative markers of behavior and function [15,16]. In mental health, this approach may be particularly feasible; for example, speech can be a key distinguishing characteristic for the diagnosis and monitoring of mental health disorders [17]. In this regard, diagnosis and monitoring are overlapping concepts as changes to mental state can be monitored and therefore prospectively tracked for diagnostic purposes. Current clinical measures such as the Young Mania Rating Scale for the diagnosis of mania [18] and the Hamilton Depression Scale for depression [19] both use clinical observations of speech to aid diagnosis. In bipolar disorder (BD), pressure of speech is a key diagnostic symptom in mania, and poverty of speech in depression. Evidence suggests that these speech differences can be quantified via measurement of verbal fluency (eg, word and error count, switching, and clustering abnormalities) [20]. With the emergence of machine learning approaches [21], the automatic classification of speech as an objective measure for mood disorders is becoming more feasible. Smartphones may therefore offer a unique opportunity to augment current mental health assessment methods or bypass many of the limitations associated with them [22].

In this review, speech/voice patterns or features refer to measurable and objective aspects of speech that affect the acoustic quality of speech production (eg, prosodic features such as pitch). The reader is referred to the review by Malhi et al [23] which covers several aspects of these features. Classifiers can be used to investigate mood states, whereby a classifier refers to a hypothesis or discrete-valued function that is used to assign (categorical) class labels to particular data points [24]. Studies have classified people according to presence/absence, severity, or score-level prediction based on brain, wearable, and Twitter activity using machine learning [25-27]. However, the well-established relationship between voice and mood disorders [25] has been under-investigated—the emergence of machine learning approaches [21] leads to the question of whether smartphone voice data could provide clinical insight into mood symptoms in real time. In recent years, studies have discussed the promise of smartphone voice data to diagnose mood disorders [28,29]. However, fundamental scientific questions remain before smartphones can be used as validated and objective clinical tools [30]. Although there is an ever-growing number of studies focusing on the collection of objective data from smartphone or external sensors to diagnose and monitor mood disorders, only a small portion of these have included speech features as a key objective marker. Considering the importance of this emerging field, the speed of innovations, and new developments [28], it was our aim to synthesize the literature on the use of speech patterns from smartphones in the diagnosis and monitoring of mood disorders, and the accuracy and technical feasibility of this approach.

Objectives

The aim of this review was to evaluate the current state of research on the use of speech patterns from smartphones to diagnose and monitor mood disorders. Specifically, objectives of this review are to (1) characterize studies that have been conducted on speech patterns to diagnose and monitor mood disorders using smartphone devices and (2) provide details on the technical feasibilities of smartphones to achieve this, such as their ability to control ambient noise and how privacy was managed. "Speech features and patterns" referred to in this review describe objective markers such as the acoustics of, rather than behavioral patterns collected from smartphone use (eg, the length of time spent on the phone).

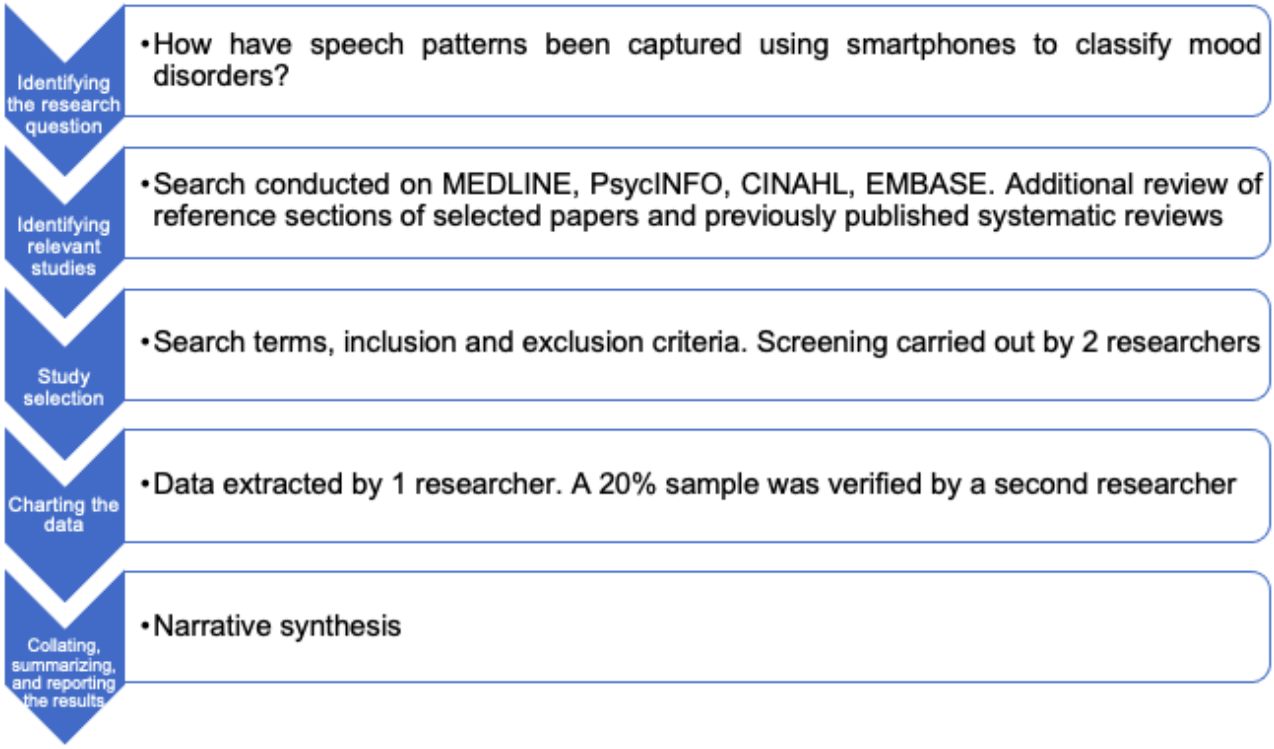
Methods

Design

A scoping approach was adopted for this review which according to Nicholas and colleagues [31] aims "to map rapidly the key concepts underpinning a research area and the main sources and types of evidence available, and can be undertaken as stand-alone projects in their own right, especially where an area is complex or has not been reviewed comprehensively before." This method was chosen because the field of machine learning in mood disorders is advancing exponentially; therefore, it was deemed appropriate to focus specifically on exploring broadly the nature of research activity, as per Arksey and O'Malley's [32] first goal of scoping reviews. This study was guided by the methodological framework proposed by Arksey

and O'Malley's [32] which involves a 5-stage process (Figure 1) that was benchmarked against the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines [33] to ensure rigor.

Figure 1. Methodological framework used in this scoping review as per Arksey and O'Malley [32].



Search Strategy

MEDLINE (ProQuest), PsycINFO (ProQuest), EMBASE (Elsevier), and CINAHL (EBSCO) databases were used to search for studies published from the date of database conception to November 24, 2020. The following search terms and their variants were used in each database: "mood disorders", "smartphone", and "voice analysis", using the Boolean search operator "OR" and "*" where appropriate, and combined using the Boolean operator "AND." Multimedia Appendix 1 presents the full search strategy. To capture appropriate studies, the search was limited to English language publications only. Textbox 1 describes the full inclusion and exclusion criteria.

Textbox 1. Inclusion and exclusion criteria. IVR: interactive voice response; EMA: ecological momentary assessment.

Inclusion criteria

- Use of smartphone
- Focus on diagnosing/monitoring of a mood disorder including depression, mania, and bipolar affective disorder
- Clinical populations recruited prospectively
- Captures voice data
- English language

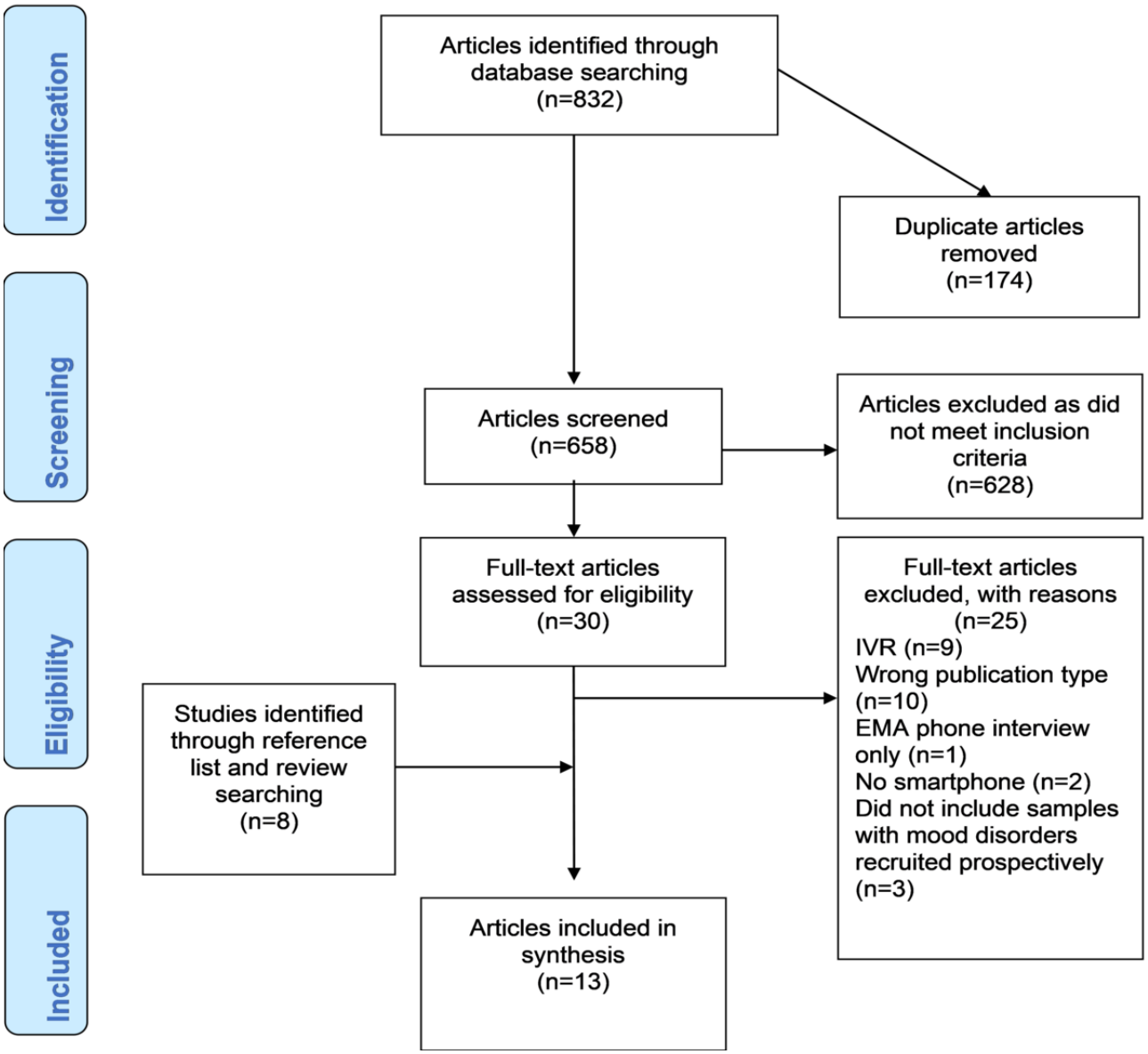
Exclusion criteria

- Not using a smartphone device (eg, laptop)
- Focus on other health conditions rather than mood disorders (eg, Parkinson disease), or focus on mood disorder treatment/intervention rather than diagnosis/monitoring; or examined effects of speech patterns, or smartphone use in general without reference to mood; or collected IVR/EMA data only
- Does not capture voice data
- Non-English language publications

Search Outcomes

Figure 2 details the process of study selection using the PRISMA flow diagram [33]. After duplicates were removed, articles were downloaded into Rayyan [34], a systematic review web application, where inclusion/exclusion decisions were made. Screening of all titles and abstracts was undertaken by the lead author (OF). A second reviewer, blind to the inclusion/exclusion decisions of the articles, randomly screened 20% of titles and abstracts, with agreement on 128 out of 132 articles (96.9%) screened for inclusion/exclusion and all conflicts resolved by consensus following discussion between both raters. Both OF and the second reviewer read all articles selected for full-text review. Reference lists of articles included in the review were also manually screened to identify any relevant studies that were not identified through database searching, and systematic reviews that were identified during the search process were also screened and relevant studies extracted.

Figure 2. PRISMA flowchart demonstrating search process. EMA: ecological momentary assessment; IVR: interactive voice response; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses.



Data Extraction and Analysis

Data were charted by OF and entered onto a data charting form using Microsoft Excel. To ensure accuracy and consistency of the process, a sample of 20% of the information being entered into Excel was verified by a second reviewer [35]. No significant discrepancies or errors were detected. The charting process allowed the researcher to describe the characteristics of the studies included in the review and prepared studies for analysis [36]. Textbox 2 describes a summary of the data captured from the studies included in the review. To analyze the data, a narrative review synthesis method [37] was selected to capture each study's characteristics and methods to investigate voice analysis in mood disorder diagnosis and monitoring.

Textbox 2. Summary of data captured.

Summary of data captured

- Author, year, location, study design
- Sample size and diagnosis
- Aim of studies
- Methods of data capture

Length of audio data capture

- Timing of data capture
- Type of audio data captured
- Ambient noise control
- Privacy
- Clinical outcome measurement
- Classifier used
- Key findings

Results

Characteristics of Included Studies

A total of 13 out of the 832 studies initially identified were included in this scoping review (studies were mostly excluded due to reasons of not involving a smartphone device and lack of focus on monitoring or detecting a mood disorder). The publication year of included studies ranged between 2011 and 2020 with 77% (10/13) of articles published from 2015 onward, highlighting the increase in interest and recent advancements in this area. Included studies used single-arm observational designs [38-45], quasi-experimental designs [46-48], and observational case reports [49,50].

Each study's aims, key characteristics, and findings are detailed in Table 1. Studies reported a variety of speech features analyzed and these features are summarized in Textbox 3. Additionally, an overview considering overlaps and differences across the included studies is provided covering these aspects: smartphone device/platform/apps and data storage; characteristics of data capture; noise and privacy; and clinical outcome measurement.

Table 1. Results of included studies.

Study; year; location; design; sample size and diagnosis	Study aim; methods of data capture; duration; clinical evaluations	Audio data captured; classifier used	Assessment; key finding
Abdullah et al [38]; 2016; USA; single-arm observational design; 7 patients with BD[a]	Assessed stability and rhythmicity for individuals with BD; microphone activated during daily conversations; 4 weeks; ratings via Social Rhythm Metric-5 (SRM-5) via initial questionnaire, poststudy questionnaire, and interview	Speaking rate and variations in pitch; SVM[b]	Classification based on the same individual's data; classified individuals into stable or unstable mood states with high accuracy (precision=0.85 and recall=0.86).
Dickerson et al [50]; 2011; USA; observational case report; 1 participant with major depressive disorder	Created a real-time depression monitoring system for the home; free speech response to daily questions; 2 weeks; scores from Centre for Epidemiological Studies Depression (CES-D) subjectively provided via touchpad	Fundamental frequency and speech pause time; pitch detection algorithm	Classification based on the same individual's data; the model fit the data with a residual error of 0.0916 on 12 degrees of freedom (P<.011).
Faurholt-Jepsen et al [39]; 2016; Denmark; single-arm observational design; 28 patients with BD	Investigated voice features collected during phone calls as objective markers of affective states in BD; natural phone calls; 12 weeks; fortnightly clinical interviews using the HAMD[c] and YMRS[d]	openSMILE toolkit with 6552 numerical features including pitch, variance, etc.; random forest algorithms	Classification based on the same individual's data; phone calls could classify manic states with an AUC[e] of 0.89 compared with an AUC of 0.78 for depressive states.
Gideon et al [46]; 2016; USA; quasi-experimental design; 37 patients with BD with rapid cycling	Investigated acoustic variations with different types of phones and the preprocessing and modeling changes necessary to detect mood; natural phone calls; 6-12 months; weekly calls with clinicians to conduct HAMD and YMRS interviews	Spectral power ratio and spectral centroid; SVM	Comparison between the Galaxy S3 and Galaxy S5 groups based on their individualized data; preprocessing, feature extraction, and data modeling improve the performance of mixed device systems (AUCs of 0.57 and 0.64 for manic and depressed states, respectively, to 0.72 and 0.75).
Grünerbl et al [40]; 2015; Austria; single-arm observational design; 10 patients with BD	Introduced a system which can recognize depressive and manic states and detect state changes of patients with BD; natural phone calls; 12 weeks; HAMD and YMRS examinations performed every 3 weeks by clinicians over the phone	Kurtosis energy, mean second and mean third MFCC[f], mean fourth delta MFCC, maximum ZCR[g] and mean harmonic-to-noise ratio, SD, and range F0; naïve Bayes, k-nearest neighbors, j48 search tree, and conjunctive rule learner algorithms	Classification based on the same individual's data; phone call behavior did not provide as high a recognition rate as voice features. A fusion of 4 different sensor modalities achieved the highest recognition accuracies of 76% and state change detection precision and recall of over 97%.
Guidi et al [49]; 2015; France; observational case report; 1 patient with BD	Collected and analyzed prosodic features via an Android app; structured tasks (eg, reading, counting), commenting on a picture performed 15 times; 14 weeks; Quick Depression Inventory (QID) and YMRS assessments conducted by a clinician during the day before each voice recording session	Mean F0, jitter, and F0 SD; the SWIPE algorithm (pitch estimator algorithm)	Classification based on the same individual's data; mean F0 from each voiced segment can be reliably estimated, but weak correlations were reported between audio features and mood.
Karam et al [41]; 2014; USA; single-arm observational design; 6 patients with BD	Investigated whether speech collected in an unstructured setting can be used to assess underlying mood state; phone calls as well as speech in clinical interviews; 6 months to 1 year; weekly phone-based HAMD and YMRS assessments with a clinician	23 Low-level features extracted using the openSMILE toolkit including pitch; root mean square energy, zero-crossing rate, and the amplitude of the speech waveform; SVM	Classification based on the same individual's data; hypomania and depression can be differentiated from euthymia using speech-based classifiers trained on both structured and unstructured cell phone recordings.

Study; year; location; design; sample size and diagnosis	Study aim; methods of data capture; duration; clinical evaluations	Audio data captured; classifier used	Assessment; key finding
Maxhuni et al [42]; 2016; Austria; single-arm observational design; 10 patients with BD	Evaluated the performance of several classifiers, different sets of features, and the role of questionnaires for classifying BD episodes; natural phone calls; 12 weeks; scheduled interviews with a clinician by phone every 3 weeks using the HAMD, ADS (Allgemeine Depressionsskala/Common Depression Scale), YRMS, and Mania Self-Rating Scale (MSS)	Features extracted using openEAR and Praat. LOG energy, ZCR, probability of voicing, F0[h], MFCC, MEL spectrum, spectral energy in bands, spectral roll-off point, spectral flux, spectral centroid, and spectral max and min with discrete Fourier transform; comparison of C4.5, random forest, SVM, k-nearest neighbors, naïve Bayes, AdaBoost, and bagging algorithms	Classification based on the same individual's data; classification accuracy using spectral characteristics=82% or emotional characteristics=82%. Decision trees performed best.
Muaremi et al [43]; 2014; Austria; single-arm observational design; 12 patients with BD	Explored the feasibility of voice analysis during phone conversation to predict BD episodes; natural phone calls; 12 weeks; HAMD and YMRS interviews conducted by a clinician every 3 weeks at the hospital	openSMILE; Kurtosis energy, mean second MFCC, mean third MFCC, mean fourth delta MFCC, maximum ZCR, mean harmonic-to-noise ratio, F0 SD, range F0; comparison of SVM, logistic regression, random forest, and neural networks	Classification based on the same individual's data; classification accuracy using patient voice characteristics=80%. Combination of all data streams=83%. Random forest algorithms performed the best.
Osmani [44]; 2015; Italy; single-arm observational design; 9 patients with BD	Investigated whether data from smartphone sensors could be used to recognize BD episodes and to detect behavior changes that could signal the onset of an episode using objective, sensor data; natural phone calls; 12 weeks; every 3 weeks by a clinician using the HAMD and YMRS	Phone calls and sound analysis; comparison of naïve Bayes, k-nearest neighbors, search tree, and a conjunctive rule learner algorithm	Classification based on the same individual's data; sound analysis accuracy=70%, recall=60%, precision=59%. Best accuracy was achieved through a combination of modalities (accelerometer, location, phone and sound recall=97.4%, precision=97.2%).
Pan et al [47]; 2018; China; quasi-experimental design; 21 hospitalized patients with BD	Compared the accuracy of SVM and GMM in the detection of manic state of BD of single patients (smaller sample size) and multiple patients (larger sample size); free open phone calls with clinicians; 2 days; Bech–Rafaelsen Mania Rating Scale (BRMS) used by a clinician while patients were in the hospital	openSMILE; pitch, formants, MFCC, LPCC[i], gammatone frequency cepstral coefficients, etc. were preprocessed and extracted; comparison of SVM and GMM[j]	Comparison between single-patient experiments (n=3) and multiple patients experiments (n=21) based on their individualized data; LPCC demonstrated the best discrimination efficiency. The accuracy of manic state detection for single patients was better using SVM than GMM methods. Detection accuracy for multiple patients was higher using GMM than SVM methods.
Place et al [45]; 2017; USA; single-arm observational design; 73 patients with a symptom of post-traumatic stress disorder or depression	Reported on models of clinical symptoms for post-traumatic stress disorder and depression derived from a scalable mobile sensing platform; daily audio diary entries; 12 weeks; baseline questionnaire at initial visit, at the end of the study patients completed a semistructured clinical interview on-site with a trained clinician, and completed a close-out survey	Mean speaking fraction, mean speaking rate, mean harmonicity, SD of harmonicity, mean vocal effort, SD of vocal effort, mean pitch variation, SD of pitch variation; comparison of 5 different algorithms for speaking fraction, speaking rate, harmonicity, vocal effort, and pitch variation	Classification based on the same individual's data; depressed mood was predicted from audio data with an AUC of 0.74.

Study; year; location; design; sample size and diagnosis	Study aim; methods of data capture; duration; clinical evaluations	Audio data captured; classifier used	Assessment; key finding
Gideon et al [48]; 2020; USA; quasi-experimental design; 51 patients with BD	Expanded clinical mood monitoring to predict when interventions are necessary using an anomaly detection framework; natural phone calls and phone calls with clinicians; 6-12 months; calls with clinicians using the HAMD and YMRS to retrospectively rate their mood each week	Emotion features (eg, Mel Filter Banks) using MADDoG[k] and transcript features (eg, speaker timing); automatic speech recognition model	High YMRS or HDRS compared with personal baseline; TempNorm can be used to transform the symptom severity ratings to effectively predict if an intervention should occur. Transcript features performed best for the clinical calls, while both transcript and emotion features worked well for natural speech.

[a]BD: bipolar disorder

[b]SVM: support vector machine

[c]HAMD: Hamilton Depression Rating Scale

[d]YMRS: Young Mania Rating Scale

[e]AUC: area under the curve

[f]MFCC: Mel-frequency cepstral coefficient

[g]ZCR: zero crossing rate

[h]F0: fundamental frequency

[i]LPCC: linear prediction cepstral coefficients

[j]GMM: Gaussian mixture model

[k]MADDoG: multiclass adversarial discriminative domain generalization

Textbox 3. Most common features used within included studies to analyze vocal aspects of speech.

Prosodic features

These include pitch (F0), speaking rate, jitter, shimmer, loudness, harmonic-to-noise ratio (HNR), log of energy, and Teager energy operation (TEO). Prosodic features represent the long-time (phoneme level) variations in perceived intonation, stress, and rhythm of speech.

- F0 refers to rate of vocal fold vibration.
- Jitter refers to the short-term fluctuations in pitch.
- Shimmer refers to the period-to-period variability of the signal peak-to-peak amplitude.
- Loudness refers to the intensity of auditory sensation produced.
- HNR refers to the average ratio of harmonic energy to inharmonic energy in a voice signal.
- Log of energy refers to the logarithmic short-term energy within a frame.
- TEO refers to amplitude and frequency modulations of vocal tract resonances generated by nonlinear airflows within the vocal tract.

Spectral and cepstral features

These include spectral flux (SF), spectral centroid (SC), Mel-frequency cepstral coefficients (MFCCs), linear prediction cepstral coefficient (LPCC), and gammatone frequency cepstral coefficients (GFCCs). These features characterize the speech spectrum, the average sound spectrum for the human voice.

- SF refers to the measure of the amount of frame-to-frame variance in the spectral shape.
- SC is a measure to characterize a spectrum.
- MFCCs are based on the Mel Filter Bank and describe the overall shape of a spectral envelope.
- LPCC models the human vocal tract as an infinite impulse response system that produces the speech signal.
- GFCCs are based upon the Gammatone Filter Bank where the filters model physiological changes in the inner ear and middle ear.

Smartphone Device/Platform/Applications and Data Storage

The majority of studies provided participants with an Android smartphone as a data collection tool. Dickerson et al [50] provided their participants with an iPhone and Faurholt-Jepsen et al [39] allowed study participants to use their own Android smartphone or were offered to loan an Android smartphone. To facilitate the collection of audio data, all studies, except Pan et al [47], used a cloud database. A variety of downloaded mobile apps were used, such as MoodRhythm [38], Empath [50] MONARCA [39,40,42-44], PRIORI [41,46,48], PSYCHE [49], and a Defense Advanced Research Projects Agency–funded app [45]. In most studies, data were captured locally on the device and then securely transmitted to a server periodically [38,39,41-43,45,46,48-50]. One study temporarily stored data

locally on the phone and then uploaded data to the cloud when the phone was being recharged and connected to Wi-Fi [47]. Grünerbl et al [40] stored data on an SD (secure digital) card at the end of everyday (for data security issues), while Osmani [44] did not mention how data were stored in his study.

Characteristics of Data Capture

Length of Audio Data Capture

The length of time spent capturing audio data ranged from 2 days [47] to 12 months [48].

Methods of Data Capture

Audio data were captured from participants when they read, counted, or commented on a picture aloud [49]; during daily conversation [38]; natural phone calls [39,40,42-44,46,48,49]; phone calls with clinicians [41,47,48]; daily audio diary entries [45]; or from responding to questions aloud such as "How was your day today" [50]. The frequency of evaluations varied greatly between studies; for instance, studies collected data daily [38-41,43,44,47,50], weekly [45,46,48,49], or were dependent on when phone calls were made [42].

Timing of Data Capture

Data were captured either during an acute episode of BD [38-41,43,44,48,49] or depression [50]; or in one study during the daily life of veterans with symptom(s) of post-traumatic stress disorder (PTSD) or depression [45].

Audio Data Captured

See Textbox 3 for a description of common audio features captured. Some studies also made use of feature extractors for signal processing and machine learning applications such as openSMILE [40,41,43,47], openEAR, and Praat [42].

Noise

Only one-third of the included studies referred to a method to control for ambient noise. The methods varied and included using energy intensity and distribution likelihood [38], using a "guard zone"/threshold to filter out noise [50], using a segmentation algorithm robust to variation in noise [46], and using a double-layer sound-insulated glass room when talking [39]. Gideon et al [48] stated that their data consisted of unconstrained natural speech in the presence of noise, so imperfect transcriptions were expected (evident by the 39.7% word count error). However, they note that their previous work reveals that mood recognition (especially mania) is improved by addressing variability in clinical recordings due to device differences [46].

Privacy

In terms of protecting participants' privacy, no study evaluated speech content—only speech features were evaluated. Four studies did not report on the measures taken to protect participant privacy [42,44,48,50].

Clinical Outcome Measurement

Most studies used the Hamilton Depression Rating Scale [39-44,46,48] and the Young Mania Rating Scale [39-44,46,48,49] for assessment of mood. Studies also used the Social Rhythm Metric [38], the Centre for Epidemiological Studies Depression Scale [50], the Quick Inventory of Depressive Symptomatology [49], the Bech–Rafaelsen Mania Scale [47], the Structured Clinical Interview for DSM-5 [45], Primary Care PTSD Screen for DSM-5 [45], and the Patient Health Questionnaire-2 [45].

Discussion

Principal Findings

This scoping review evaluates the current state of research on the use of speech patterns from smartphones to diagnose and monitor mood disorders. We found robust evidence that demonstrates a high potential to use smartphone voice data to monitor/detect mood disorders in individuals in real time. These voice analyses can be used to detect changes in mood at the different stages of mental health presentation [25]—first at the onset where acute changes in speech patterns can occur and during remission, as speech patterns return to the individual's baseline level; and then later to monitor for early warning signs that may predict relapse [39]. There is also potential for these voice data to be used to distinguish between clinical conditions such as BD and schizophrenia [51], and within a disorder, between different clinical states such as mania, hypomania, and mixed states for bipolar [20,46]. This section discusses the key findings from this review (most common speech features, classifiers, and audio capture methods used and smartphone device technical considerations) and the various challenges that remain with regard to accuracy, feasibility, and practical considerations and identification of gaps and future research implications.

Accuracy, Feasibility, and Practical Considerations

With regard to feature extraction, there are many speech features that have been found to be related to depression and BD [52]. Within the included studies, the most common speech features analyzed included prosodic (fundamental frequency, speaking rate, and energy), spectral (spectral centroid), and cepstral features (Mel-frequency cepstral coefficients). Karam et al [41] revealed that the most informative features for classification of bipolar states are the average binary voiced activity detection, SD of pitch, segment average of the zero-crossing rate, and segment average of the smoothed voiced activity detection. Muaremi et al [43] showed that the most important speech features for prediction of bipolar states were harmonic-to-noise ratio (HNR) value, the number of short turns, and the variance of pitch F0. Moreover, Pan et al [47] found that linear prediction cepstral coefficient and gammatone frequency cepstral coefficient contain important mood information for manic state than other features. Overall, all studies analyzed prosodic features of speech, with F0 being the most common feature. However, due to the natural variations in individual speaking styles and the wide clinical profile of BD and depression, a single-dimensional prosodic feature does not contain sufficient discriminatory information for use as a clinical marker, and a multivariate approach is required. In addition, further research is required to verify whether other features, such as glottal features, can be utilized to monitor and diagnose mood disorders.

Given the current lack of a reliable speech feature or clarity around multivariate features for mood disorder classification or prediction, fusion of objective data measures acquired from multiple sensors (eg, GPS, voice, and acceleration) or a combination of physiological (eg, heart rate variability) and behavioral parameters is a promising approach moving forward. This is reflected in the current work whereby studies that combined data on voice features with other automatically generated objective data increased the accuracy, sensitivity, and specificity of classifying affective states [39,40,43,44].

The 2 most popular modeling and classification techniques include support vector machines (SVMs) and Gaussian mixture models (GMMs). The most common classifier used in this study was SVM [38,41,46,47]. For instance, Pan et al [47] compared SVM with GMM in the detection of a manic state associated with BD of individual and multiple patients. They found SVM provided an appropriate tool for detecting manic states for individual patients, whereas GMM worked better when detecting manic states for multiple patients. Studies that have also compared multiple classifiers [40,42-45] found high promise for the use of random forest and other decision tree classification models in the detection of mood disorders [42,43]. The majority of studies reviewed in this study utilized supervised classification techniques [38,41,46,47] (ie, learning from labeled data to predict the class label of unlabeled input data [53]) rather than other machine learning techniques. This is most likely a result of the focus being on detection and diagnosis. Although SVM and GMM have been widely utilized, results hold promise for decision tree classification methods, which are able to assess the importance of the variables during the training process. This knowledge helps us to discover which nonrelevant parameters can be ignored, potentially resulting in a reduced computational effort on the smartphone.

Included studies in this review mostly used Android smartphones, which is unsurprising given their global market dominance [54]. However, despite their popularity, previous research has indicated less acoustic signal conformity in Android devices [55], attributed to the nonstandard hardware and software designs across manufacturers. Included in this review, Gideon et al [46] compared 2 different phones with various amounts of clipping, loudness, and noise and described methodologies to use during preprocessing, feature extraction, and data modeling to correct these differences and make the devices more comparable. Such methods were found to significantly increase the performance of mixed device systems. Given the increasing global popularity of smartphones, proper processing of acoustic data from multiple types of smartphones will be necessary to increase reliability and accuracy and mitigate the effects of differing amounts of clipping, loudness, and noise. This finding has important implications for engineers who create speech-based mood classification systems for smartphones, as they will have to optimize their design for a wide number of handset models.

In terms of what audio data were captured, the methods varied between using fixed or spontaneous speech. However, the evidence suggests that spontaneous speech such as free conversation or interviews contain more variability and can increase depressive and manic mood-state detection accuracies than using fixed speech (eg, reading text) [52,56,57]. Speech collection in natural environments highlights the applicability for autonomous ecologically valid monitoring of mood disorders. Future research therefore would benefit from adopting an unscripted setup, which preserves naturally expressed emotion. The length of data collection within the studies included in this review varied but were mostly of short duration, resulting in some studies having to exclude participants from final analyses as they did not exhibit recognizable changes in mood state [38,43]. To identify individual patterns that predict state changes, longer monitoring durations (greater than 12 weeks) may be required.

Gaps and Future Work

Spontaneous speech brings a greater need to handle ambient noise. Less than half of the studies included in the review described how noise was handled. The most practical method used was a “guard zone”/threshold to filter out noise [50] or the use of an algorithm that is robust to noise variation [46]. Future research needs to compare and investigate robust features and modeling techniques to mitigate the effects of noise. For example, a recent study by Mitra et al [58] found that using suitable and robust features and modeling strategies mitigated the performance degradation from varying background conditions. In their case they used damped oscillator cepstral coefficients instead of standard Mel-frequency cepstral coefficients and compared support vector regression and artificial neural networks for depression score prediction, revealing artificial neural networks to be more robust to support vector regressions.

Future research will also need to address technical, acceptability, and ethical issues of smartphone-based monitoring in order for this method to be reliably used in clinical practice. For instance, technical factors such as battery lifetime or individual usage (some individuals bring their smartphones everywhere they go, others do not) of the smartphone may serve as obstacles. Similarly, ethical issues remain such as how an individual’s privacy is preserved, how to mitigate the acceptability concerns (eg, unease or increased anxiety that constant surveillance and monitoring may cause), and how sensitive data concerning mental health are protected. None of the studies included in this review collected data on speech content but only speech features; however, if these systems are to be used in routine clinical care, a high standard of protection from security breaches is required.

Lastly, it is important that future research investigates which combination of speech features are the most accurate for diagnostic and monitoring purposes. Cummins et al [59] have called for greater research collaboration and cooperation in order to progress the field, and more recently, Barnett et al [60] have called for a complete and comprehensive data platform to capture the breadth of available sensor data in a meaningful way. Moving in these directions to find valid clinical speech–based markers for mood disorders will help to ensure the ongoing development of this field and mitigate some of the risks and challenges highlighted from this review.

Implications for Practice

The findings in this review suggest there are key opportunities for smartphone-based voice monitoring systems in the assessment and management of mood disorders. By linking the data generated by these monitoring systems, we may be able to deliver interventions at the right time, when care is most useful and crucial for the individual. This would prove beneficial as face-to-face therapeutic interventions are primarily based on retrospective and subjective information, and evidence suggests that mental health disorders can become increasingly difficult to treat the longer it is left untreated. However, there is a need to consider the limitations of the current technology. A review by Dogan et al [28] stated that relying on mental health apps for disorder management and therapy would be placing false trust and confidence in a young technology, and that a broader empirical database is needed regarding effectiveness and potential adverse effects of continuous monitoring of physiological and behavioral data using smartphone devices.

Whilst smartphone-based voice data collection provides a level of objectivity in the detection and monitoring of mood disorders, these data cannot currently be used alone in clinical management—these technological tools should be considered as "add-ons" that support practitioners to detect early signs of relapse and remission.

Although there is still skepticism about the potential of smartphones to provide meaningful data to help detect and monitor mental illness, uncertainties are starting to reduce due to the success of modern machine learning methods [13]. Further research demonstrating whether this can be a robust, cost-effective, and acceptable approach is needed before a clear transition into clinical practice can be made.

Additionally, despite the high prevalence of depression, mental health service access remains suboptimal and there remain gaps between service capacity and the needs of the general population. This is likely to be exacerbated by the increasing psychological distress reported globally [61], which has posed considerable pressures on the health care system. New methods of diagnosing and monitoring mood disorders will not only ameliorate the considerable demand placed on mental health services but also potentially allow wider access to mental health interventions [62].

Limitations

This review has 4 key limitations. First, this review did not focus on the ethical and acceptability aspects of smartphone-based monitoring due to the limited data available on these aspects. This is a key area that future research should focus on as it affects the feasibility, acceptability, adherence, and ultimately uptake of these technologies in practice, and thus are crucial barriers to the successful implementation of smartphone-based monitoring into routine practice. As more data on acceptability are reported, future reviews should focus on this to aid decision makers on the clinical translation of these advances.

Second, restrictions in the search methodology may have resulted in relevant articles being missed, for example, the exclusion of gray literature and broad search terms. This is a common limitation reported in scoping reviews, attributable to the balance between achieving both breadth and depth of analysis within a rapid timeframe [63]. This review was successfully able to map a broad cross-section of the literature and provide a useful synthesis for researchers, engineers, and clinicians to understand the potential and technical feasibility of smartphone use and machine learning within their respective fields. Although a more comprehensive systematic review would provide greater clarity on gaps in the literature (in terms of possibilities of this methodology to differentiate mood states and the accuracy/practicalities/feasibility to implement them in real-world clinical practice), such a review would be less feasible to complete and would quickly be out of date given the rapidly evolving nature of the field. Further to this point, the search string used to identify relevant articles was too broad, as most of the included studies were identified through reference lists and review searching. This could be attributed to the nonstandardized definition of the concept of speech patterns. For instance, while this review refers to "speech features" or "patterns," the term varies across the literature, for example, "vocal cues" [23] "the acoustics of speech," [48] and "voice features" [28]. As the field continues to develop, this concept will need to be homogenized in order to improve the quality of review findings.

Third, the review was limited only to one possible digital measure of mood disorders—voice data. There are other features such as heart rate variability and physical activity that can be used to detect mood changes which were not explored in this review. Speech characteristics is however one of the key symptoms of mood disorders. Yet, speech as a digital domain has received relatively less attention than others. This review synthesizes the current evidence to provide clinicians and researchers a summary of which speech features are measurable and the technical considerations in assessing these, which can be used to inform future software development for voice analysis. There remain information gaps and challenges to enable transition of this technology into clinical practice.

Conclusions

The aim of this review was to synthesize the state of research on voice analysis from smartphones to diagnose and monitor mood disorders. Findings from this synthesis may have implications for the development of speech-based classification systems for smartphones which may allow early identification of behavioral markers of mental health disorders so that health care providers can react early to patients' needs and deliver timely and personalized treatment. While several research groups have started developing smartphone-based tools for the diagnosis and monitoring of mood disorders and have produced promising tests of feasibility, this review highlights that only a small number of systems that are currently available or are in preparation have been subjected to empirical studies. Nonetheless, smartphone-based monitoring of objective data in mood disorders is a rapidly growing approach and a highly innovative research field. This is evident in a number of study protocols stating ambitions to expand and intensify research in the field [64,65]. Although promising, a much larger evidence base is required to fully realize the potential, as well as the risks, of these approaches.

Acknowledgments

We thank Sydney Mingle, a Psychology Student from The Pennsylvania State University, undertaking the Northern Hemisphere Summer Research Scholarship at the University of Auckland, for her contribution to the screening of articles and verification of extracted data.

Conflicts of Interest

AC reports grants and consultancy fees from Janssen-Cilag, consultancy fees from Spoonful of Sugar Ltd, grants from A+ charitable trust (Auckland District Health Board), Maurice and Phyllis Paykel trust, Universitas 21, New Zealand Pharmacy Education Research Fund, Auckland Academic Health Alliance, Asthma UK, Health Research Council, Oakley Mental Health Foundation, Chorus, the University of Auckland, and is the recipient of the Robert Irwin Postdoctoral Fellowship, outside the submitted work.

Multimedia Appendix 1
Detailed search strategy applied to all Databases.
[DOCX File , 13 KB - mhealth_v9i9e24352_app1.docx]

References

1. Baumeister H, Härter M. Prevalence of mental disorders based on general population surveys. Soc Psychiatry Psychiatr Epidemiol 2007 Jul 21;42(7):537-546. [doi: 10.1007/s00127-007-0204-1] [Medline: 17516013]
2. Bonnín CM, Torrent C, Goikolea JM, Reinares M, Solé B, Valentí M, et al. The impact of repeated manic episodes and executive dysfunction on work adjustment in bipolar disorder. Eur Arch Psychiatry Clin Neurosci 2014 Apr 3;264(3):247-254. [doi: 10.1007/s00406-013-0431-2] [Medline: 23912643]
3. Chisholm D, Sweeny K, Sheehan P, Rasmussen B, Smit F, Cuijpers P, et al. Scaling-up treatment of depression and anxiety: a global return on investment analysis. The Lancet Psychiatry 2016 May;3(5):415-424. [doi: 10.1016/s2215-0366(16)30024-4]
4. Whiteford HA, Degenhardt L, Rehm J, Baxter AJ, Ferrari AJ, Erskine HE, et al. Global burden of disease attributable to mental and substance use disorders: findings from the Global Burden of Disease Study 2010. The Lancet 2013 Nov;382(9904):1575-1586. [doi: 10.1016/s0140-6736(13)61611-6]
5. Stafford MR, Jackson H, Mayo-Wilson E, Morrison AP, Kendall T. Early interventions to prevent psychosis: systematic review and meta-analysis. BMJ 2013 Jan 18;346(jan18 1):f185-f185 [FREE Full text] [doi: 10.1136/bmj.f185] [Medline: 23335473]
6. Ben-Zeev D, Young M. Accuracy of hospitalized depressed patients' and healthy controls' retrospective symptom reports: an experience sampling study. Journal of Nervous & Mental Disease 2010;198(4):280-285. [doi: 10.1097/nmd.0b013e3181d6141f]
7. Trull TJ, Ebner-Priemer U. Ambulatory assessment. Annu Rev Clin Psychol 2013 Mar 28;9(1):151-176 [FREE Full text] [doi: 10.1146/annurev-clinpsy-050212-185510] [Medline: 23157450]
8. Komatsu H, Sekine Y, Okamura N, Kanahara N, Okita K, Matsubara S, et al. Effectiveness of Information Technology Aided Relapse Prevention Programme in Schizophrenia excluding the effect of user adherence: a randomized controlled trial. Schizophr Res 2013 Oct;150(1):240-244 [FREE Full text] [doi: 10.1016/j.schres.2013.08.007] [Medline: 23998952]
9. Hisham IN, Townsend G, Gillard S, Debnath B, Sin J. COVID-19: the perfect vector for a mental health epidemic. BJPsych Bull 2020 Jun 01:1-7 [FREE Full text] [doi: 10.1192/bjb.2020.60] [Medline: 32475375]
10. Banerjee D. The impact of Covid-19 pandemic on elderly mental health. Int J Geriatr Psychiatry 2020 Dec;35(12):1466-1467 [FREE Full text] [doi: 10.1002/gps.5320] [Medline: 32364283]
11. Nguyen HC, Nguyen MH, Do BN, Tran CQ, Nguyen TTP, Pham KM, et al. People with Suspected COVID-19 Symptoms Were More Likely Depressed and Had Lower Health-Related Quality of Life: The Potential Benefit of Health Literacy. J Clin Med 2020 Mar 31;9(4):965 [FREE Full text] [doi: 10.3390/jcm9040965] [Medline: 32244415]
12. O'Donoghue B, O'Connor K, Thompson A, McGorry P. The need for early intervention for psychosis to persist throughout the COVID-19 pandemic and beyond. Ir J Psychol Med 2020 May 21:1-6 [FREE Full text] [doi: 10.1017/ipm.2020.56] [Medline: 32434611]
13. Rutledge RB, Chekroud AM, Huys QJ. Machine learning and big data in psychiatry: toward clinical applications. Curr Opin Neurobiol 2019 Apr;55:152-159. [doi: 10.1016/j.conb.2019.02.006] [Medline: 30999271]
14. Torous J, Baker JT. Why Psychiatry Needs Data Science and Data Science Needs Psychiatry: Connecting With Technology. JAMA Psychiatry 2016 Jan 01;73(1):3-4. [doi: 10.1001/jamapsychiatry.2015.2622] [Medline: 26676879]
15. Gaggioli A, Riva G. From mobile mental health to mobile wellbeing: opportunities and challenges. Stud Health Technol Inform 2013;184:141-147. [Medline: 23400146]

16. Mohr DC, Zhang M, Schueller SM. Personal Sensing: Understanding Mental Health Using Ubiquitous Sensors and Machine Learning. Annu Rev Clin Psychol 2017 May 08;13(1):23-47 [FREE Full text] [doi: 10.1146/annurev-clinpsy-032816-044949] [Medline: 28375728]
17. Park C, Lee S, Park H, Baik Y, Park Y, Park Y. Autonomic function, voice, and mood states. Clin Auton Res 2011 Apr 6;21(2):103-110. [doi: 10.1007/s10286-010-0095-1] [Medline: 21057851]
18. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. Br J Psychiatry 1978 Nov 29;133(5):429-435. [doi: 10.1192/bjp.133.5.429] [Medline: 728692]
19. Hamilton M. Hamilton Depression Scale. ECDEU Assessment Manual For Psychopharmacology. Rockville, MD: National Institute of Mental Health; 1976:179-192.
20. Weiner L, Doignon-Camus N, Bertschy G, Giersch A. Thought and language disturbance in bipolar disorder quantified via process-oriented verbal fluency measures. Sci Rep 2019 Oct 03;9(1):14282 [FREE Full text] [doi: 10.1038/s41598-019-50818-5] [Medline: 31582814]
21. Shatte ABR, Hutchinson DM, Teague SJ. Machine learning in mental health: a scoping review of methods and applications. Psychol. Med 2019 Feb 12;49(09):1426-1448. [doi: 10.1017/s0033291719000151]
22. Luxton DD, McCann RA, Bush NE, Mishkind MC, Reger GM. mHealth for mental health: Integrating smartphone technology in behavioral healthcare. Professional Psychology: Research and Practice 2011 Dec;42(6):505-512. [doi: 10.1037/a0024485]
23. Malhi GS, Hamilton A, Morris G, Mannie Z, Das P, Outhred T. The promise of digital mood tracking technologies: are we heading on the right track? Evid Based Ment Health 2017 Nov 30;20(4):102-107. [doi: 10.1136/eb-2017-102757] [Medline: 28855245]
24. Pereira F, Mitchell T, Botvinick M. Machine learning classifiers and fMRI: a tutorial overview. Neuroimage 2009 Mar;45(1 Suppl):S199-S209 [FREE Full text] [doi: 10.1016/j.neuroimage.2008.11.007] [Medline: 19070668]
25. de Oliveira L, Portugal LC, Pereira M, Chase HW, Bertocci M, Stiffler R, et al. Predicting Bipolar Disorder Risk Factors in Distressed Young Adults From Patterns of Brain Activation to Reward: A Machine Learning Approach. Biol Psychiatry Cogn Neurosci Neuroimaging 2019 Aug;4(8):726-733 [FREE Full text] [doi: 10.1016/j.bpsc.2019.04.005] [Medline: 31201147]
26. Ghandeharioun A, Fedor S, Sangermano L. Objective assessment of depressive symptoms with machine learning wearable sensors data. New York, NY: IEEE; 2017 Presented at: International Conference on Affective Computing Intelligent Interaction Workshops (ACII); 2017; San Antonio, TX p. 325-332. [doi: 10.1109/acii.2017.8273620]
27. Tsugawa S, Kikuchi Y, Kishino F. Recognizing Depression from Twitter Activity. New York, NY: Association for Computing Machinery; 2015 Apr Presented at: CHI '15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems; April 2015; Seoul, Republic of Korea p. 3187-3196. [doi: 10.1145/2702123.2702280]
28. Dogan E, Sander C, Wagner X, Hegerl U, Kohls E. Smartphone-Based Monitoring of Objective and Subjective Data in Affective Disorders: Where Are We and Where Are We Going? Systematic Review. J Med Internet Res 2017 Jul 24;19(7):e262 [FREE Full text] [doi: 10.2196/jmir.7006] [Medline: 28739561]
29. Or F, Torous J, Onnela J. High potential but limited evidence: Using voice data from smartphones to monitor and diagnose mood disorders. Psychiatr Rehabil J 2017 Sep;40(3):320-324. [doi: 10.1037/prj0000279] [Medline: 28891659]
30. Areàn PA, Ly KH, Andersson G. Mobile technology for mental health assessment. Dialogues Clin Neurosci 2016 Jul;18(2):163-169. [doi: 10.31887/dcns.2016.18.2/parean]
31. Nicholas M, Roberts E, Popay J. Synthesising research evidence. In: Allen P, Black N, Clarke A, Fulop N, Anderson S, editors. Studying the Organisation and Delivery of Health Services. London, UK: Routledge; 2001:188-220.
32. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal of Social Research Methodology 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
33. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Int J Surg 2010;8(5):336-341 [FREE Full text] [doi: 10.1016/j.ijsu.2010.02.007] [Medline: 20171303]
34. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan-a web and mobile app for systematic reviews. Syst Rev 2016 Dec 05;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
35. Nelson C, Hartling L, Campbell S, Oswald AE. The effects of audience response systems on learning outcomes in health professions education. A BEME systematic review: BEME Guide No. 21. Medical Teacher 2012 May 11;34(6):e386-e405. [doi: 10.3109/0142159x.2012.680938]
36. Sutcliffe K, Olivier S, Richardson M. Describing and analysing studies. In: Gough D, Olivier S, Thomas J, editors. An Introduction to Systematic Reviews (2nd Edition). London, UK: Sage Publications; 2017:123-145.
37. Greenhalgh T, Robert G, Bate P. Diffusion of Innovations in Service Organisations: A Systematic Literature Review. Malden, MA: Blackwell Publishing; 2005.
38. Abdullah S, Matthews M, Frank E, Doherty G, Gay G, Choudhury T. Automatic detection of social rhythms in bipolar disorder. J Am Med Inform Assoc 2016 May;23(3):538-543. [doi: 10.1093/jamia/ocv200] [Medline: 26977102]
39. Faurholt-Jepsen M, Busk J, Frost M, Vinberg M, Christensen EM, Winther O, et al. Voice analysis as an objective state marker in bipolar disorder. Transl Psychiatry 2016 Jul 19;6(7):e856-e856 [FREE Full text] [doi: 10.1038/tp.2016.123] [Medline: 27434490]

40. Grunerbl A, Muaremi A, Osmani V, Bahle G, Ohler S, Troster G, et al. Smartphone-Based Recognition of States and State Changes in Bipolar Disorder Patients. IEEE J. Biomed. Health Inform 2015 Jan;19(1):140-148. [doi: 10.1109/jbhi.2014.2343154]
41. Karam Z, Provost E, Singh S, Montgomery J, Archer C, Harrington G, et al. Ecologically valid long-term mood monitoring of individuals with bipolar disorder using speech. Proc IEEE Int Conf Acoust Speech Signal Process 2014 May;2014:4858-4862 [FREE Full text] [doi: 10.1109/ICASSP.2014.6854525] [Medline: 27630535]
42. Maxhuni A, Muñoz-Meléndez A, Osmani V, Perez H, Mayora O, Morales EF. Classification of bipolar disorder episodes based on analysis of voice and motor activity of patients. Pervasive and Mobile Computing 2016 Sep;31:50-66. [doi: 10.1016/j.pmcj.2016.01.008]
43. Muaremi A, Gravenhorst F, Grünerbl A, Arnrich B, Tröster G. Assessing Bipolar Episodes Using Speech Cues Derived from Phone Calls. In: Cipresso P, Matic A, Lopez G, editors. Pervasive Computing Paradigms for Mental Health. MindCare 2014. Cham, Switzerland: Springer; 2014:103-114.
44. Osmani V. Smartphones in Mental Health: Detecting Depressive and Manic Episodes. IEEE Pervasive Comput 2015 Jul;14(3):10-13. [doi: 10.1109/mprv.2015.54]
45. Place S, Blanch-Hartigan D, Rubin C, Gorrostieta C, Mead C, Kane J, Pentland, et al. Behavioral Indicators on a Mobile Sensing Platform Predict Clinically Validated Psychiatric Symptoms of Mood and Anxiety Disorders. J Med Internet Res 2017 Mar 16;19(3):e75 [FREE Full text] [doi: 10.2196/jmir.6678] [Medline: 28302595]
46. Gideon J, Provost EM, McInnis M. MOOD STATE PREDICTION FROM SPEECH OF VARYING ACOUSTIC QUALITY FOR INDIVIDUALS WITH BIPOLAR DISORDER. Proc IEEE Int Conf Acoust Speech Signal Process 2016 Mar;2016:2359-2363 [FREE Full text] [doi: 10.1109/ICASSP.2016.7472099] [Medline: 27570493]
47. Pan Z, Gui C, Zhang J, Zhu J, Cui D. Detecting Manic State of Bipolar Disorder Based on Support Vector Machine and Gaussian Mixture Model Using Spontaneous Speech. Psychiatry Investig 2018 Jul;15(7):695-700 [FREE Full text] [doi: 10.30773/pi.2017.12.15] [Medline: 29969852]
48. Gideon J, Matton K, Anderau S. When to intervene: predicting individual outcomes for rapid intervention (PRIORI) using passively collected speech from smartphone conversations. Bipolar Disorders. Hoboken, NJ: Wiley; 2020. URL: https://onlinelibrary.wiley.com/doi/full/10.1111/bdi.12934 [accessed 2021-08-18]
49. Guidi A, Salvi S, Ottaviano M, Gentili C, Bertschy G, de Rossi D, et al. Smartphone Application for the Analysis of Prosodic Features in Running Speech with a Focus on Bipolar Disorders: System Performance Evaluation and Case Study. Sensors (Basel) 2015 Nov 06;15(11):28070-28087 [FREE Full text] [doi: 10.3390/s151128070] [Medline: 26561811]
50. Dickerson RF, Gorlin EI, Stankovic JA. Empath: A Continuous Remote Emotional Health Monitoring System for Depressive Illness. 2011. URL: https://www.cs.virginia.edu/~stankovic/psfiles/robempath.pdf [accessed 2021-08-18]
51. Taylor MA, Reed R, Berenbaum S. Patterns of speech disorders in schizophrenia and mania. J Nerv Ment Dis 1994 Jun;182(6):319-326. [doi: 10.1097/00005053-199406000-00002] [Medline: 8201303]
52. Low LA, Maddage NC, Lech M, Sheeber L, Allen N. Influence of acoustic low-level descriptors in the detection of clinical depression in adolescents. New York, NY: IEEE; 2010 Presented at: 2010 IEEE International Conference on Acoustics, Speech and Signal Processing; March 14-19, 2010; Dallas, TX p. 5154-5157. [doi: 10.1109/ICASSP.2010.5495018]
53. Naqa I, Murphy M. What is machine learning? In: Naqa IEI, Ruijang L, Murphy MJ, editors. Machine Learning in Radiation Oncology. Berlin, Germany: Springer; 2015:3-11.
54. Ventola CL. Mobile devices and apps for health care professionals: uses and benefits. P T 2014 May;39(5):356-364 [FREE Full text] [Medline: 24883008]
55. Kardous CA, Shaw PB. Evaluation of smartphone sound measurement applications. J Acoust Soc Am 2014 Apr;135(4):EL186-EL192 [FREE Full text] [doi: 10.1121/1.4865269] [Medline: 25236152]
56. Alghowinem S. From Joyous to Clinically Depressed: Mood Detection Using Multimodal Analysis of a Person's Appearance and Speech. New York, NY: IEEE; 2013 Presented at: 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction; September 2-5, 2013; Geneva, Switzerland p. 648-654. [doi: 10.1109/ACII.2013.113]
57. Moore E, Clements M, Peifer J, Weisser L. Critical Analysis of the Impact of Glottal Features in the Classification of Clinical Depression in Speech. IEEE Trans. Biomed. Eng 2008 Jan;55(1):96-107. [doi: 10.1109/tbme.2007.900562]
58. Mitra V, Tsiartas A, Shriberg E. Noise and reverberation effects on depression detection from speech. New York, NY: IEEE; 2016 Mar Presented at: IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2016); March 20-25, 2016; Shanghai, China URL: https://www.researchgate.net/publication/294998720_Noise_and_Reverberation_Effects_on_Depression_Detection_from_Speech [doi: 10.1109/ICASSP.2016.7472788]
59. Cummins N, Scherer S, Krajewski J, Schnieder S, Epps J, Quatieri TF. A review of depression and suicide risk assessment using speech analysis. Speech Communication 2015 Jul;71:10-49. [doi: 10.1016/j.specom.2015.03.004]
60. Barnett S, Huckvale K, Christensen H, Venkatesh S, Mouzakis K, Vasa R. Intelligent Sensing to Inform and Learn (InSTIL): A Scalable and Governance-Aware Platform for Universal, Smartphone-Based Digital Phenotyping for Research and Clinical Applications. J Med Internet Res 2019 Nov 06;21(11):e16399 [FREE Full text] [doi: 10.2196/16399] [Medline: 31692450]

61. Butterworth P, Watson N, Wooden M. Trends in the Prevalence of Psychological Distress Over Time: Comparing Results From Longitudinal and Repeated Cross-Sectional Surveys. Front Psychiatry 2020 Nov 26;11:595696 [FREE Full text] [doi: 10.3389/fpsyt.2020.595696] [Medline: 33324261]
62. Amsalem D, Dixon L, Neria Y. The Coronavirus Disease 2019 (COVID-19) Outbreak and Mental Health: Current Risks and Recommended Actions. JAMA Psychiatry 2021 Jan 01;78(1):9-10. [doi: 10.1001/jamapsychiatry.2020.1730] [Medline: 32579160]
63. Pham MT, Rajić A, Greig JD, Sargeant JM, Papadopoulos A, McEwen SA. A scoping review of scoping reviews: advancing the approach and enhancing the consistency. Res Synth Methods 2014 Dec 24;5(4):371-385 [FREE Full text] [doi: 10.1002/jrsm.1123] [Medline: 26052958]
64. Tonn P, Degani Y, Hershko S, Klein A, Seule L, Schulze N. Development of a Digital Content-Free Speech Analysis Tool for the Measurement of Mental Health and Follow-Up for Mental Disorders: Protocol for a Case-Control Study. JMIR Res Protoc 2020 May 14;9(5):e13852 [FREE Full text] [doi: 10.2196/13852] [Medline: 32406862]
65. Matcham F, Barattieri di San Pietro C, Bulgari V, de Girolamo G, Dobson R, Eriksson H, RADAR-CNS consortium. Remote assessment of disease and relapse in major depressive disorder (RADAR-MDD): a multi-centre prospective cohort study protocol. BMC Psychiatry 2019 Feb 18;19(1):72 [FREE Full text] [doi: 10.1186/s12888-019-2049-z] [Medline: 30777041]

Abbreviations

BD: bipolar disorder
EMA: ecological momentary assessment
GFCC: gammatone frequency cepstral coefficient
GMM: Gaussian mixture model
HNR: harmonic-to-noise ratio
IVR: interactive voice response
LPCC: linear prediction cepstral coefficient
MFCC: Mel-frequency cepstral coefficients
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PTSD: post-traumatic stress disorder
SC: spectral centroid
SD: secure digital
SF: spectral flux
SVM: support vector machine
TEO: Teager energy operation

Edited by L Buis; submitted 15.09.20; peer-reviewed by D Hidalgo-Mazzei, S Al-Arkee; comments to author 10.11.20; revised version received 04.01.21; accepted 23.07.21; published 17.09.21.

Please cite as:
Flanagan O, Chan A, Roop P, Sundram F
Using Acoustic Speech Patterns From Smartphones to Investigate Mood Disorders: Scoping Review
JMIR Mhealth Uhealth 2021;9(9):e24352
URL: https://mhealth.jmir.org/2021/9/e24352
doi:10.2196/24352
PMID:34533465

Review

Effectiveness of Text Messaging Interventions on Blood Pressure Control Among Patients With Hypertension: Systematic Review of Randomized Controlled Trials

Hon Lon Tam[1,2*], MSc; Eliza Mi Ling Wong[3*], PhD; Kin Cheung[1], PhD; Siu Fung Chung, PhD

[1]School of Nursing, Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[2]Kiang Wu Nursing College of Macau, Macau, Macao

[3]School of Nursing, Tung Wah College, Hong Kong, China (Hong Kong)

[*]these authors contributed equally

Corresponding Author:
Hon Lon Tam, MSc
School of Nursing
Hong Kong Polytechnic University
Hung Hom, Kowloon
Hong Kong
China (Hong Kong)
Phone: 852 34003804
Email: hon-lon.tam@connect.polyu.hk

Abstract

Background: Controlling blood pressure (BP) is an international health concern, and high BP is a major contributor to cardiovascular disease mortality. Evidence has shown that educational interventions directed at patients potentially improve BP control and adherence to medications and lifestyle modifications. In addition, a text messaging intervention has a potential effect on BP control; however, the dosage of a text messaging intervention has not been determined in previous reviews, resulting in difficult application in practice.

Objective: This review aimed to identify the effectiveness of a text messaging intervention on hypertension management with a specific focus on the dosage of text messaging and the type of additional interventions with text messaging.

Methods: A systematic review was conducted and reported on in accordance with PRISMA guideline. Participants were aged 18 years and older and diagnosed with primary hypertension. The included studies used text messaging as a component of the intervention. We searched for randomized controlled trials published until June 30, 2020, from the following health-related electronic databases: Embase, Medline, CINAHL Complete, PsycINFO, and Scopus. Data were extracted for qualitative synthesis and meta-analysis. The Physiotherapy Evidence Database Scale was used to assess the methodological quality of each study, and the quality of the included studies was assessed independently by two authors.

Results: Twelve studies met the inclusion criteria. The overall methodological quality was fair (mean score 5.75). The frequency of text message delivery varied from daily to biweekly. Health education was identified in 4 studies as an additional intervention with text messaging. The overall results showed that the text messaging intervention significantly reduced systolic BP (SBP) but not diastolic BP (DBP). There was no significant difference in BP reduction between studies that lasted 6 months or less and those that lasted more than 7 months. Seven studies that lasted 6 months or less involving 1428 patients with hypertension were pooled for further meta-analysis. Text messages delivered at a lower frequency (once per week or less) had a small effect on SBP reduction (effect size 0.35, $P<.01$) and DBP reduction (effect size 0.28, $P=.01$). In addition, the use of a text messaging intervention halved the odds of uncontrolled BP among patients with hypertension in 6 months (odds ratio 0.46, $P=.02$).

Conclusions: This review found that a text messaging intervention was effective in BP control. One-way text messaging delivered in a weekly manner was suggested to be effective and required fewer resources. Future studies should use different forms of text message and be integrated into other interventions to improve adherence behaviors and BP control among patients with hypertension.

(JMIR Mhealth Uhealth 2021;9(9):e24527) doi:10.2196/24527

KEYWORDS

text messaging; hypertension; blood pressure; mHealth; meta-analysis

Introduction

Background

For a decade, hypertension (HTN) has been the leading risk factor for global disease burden [1]. Over 1 billion people are estimated to have HTN worldwide [2], which necessitates the development of national and international guidelines that provide scientific evidence for the control of blood pressure (BP) [3-8]. Persistent adherence to medications and lifestyle modifications is emphasized in the guidelines to control BP effectively; however, a low adherence rate has been noted in reviews and studies. Two systematic reviews have revealed that medication adherence among patients with HTN is 55% [9,10]. Regarding adherence to lifestyle modifications, studies have shown that adherence to one component of lifestyle modifications ranges from 14% to 85% [11-14], while only 1.7% to 23.6% of patients with HTN adhered to all components of lifestyle modifications: smoking cessation, limited alcohol consumption, regular exercise, maintenance of optimal body weight, and healthy diet [14,15].

Different interventions have been developed such as self BP monitoring, educational interventions, and health professional–led care to improve adherence to HTN management. Reviews suggested that educational interventions directed at patients not only improved BP significantly [16] but also had a significant effect on adherence to medications and lifestyle modifications [17,18]. The reviews further indicated that adding digital components such as text messaging potentially enhances the effect of educational interventions, resulting in improved BP control.

Text Messaging Intervention

Systematic reviews of the effect of text messaging intervention on HTN management were not lacking; nonetheless, the results were inconclusive. Vargas and colleagues [19] included both quasi-experimental and randomized controlled trials (RCTs) in the review. They searched for articles published until July 2015, and 6 studies were included in the meta-analysis. Although the review revealed that the use of 2-month to 12-month text messaging interventions potentially decreases both systolic blood pressure (SBP) and diastolic blood pressure (DBP), the effect size was not determined because of insufficient data and high heterogeneity was noted because of different study designs [19]. In another systematic review, Islam et al [20] included RCTs published between January 1990 and July 2016, and the interventions lasted at least 6 months, with 70% of participants with cardiovascular diseases completing the study. Nine studies were included in the meta-analysis, and the results showed that interventions using a 6-month to 12-month text messaging intervention significantly reduced SBP and DBP among patients with cardiovascular diseases; however, high heterogeneity was observed as different dosages of text messaging intervention were pooled for meta-analysis [20]. Kassavou and Sutton [21] reviewed the use of voice messaging and text messaging interventions on medication adherence among patients with cardiometabolic diseases. They searched for RCTs published between January 1992 and April 2016. Of the 17 included studies, 9 used text messaging. The results of the meta-analysis indicated that the 25-day to 12-month interventions could improve medication adherence significantly; however, high heterogeneity was noted again because of wide variety of interventions used with text messaging. In addition, the authors did not differentiate the effects of text messaging from those of voice messaging on medication adherence [21]. Although the use of a text messaging intervention potentially improves BP control and medication adherence, the frequency of using a text messaging intervention varies from multiple messages per day to fewer than one message per week [19-21]. A review suggested that the use of 2-way text messaging, which required the participants to reply to the received text message, could improve BP control [19]. However, this suggestion was not based on the effect of pooled data. Thus, the directionality of effective text messaging remains inconclusive.

Other than a text messaging intervention, the use of smartphone apps and websites was reviewed and found to have a significant effect on BP control [22]. The interventions required participants to download a HTN smartphone app or access specific websites for HTN management. The ownership of a smartphone is the basic requirement for smartphone apps, and internet access is a prerequisite for both the apps and websites. However, concerns regarding the required technological competency and data protection have been raised [23]. Alternatively, a text messaging intervention can be delivered to recipients via a telecommunication network without any specific apps or internet access. Text message is compatible with being delivered and received between mobile phones and smartphones. Thus, the use of a test messaging intervention potentially covers more people than that of smartphone apps or websites.

Research Gap

In summary, a text messaging intervention may improve HTN management; however, effective additional interventions with text messages have not been identified in previous reviews. A significant effect on BP reduction and medication adherence on HTN-related diseases could be noted if the intervention lasted more than 6 months, but the effectiveness of text messaging interventions lasting 6 months or less was unclear. Also, the dosage of text messaging in terms of frequency and directionality was inconclusive in previous reviews. Regarding the inclusion and exclusion criteria of previous reviews, the included studies were published until July 2016, which might not have reflected the increased use of mobile phones in recent years. As a text messaging intervention is simple to use and widely accepted by both mobile phone and smartphone users, it is worth including recent evidence to review the use of text messaging in HTN management.

Aims

This review aimed to identify the effectiveness of a text messaging intervention in HTN management, with a specific focus on the dosage of text messaging interventions lasting 6

months or less in terms of frequency and directionality and the type of additional interventions with text messages.

Methods

Study Design

A systematic review and meta-analysis were conducted and reported in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) statement [24], and the search was guided by the PRISMA framework:

- Participants: adults with HTN
- Interventions: text messaging
- Comparisons: standard care or usual care
- Outcomes: BP
- Study design: RCT

Search Strategy

Keywords used in the search included hypertension, high blood pressure, adult, text messaging, text message, sms, short message service, and texting. The electronic databases related to health sciences used were Embase, Medline, CINAHL Complete (via EbscoHost), PsycINFO, and Scopus. Since the number of included studies in the previous review of the use of text messaging interventions was small [19-21], we searched for articles available in the databases until June 30, 2020 for screening. A manual search was conducted to identify potentially eligible studies from the reference lists of previous reviews [19-21].

Selection Criteria

Participants

Adults aged 18 years and over who had been diagnosed with primary HTN were included. Patients diagnosed with secondary HTN or pregnancy-related HTN were excluded as they required different management strategies [3,8].

Interventions

Studies that used text messaging as a single or combined intervention were included. Studies using any specific HTN smartphone apps were excluded.

Comparisons and Design

Standard care or usual care provided to the control group was set as a comparator. For the 3-arm RCTs, a comparison was made between the control group and the text messaging intervention group. The comparison between the text messaging intervention group and other active intervention group was not conducted. The study design was limited to RCT to yield reliable evidence on the effectiveness of the interventions. All unpublished theses, conference papers, and non-English articles were excluded.

Outcome

The mean and standard deviation of both SBP and DBP were used as continuous data for the meta-analysis. The number of participants with uncontrolled BP in each group was used as dichotomous data for the meta-analysis.

Quality Assessment

The methodological quality of the RCTs was assessed using the Physiotherapy Evidence Database (PEDro) scale. The scale was designed to assess the quality of an RCT and assesses 11 items including randomization, allocation concealment, blinding, treatment of data, and dropout rate [25]. Among the items, 1 point was awarded to each item if the criteria were fulfilled; 10 of the items were rated, resulting in a total score range of 0-10. The total score <4 was considered low quality, which affected the applicability of the evidence [26]. A cutoff of 4 was used to determine the selection of each study. The quality of the RCTs was assessed independently by the first and fourth authors, and the disagreements were discussed and resolved by the second author.

Data Synthesis and Meta-analysis

Data from eligible studies were extracted into a form containing the following information: authors, year of publication, type of RCT (2-arm or 3-arm), guiding theory, age of participants, setting, dosage of text messaging intervention, and outcomes.

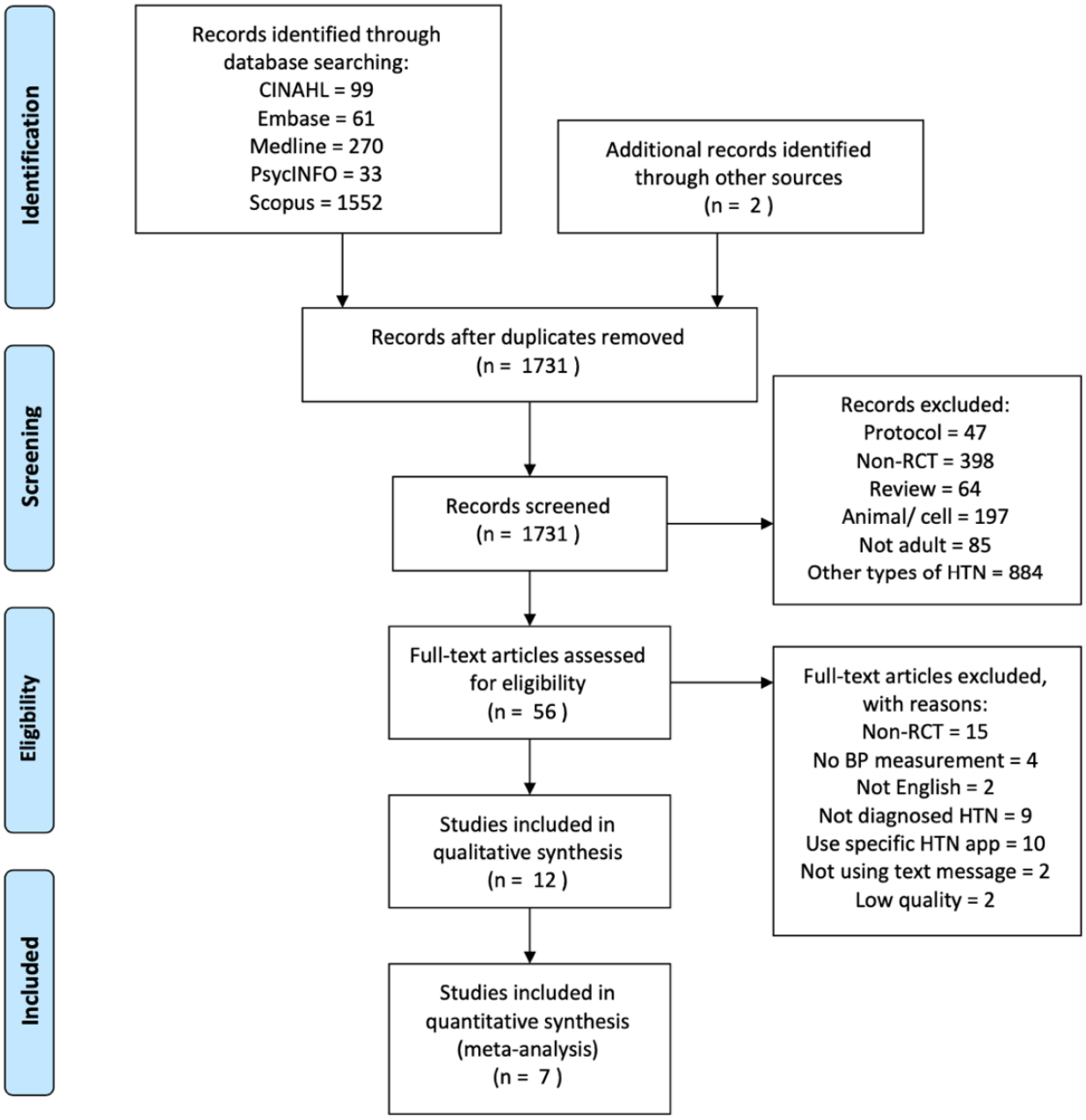
Review Manager 5.4 (Cochrane Collaboration) was used in the meta-analysis to extract data regarding changes in SBP and DBP between baseline and final assessment. In accordance with the aim of this review, 3 subgroup analyses were conducted according to directionality, frequency, and combined interventions. Standardized mean difference (SMD) with a 95% confidence interval was used for the pooled effect of continuous data, SBP, and DBP. The I^2 statistic was used to detect heterogeneity. A random effects model was used as it allows for different true effect sizes across studies [27]. SMDs equal to 0.2, 0.5, and 0.8 represented small, moderate, and large effects, respectively [28]. Regarding the dichotomous data on BP control, the number of participants with uncontrolled BP and total sample size in each of the intervention and comparator groups were used to determine the effect size and reported as odds ratio (OR). A sensitivity analysis was conducted to explore possible differences between the intention-to-treat analysis and others.

Results

Search Results

A total of 1731 articles were initially identified from database searches, and 2 additional articles were identified in the manual search, as shown in Figure 1. A web-based application for systematic reviews, Rayyan, was used to remove duplicates and screen titles and abstracts [29]. Among the 56 articles assessed in full text, 2 were excluded due to low quality (PEDro score <4). The reasons for excluding those studies were that they did not conceal the group allocation process from the participants, therapists, and assessors; furthermore, only the participants from the intervention group withdrew from the study, and the data management and analysis methods were not stated clearly. Finally, 12 studies met the selection criteria. The characteristics of the studies and dosage of the text messaging intervention are described below.

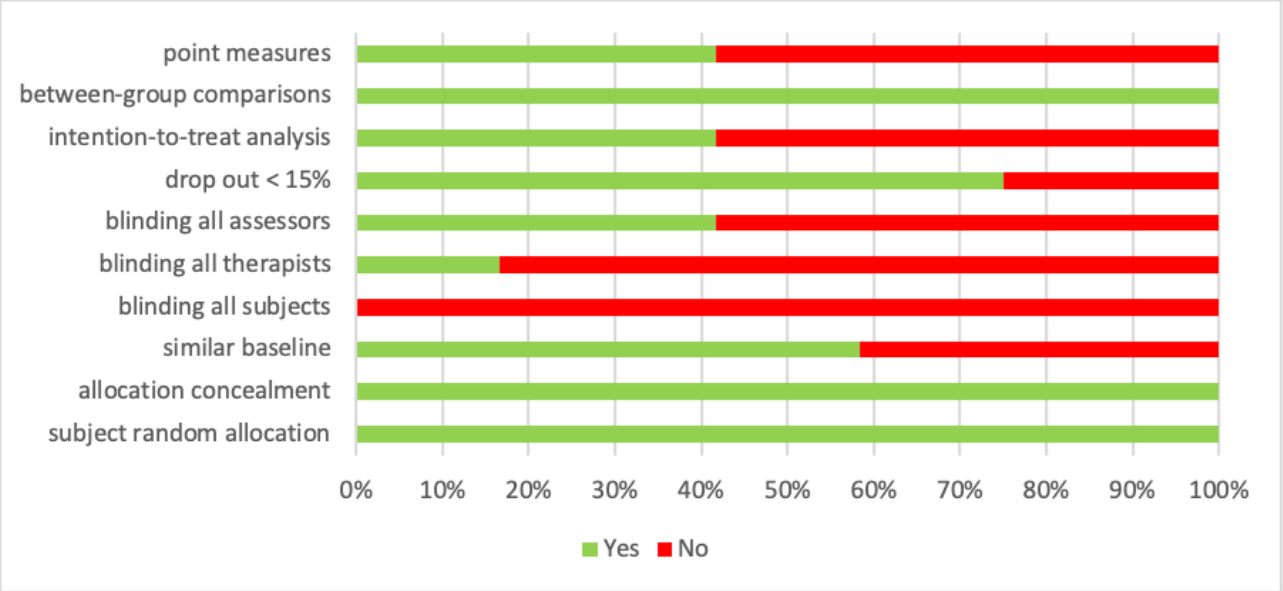
Figure 1. Flow diagram of literature selection process. BP: blood pressure; HTN: hypertension; RCT: randomized controlled trial.



Quality of Studies

Of the 12 included studies, the overall methodological quality was fair, and the PEDro score ranged from 4 to 8, with a mean score of 5.75 (Table 1). None of the studies was able to fulfill the participant-blinding criterion due to the nature of the text messaging intervention (Figure 2). Participants automatically knew their group allocation because no text messages were delivered to the control group. In addition, blinding of therapists was lacking in most studies in that the therapists knew the participant group allocation when providing care. The assessors in some studies knew the group allocation during follow-up data collection. Regarding the treatment of data, intention-to-treat was not followed in some studies, and clinical significance in each group was not assessed.

Figure 2. Methodological quality as percentage across all included studies.



Characteristics of Studies

The included studies were published between 2004 and 2020; 9 were published after 2016 [30-38]. The characteristics of the included studies are summarized in Table 1. Of the 12 included studies, 2 were 3-arm RCTs [34,39], and 4 studies provided the interventions lasted more than 7 months [31,36,37,39]. The number of participants in each study varied from 67 to 1432, with a mean age of 58.97 years. Participants mean age was over 60 years in 3 studies [32,38,40]. Four studies were conducted in the United States; 2 each in China and Spain; and one each in Argentina, South Africa, Chile, and Finland.

Regarding the additional interventions with text messaging, 4 studies integrated face-to-face health education with text messaging intervention to reinforce the effect on HTN management [31,33,38,41]. Only 4 studies were guided by theory, 2 of which used the health belief model [30,33], with the social cognitive theory and information-motivation-behavioral skills model used in 2 other studies [32,37].

Table 1. Characteristics of included studies.

Author	Duration of intervention (study design)	Frequency (directionality)	Theory	Additional intervention	PEDro[a] score
Márquez Contreras et al [41]	6 m[b] (2-arm)	2/w[c] (1-way)	No	Health education, print materials	5
Carrasco et al [40]	6 m (2-arm)	4/w (1-way)	No	No	6
Bobrow et al [39]	12 m (3-arm)	1/w (1-way & 2-way)	No	No	8
Buis et al [30]	1 m (2-arm)	1/d[d] (1-way)	Health belief model	No	4
He et al [31]	18 m (2-arm)	1/w (1-way)	No	Health education	6
Varleta et al [32]	6 m (2-arm)	1/14 days (1-way)	Social cognitive theory	No	5
Wan et al [33]	3 m (2-arm)	1/w (1-way)	Health belief model	Health education, phone call, booklet	7
Mehta et al [34]	4 m (3-arm)	1/d (2-way)	No	No	7
Zahr et al [35]	6 m (2-arm)	2/d (2-way)	No	No	4
Schroeder et al [36]	12 m (2-arm)	1/w (1-way)	No	No	7
Tahkola et al [37]	12 m (2-arm)	Varied (1-way)	Information-motivation-behavioral skills model	No	5
Zhai et al [38]	3 m (2-arm)	1/3 d (1-way)	No	Health education	5

[a]PEDro: Physiotherapy Evidence Database.

[b]m: month.

[c]w: week.

[d]d: day.

Dosage of Text Messaging Interventions

The frequency of delivering text messages to patients with HTN varied from daily to biweekly among the included studies. Only one study used an individualized frequency to deliver text messages to the participants [37]. Text messages were delivered every day for the first 2 weeks, after which the frequency decreased from the third week onward. The first telephone follow-up was conducted in the fourth week to evaluate the participants' BP levels. The frequency of text messages would increase thereafter if BP was not controlled at the follow-up. In addition, He et al [31] and Tahkola et al [37] tailored the content of text messages in accordance with the barriers identified in the assessment and follow-up.

Regarding the directionality, 2-way text messaging was seldom used in the studies. Zahr et al [35] required the participants to report SBP and DBP via the text messages. In the Mehta et al [34] 3-arm RCT, one intervention arm received an electronic pill box that was connected to an internet platform to monitor the use of medication, while another intervention arm required the participants to respond to text messages if they had taken the medication. The Bobrow et al [39] study was a 3-arm RCT that compared the effects of 1-way and 2-way text messaging on HTN management with the control group. Participants replied to text messages to change the medical appointments, delivery times, and language of the text message delivery. In comparison with the control group, the results of the study by Bobrow et al [39] showed that the use of 1-way text messaging decreased SBP significantly, whereas the decrease in SBP in the 2-way text messaging arm failed to achieve statistical significance. Bobrow et al [39] suggested that the nonsignificant findings might have been related to the older age of the participants or their experience in using technology.

Other Outcomes

In addition to BP measurement, medication adherence was commonly assessed (Table 2). However, the methods of assessment differed among the studies. Six studies used self-reported measures to assess the medication adherence [30-33,36,38]. One study compared medication adherence between electronic pill boxes and text message responses [34] and another used the pill count [39]. Among the studies using self-reported measures, the Morisky Medication Adherence Scale was used in 3 studies. Most studies demonstrated an improvement in medication adherence after the use of text messaging. However, the diverse assessment methods resulted in difficulty in comparing medication adherence between studies. It is noteworthy that only one study used a self-reported approach to assess the adherence to lifestyle modifications among the patients with HTN [33].

Table 2. Outcome measures of included studies.

Author	Blood pressure	Medication adherence	Adherence to lifestyle modifications	Appointment adherence
Márquez Contreras et al [41]	+[a]	N[b]	N	–[c]
Carrasco et al [40]	–	N	N	N
Bobrow et al [39]	+	–	N	N
Buis et al [30]	–	–	N	N
He et al [31]	+	+	N	N
Varleta et al [32]	–	+	N	N
Wan et al [33]	+	+	+	N
Mehta et al [34]	–	–	N	N
Zahr et al [35]	–	N	N	N
Schroeder et al [36]	–	–	N	–
Tahkola et al [37]	–	N	N	N
Zhai et al [38]	+	+	N	N

[a]+: significant differences between intervention group (text messaging alone or with additional interventions) and control group (standard care or usual care; $P<.05$).

[b]N: not measured.

[c]–: nonsignificant differences between intervention and control groups.

Meta-analysis

Four studies used interventions that lasted 12 months or more, while the interventions in 8 studies lasted 6 months or less. The corresponding authors of 2 studies were contacted via email to obtain unreported data, but only 1 of them provided the requested data [32]. The study from which unreported data could not be obtained was excluded from the meta-analysis. In the Bobrow et al [39] 3-arm RCT, text messaging was used in 2 intervention groups, and 2 comparisons with control group were extracted for meta-analysis. In the Mehta et al [34] 3-arm RCT, text messaging was used in one intervention group, while an electronic pill box was used in another intervention group. Thus, only one comparison was performed. Although some long-term studies, such as those by Bobrow et al [39] and He et al [31], provided midprocess data at 6 months, the interventions were not delivered to completion, and the 6-month data failed to reflect the holistic effect of the interventions. The midprocess data at 6 months was not extracted for meta-analysis. Figures

3 and 4 illustrate the durations of intervention lasting 6 months or less and those lasting more than 7 months made no statistical differences in SBP and DBP reduction (subgroup differences, *P*>.05). The overall results revealed that the text messaging intervention significantly reduced SBP (Figure 3, SMD=.13, *P*=.01) but not DBP (Figure 4, SMD=.06, *P*=.56). To explore the effectiveness of text messaging interventions lasting 6 months or less, we excluded all studies with interventions lasted more than 7 months. Also, the study without obtaining unreported data was excluded. As a result, 7 studies lasting 6 months or less were pooled for meta-analysis in terms of directionality, frequency, and type of intervention.

Figure 3. The effect of study duration on systolic blood pressure reduction.

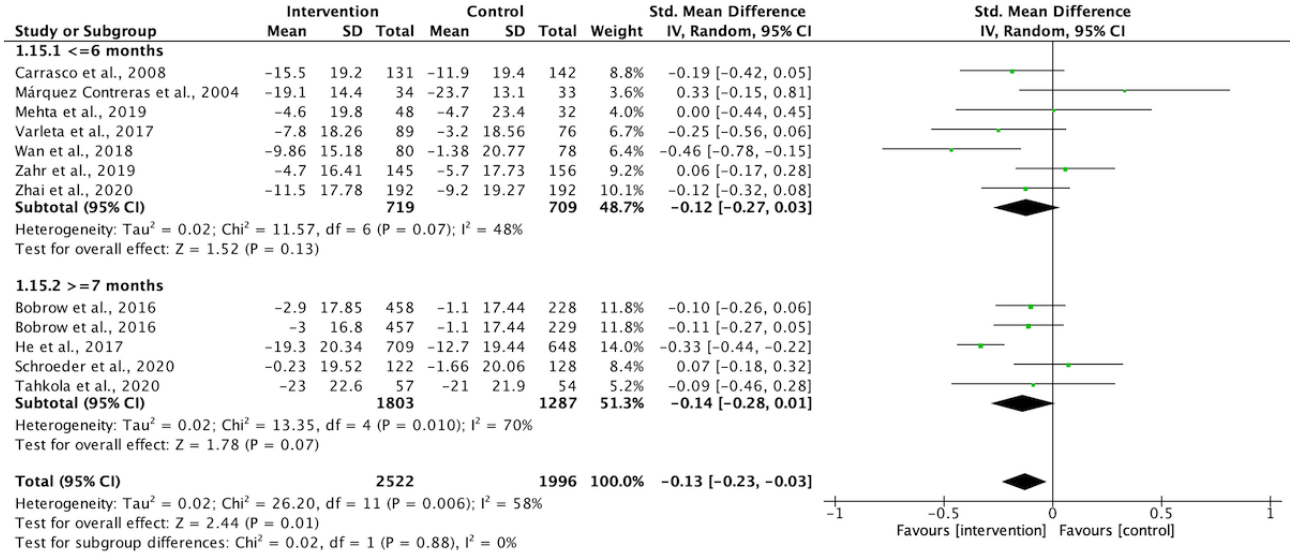
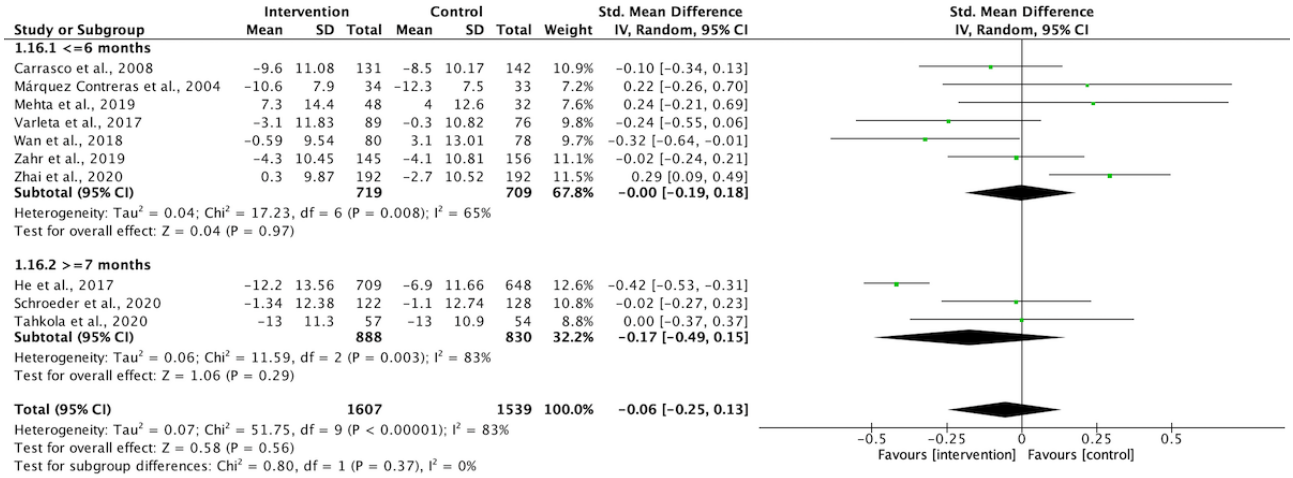


Figure 4. The effect of study duration on diastolic blood pressure reduction.



Tables 3 and 4 show the effectiveness of a text messaging intervention on SBP and DBP reduction with interventions lasted 6 months or less. Regarding the directionality of text messaging, neither 1-way nor 2-way text messaging had a significant effect on SBP or DBP reduction. However, a small effect on SBP and DBP reduction was noted when text messages were delivered in a weekly manner (ie, 1 text message per week). However, the use of a text messaging intervention alone or in combination with health education did not significantly affect SBP or DBP reduction. Six studies reported the number of participants with uncontrolled BP (SBP ≥140 mm Hg or DBP ≥90 mm Hg) at the end of the study. The data were pooled as shown in Figure 5, and the use of a text messaging intervention helped patients with HTN achieve a controlled BP with SBP <140 mm Hg and DBP <90 mm Hg (OR 0.46, *P*=.02).

Table 3. Effectiveness of text messaging on systolic blood pressure reduction with interventions lasting 6 months or less.

Subgroup analysis	Number of studies	Effect size (95% CI)	Heterogeneity I^2 (%)	*P* value	Significance of subgroup differences (*P* value)
Directionality of text messaging					.10
1-way	5	0.18 (0.00, 0.36)	49	.05	—[a]
2-way	2	–0.05 (–0.25, 0.15)	0	.65	—
Frequency of text messaging					.02
>1 per week	5	0.04 (–0.09, 0.18)	22	.53	—
≤1 per week	2	0.35 (0.13, 0.57)	0	.002	—
Type of intervention					.87
With health education	3	0.13 (–0.23, 0.48)	74	.49	—
Text messaging only	4	0.09 (–0.06, 0.24)	14	.23	—

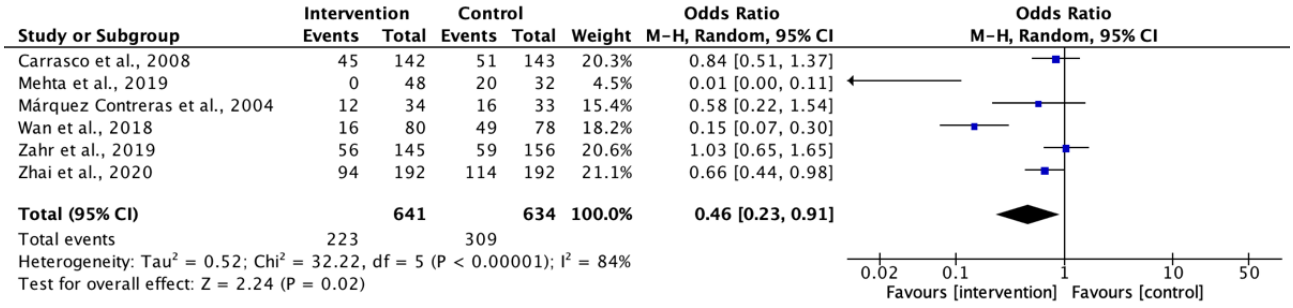
[a]Not applicable.

Table 4. Effectiveness of text messaging on diastolic blood pressure reduction with interventions lasting 6 months or less.

Subgroup analysis	Number of studies	Effect size (95% CI)	Heterogeneity I^2 (%)	*P* value	Significance of subgroup differences (*P* value)
Directionality of text messaging					.58
1-way	5	0.03 (–0.22, 0.29)	75	.79	—[a]
2-way	2	–0.03 (–0.24, 0.17)	0	.75	—
Frequency of text messaging					.01
>1 per week	5	–0.10 (–0.28, 0.08)	50	.27	—
≤1 per week	2	0.28 (0.06, 0.50)	0	.01	—
Type of intervention					.97
With health education	3	–0.06 (–0.48, 0.36)	81	.77	—
Text messaging only	4	0.07 (–0.08, 0.21)	10	.37	—

[a]Not applicable.

Figure 5. Odds ratio of text messaging on blood pressure control.



Sensitivity Analysis

Among the 11 studies with sufficient data pooled in the meta-analysis, 5 used intention-to-treat analysis and the rest did not clearly state their method of analysis. No significant subgroup difference in SBP and DBP reduction was noted between the intention-to-treat analysis and others (Multimedia Appendix 1). Seven studies with text messaging interventions lasted 6 months or less, 2 of which used intention-to-treat analysis. Figure S4 showed a significant subgroup difference in DBP reduction (Multimedia Appendix 1). The pooled results of studies using intention-to-treat analysis showed a favorable effect in the control group, in which usual care reduced DBP more effectively than the text messaging intervention [34,38]. No significant subgroup difference in SBP reduction was noted among interventions that lasted 6 months or less (Multimedia Appendix 1).

Discussion

Summary

This review identified 12 RCTs from 5 electronic databases. The study duration varied from 1 to 18 months. Five studies were published between January 2019 and June 2020, showing that the interest in using text messaging has increased in HTN research. In the meta-analysis, the use of a text messaging intervention significantly reduced SBP but not DBP. Seven studies with a text messaging intervention lasting 6 months or less, involving 1428 patients with HTN, were pooled for further analysis. The results showed that the delivery of weekly text messages significantly improved both SBP and DBP. In addition, the use of a text messaging intervention halved the odds of uncontrolled BP in patients with HTN in 6 months. This review provided information regarding the dosage of a text messaging intervention and type of additional interventions in HTN research using text messages.

Directionality and Frequency

A systematic review of patients with HTN claimed that the use of 2-way text messaging potentially reduces SBP and DBP [19]; however, our findings did not reveal any significant effect relating to the directionality of text messaging on SBP and DBP reduction with an intervention lasting 6 months or less. Noteworthily, participants were required to respond when 2-way text messaging was used. In terms of simple yes/no responses, a session on how to reply to text messages was necessary if specific information, such as BP level, was required. The use of 2-way text messaging may potentially require additional resources. Head et al [42] found that the use of 1-way text messaging significantly increased the likelihood of healthy behaviors. The Bobrow et al [39] 3-arm RCT also showed that the use of 1-way text messaging decreased SBP significantly. Having the same tendency, our findings revealed that 1-way text messaging might potentially reduce SBP (Table 3, SMD=.18, *P*=.05). Therefore, 1-way text messaging might be more feasible and effective for patients with HTN.

Regarding frequency, a decreasing frequency was found to have a moderate effect on health promotion in a review [42]. Since only 1 included study used a decreasing frequency of text messaging on HTN management [37], the effect was not examined in this review. Accordingly, a previous review revealed that the daily or weekly delivery of text messages had a minute effect on promoting health behaviors [42]. Our findings revealed that text messaging could reduce SBP as well as the odds of uncontrolled BP given a target population of patients with HTN and the weekly delivery of text messages.

Types of Intervention

Health education was a commonly used intervention in combination with text messages among the included studies. In this review, a text messaging intervention alone and the combined use of health education had no significant effect on SBP and DBP reduction. The findings contradicted those of a recent review, in which the use of supportive methods, such as text messages and take-home reading materials, with health education yielded a moderately significant effect in improving both SBP and DBP among patients with HTN [18]. In addition, a review by Head et al [42] found that a text messaging intervention alone and the combination of text messages with other interventions could have a significantly small effect on health promotion; however, the duration of their included studies was not limited, and a wide variety of combined interventions was noted such as websites, health education, print materials, pedometers, and daily records of health behavior. Hence, the effectiveness of a text messaging intervention alone or in combination with other interventions to manage HTN remains inconclusive.

On the other hand, tailored content of text messages showed a significant effect on health promotion and medication adherence in previous reviews [21,42]. Two studies in this review were found to tailor their text messages to the participants [31,37], but the interventions lasted 12 to 18 months, resulting in an effect that was not analyzed in the meta-analysis. He et al [31] recruited 1432 patients with HTN in their study and revealed that the tailored content reduced SBP and DBP significantly; however, a significant reduction in SBP and DBP was not noted in the study by Tahkola et al [37]. Another study recruiting people with uncontrolled BP showed that tailored content did not make a significant difference in the reduction of SBP and DBP [43]. The effect of tailored content of text messages on SBP and DBP in patients with HTN was inconclusive. In addition, tailored content requires additional resources, as He et al [31] estimated that the use of tailored content would cost an extra US $6.36 per patient per month. Therefore, tailored content of text messages may not be beneficial to patients with HTN, especially in health care systems with limited resources.

Assessment of Adherence to HTN Management

Medication adherence was assessed in most of the included studies. However, adherence to lifestyle modifications is the core element in HTN management guidelines for patients with HTN and those at risk of HTN [3-8]. The only included study that assessed adherence to lifestyle modifications was conducted by Wan et al [33] in 2018. They revealed that the use of standardized content of text messages combined with health education and leaflet intervention could improve adherence behaviors in 3 months [33]. Another 2-arm RCT recruited 710 patients with coronary heart diseases and showed that adherence to lifestyle modifications improved significantly when tailored content of text messages was delivered 4 times a week for 6 months [44]. These studies suggest that a text messaging intervention may be an effective method to improve adherence to lifestyle modifications. Future studies should examine the effect of text messaging on adherence to both medications and lifestyle modifications among patients with HTN. Regarding the choice of adherence scale, the Treatment Adherence Questionnaire for Patients with Hypertension [45] was suggested in 2 systematic reviews, as it is a comprehensive measure that covers adherence to both medication and lifestyle modifications and is designated for patients with HTN [46,47].

Intention-to-Treat Analysis

A significant difference of DBP reduction in control group was found in the sensitivity analysis (Multimedia Appendix 1). The significant result was from studies that delivered text messages

more frequently (more than once per week) [34,38]. The significant DBP reduction in the control group may indicate information overload by means of high frequency of text messaging [48].

Limitations

Several limitations were noted in this review. We searched for and included only English articles, thus reducing the diversity of the analyzed studies. Only 7 studies were included in the subgroup analysis. Hence, the findings should be interpreted with caution, since the subgroup analyses consisted of a very small number of studies and the variations in dosage of text messaging and combination with health education caused high heterogeneity in the meta-analysis. The decision to only include RCTs in the meta-analysis increased the internal validity of the findings; nonetheless, the external validity might have decreased. The scope of the review was limited to the text messaging, and advanced features added to today's mobile phones or smartwatches to enhance HTN management were not explored.

Implications

We have provided advice for the use of text messaging in HTN management in practice and research. The use of text messaging could help patients with HTN return to controlled BP levels in 6 months or less, and 1-way text messaging could be useful if delivered weekly to patients with HTN. The standardized content of text messages can be stored in a preprogrammed database. Once clients agree to receive text messages, their mobile phone numbers can be entered into the program, thus facilitating the regular delivery of text messages to the clients.

The rapid growth of smartphone ownership and advancing technologies in recent decades has evolved a new mode of delivering text messages. Data from the Pew Research Center showed that smartphone ownership in advanced economies increased from 68% to 76% between 2015 and 2018 [49,50]. Instead of using a traditional short message service, text messages can now be delivered in the form of voice, images, and videos via general smartphone communication apps such as Telegram and WhatsApp. The advanced features of text messaging may enhance adherence behaviors, resulting in an improvement in BP control [51,52]. Hence, future studies should examine the effect of informative images and videos on HTN management, including BP regulation, medication adherence, and adherence to lifestyle modifications. The integration of text messaging with advanced technology should be explored in future studies.

Conclusions

This review updates the information from previous reviews with a focus on patients with HTN. The meta-analysis provides evidence for the use of a text messaging intervention in BP control. As indicated in the meta-analysis, the use of a text messaging intervention lowers the odds of uncontrolled BP. Text messages delivered at a lower frequency had a small effect on the reduction of SBP and DBP. Although the effect of a text messaging intervention on medication adherence was not examined, most included studies showed improvement after the intervention. Thus, text messaging is a potentially useful intervention for HTN management. Regarding the implications of this review, weekly 1-way text messaging is recommended in practice and research. The use of text messages should be incorporated into different interventions in future studies to further improve adherence behaviors and BP control among patients with HTN. There is great potential for professional development in the area of using advanced features of text messages and the more feasible use of these features in delivering messages to clients effectively.

Authors' Contributions

HLT contributed to data search, extraction, and analysis and then drafted and revised this paper. EMLW advised on data analysis, revised the paper, and approved the final manuscript. KC revised the paper and approved the final manuscript. SFC contributed to data search and extraction and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Sensitivity analysis.
[DOC File , 947 KB - mhealth_v9i9e24527_app1.doc]

References

1. GBD 2019 Diseases and Injuries Collaborators. Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 2020 Oct 17;396(10258):1204-1222 [FREE Full text] [doi: 10.1016/S0140-6736(20)30925-9] [Medline: 33069326]
2. World Health Organization. Hypertension. Geneva: World Health Organization URL: https://www.who.int/news-room/fact-sheets/detail/hypertension
3. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart

Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol 2018 May 15;71(19):e127-e248 [FREE Full text] [doi: 10.1016/j.jacc.2017.11.006] [Medline: 29146535]
4. Williams B, Mancia G, Spiering W, Agabiti Rosei E, Azizi M, Burnier M, ESC Scientific Document Group. 2018 ESC/ESH Guidelines for the management of arterial hypertension. Eur Heart J 2018 Sep 01;39(33):3021-3104. [doi: 10.1093/eurheartj/ehy339] [Medline: 30165516]
5. Joint Committee for Guideline Revision. 2018 Chinese Guidelines for Prevention and Treatment of Hypertension: a report of the Revision Committee of Chinese Guidelines for Prevention and Treatment of Hypertension. J Geriatr Cardiol 2019 Mar;16(3):182-241 [FREE Full text] [doi: 10.11909/j.issn.1671-5411.2019.03.014] [Medline: 31080465]
6. Umemura S, Arima H, Arima S, Asayama K, Dohi Y, Hirooka Y, et al. The Japanese Society of Hypertension Guidelines for the Management of Hypertension (JSH 2019). Hypertens Res 2019 Sep;42(9):1235-1481. [doi: 10.1038/s41440-019-0284-9] [Medline: 31375757]
7. Lee H, Shin J, Kim G, Park S, Ihm S, Kim HC, et al. 2018 Korean Society of Hypertension Guidelines for the management of hypertension: part II-diagnosis and treatment of hypertension. Clin Hypertens 2019;25:20 [FREE Full text] [doi: 10.1186/s40885-019-0124-x] [Medline: 31388453]
8. Unger T, Borghi C, Charchar F, Khan NA, Poulter NR, Prabhakaran D, et al. 2020 International Society of Hypertension Global Hypertension Practice Guidelines. Hypertension 2020 Jun;75(6):1334-1357. [doi: 10.1161/HYPERTENSIONAHA.120.15026] [Medline: 32370572]
9. Abegaz TM, Shehab A, Gebreyohannes EA, Bhagavathula AS, Elnour AA. Nonadherence to antihypertensive drugs: a systematic review and meta-analysis. Medicine (Baltimore) 2017 Jan;96(4):e5641 [FREE Full text] [doi: 10.1097/MD.0000000000005641] [Medline: 28121920]
10. Carvalho AS, Santos P. Medication adherence in patients with arterial hypertension: the relationship with healthcare systems' organizational factors. Patient Prefer Adherence 2019;13:1761-1774 [FREE Full text] [doi: 10.2147/PPA.S216091] [Medline: 31802854]
11. Alefan Q, Huwari D, Alshogran O, Jarrah MI. Factors affecting hypertensive patients' compliance with healthy lifestyle. Patient Prefer Adherence 2019;13:577-585 [FREE Full text] [doi: 10.2147/PPA.S198446] [Medline: 31114171]
12. Selçuk KT, Çevik C, Mercan Y, Koca H. Hypertensive patients' adherence to pharmacological and non-pharmacological treatment methods, in Turkey. Int J Community Med Public Health 2017 Jul 22;4(8):2648-2657. [doi: 10.18203/2394-6040.ijcmph20173308]
13. Tibebu A, Bekele DM, Negesa L. Adherence to recommended lifestyle modifications and factors associated for hypertensive patients attending chronic follow-up units of selected public hospitals in Addis Ababa, Ethiopia. Patient Prefer Adherence 2017 Feb;11:323-330. [doi: 10.2147/ppa.s126382] [Medline: 28280305]
14. Fang J, Moore L, Loustalot F, Yang Q, Ayala C. Reporting of adherence to healthy lifestyle behaviors among hypertensive adults in the 50 states and the District of Columbia, 2013. J Am Soc Hypertens 2016 Mar;10(3):252-262.e3 [FREE Full text] [doi: 10.1016/j.jash.2016.01.008] [Medline: 26851000]
15. Andualem A, Gelaye H, Damtie Y. Adherence to lifestyle modifications and associated factors among adult hypertensive patients attending chronic follow-up units of Dessie Referral Hospital, North East Ethiopia, 2020. Integr Blood Press Control 2020;13:145-156 [FREE Full text] [doi: 10.2147/IBPC.S275575] [Medline: 33122940]
16. Glynn L, Murphy AW, Smith SM, Schroeder K, Fahey T. Interventions used to improve control of blood pressure in patients with hypertension. Cochrane Database Syst Rev 2010 Mar 17(3):CD005182. [doi: 10.1002/14651858.CD005182.pub4] [Medline: 20238338]
17. Conn VS, Ruppar TM, Chase JD, Enriquez M, Cooper PS. Interventions to improve medication adherence in hypertensive patients: systematic review and meta-analysis. Curr Hypertens Rep 2015 Dec 11;17(12):94 [FREE Full text] [doi: 10.1007/s11906-015-0606-5] [Medline: 26560139]
18. Tam HL, Wong EML, Cheung K. Effectiveness of educational interventions on adherence to lifestyle modifications among hypertensive patients: an integrative review. Int J Environ Res Public Health 2020 Apr 07;17(7):2513 [FREE Full text] [doi: 10.3390/ijerph17072513] [Medline: 32272601]
19. Vargas G, Cajita MI, Whitehouse E, Han H. Use of short messaging service for hypertension management: a systematic review. J Cardiovasc Nurs 2017;32(3):260-270 [FREE Full text] [doi: 10.1097/JCN.0000000000000336] [Medline: 27111819]
20. Shariful Islam SM, Farmer AJ, Bobrow K, Maddison R, Whittaker R, Pfaeffli Dale LA, et al. Mobile phone text-messaging interventions aimed to prevent cardiovascular diseases (Text2PreventCVD): systematic review and individual patient data meta-analysis. Open Heart 2019;6(2):e001017 [FREE Full text] [doi: 10.1136/openhrt-2019-001017] [Medline: 31673381]
21. Kassavou A, Sutton S. Automated telecommunication interventions to promote adherence to cardio-metabolic medications: meta-analysis of effectiveness and meta-regression of behaviour change techniques. Health Psychol Rev 2017 Sep 12:1-18. [doi: 10.1080/17437199.2017.1365617] [Medline: 28805162]
22. Ma Y, Cheng HY, Cheng L, Sit JW. The effectiveness of electronic health interventions on blood pressure control, self-care behavioural outcomes and psychosocial well-being in patients with hypertension: a systematic review and meta-analysis. Int J Nurs Stud 2019 Apr;92:27-46. [doi: 10.1016/j.ijnurstu.2018.11.007] [Medline: 30690164]

23. Morrissey EC, Casey M, Glynn LG, Walsh JC, Molloy GJ. Smartphone apps for improving medication adherence in hypertension: patients' perspectives. Patient Prefer Adherence 2018;12:813-822 [FREE Full text] [doi: 10.2147/PPA.S145647] [Medline: 29785096]
24. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009 Jul 21;339(1):b2535 [FREE Full text] [doi: 10.1136/bmj.b2535] [Medline: 19622551]
25. Maher C, Sherrington C, Herbert R, Moseley A, Elkins M. Reliability of the PEDro scale for rating quality of randomized controlled trials. Phys Ther 2003 Aug;83(8):713-721. [Medline: 12882612]
26. Cashin AG, McAuley JH. Clinimetrics: Physiotherapy Evidence Database (PEDro) Scale. J Physiother 2020 Jan;66(1):59 [FREE Full text] [doi: 10.1016/j.jphys.2019.08.005] [Medline: 31521549]
27. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med 2002 Jun 15;21(11):1539-1558. [doi: 10.1002/sim.1186] [Medline: 12111919]
28. Faraone SV. Interpreting estimates of treatment effects: implications for managed care. Pharma Therapeutics 2008 Dec;33(12):700-711 [FREE Full text] [Medline: 19750051]
29. Ouzzani M, Hammady H, Fedorowicz Z, Elmagarmid A. Rayyan—a web and mobile app for systematic reviews. Syst Rev 2016 Dec 05;5(1):210 [FREE Full text] [doi: 10.1186/s13643-016-0384-4] [Medline: 27919275]
30. Buis L, Hirzel L, Dawood RM, Dawood KL, Nichols LP, Artinian NT, et al. Text messaging to improve hypertension medication adherence in African Americans from primary care and emergency department settings: results from two randomized feasibility studies. JMIR Mhealth Uhealth 2017 Feb 01;5(2):e9 [FREE Full text] [doi: 10.2196/mhealth.6630] [Medline: 28148474]
31. He J, Irazola V, Mills KT, Poggio R, Beratarrechea A, Dolan J, HCPIA Investigators. Effect of a community health worker-led multicomponent intervention on blood pressure control in low-income patients in Argentina: a randomized clinical trial. JAMA 2017 Sep 19;318(11):1016-1025 [FREE Full text] [doi: 10.1001/jama.2017.11358] [Medline: 28975305]
32. Varleta P, Acevedo M, Akel C, Salinas C, Navarrete C, García A, et al. Mobile phone text messaging improves antihypertensive drug adherence in the community. J Clin Hypertens (Greenwich) 2017 Dec 21;19(12):1276-1284 [FREE Full text] [doi: 10.1111/jch.13098] [Medline: 28941056]
33. Wan LH, Zhang XP, You LM, Ruan HF, Chen SX. The efficacy of a comprehensive reminder system to improve health behaviors and blood pressure control in hypertensive ischemic stroke patients: a randomized controlled trial. J Cardiovasc Nurs 2018;33(6):509-517. [doi: 10.1097/JCN.0000000000000496] [Medline: 29901484]
34. Mehta SJ, Volpp KG, Troxel AB, Day SC, Lim R, Marcus N, et al. Electronic pill bottles or bidirectional text messaging to improve hypertension medication adherence (Way 2 Text): a randomized clinical trial. J Gen Intern Med 2019 Nov 08;34(11):2397-2404 [FREE Full text] [doi: 10.1007/s11606-019-05241-x] [Medline: 31396815]
35. Zahr RS, Anthony CA, Polgreen PM, Simmering JE, Goerdt CJ, Hoth AB, et al. A texting-based blood pressure surveillance intervention. J Clin Hypertens (Greenwich) 2019 Oct 10;21(10):1463-1470 [FREE Full text] [doi: 10.1111/jch.13674] [Medline: 31503408]
36. Schroeder EB, Moore KR, Manson SM, Baldwin MA, Goodrich GK, Malone AS, et al. A randomized clinical trial of an interactive voice response and text message intervention for individuals with hypertension. J Clin Hypertens (Greenwich) 2020 Jul 09;22(7):1228-1238 [FREE Full text] [doi: 10.1111/jch.13909] [Medline: 32516486]
37. Tahkola A, Korhonen P, Kautiainen H, Niiranen T, Mäntyselkä P. Personalized text message and checklist support for initiation of antihypertensive medication: the cluster randomized, controlled check and support trial. Scand J Prim Health Care 2020 Jun 02;38(2):201-209. [doi: 10.1080/02813432.2020.1753380] [Medline: 32362175]
38. Zhai P, Hayat K, Ji W, Li Q, Shi L, Atif N, et al. Efficacy of text messaging and personal consultation by pharmacy students among adults with hypertension: randomized controlled trial. J Med Internet Res 2020 May 20;22(5):e16019 [FREE Full text] [doi: 10.2196/16019] [Medline: 32432556]
39. Bobrow K, Farmer AJ, Springer D, Shanyinde M, Yu L, Brennan T, et al. Mobile phone text messages to support treatment adherence in adults with high blood pressure (SMS-Text Adherence Support [StAR]): a single-blind, randomized trial. Circulation 2016 Feb 09;133(6):592-600 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.017530] [Medline: 26769742]
40. Carrasco MP, Salvador CH, Sagredo PG, Márquez-Montes J, González DMMA, Fragua JA, et al. Impact of patient-general practitioner short-messages-based interaction on the control of hypertension in a follow-up service for low-to-medium risk hypertensive patients: a randomized controlled trial. IEEE Trans Inf Technol Biomed 2008 Nov;12(6):780-791 [FREE Full text] [doi: 10.1109/TITB.2008.926429] [Medline: 19000959]
41. Márquez Contreras E, de la Figuera von Wichmann M, Gil Guillén V, Ylla-Catalá A, Figueras M, Balaña M, et al. [Effectiveness of an intervention to provide information to patients with hypertension as short text messages and reminders sent to their mobile phone (HTA-Alert)]. Aten Primaria 2004 Nov 15;34(8):399-405 [FREE Full text] [doi: 10.1016/s0212-6567(04)78922-2] [Medline: 15546536]
42. Head KJ, Noar SM, Iannarino NT, Grant HN. Efficacy of text messaging-based interventions for health promotion: a meta-analysis. Soc Sci Med 2013 Nov;97:41-48. [doi: 10.1016/j.socscimed.2013.08.003] [Medline: 24161087]

43. Tobe SW, Yeates K, Campbell NRC, Maar MA, Perkins N, Liu PP, et al. Diagnosing hypertension in Indigenous Canadians (DREAM-GLOBAL): a randomized controlled trial to compare the effectiveness of short message service messaging for management of hypertension: Main results. J Clin Hypertens (Greenwich) 2019 Jan 26;21(1):29-36 [FREE Full text] [doi: 10.1111/jch.13434] [Medline: 30474909]
44. Chow CK, Ariyarathna N, Islam SMS, Thiagalingam A, Redfern J. mHealth in cardiovascular health care. Heart Lung Circ 2016 Aug;25(8):802-807. [doi: 10.1016/j.hlc.2016.04.009] [Medline: 27262389]
45. Ma C, Chen S, You L, Luo Z, Xing C. Development and psychometric evaluation of the Treatment Adherence Questionnaire for Patients with Hypertension. J Adv Nurs 2012 Jun;68(6):1402-1413. [doi: 10.1111/j.1365-2648.2011.05835.x] [Medline: 21954893]
46. Pérez-Escamilla B, Franco-Trigo L, Moullin JC, Martinez-Martinez F, Garcia-Corpas JP. Identification of validated questionnaires to measure adherence to pharmacological antihypertensive treatments. Patient Prefer Adherence 2015 Apr:569. [doi: 10.2147/ppa.s76139]
47. Pareja-Martínez E, Esquivel-Prados E, Martínez-Martínez F, García-Corpas JP. Questionnaires on adherence to antihypertensive treatment: a systematic review of published questionnaires and their psychometric properties. Int J Clin Pharm 2020 Apr 05;42(2):355-365. [doi: 10.1007/s11096-020-00981-x] [Medline: 32026354]
48. Liu C, Kuo K. Does information overload prevent chronic patients from reading self-management educational materials? Int J Med Inform 2016 May;89:1-8. [doi: 10.1016/j.ijmedinf.2016.01.012] [Medline: 26980353]
49. Poushter J. Smartphone ownership and internet usage continues to climb in emerging economies. Washington: Pew Internet and American Life Profect; 2016 Feb 22. URL: https://www.pewresearch.org/global/wp-content/uploads/sites/2/2016/02/pew_research_center_global_technology_report_final_february_22__2016.pdf [accessed 2021-01-31]
50. Taylor K, Silver L. Smartphone ownership is growing rapidly around the world, but not always equally. Washington: Pew Internet and American Life Project; 2019 Feb 05. URL: https://www.pewresearch.org/global/wp-content/uploads/sites/2/2019/02/Pew-Research-Center_Global-Technology-Use-2018_2019-02-05.pdf [accessed 2021-01-31]
51. Ng WWM, Wong ASW, Cheung K. A cluster randomized controlled trial feasibility study of a WhatsApp-delivered intervention to promote healthy eating habits in male firefighters. Int J Environ Res Public Health 2021 Jun 20;18(12):6633 [FREE Full text] [doi: 10.3390/ijerph18126633] [Medline: 34203034]
52. Xiong S, Berkhouse H, Schooler M, Pu W, Sun A, Gong E, et al. Effectiveness of mHealth interventions in improving medication adherence among people with hypertension: a systematic review. Curr Hypertens Rep 2018 Aug 07;20(10):86. [doi: 10.1007/s11906-018-0886-7] [Medline: 30088110]

Abbreviations

BP: blood pressure
DBP: diastolic blood pressure
HTN: hypertension
OR: odds ratio
PEDro: Physiotherapy Evidence Database
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analyses
RCT: randomized controlled trial
SBP: systolic blood pressure
SMD: standardized mean difference

Edited by L Buis; submitted 23.09.20; peer-reviewed by R Krukowski, N Agarwal; comments to author 12.11.20; revised version received 04.01.21; accepted 05.08.21; published 22.09.21.

Please cite as:
Tam HL, Wong EML, Cheung K, Chung SF
Effectiveness of Text Messaging Interventions on Blood Pressure Control Among Patients With Hypertension: Systematic Review of Randomized Controlled Trials
JMIR Mhealth Uhealth 2021;9(9):e24527
URL: https://mhealth.jmir.org/2021/9/e24527
doi:10.2196/24527
PMID:34550078

Original Paper

Exploring the Shift in International Trends in Mobile Health Research From 2000 to 2020: Bibliometric Analysis

Jianfei Cao[1], MSc; Yeongjoo Lim[2], PhD; Shintaro Sengoku[3], PhD; Xitong Guo[4], PhD; Kota Kodama[1], PhD

[1]Graduate School of Technology Management, Ritsumeikan University, Ibaraki, Japan

[2]Department of Business Administration, Ritsumeikan University, Ibaraki, Japan

[3]Department of Innovation Science, Tokyo Institute of Technology, Tokyo, Japan

[4]Institute at School of Management, Harbin Institute of Technology, Harbin, China

Corresponding Author:
Kota Kodama, PhD
Graduate School of Technology Management
Ritsumeikan University
Ibaraki
Japan
Phone: 81 0726652448
Email: kkodama@fc.ritsumei.ac.jp

Abstract

Background: Smartphones have become an integral part of our lives with unprecedented popularity and a diverse selection of apps. The continuous upgrading of information technology has also enabled smartphones to display great potential in the field of health care.

Objective: We aimed to determine the future research direction of mobile health (mHealth) by analyzing its research trends and latest research hotspots.

Methods: This study collected mHealth-related literature published between 2000 and 2020 from the Web of Science database. Descriptive statistics of publication trends of mHealth research were determined by analyzing the annual number of publications in the literature and annual number of publications by country. We constructed visualization network maps of country (or regional) collaborations and author-provided keyword co-occurrences, as well as overlay visualization maps of the average publication year of author-provided keywords to analyze the hotspots and research trends in mHealth research.

Results: In total, 12,593 mHealth-related research papers published between 2000 and 2020 were found. The results showed an exponential growth trend in the number of annual publications in mHealth literature. JMIR mHealth and uHealth, the Journal of Medical Internet Research, and JMIR Research Protocols were the 3 top journals with respect to number of publications. The United States remained the leading contributor to the literature in this area (5294/12,593, 42.0%), well ahead of other countries and regions. Other countries and regions also showed a clear trend of annual increases in the number of mHealth publications. The 4 countries with the largest number of publications—the United States, the United Kingdom, Canada, and Australia—were found to cooperate more closely. The rest of the countries and regions showed a clear geographic pattern of cooperation. The keyword co-occurrence analysis of the top 100 authors demonstrated 5 clusters, namely, development of mHealth medical technology and its application to various diseases, use of mHealth technology to improve basic public health and health policy, mHealth self-health testing and management in daily life, adolescent use of mHealth, and mHealth in mental health. The research trends revealed a gradual shift in mHealth research from health policy and improving public health care to the development and social application of mHealth technologies.

Conclusions: To the best of our knowledge, the most current bibliometric analysis dates back to 2016. However, the number of mHealth research published between 2017 and 2020 exceeds the previous total. The results of this study shed light on the latest hotspots and trends in mHealth research. These findings provide a useful overview of the development of the field; they may also serve as a valuable reference and provide guidance for researchers in the digital health field.

(JMIR Mhealth Uhealth 2021;9(9):e31097) doi:10.2196/31097

KEYWORDS

mobile health; digital health; digital medicine; bibliometric analysis; journalology; data visualization; co-occurrence analysis; research trends; mental health; mHealth; paradigm; innovation; smartphone; research; trend; literature; bibliometric; review; app; cooperation; development; public health; health policy; self-management; adolescent

Introduction

In recent years, smartphones have become popular in many countries; especially in high-income countries such as the United Kingdom and the United States—as of September 2019, the smartphone penetration rate is as high as 80% [1]. With the popularity of smartphones, the richness of smartphone app functions and the anytime-anywhere operability provide more opportunities for health promotion, especially in the medical field [2,3]. Services for medical and public health supported by mobile devices is defined as mobile health (mHealth). The outbreak of coronavirus disease 2019 (COVID-19) in 2020 has exposed the lack of medical resources in many countries [4-7]. In this context, mHealth apps can monitor body information, such as heart rate, as well as behavioral information, such as real-time acceleration, through smartphones, smartwatches, and other mobile devices. This can enable people to check their health status at any time and provide medical staff with more reference data [8-10]. Therefore, the development of mHealth can alleviate the shortage of medical resources to a certain extent [11,12]. The great potential shown by mHealth in the medical field has received attention from researchers in many countries [13]. A focus for an increasing number of researchers is to determine how further developments in the mHealth field can reasonably create more social value; therefore, it is necessary to have an in-depth understanding of current research trends and hot spots in mHealth.

Bibliometrics can quantify comprehensive textual information to provide numerical statistics on the development process of a particular topic [14]. The quantified numerical information can also help scholars identify the future trends of a subject [15]. Bibliometrics is widely used in academics, specifically for the in-depth analysis of journal papers [16,17]. Recently, researchers have developed many tools that meet the needs of bibliographic analysis and enrich the bibliographic treatment, such as for the analysis of co-authors' countries (or regions) and research institutions to elucidate the collaboration between different regions or research institutions [18-20], the extraction of keywords for co-occurrence analysis to identify research hotspots [21,22], and keyword clustering to identify the main research directions in a field [23,24]. Thus, bibliometrics plays an important role, both in providing an overview of the past and to provide predictive information.

Currently, there are only a few papers on bibliometric analyses of mHealth literature. Sweileh et al [13] searched Scopus for mHealth papers between 2006 and 2016 and found that most keywords were related to diabetes, medication adherence, and obesity. This study also found an exponential growth in mHealth literature.

Shen et al [25] collected 2704 papers related to mHealth from the Web of Science database as of 2016. Although different from the database searched by Sweileh et al [13], the results of the 2 studies were similar in that both found the United States to be the most active country in mHealth research worldwide; they also showed an exponential growth trend in publications on mHealth in the Web of Science. By identifying the keywords, Shen et al [25] classified the research hotspots in mHealth research into the following 4 main areas: (1) patient engagement and patient intervention, (2) health monitoring and self-care, (3) mobile device and mobile computing, and (4) security and privacy.

Another bibliometric analysis [26] of mHealth literature, published in 2020, focused on papers related to mHealth apps. A total of 2802 papers published between 2000 and 2019 were collected from the Web of Science. The current state of research, research trends, hotspots, and coauthorship networks showed that the United States, England, Australia, and Canada were the most productive countries for mHealth apps research and the hot topics of mHealth apps research formed 5 clusters: (1) technology and system development of mobile health apps, (2) mobile health apps used in mental health, (3) mobile health apps used as mobile health tools in telemedicine, chronic disease, and medication adherence management, (4) mobile health apps used in health behavior and health promotion, and (5) mobile health apps used in disease prevention via the internet.

However, a gap—between 2017 and 2020—in bibliometric analysis of mHealth research remains. Both Sweileh et al [13] and Shen et al [25] found that there was an exponential growth trend up until 2016; therefore, it can be expected that the number would have grown substantially from 2017 to 2020. In fact, the number of publications in the mid-2017 to 2020 period surpasses the previous total. Therefore, a renewed bibliometric analysis of mHealth research from 2000 to 2020 was necessary. The period 2000 to 2020, instead of only 2017 to 2020, was chosen to facilitate the calculation of logical growth curves for publications and the visualization of trends in research hotspots.

Methods

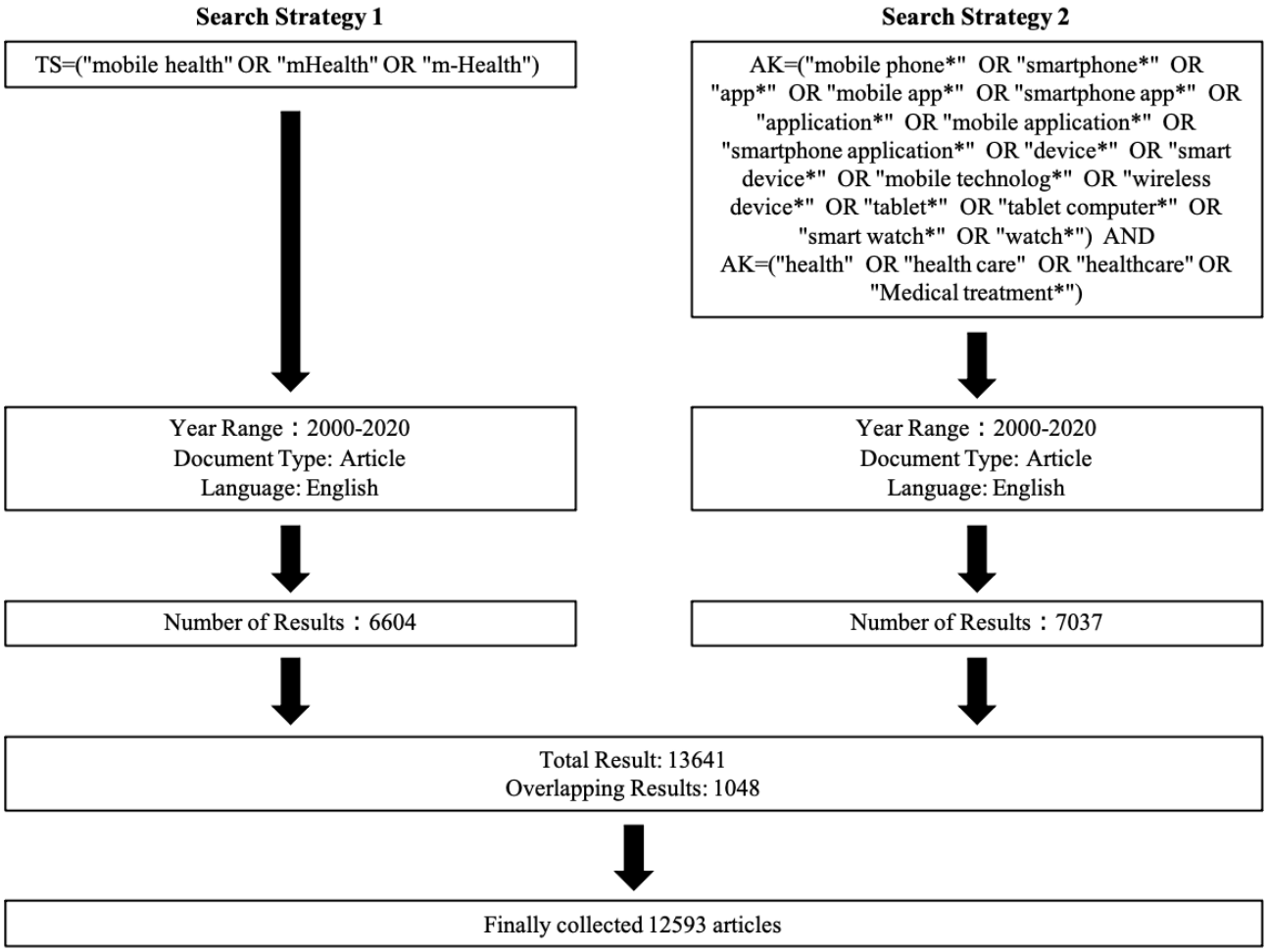
Data Collection

We collected metadata (paper title, abstract, author keywords, author information, country, and references [27]) on papers related to mHealth published between 2000 and 2020 from the Web of Science database. The Web of Science database was chosen because it covers a wide range of fields of study and includes 21,000 peer reviewed and high-quality journals. In addition, the Web of Science includes 6 high-impact citation databases, the Science Citation Index extension, Social Science Citation Index, and many regional databases [28] in its core collection. Thus, the Web of Science database was considered to be appropriate for the bibliometric analysis.

We conducted searches using mHealth and its synonyms as search-topic keywords (in titles, abstracts, and author-provided keywords) to find potential publications related to mHealth;

however, this simple approach has a major limitation. As Sweileh et al [13] suggested, many researchers might not identify their papers as focusing on mHealth though the papers are mHealth-related. Therefore, a second search strategy was also used. Given that mHealth depends on mobile devices, we searched for author-provided keywords related to both mobile devices and mHealth (smartphone, mobile phone, etc) and general health (health, health care, etc). Author-provided keyword searches were used instead of topic searches because the latter may have led to the inclusion of papers that did not emphasize the study of mobile devices and health, whereas the former represents keywords chosen by authors to highlight the contents of their papers. Thus, we determined that searching for author-provided keywords would be more appropriate to collect articles related to mHealth. Both search strategies were conducted for the period from 2000 to 2020, and only papers published in English were retrieved (Figure 1). We implemented the search on March 2, 2021. The results from both strategies were aggregated, and duplicates were removed.

Figure 1. Data collection strategy for mHealth research bibliometric analysis. AK: author-provided keywords; TS: topic search.



Data Analysis

We used VOSviewer (version 1.6.15) for data analysis. In bibliometric analysis, mapping and clustering techniques can provide insight into network structure and are usually used together [28,29]; however, these 2 techniques were developed independently and rely upon different ideas and assumptions. Waltman et al [30] proposed a unified mapping approach and clustering, which is used in VOSviewer [31]. This tool has been used in bibliometric analyses in many fields [32,33].

The annual number of publications, the annual growth rate, *AGR*; relative growth rate, *RGR*; doubling time, *DT*; and the growth curve of publications were calculated to observe publication trends in mHealth literature using Excel (version 2013; Microsoft Inc). In the growth curve, x is the number of years of growth since 2000, and y is the cumulative number of publications. We examined the coefficient of determination (R^2) to confirm the explanatory power of the growth curve. *AGR* was defined as the percentage change in the number of publications per year and is calculated with the following formula: $AGR = [(N_2 - N_1) / N_1] * 100$, where N is the annual number of publications. *RGR* was defined as the growth rate of the cumulative number of publications per unit of time and was calculated with the following formula [13,34]: $RGR = [(lnTN_2 - lnTN_1) / (T_2 - T_1)] * 100$, where T is the year and TN is the cumulative number of publications. *DT* was defined as the number of times the number of publications double in 1 year and was calculated with the following formula [13,34]: $DT = 0.693/RGR$.

In addition, we analyzed the publication trends by country (or region) and the distribution of publications by journal. Using

VOSviewer, we created bibliometric maps for social networks, based on countries and regions, to identify international partnerships in the mHealth field.

In this study, we did a co-occurrence analysis using author-provided keywords in VOSviewer to elucidate research hotspots in the mHealth field. We set the minimum number of co-occurrences to 50. The keywords *mHealth* and *smartphone* (as well as keywords with a similar meaning) appeared more frequently because of the search strategy and took up a large weight in the co-occurrence network graph. Such keywords were considered to influence the distribution of the remaining keywords; hence, we removed the keywords used in the search strategy that appeared in the results, to focus the results on valuable research-topic buzzwords. We then extracted the top 100 keywords and mapped them into a keyword co-occurrence network. The top 100 author-provided keywords were superimposed and visualized according to the average publication year to determine the changes in research hotspots of mHealth over time. The node size indicates the number of times the author's keyword appeared, and the color of the node changes gradually, according to the average publication year.

Results

mHealth Research Publications

Through the first search strategy, 6604 search results were obtained, and through the second strategy, 7037 search results were obtained. After removing 1048 documents; there were 12,593 remaining (Figure 1). The number of publications related to mHealth has been increasing since 2004 (Table 1, Figure 2, and Figure 3) and has demonstrated an approximately exponential growth trend. By fitting an exponential function equation, the growth curve can be represented by $y=37e^{0.3062x}$, with $R^2=0.9935$. Specifically, the year 2015 was a flashpoint. The number of documents published in 2015 increased by 366 compared to 2014, and the annual growth rate reached 61%, becoming the highest annual growth rate in 20 years. *RGR* dropped from 58% in 2001 to 30% in 2003 and then stabilized at 28% (SD 5%). *DT* increased from 1.2 in 2001 to 2.3 in 2003 and then stabilized at 2.6 (SD 0.5). The stability of *RGR* and *DT* demonstrates the exponential growth trend [13,34] of the number of publications and confirms that the curve in Figure 3 is exponential, which indicates that the field of mHealth is increasingly receiving attention from scholars.

Table 1. Descriptive statistics of the collected mHealth literature.

Year	Publications, n	Annual growth, n	*AGR*[a] (%)	*RGR*[b] (%)	*DT*[c]	Cumulative total, n
2000	37	N/A[d]	N/A	N/A	N/A	37
2001	29	–8	–22	58	1.2	66
2002	41	12	41	48	1.4	107
2003	37	–4	–10	30	2.3	144
2004	50	13	35	30	2.3	194
2005	77	27	54	33	2.1	271
2006	86	9	12	28	2.5	357
2007	96	10	12	24	2.9	453
2008	123	27	28	24	2.9	576
2009	158	35	28	24	2.9	734
2010	184	26	16	22	3.1	918
2011	236	52	28	23	3.0	1154
2012	302	66	28	23	3.0	1456
2013	437	135	45	26	2.6	1893
2014	603	166	38	28	2.5	2496
2015	970	367	61	33	2.1	3466
2016	1206	236	24	30	2.3	4672
2017	1383	177	15	26	2.7	6055
2018	1725	342	25	25	2.8	7780
2019	2132	407	24	24	2.9	9912
2020	2681	549	26	24	2.9	12,593

[a]*AGR*: annual growth rate.

[b]*RGR*: relative growth rate.

[c]*DT*: doubling time.

[d]N/A: not applicable.

Figure 2. Number of publications in mHealth literature between 2000 and 2020.

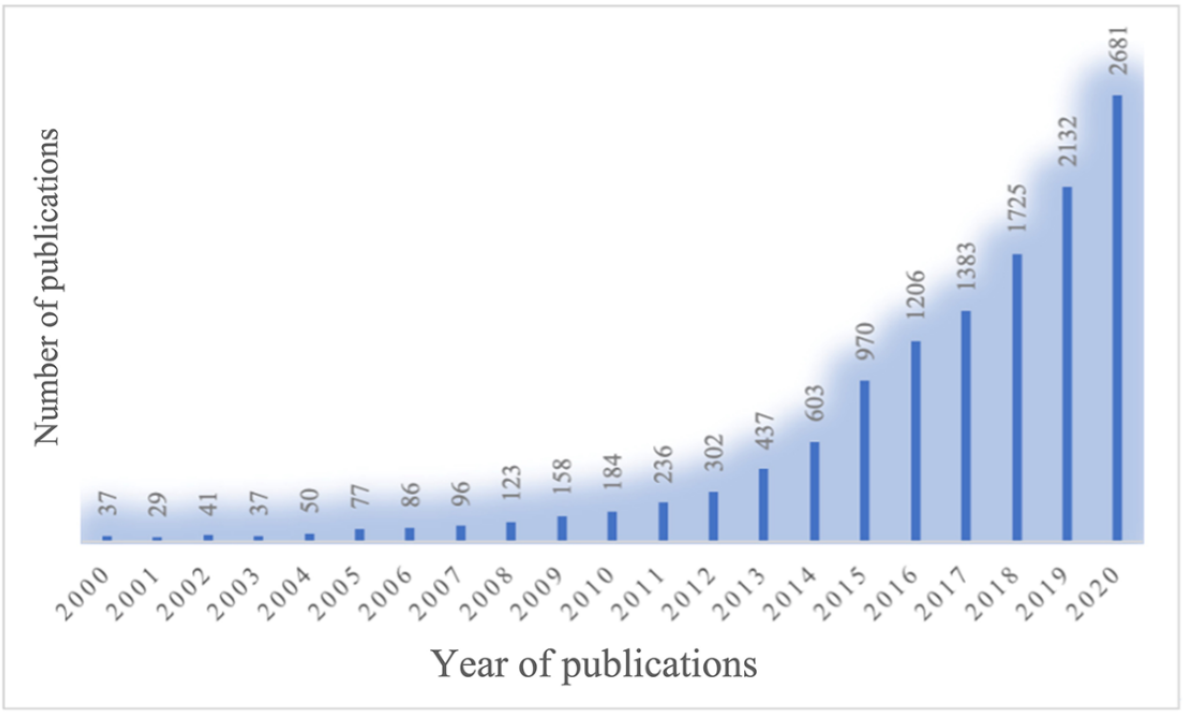
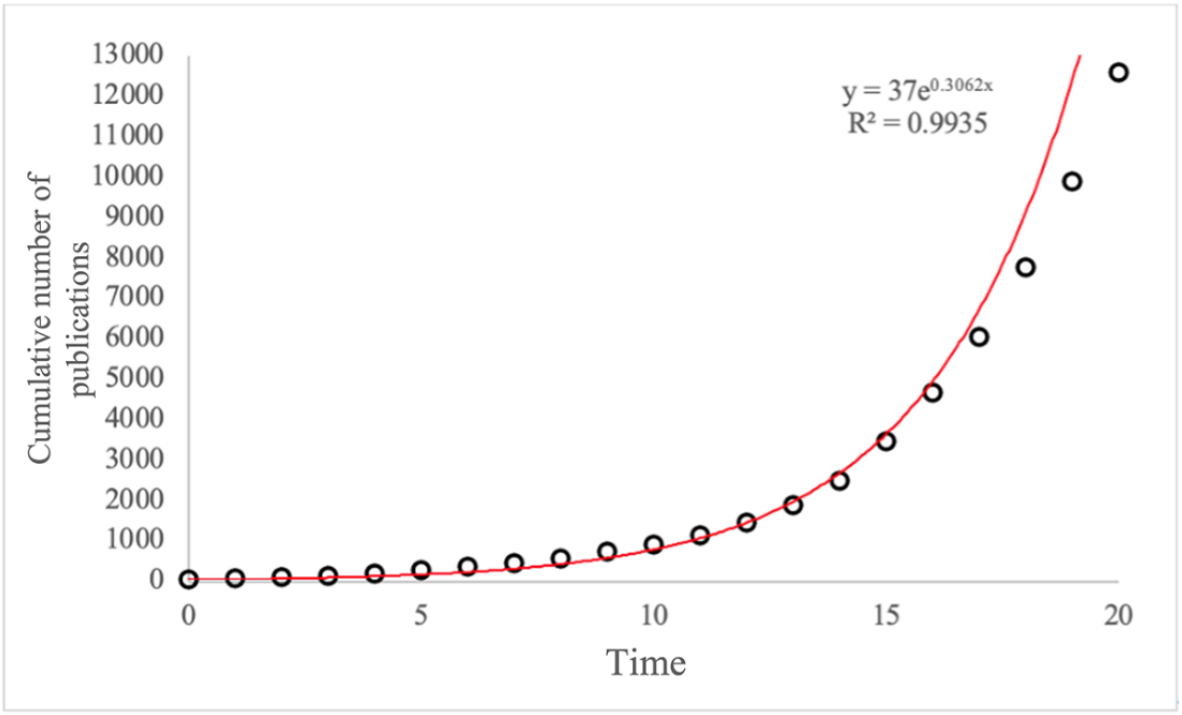


Figure 3. Growth curve of the cumulative number of publications in mHealth literature.



Publishing Trends and Cooperation Among Countries and Regions

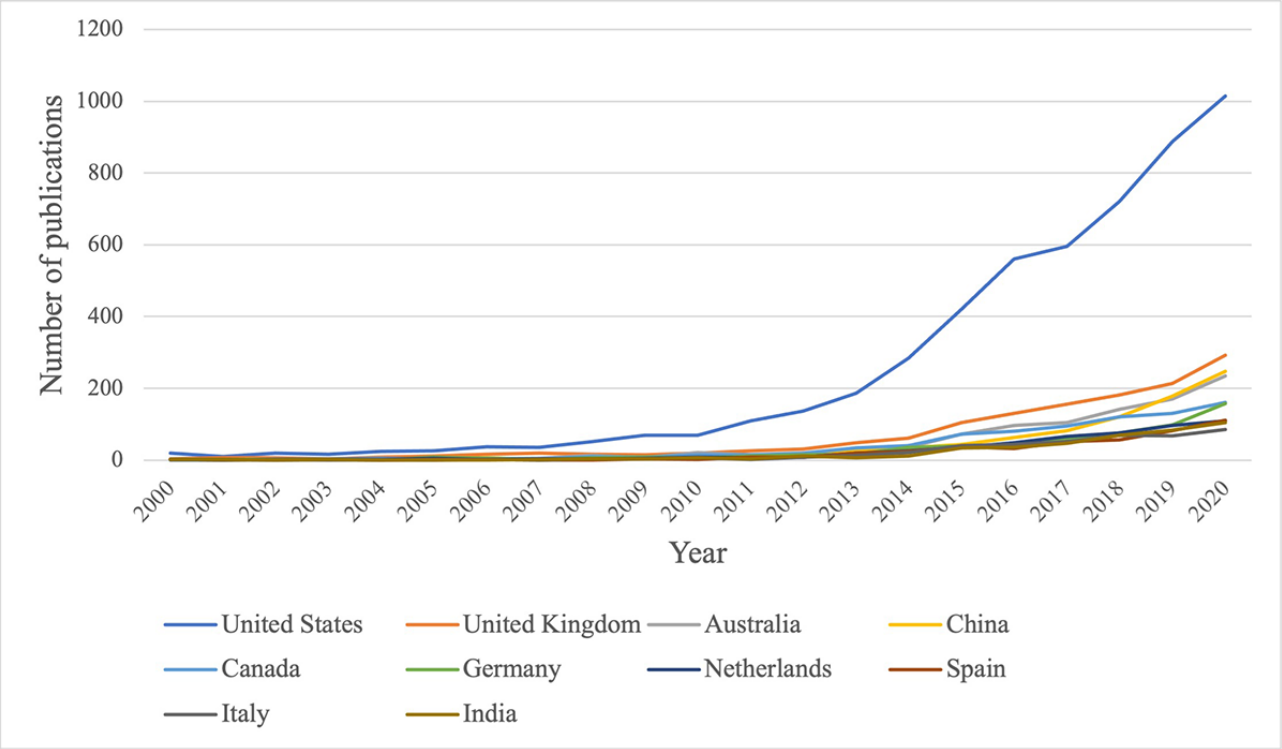
We found that scholars from 166 countries and regions contributed to publications on mHealth (Multimedia Appendix 1). The United States had the largest number of publications (5294/12,593, 42.0%), followed by the United Kingdom (1372/12,593, 10.9%), and then Australia (979/12,593, 7.8%), China (842/12,593, 6.7%), and Canada (828/12,593, 6.6%) (Table 2). Compared with that of other countries, the growth curve of the United States shows explosive growth (Figure 4); mHealth received more attention, early on, from scholars in the United States which continued throughout the period. All countries and regions show growth, though not as high as that of the United States.

Table 2. Top 10 contributing countries in mHealth literature between 2000 and 2020.

Rank	Country and territory	Publications (n=12,593), n (%[a])
1	United States	5294 (42.0)
2	United Kingdom	1372 (10.9)
3	Australia	979 (7.8)
4	China	842 (6.7)
5	Canada	828 (6.6)
6	Germany	583 (4.6)
7	The Netherlands	526 (4.2)
8	Spain	445 (3.5)
9	Italy	426 (3.4)
10	India	424 (3)

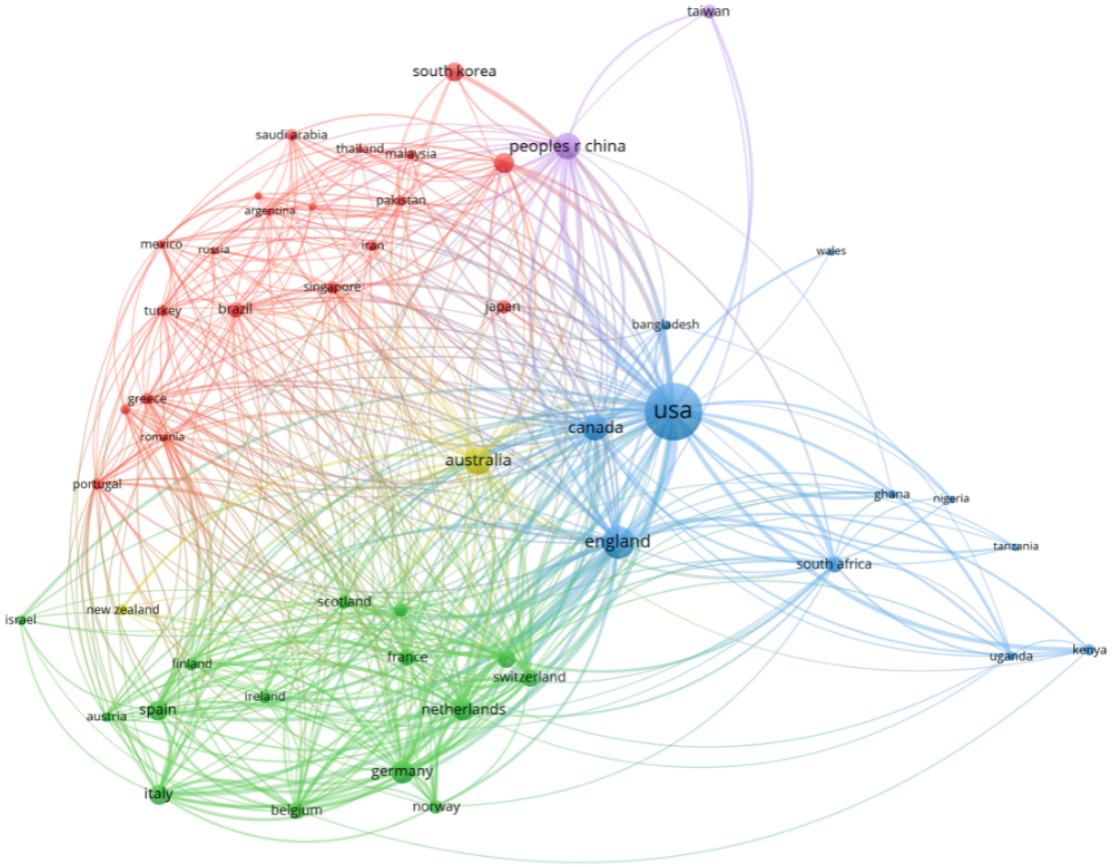
[a]Due to research cooperation between scholars of different nationalities, some papers have been counted more than once.

Figure 4. Comparison of the growth trends of mHealth-related research publications in various countries between 2000 and 2020. Due to research cooperation between scholars of different nationalities, some papers have been counted more than once.



Usually, the closer the two circles, the thicker the links and the stronger the relationship (between the countries). Different colors indicate different clusters, and circles belonging to the same cluster usually have similar properties or characteristics [31]. All countries had a cooperative relationship with the United States (Figure 5). Of the top 5 countries, in terms of the number of publications, United States, the United Kingdom, Canada, and Australia occupy the center of the network diagram with similar distances between the nodes; these 4 productive countries have strong collaborative relationships. Furthermore, it is evident from the location of the countries' nodes that the cooperation between countries and regions have geographic characteristics.

Figure 5. Visual network diagram of cooperation between countries or regions. The size of the circles indicates the number of publications. The larger the circle, the greater the number of publications. The length and thickness of the links between the circles indicate the strength of partnerships between countries. Asian countries and regions represented by the red cluster and the European countries and regions represented by the green cluster.



Journal Distribution

Literature related to mHealth was distributed among 3268 journals (Table 3). The Canadian *Journal of Medical Internet Research* and its sister journals *JMIR mHealth and uHealth*, *JMIR Research Protocols*, and *JMIR Mental Health* were in the top 10 journals, with respect to number of publications, and together represented 14% of all publications (1763/12,593). In addition, all of the top 10 journals, except *JMIR Research Protocols*, have an impact factor above 2.

Table 3. Top 10 journals, in terms of the number of mHealth publications, between 2000 and 2020.

Rank	Journal	Country	2-year impact factor (in 2019)	Publications (n=12,593), n (%)
1	JMIR mHealth and uHealth	Canada	4.31	956 (7.6)
2	Journal of Medical Internet Research	Canada	5.03	463 (3.7)
3	JMIR Research Protocols	Canada	__[a]	235 (1.9)
4	Telemedicine and Health	The United States	2.841	202 (1.6)
5	International Journal of Environmental Research and Public Health	Switzerland	2.849	145 (1.2)
6	BMC Public Health	The United Kingdom	2.69	139 (1.1)
7	JMIR Mental Health	Canada	3.54	109 (0.87)
8	International Journal of Medical Informatics	Ireland	3.025	106 (0.84)
9	BMC Medical Informatics and Decision Making	The United Kingdom	2.317	101 (0.80)
10	Sensors	Switzerland	3.275	99 (0.79)

[a]Not available.

Author Keywords Co-occurrence Analysis

The top 100 keywords (Multimedia Appendix 2) were classified into 5 clusters using keyword clustering analysis (Figure 6), and the top 10 keywords by co-occurrence frequency are shown (Table 4). The average year of publication for the keywords shown in Table 4 ranged from 2015.26 to 2017.90, and the average number of citations ranged from 10.75 to 17.98. The most frequently occurring keyword was *mental health*, with a co-occurrence frequency of 449, followed by *physical activity*, with a co-occurrence frequency of 285.

Figure 6. Co-occurrence network diagram of the top 100 author keywords in mHealth research between 2000 and 2020.

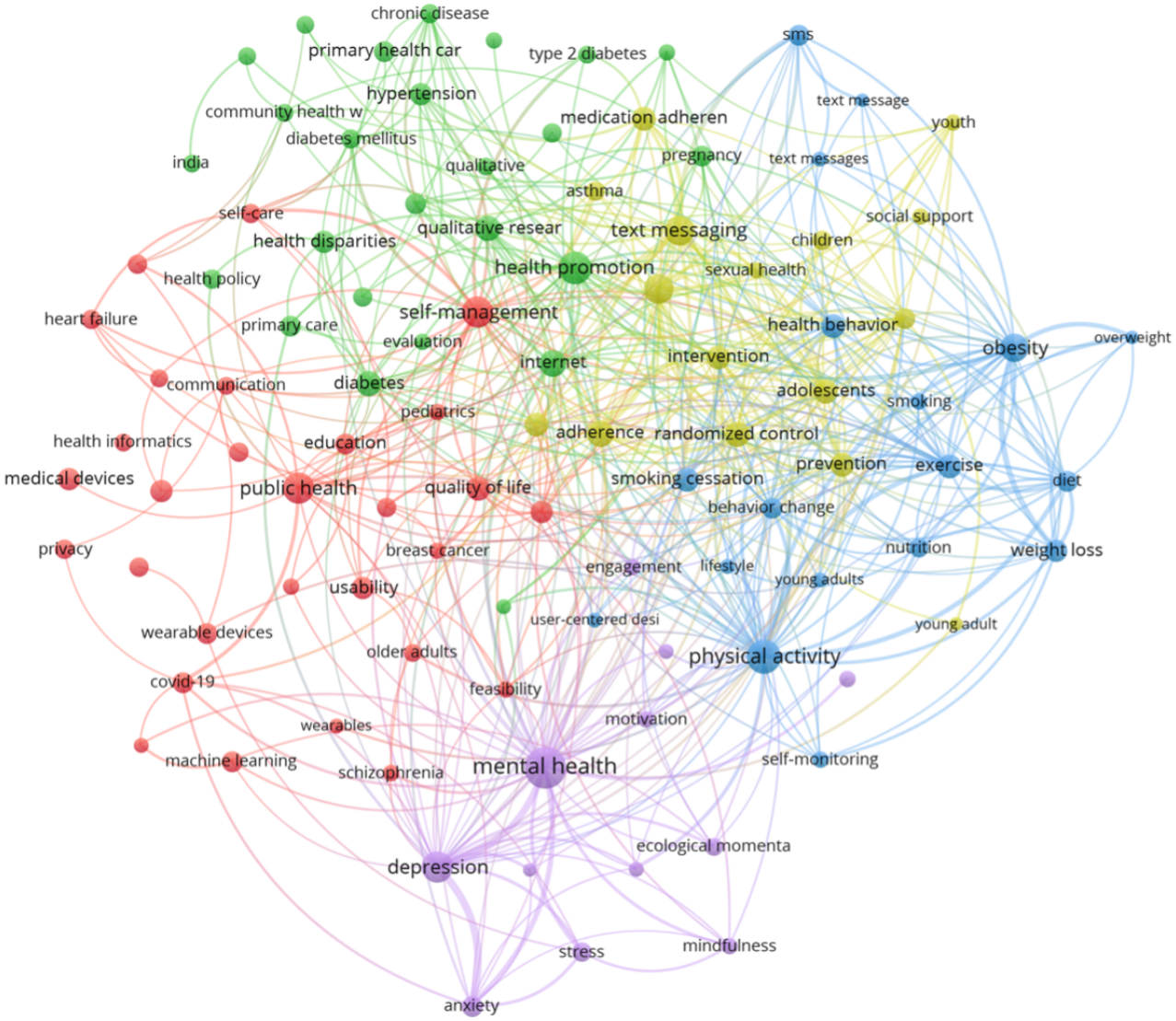


Table 4. Top 10 author-provided keywords of mHealth research between 2000 and 2020.

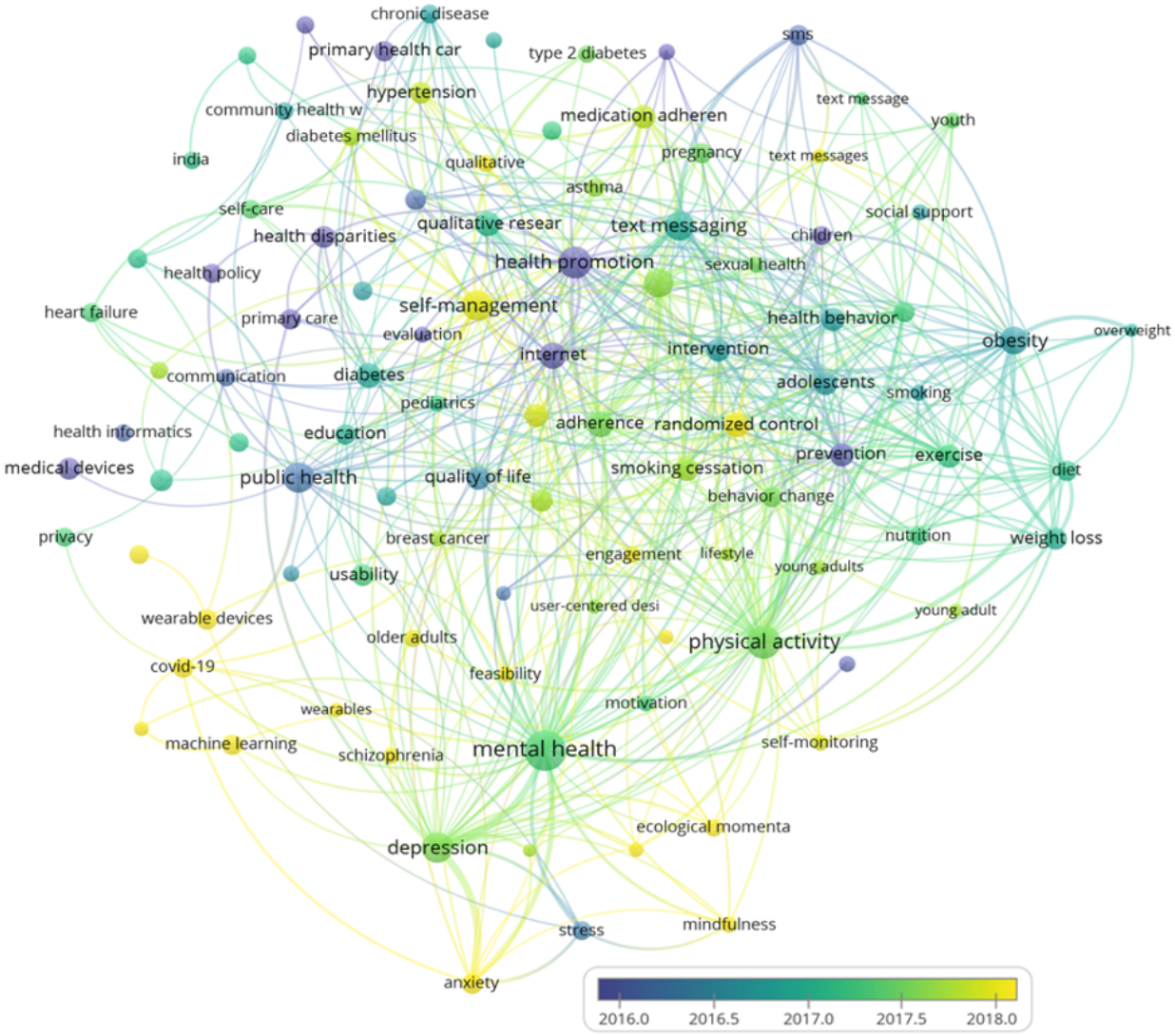
Rank	Keyword	Cluster	Occurrences, n	Average year of publication	Average number of citations
1	mental health	Purple	449	2017.30	12.48
2	physical activity	Blue	285	2017.46	14.00
3	health promotion	Green	243	2015.26	14.97
4	self-management	Red	234	2017.90	10.75
5	public health	Red	232	2016.29	13.41
6	depression	Purple	227	2017.51	17.98
7	HIV	Yellow	208	2017.57	11.37
8	text messaging	Yellow	207	2016.90	13.22
9	obesity	Blue	173	2016.65	13.81
10	adherence	Yellow	157	2017.48	13.85

The average publication year range of the top 15 author-provided keywords was 2017.98 to 2020.05 (Table 5), and the occurrence range was 41 to 135 (Figure 7). Among the top 15 keywords, 8 belonged to cluster *red*, 5 belonged to cluster *purple*, 1 belonged to cluster *yellow*, and 1 belonged to cluster *green*. The average publication year range of the bottom 15 author-provided keywords was 2015.26 to 2016.19, and the occurrence range was 46 to 243. Among the bottom 15 keywords, 10 belonged to cluster *green*, 2 belonged to cluster *red*, 2 belonged to cluster *yellow*, and 1 belonged to cluster *purple*.

Table 5. Comparison of the top 15 and bottom 15 author-provided keywords.

Cluster	Occurrences, n	Average publication year	Keyword
Top 15			
Red	86	2020.05	covid-19
Red	41	2019.05	artificial intelligence
Red	43	2018.79	wearables
Red	85	2018.55	machine learning
Purple	42	2018.54	gamification
Red	47	2018.48	feasibility
Red	87	2018.31	wearable devices
Purple	67	2018.26	ecological momentary assessment
Yellow	135	2018.24	randomized controlled trial
Red	69	2018.22	internet of things
Purple	54	2018.16	mindfulness
Purple	45	2018.09	sleep
Purple	93	2018.05	anxiety
Green	59	2018.04	qualitative
Red	55	2017.98	schizophrenia
Bottom 15			
Green	243	2015.26	health promotion
Green	95	2015.35	primary health care
Green	81	2015.40	health policy
Green	55	2015.45	evaluation
Yellow	80	2015.53	children
Red	103	2015.58	medical devices
Yellow	117	2015.79	prevention
Green	97	2015.82	health disparities
Green	152	2015.84	internet
Green	50	2015.94	focus groups
Green	87	2015.97	primary care
Green	67	2016.00	developing countries
Purple	46	2016.07	well-being
Red	59	2016.15	health informatics
Green	84	2016.19	health education

Figure 7. Overlay visualization maps of the average publication year of the top 100 author keywords. The more the node displays a yellow gradient, the later the average publication year of the keyword.



Discussion

Principal Results

Publishing Trends of mHealth Literature

The emergence of mHealth is a great innovation in the rapid development of information technology. It has circumvented the obstacles of location and medical resources of traditional health care, making health care more accessible to a wider range of people. The growth trend for mHealth literature published between 2000 and 2020 was exponential, which suggests that, when mHealth first started, acceptance was low, the number of users was small, and research on mHealth progressed relatively slowly. As the number of users of mHealth gradually increased, more and more researchers focused on this area, and the number of mHealth publications showed an increasing trend. Based on the theory of diffusion of innovation [35], the growth curve (Figure 3) coincides with the early part of the diffusion model of innovations; we can surmise that the development of mHealth technology is currently in the early stages of rapid growth.

International Trends

A comparison with the bibliometric analysis [25] of mHealth research up to 2016 shows that the United States remains the most productive country in this field. The number of annual publications in the United States continues to show a steady growth trend. This is followed by the United Kingdom, China, Australia, and Canada, which are also experiencing rapid growth in their publication trends.

The 4 most productive countries—the United States, the United Kingdom, Canada, and Australia—had close cooperative relationships with each other. In contrast, the rest of the countries and regions showed a clear geographic pattern. Cluster *red* contains mainly of Asian countries such as Japan, South Korea, Russia, Malaysia, Thailand, and Singapore (Figure 5), and cluster *green* is composed mainly of European countries such as France, Netherlands, Germany, Spain, and Italy. It is not difficult to speculate that the specificity of the EU has led to closer research cooperation among EU countries. Cluster blue comprises African countries such as Kenya, South Africa, Ghana, Nigeria, Tanzania and 3 of the most productive countries—the United States, Canada, and United Kingdom. It

can be presumed that African countries establish cooperation based on geography and have a major cooperation relationship with these 3 countries. In addition, China and Taiwan may be grouped in cluster *purple* because of the same language. Australia and New Zealand belong to cluster *yellow*, due to their close geographic locations. Therefore, we conjecture that international partnerships may be influenced by geography, regional characteristics, language, international relations, political, and economic alliances.

Research Hotspots

Cluster *red* contains the most author-provided keywords, comprising 29 keywords such as *artificial intelligence*, *electronic health records*, *global health*, *health informatics*, *health information technology*, *machine learning*, *medical devices*, *self-care*, and *wearable devices*. The keywords *breast cancer*, *cancer*, *covid-19*, *heart failure*, and other diseases also appeared in the list. This cluster focuses on the development of mHealth technologies and their application to various diseases. Globally, health issues such as aging populations and cancer pose a serious challenge to health care providers [36-38]. Researchers are increasingly trying to address many health issues with the use of mHealth technologies. *COVID-19* also appears among the high-frequency keywords. Importantly, the COVID-19 pandemic exposed the shortage of health care resources in several countries. The demand for telemedicine, including mHealth, has also been indirectly increased by countries promoting policies to prevent their population from going outside under social isolation measures adopted to tackle COVID-19 [39]. It is also worth noting that the keyword *privacy* appears in this cluster. Patient privacy, security in data transmission, and privacy-related health policy issues remain major barriers to the development of mHealth in both high-income and low- to middle-income countries [40].

Cluster *green* focuses on the use of mHealth technologies to improve basic public health and health policies. Some of the 25 keywords under this cluster include *health promotion*, *primary health care*, *health education*, *health policy*, *health communication*, and *community health workers*. Health care is one of the largest industries in the world. According to the World Health Organization, global health expenditure in 2017 was US $7.8 trillion, or approximately 10% of the total gross domestic product [41]. Compared to traditional health services, mHealth, which relies upon mobile devices such as smartphones, provides timely health information and fast, inexpensive access to primary care. As of 2017, mobile phone apps related to mHealth exceeded 325,000 [42]. Therefore, it is necessary to formulate corresponding health policies to ensure that mHealth technology can serve society more effectively and to provide direction for future health initiatives. The author-provided keyword *developing countries* appeared in this cluster. The development of mHealth in low- and middle-income countries faces more serious challenges than those faced in high-income countries. Although smartphones have become commonplace globally, challenges exist in terms of the cost of owning and using smartphones in low- and middle-income countries. For example, resource scarcity and other issues have forced low- and middle-income countries to reduce the budget for building mHealth and related infrastructure to allocate resources to other necessities such as potable water and food. The shortage of trained medical professionals and technical skills in low- and middle-income countries has also made the development of mHealth difficult [38]. Therefore, research focused on low- and middle-income countries remains a key research priority for the future of the field.

Cluster *blue* focuses on self-health testing and management in daily life. This cluster comprises 18 keywords such as *behavior change*, *diet*, *exercise*, *health behavior*, *lifestyle*, and *self-monitoring*. An increasing number of people are using emerging mHealth apps to improve their lifestyles and manage their health; these apps have a variety of functions. For example, people can control their daily calorie intake by recording their diet [43] or detect changes in their health by recording their weight, heart rate, and breathing rate [44,45]. In fact, the emergence of such apps has played a positive role in the popularization of mHealth. For example, mobile phone apps related to physical exercise have been combined with users' social networks. People are more willing to use the tracking function of such apps to record their physical changes and share their exercise status with others, thereby increasing their social contacts' motivation to exercise [46,47].

Cluster *yellow* focuses on the use of mHealth among adolescents. This cluster contains 16 keywords such as *adolescent*, *adherence*, *children*, *HIV*, *intervention*, *sexual health*, *social media*, *social support*, and *youth*. Research shows that the youth are the most prone to smartphone addiction [48,49]. There has been considerable research on the negative effects of smartphone addiction on health [50-52]. Excessive smartphone use affects sleep quality, and thus, other daily activities [48,53]. Adolescents are also a priority group for HIV prevention. mHealth apps that use social media technology make it easier for health workers to spread sexual health information more effectively, and thus, reduce the risk of HIV infection among adolescents [54]. Therefore, mHealth research focusing on adolescents is essential.

Cluster *purple* focuses on the use of mHealth in the context of mental health. It contains 12 keywords, including *mental health*, *anxiety*, *mindfulness*, *stress*, and *well-being*. The keyword *mental health* has the most frequent co-occurrence. Therefore, it can be assumed that this topic is the primary focus of researchers. Various factors influence mental health, such as past experiences [55], social stress [56], and interpersonal relationships [57]. People with mental health problems often resist talking to others [58], and even those who have undergone psychotherapy and have recovered are at high risk of reoccurrence [59]. Mental illness is a severe social problem, especially in high-income countries. For example, in Japan, the suicide rate due to depression has been high, and it has been increasing among youth in recent years [60,61]. For a country with a serious aging problem, an increase in the suicide rate among young people can incur a huge cost to the national economy. Moreover, people with depression can have poor physical health compared with that of individuals in the general population [62]. Timely intelligence technology that captures body information provided by mHealth can provide psychologists with more reference data to detect physical changes in patients through ecological

momentary assessment, thus providing more guidance to patients.

Research Trends

Based on the clusters to which these keywords belong, we can speculate that mHealth research hotspots have gradually shifted from research on mHealth policy and the improvement of public health care to the development of mHealth technology and social apps (cluster *green* to cluster *red* and cluster *purple*). Thus, we find that the development of mHealth requires appropriate health policy as a cornerstone. However, individual governments usually develop health policies, leading to national and regional limitations in the scope of policy application. In contrast, the scope of web-based mHealth services can be global. This may also make it more difficult to regulate mHealth services; therefore, it is still necessary to continue to explore how to establish regulations for cross-border telehealth in the future. Furthermore, we note that in high-income countries, especially in the health care field, government regulatory formation is critical to the growth of the mHealth market [63]. Governmental oversight measures often limit the development of mHealth technologies and services [64]. Although the United States is absolutely central to mHealth research, health care regulations in the country may be more conservative and less susceptible to change due to the huge health care infrastructure. Conversely, mHealth policy reforms are likely to be smoother in low- and middle-income countries because they are met with less opposition and fewer infrastructural barriers [65]. Therefore, effective strategies are needed to advance regulatory reforms related to mHealth.

Limitations

To the best of our knowledge, the results obtained in this study are the most recent available for mHealth bibliometric analysis; however, this study has some limitations. First, we developed a search strategy that included as many mHealth-related studies as possible, but we still could not guarantee the inclusion of all mHealth-related studies. Second, our search strategy collected only English-language literature, which narrowed the scope. Hence, the data results are not representative of papers and conference papers published in other languages. Finally, the data used in this study were extracted only from the Web of Science and did not include other search engines such as Scopus and PubMed. Although the Web of Science has a large enough database to ensure the accuracy of the data to a certain extent, there are still many papers that are included only in the other databases, which may have impacted the study results. For example, our finding suggest that only 175 mHealth papers were in Japan (Multimedia Appendix 1); however, many mHealth papers published in Japanese are included in the CiNii database maintained by the National Institute of Informatics in Japan. The Chinese Science Citation Database in China also contains many papers published in Chinese; therefore, future studies can include more databases and languages to make the research results more accurate and rigorous.

Conclusions

This study reveals the latest research trends and hotspots and the current state of international collaboration in mHealth research. As previously suggested, mHealth has shown great potential in recent years for use in all aspects of our lives; however, the development of mHealth faces challenges from regulatory policies, national economies, and personal privacy. Therefore, we advise researchers in this field to work on these issues to further develop the mHealth field. We also hope that the results of this study provide valuable guidance for future mHealth research.

Acknowledgments

We would like to express our gratitude to Ritsumeikan University for providing access to the Web of Science database. We would also like to thank all study participants for their constructive advice and guidance for this research. This work was supported by the *Foundation France-Japon*/Air Liquide Fellowship. The authors gratefully acknowledge the generous support and assistance of the *Fondation France-Japon de École des Hautes Études en Sciences Sociale*s and Air Liquide.

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of countries and regions that have contributed to publications on mHealth.
[XLS File (Microsoft Excel File), 41 KB - mhealth_v9i9e31097_app1.xls]

Multimedia Appendix 2
Details of the top 100 author keywords in mHealth research between 2000 and 2020.
[XLS File (Microsoft Excel File), 42 KB - mhealth_v9i9e31097_app2.xls]

References

1. Penetration rate of smartphones in selected countries 2020. Statista. URL: https://www.statista.com/statistics/539395/smartphone-penetration-worldwide-by-country/ [accessed 2021-03-02]

2. Guo X, Chen S, Zhang X, Ju X, Wang X. Exploring patients' intentions for continuous usage of mhealth services: elaboration-likelihood perspective study. JMIR Mhealth Uhealth 2020 Apr 06;8(4):e17258 [FREE Full text] [doi: 10.2196/17258] [Medline: 32250277]
3. Fu Y, Guo J. Blood cholesterol monitoring with smartphone as miniaturized electrochemical analyzer for cardiovascular disease prevention. IEEE Trans Biomed Circuits Syst 2018 Aug;12(4):784-790. [doi: 10.1109/tbcas.2018.2845856]
4. Nerminathan A, Harrison A, Phelps M, Alexander S, Scott KM. Doctors' use of mobile devices in the clinical setting: a mixed methods study. Intern Med J 2017 Dec;47(3):291-298. [doi: 10.1111/imj.13349] [Medline: 27925381]
5. Rowan NJ, Laffey JG. Challenges and solutions for addressing critical shortage of supply chain for personal and protective equipment (PPE) arising from Coronavirus disease (COVID19) pandemic - Case study from the Republic of Ireland. Sci Total Environ 2020 Jul 10;725:138532 [FREE Full text] [doi: 10.1016/j.scitotenv.2020.138532] [Medline: 32304970]
6. Sun S, Xie Z, Yu K, Jiang B, Zheng S, Pan X. COVID-19 and healthcare system in China: challenges and progression for a sustainable future. Global Health 2021 Jan 21;17(1):14 [FREE Full text] [doi: 10.1186/s12992-021-00665-9] [Medline: 33478558]
7. Sen-Crowe B, Sutherland M, McKenney M, Elkbuli A. A closer look into global hospital beds capacity and resource shortages during the COVID-19 pandemic. J Surg Res 2021 Apr;260:56-63 [FREE Full text] [doi: 10.1016/j.jss.2020.11.062] [Medline: 33321393]
8. Tarricone R, Petracca F, Ciani O, Cucciniello M. Distinguishing features in the assessment of mHealth apps. Expert Rev Pharmacoecon Outcomes Res 2021 Aug 04;21(4):521-526. [doi: 10.1080/14737167.2021.1891883] [Medline: 33663324]
9. Byambasuren O, Sanders S, Beller E, Glasziou P. Prescribable mHealth apps identified from an overview of systematic reviews. NPJ Digit Med 2018 May 9;1(1):12 [FREE Full text] [doi: 10.1038/s41746-018-0021-9] [Medline: 31304297]
10. Wu X, Guo X, Zhang Z. The efficacy of mobile phone apps for lifestyle modification in diabetes: systematic review and meta-analysis. JMIR Mhealth Uhealth 2019 Jan 15;7(1):e12297 [FREE Full text] [doi: 10.2196/12297] [Medline: 30664494]
11. Xu H, Huang S, Qiu C, Liu S, Deng J, Jiao B, et al. Monitoring and management of home-quarantined patients with COVID-19 using a WChat-based telemedicine system: retrospective cohort study. J Med Internet Res 2020 Jul 02;22(7):e19514 [FREE Full text] [doi: 10.2196/19514] [Medline: 32568727]
12. Baumgart DC. Digital advantage in the COVID-19 response: perspective from Canada's largest integrated digitalized healthcare system. NPJ Digit Med 2020 Aug 31;3(1):114 [FREE Full text] [doi: 10.1038/s41746-020-00326-y] [Medline: 32923691]
13. Sweileh WM, Al-Jabi SW, AbuTaha AS, Zyoud SH, Anayah FMA, Sawalha AF. Bibliometric analysis of worldwide scientific literature in mobile-health: 2006-2016. BMC Med Inform Decis Mak 2017 May 30;17(1):72 [FREE Full text] [doi: 10.1186/s12911-017-0476-7] [Medline: 28558687]
14. Daim TU, Rueda G, Martin H, Gerdsri P. Forecasting emerging technologies:use of bibliometrics and patent analysis. Technol Forecast Soc Change 2006 Oct;73(8):981-1012. [doi: 10.1016/j.techfore.2006.04.004]
15. Morris S, DeYong C, Wu Z, Salman S, Yemenu D. Diva: a visualization system for exploring document databases for technology forecasting. Comput Ind Eng 2002 Sep;43(4):841-862. [doi: 10.1016/s0360-8352(02)00143-2]
16. Fahimnia B, Sarkis J, Davarzani H. Green supply chain management: a review and bibliometric analysis. Int J Prod Econ 2015 Apr;162:101-114. [doi: 10.1016/j.ijpe.2015.01.003]
17. Zupic I, Čater T. Bibliometric methods in management and organization. Organ Res Methods 2014 Dec 22;18(3):429-472. [doi: 10.1177/1094428114562629]
18. Kato M, Ando A. The relationship between research performance and international collaboration in chemistry. Scientometrics 2013 Apr 10;97(3):535-553. [doi: 10.1007/s11192-013-1011-y]
19. Chen K, Zhang Y, Fu X. International research collaboration:an emerging domain of innovation studies? Research Policy 2019 Feb;48(1):149-168. [doi: 10.1016/j.respol.2018.08.005]
20. Li X, Du J, Long H. Dynamic analysis of international green behavior from the perspective of the mapping knowledge domain. Environ Sci Pollut Res Int 2019 Mar 7;26(6):6087-6098. [doi: 10.1007/s11356-018-04090-1] [Medline: 30617886]
21. Liu Z, Yin Y, Liu W, Dunford M. Visualizing the intellectual structure and evolution of innovation systems research: a bibliometric analysis. Scientometrics 2015 Jan 22;103(1):135-158. [doi: 10.1007/s11192-014-1517-y]
22. Liao H, Tang M, Luo L, Li C, Chiclana F, Zeng X. A bibliometric analysis and visualization of medical big data research. Sustainability 2018 Jan 11;10(2):166. [doi: 10.3390/su10010166]
23. Servantie V, Cabrol M, Guieu G, Boissin J. Is international entrepreneurship a field? a bibliometric analysis of the literature (1989–2015). J Int Entrep 2016 Mar 8;14(2):168-212. [doi: 10.1007/s10843-015-0162-8]
24. Donthu N, Kumar S, Pattnaik D. Forty-five years of journal of business research: a bibliometric analysis. J Bus Res 2020 Mar;109:1-14. [doi: 10.1016/j.jbusres.2019.10.039]
25. Shen L, Xiong B, Li W, Lan F, Evans R, Zhang W. Visualizing collaboration characteristics and topic burst on international mobile health research: bibliometric analysis. JMIR Mhealth Uhealth 2018 Jul 05;6(6):e135 [FREE Full text] [doi: 10.2196/mhealth.9581] [Medline: 29871851]
26. Peng C, He M, Cutrona SL, Kiefe CI, Liu F, Wang Z. Theme trends and knowledge structure on mobile health apps: Bibliometric analysis. JMIR Mhealth Uhealth 2020 Jul 27;8(7):e18212 [FREE Full text] [doi: 10.2196/18212] [Medline: 32716312]

27. Agarwal A, Durairajanayagam D, Tatagari S, Esteves S, Harlev A, Henkel R, et al. Bibliometrics: tracking research impact by selecting the appropriate metrics. Asian J Androl 2016;18(2):296. [doi: 10.4103/1008-682x.171582]
28. He Y, Cheung Hui S. Mining a web citation database for author co-citation analysis. Inf Process Manag 2002 Jul;38(4):491-508. [doi: 10.1016/s0306-4573(01)00046-2]
29. Uribe-Toril J, Ruiz-Real JL, Nievas-Soriano BJ. A study of eHealth from the perspective of social sciences. Healthcare 2021 Jan 21;9(2):108. [doi: 10.3390/healthcare9020108]
30. Waltman L, van Eck NJ, Noyons EC. A unified approach to mapping and clustering of bibliometric networks. J Informetr 2010 Oct;4(4):629-635. [doi: 10.1016/j.joi.2010.07.002]
31. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics 2009 Dec 31;84(2):523-538. [doi: 10.1007/s11192-009-0146-3]
32. Merediz-Solà I, Bariviera AF. A bibliometric analysis of bitcoin scientific production. Res Int Bus Finance 2019 Dec;50:294-305. [doi: 10.1016/j.ribaf.2019.06.008]
33. Zyoud SH, Sweileh WM, Awang R, Al-Jabi SW. Global trends in research related to social media in psychology: mapping and bibliometric analysis. Int J Ment Health Syst 2018;12:4 [FREE Full text] [doi: 10.1186/s13033-018-0182-6] [Medline: 29387147]
34. Kumar R, Kaliyaperumal K. Scientometric analysis of global publication output in mobile technology. DESIDOC J Library Inf Technol 2015 Jul 30;35(4):287-292. [doi: 10.14429/djlit.35.4.7884]
35. Rogers E. Diffusion of Innovations: Fifth Edition. New York, NY: Free Press; 2003.
36. Köppen J, Maier CB, Busse R, MUNROS team. What are the motivating and hindering factors for health professionals to undertake new roles in hospitals? A study among physicians, nurses and managers looking at breast cancer and acute myocardial infarction care in nine countries. Health Policy 2018 Oct 25;122(10):1118-1125. [doi: 10.1016/j.healthpol.2018.07.018] [Medline: 30097353]
37. Kurpas D, Gwyther H, Szwamel K, Shaw RL, D'Avanzo B, Holland CA, et al. Patient-centred access to health care: a framework analysis of the care interface for frail older adults. BMC Geriatr 2018 Nov 12;18(1):273 [FREE Full text] [doi: 10.1186/s12877-018-0960-7] [Medline: 30419817]
38. Kanchanachitra C, Lindelow M, Johnston T, Hanvoravongchai P, Lorenzo FM, Huong NL, et al. Human resources for health in southeast Asia: shortages, distributional challenges, and international trade in health services. Lancet 2011 Mar 26;377(9767):769-781. [doi: 10.1016/S0140-6736(10)62035-1] [Medline: 21269674]
39. Rauschenberg C, Schick A, Hirjak D, Seidler A, Paetzold I, Apfelbacher C, et al. Evidence synthesis of digital interventions to mitigate the negative impact of the COVID-19 pandemic on public mental health: rapid meta-review. J Med Internet Res 2021 Mar 10;23(3):e23365 [FREE Full text] [doi: 10.2196/23365] [Medline: 33606657]
40. Sunyaev A, Dehling T, Taylor PL, Mandl KD. Availability and quality of mobile health app privacy policies. J Am Med Inform Assoc 2015 Apr;22(e1):e28-e33 [FREE Full text] [doi: 10.1136/amiajnl-2013-002605] [Medline: 25147247]
41. Global spending on health: a world in transition. World Health Organization. 2019. URL: https://www.who.int/health_financing/documents/health-expenditure-report-2019/en/ [accessed 2021-03-02]
42. Larson RS. A path to better-quality mHealth apps. JMIR Mhealth Uhealth 2018 Jul 30;6(7):e10414 [FREE Full text] [doi: 10.2196/10414] [Medline: 30061091]
43. Franco RZ, Fallaize R, Lovegrove JA, Hwang F. Popular nutrition-related mobile apps: a feature assessment. JMIR Mhealth Uhealth 2016 Aug 01;4(3):e85 [FREE Full text] [doi: 10.2196/mhealth.5846] [Medline: 27480144]
44. Coorey GM, Neubeck L, Mulley J, Redfern J. Effectiveness, acceptability and usefulness of mobile applications for cardiovascular disease self-management: Systematic review with meta-synthesis of quantitative and qualitative data. Eur J Prev Cardiol 2018 Dec;25(5):505-521. [doi: 10.1177/2047487317750913] [Medline: 29313363]
45. Coppetti T, Brauchlin A, Müggler S, Attinger-Toller A, Templin C, Schönrath F, et al. Accuracy of smartphone apps for heart rate measurement. Eur J Prev Cardiol 2017 Aug 02;24(12):1287-1293. [doi: 10.1177/2047487317702044] [Medline: 28464700]
46. Yu N, Huang Y. Important factors affecting user experience design and satisfaction of a mobile health app: A case study of daily yoga app. Int J Environ Res Public Health 2020 Sep 23;17(19):6967 [FREE Full text] [doi: 10.3390/ijerph17196967] [Medline: 32977635]
47. Hamari J, Koivisto J. “Working out for likes”: an empirical study on social influence in exercise gamification. Comput Human Behav 2015 Sep;50:333-347. [doi: 10.1016/j.chb.2015.04.018]
48. CAO J, Lim Y, Kodama K. Smartphone addiction and life satisfaction: mediating effects of sleep quality and self-health. Glob J Health Sci 2021 Jan 13;13(3):8. [doi: 10.5539/gjhs.v13n3p8]
49. Duan L, Shao X, Wang Y, Huang Y, Miao J, Yang X, et al. An investigation of mental health status of children and adolescents in China during the outbreak of COVID-19. J Affect Disord 2020 Oct 01;275:112-118 [FREE Full text] [doi: 10.1016/j.jad.2020.06.029] [Medline: 32658812]
50. Elhai JD, Dvorak RD, Levine JC, Hall BJ. Problematic smartphone use: a conceptual overview and systematic review of relations with anxiety and depression psychopathology. J Affect Disord 2017 Jan 01;207:251-259. [doi: 10.1016/j.jad.2016.08.030] [Medline: 27736736]

51. David P, Kim J, Brickman JS, Ran W, Curtis CM. Mobile phone distraction while studying. New Media Soc 2014 Apr 22;17(10):1661-1679. [doi: 10.1177/1461444814531692]
52. Lepp A, Barkley JE, Sanders GJ, Rebold M, Gates P. The relationship between cell phone use, physical and sedentary activity, and cardiorespiratory fitness in a sample of U.S. college students. Int J Behav Nutr Phys Act 2013 Jul 21;10(1):79 [FREE Full text] [doi: 10.1186/1479-5868-10-79] [Medline: 23800133]
53. Demirci K, Akgönül M, Akpinar A. Relationship of smartphone use severity with sleep quality, depression, and anxiety in university students. J Behav Addict 2015 Jun;4(2):85-92 [FREE Full text] [doi: 10.1556/2006.4.2015.010] [Medline: 26132913]
54. Hightow-Weidman LB, Muessig KE, Bauermeister J, Zhang C, LeGrand S. Youth, technology, and HIV: recent advances and future directions. Curr HIV/AIDS Rep 2015 Dec;12(4):500-515. [doi: 10.1007/s11904-015-0280-x] [Medline: 26385582]
55. Crandall AA, Miller JR, Cheung A, Novilla LK, Glade R, Novilla MLB, et al. ACEs and counter-ACEs: how positive and negative childhood experiences influence adult health. Child Abuse Negl 2019 Oct;96:104089. [doi: 10.1016/j.chiabu.2019.104089] [Medline: 31362100]
56. Wolniewicz CA, Tiamiyu MF, Weeks JW, Elhai JD. Problematic smartphone use and relations with negative affect, fear of missing out, and fear of negative and positive evaluation. Psychiatry Res 2018 Apr;262:618-623. [doi: 10.1016/j.psychres.2017.09.058] [Medline: 28982630]
57. Shin YC, Lee D, Seol J, Lim SW. What kind of stress is associated with depression, anxiety and suicidal ideation in Korean employees? J Korean Med Sci 2017 May;32(5):843-849 [FREE Full text] [doi: 10.3346/jkms.2017.32.5.843] [Medline: 28378560]
58. Capone V. Patient communication self-efficacy, self-reported illness symptoms, physician communication style and mental health and illness in hospital outpatients. J Health Psychol 2016 Jul 01;21(7):1271-1282. [doi: 10.1177/1359105314551622] [Medline: 25274717]
59. Solomon D, Keller M, Leon A, Mueller T, Lavori P, Shea M, et al. Multiple recurrences of major depressive disorder. Am J Psychiatry 2000 Mar;157(2):229-233. [doi: 10.1176/appi.ajp.157.2.229] [Medline: 10671391]
60. Tanaka T, Okamoto S. Increase in suicide following an initial decline during the COVID-19 pandemic in Japan. Nat Hum Behav 2021 Feb;5(2):229-238. [doi: 10.1038/s41562-020-01042-z] [Medline: 33452498]
61. Mizuta A, Suzuki K, Yamagata Z, Ojima T. Teachers' support and depression among Japanese adolescents: a multilevel analysis. Soc Psychiatry Psychiatr Epidemiol 2017 Feb;52(2):211-219. [doi: 10.1007/s00127-016-1320-6] [Medline: 28032137]
62. Aiba M, Matsui Y, Kikkawa T, Matsumoto T, Tachimori H. Factors influencing suicidal ideation among Japanese adults: from the national survey by the Cabinet Office. Psychiatry Clin Neurosci 2011 Aug;65(5):468-475 [FREE Full text] [doi: 10.1111/j.1440-1819.2011.02228.x] [Medline: 21679277]
63. Onodera R, Sengoku S. Innovation process of mHealth: an overview of FDA-approved mobile medical applications. Int J Med Inform 2018 Oct;118:65-71. [doi: 10.1016/j.ijmedinf.2018.07.004] [Medline: 30153924]
64. Malvey D, Slovensky D. mHealth: Transforming Healthcare. New York, NY: Springer; 2014.
65. Malvey DM, Slovensky DJ. Global mHealth policy arena: status check and future directions. Mhealth 2017 Sep 22;3:41-41 [FREE Full text] [doi: 10.21037/mhealth.2017.09.03] [Medline: 29184893]

Abbreviations

COVID-19: coronavirus disease 2019
HIV: human immunodeficiency virus
mHealth: mobile health

Edited by G Eysenbach; submitted 10.06.21; peer-reviewed by T Lefèvre; comments to author 02.07.21; revised version received 15.07.21; accepted 02.08.21; published 08.09.21.

Please cite as:
Cao J, Lim Y, Sengoku S, Guo X, Kodama K
Exploring the Shift in International Trends in Mobile Health Research From 2000 to 2020: Bibliometric Analysis
JMIR Mhealth Uhealth 2021;9(9):e31097
URL: https://mhealth.jmir.org/2021/9/e31097
doi:10.2196/31097
PMID:34494968

Original Paper

Acceptability and Potential Effectiveness of eHealth Tools for Training Primary Health Workers From Nigeria at Scale: Mixed Methods, Uncontrolled Before-and-After Study

Joseph Paul Hicks[1*], BSc, MSc, PhD; Matthew John Allsop[2*], BSc, PhD; Godwin O Akaba[3], MBBS, MSc, MPH, FWACS; Ramsey M Yalma[4], MBBS, FMCPH; Osasuyi Dirisu[5], PhD; Babasola Okusanya[6], FWACS; Jamilu Tukur[7], FWACS; Kehinde Okunade[6], MBBS, FWACS, FMCOG; David Akeju[8], PhD; Adegbenga Ajepe[6], FWACS; Okey Okuzu[9], MBA; Tolib Mirzoev[10], MD, MA, PhD; Bassey Ebenso[1], MBBS, MPH, PhD

[1]Nuffield Centre for International Health and Development, University of Leeds, Leeds, United Kingdom
[2]Academic Unit of Palliative Care, Leeds Institute of Health Sciences, University of Leeds, Leeds, United Kingdom
[3]Department of Obstetrics and Gynaecology, University of Abuja, Gwagwalada, Abuja, Nigeria
[4]Department of Community Medicine, University of Abuja, Gwagwalada, Abuja, Nigeria
[5]Population Council, Abuja, Nigeria
[6]Department of Obstetrics and Gynaecology, College of Medicine, University of Lagos, Lagos, Nigeria
[7]Aminu Kano Teaching Hospital, Kano, Nigeria
[8]Department of Sociology, College of Medicine, University of Lagos, Lagos, Nigeria
[9]Instrat Global Health Solutions, Abuja, Nigeria
[10]Department of Global Health and Development, London School of Hygiene & Tropical Medicine, London, United Kingdom
[*]these authors contributed equally

Corresponding Author:
Matthew John Allsop, BSc, PhD
Academic Unit of Palliative Care
Leeds Institute of Health Sciences
University of Leeds
Worsley Building
Clarendon Way
Leeds, LS2 9LU
United Kingdom
Phone: 44 113434185
Email: m.j.allsop@leeds.ac.uk

Abstract

Background: The in-service training of frontline health workers (FHWs) in primary health care facilities plays an important role in improving the standard of health care delivery. However, it is often expensive and requires FHWs to leave their posts in rural areas to attend courses in urban centers. This study reports the implementation of a digital health tool for providing video training (VTR) on maternal, newborn, and child health (MNCH) care to provide in-service training at scale without interrupting health services. The VTR intervention was supported by satellite communications technology and existing 3G mobile networks.

Objective: This study aims to determine the feasibility and acceptability of these digital health tools and their potential effectiveness in improving clinical knowledge, attitudes, and practices related to MNCH care.

Methods: A mixed methods design, including an uncontrolled pre- and postquantitative evaluation, was adopted. From October 2017 to May 2018, a VTR mobile intervention was delivered to FHWs in 3 states of Nigeria. We examined changes in workers' knowledge and confidence in delivering MNCH services through a pre- and posttest survey. Stakeholders' experiences with the intervention were explored through semistructured interviews that drew on the technology acceptance model to frame contextual factors that shaped the intervention's acceptability and usability in the work environment.

Results: In total, 328 FHWs completed both pre- and posttests. FHWs achieved a mean pretest score of 51% (95% CI 48%-54%) and mean posttest score of 69% (95% CI 66%-72%), reflecting, after adjusting for key covariates, a mean increase between the pre- and posttest of 17 percentage points (95% CI 15-19; $P<.001$). Variation was identified in pre- and posttest scores by the sex

and location of participants alongside topic-specific areas where scores were lowest. Stakeholder interviews suggested a wide acceptance of VTR Mobile (delivered via digital technology) as an important tool for enhancing the quality of training, reinforcing knowledge, and improving health outcomes.

Conclusions: This study found that VTR supported through a digital technology approach is a feasible and acceptable approach for supporting improvements in clinical knowledge, attitudes, and reported practices in MNCH. The determinants of technology acceptance included ease of use, perceived usefulness, access to technology and training contents, and the cost-effectiveness of VTR, whereas barriers to the adoption of VTR were poor electricity supply, poor internet connection, and FHWs' workload. The evaluation also identified the mechanisms of the impact of delivering VTR Mobile at scale on the micro (individual), meso (organizational), and macro (policy) levels of the health system. Future research is required to explore the translation of this digital health approach for the VTR of FHWs and its impact across low-resource settings to ameliorate the financial and time costs of training and support high-quality MNCH care delivery.

Trial Registration: ISRCTN Registry 32105372; https://www.isrctn.com/ISRCTN32105372

(JMIR Mhealth Uhealth 2021;9(9):e24182) doi:10.2196/24182

KEYWORDS

primary health worker training; digital health technology; eHealth; video-based training; maternal and child health; Nigeria; mobile phone

Introduction

Background

After more than a decade of rapid development, approaches using digital health technologies are gaining prominence as a means of addressing health system challenges for improving access to and the quality of health service delivery. However, a need to continue the development of evidence bases still exists [1]. Digital technologies can be used to strengthen health systems and work toward achieving universal health coverage [2]. The onus is now on governments to recognize the importance of digital technologies used in this way, with the World Health Organization's (WHO) member states endorsing this approach at the 71st World Health Assembly [3].

The WHO defines digital health technologies as a salient field of practice for using routine and innovative forms of information and communications technology (ICT) to address health needs [4]. This has been recognized through an ICT-related target in Millennium Development Goal 8, in the proceedings of the World Health Assembly resolution on eHealth (WHA58.28), alongside a recent World Health Assembly resolution passed on Digital Health (A71-R7) urging the member states "to prioritize, as appropriate, the development, evaluation, implementation, scale-up and greater use of digital technologies, as a means of promoting equitable, affordable and universal access to health for all" [5]. Regionally, in Sub-Saharan Africa, there have been diverse implementations of digital health tools, including those targeting improvements in the use of health care services via reminder text messages, teleconferencing, data management, and information dissemination [6]. A large proportion of published literature on the use of digital health tools specific to maternal, newborn, and child health (MNCH) services in low- and middle-income countries (LMICs) is based in Sub-Saharan Africa [7,8]. In Nigeria, there have been numerous government-led initiatives exploring the role of digital technologies in improving MNCH services, which continue to be the most explored aspects. A recent review of the landscape and inventory of ICTs for health in Nigeria revealed that more than 100 different ICT projects were implemented across the country to strengthen a range of health system functions [9]. The focus on the different program areas included 63 ICT projects on MNCH (eg, providing information to women on healthy living and clinical decision support to their caregivers), 36 projects on health system functions (eg, improving health information systems), 13 projects on health worker training and education, and 6 projects on health financing. This proliferation of ICT programs commenced in 2013 when the Federal Government of Nigeria prioritized ICTs as a strategy for achieving the targets of a *Saving One Million Lives* initiative that aimed at broadening universal access to essential primary health care (PHC) services for vulnerable mothers and their infants [10]. This government-led digital health coordination mechanism under the *Saving One Million Lives* initiative is an example of a national-level institutionalization of digital health [11].

In 2019, the WHO released guidelines on digital health interventions for health system strengthening [4]. Following the critical evaluation of the evidence on emerging digital health interventions for health system improvements, they made multiple recommendations for interventions. These included providing training and educational content digitally, with evidence suggesting that digital apps may increase both health workers' knowledge, the acceptability of the training and educational content to workers, and the feasibility of delivering it. Such an approach can help to address health system constraints, including training deficits and rural reach, which are known to have a detrimental impact on MNCH services. Training delivered in this way holds promise for transforming education and training for health providers and patients by virtue of its potential to (1) reach users across large geographical areas, (2) increase the speed of delivery of training content, and (3) provide learners with the flexibility to study at their own pace and convenience while adapting their learning to their needs and preferences [12]. Even with promising research, there is currently a lack of evidence on the effects of interventions that seek to leverage digital approaches to deliver training and educational content in LMIC contexts, including their impacts on outcomes such as health workers' performance, skills, and

attitudes [4]. This situation is further reflected at the national level in Nigeria, where there is a lack of empirical evaluations on the impact of digital health approaches for health tools in general [9], and in training frontline health workers (FHWs) in remote areas that lack regular telecommunications network connectivity.

Objectives

This study addresses this gap in the evidence base by outlining the evaluation of a digital health tool to extend tablet-based educational training for FHWs in rural areas of Nigeria via satellite telecommunications technology. The development of digital health tools has been reported elsewhere [13]. To the best of our knowledge, this is the first report on providing an e-learning intervention that uses satellite telecommunication to reach rural FHWs at scale in an LMIC context, with content specifically tailored to the education and training of FHWs on MNCH care. This paper defines a successful scale in digital technology as the institutionalization or embedding of a digital health product into each level of the health system rather than regarding it as a separate activity [11,14]. In this sense, the integration of e-learning into policy, practices, workflows, and daily lives of health workers in multiple states of Nigeria represents the successful scaling of digital technology. In this paper, we aim to report on the feasibility and acceptability of these tools to rural FHWs and their potential effectiveness in improving FHWs' clinical knowledge, attitudes, and reported practices related to MNCH care.

Methods

Study Design

The e-learning study reported here was embedded within a larger project that combined the video training (VTR) and digitization of health data interventions [13]. Only VTR interventions are reported here. The e-learning component of the larger project involved supplying a computer tablet–based VTR app to 126 rural PHC facilities across three Nigerian states: Federal Capital Territory (FCT), Kano state, and Ondo state [13]. The system enabled the transmission of prerecorded, high-quality training videos and other learning content from a remote server to facilitate the training of rural FHWs on MNCH care, further reducing the need for FHWs to travel to metropolitan cities for training. This larger project included a nonrandomized cluster trial examining the impacts of providing eHealth tools and facilitating infrastructure, specifically satellite communication (SatCom) equipment, to enable remote rural PHC facilities for accessing the internet. This was compared with not providing any eHealth intervention, facilitating infrastructure or any internet access, on routine health service data quality and service provision and use. The data reported in this study relate only to intervention sites that had internet access either via existing 3G mobile networks or through SatCom in facilities without 3G connectivity.

For the quantitative component of this study, we used an uncontrolled before-and-after (or pre-post) design to compare whether rural FHWs' knowledge, attitudes, and reported practices across a range of key MNCH topics changed before and after receiving access to the VTR intervention tools and associated training content. For the qualitative component of the study, we used face-to-face, semistructured in-depth interviews (IDIs) with purposefully selected stakeholders, including FHWs, heads of health facilities, and policy makers, to understand the acceptability, feasibility, and use of computer-enabled VTR to improve health care provision in participating states of Nigeria. IDIs were conducted from February 19 to March 9, 2018 (ie, 12-14 weeks into the implementation of VTR).

Setting and Participants

The study was conducted across three states in Nigeria (Kano, Ondo, and FCT), as outlined in Table 1. Within each state, health facilities were selected purposively (for the wider project) across local government areas, which were assigned as intervention local government areas. We included a total of 126 health facilities in this study, which were subcategorized according to the National PHC Development Agency criteria as PHC facilities, comprehensive health centers, health posts, and basic health centers, and were unequally distributed in number and type across the three study areas (Table 1).

Table 1. Distribution of participating health facilities by their locations in Nigeria (N=126).

Participating states	Region	Population size (2006 census figures) [15]	Participating local government areas	Facility type distribution by state				
				Primary health center (n=91), n (%)	Comprehensive health center (n=6), n (%)	Health post (n=22), n (%)	Basic health center (n=7), n (%)	Total (N=126), n (%)
Ondo state	Western Nigeria	3,460,877 (18th of 37)	Akoko South, Idanre, and Odigbo	58 (63.7)	4 (66.7)	0 (0)	0 (0)	62 (49.2)
Kano state	Northern Nigeria	9,401,288 (1st of 37)	Dawakin Tofa and Sumaila	7 (7.7)	0 (0)	21 (95.5)	7 (100)	35 (27.8)
FCT[a]	Federal territory containing Abuja, in Central Nigeria	1,406,239 (37th of 37)	Gwagwalada and Kuje	26 (28.6)	2 (33.3)	1 (4.5)	0 (0)	29 (23)

[a]FCT: Federal Capital Territory.

PHC facilities, health posts, and basic health centers all provide primary-level care, whereas comprehensive health centers provide secondary-level care. Primary-level facilities often serve as the first point of contact for patients and are mainly staffed by community health extension workers (CHEWs) but have no medical doctors or midwives, whereas secondary-level facilities serve as referral centers and are staffed by CHEWs, medical doctors, nurses, and midwives. In Nigeria, CHEWs are FHWs trained for 2 to 3 years in the schools of health technology to provide basic public health services at primary-level facilities and mainly to assist nurses and midwives in their duties [16]. Our study involved two types of FHWs who were targeted by the intervention: CHEWs, who are present in all four types of health facilities in the study, plus nurses and midwives, who are only present in PHC facilities and comprehensive health centers. In basic health centers and health posts, there was typically just one FHW available for the study (often the facility manager), whereas there were typically at least two FHWs available in PHC facilities and comprehensive health centers, as they usually have a mix of cadres present. Members of the research team recruited all selected FHWs after obtaining permission from their facility managers, explaining the study's objectives to the FHWs and obtaining their consent to participate. This was followed by an orientation on how to use eHealth interventions.

eHealth Intervention

The intervention involved providing all recruited facilities with a tablet computer containing a VTR app (VTR Mobile). VTR Mobile allows users to access video, audio, and text-based learning materials through the internet. The educational videos used for this study were developed by Medical Aid Films [17] and Global Health Media Project [18] and accessed via the ORB platform [19] developed by the mPowering FHWs Partnership [20]. The ORB platform hosts high-quality medical content that can be used under a Creative Commons License to train frontline workers via the internet or via downloads to mobile devices. The educational videos provided clear educational content and engaged clinical scenarios focused on MNCH care, specifically antenatal care, basic obstetric care, perinatal care, and postnatal care. We selected the content of the videos in consultation with the relevant state ministries of health. The videos were delivered to the users via a structured VTR mobile program. User log-ins were created and provided by the staff of the eHealth intervention provider, InStrat Global Health Solutions [21], to the study participants to enable them to log in and work their way through the program, which also tracked their progress.

Quantitative Data Collection and Outcomes

We collected all data via the tablet computers used by the FHWs, with the data automatically uploaded onto remote servers before being accessed by the research team. To assess whether there were any changes in FHWs' knowledge, attitudes, and reported practices related to MNCH care, following access to the eHealth intervention tools and information, all FHWs accessing the VTR mobile system first took a multiple-choice (48 questions) preintervention test (*pretest*) that assessed their reported MNCH knowledge, attitudes, and reported practices on the following 9 topics: (1) focused antenatal care (5 questions), (2) respectful maternity care (2 questions), (3) warning signs in pregnancy (6 questions), (4) how to use a partograph (5 questions), (5) the prevention of postpartum hemorrhage (5 questions), (6) the management of postpartum hemorrhage in a low-resource setting (8 questions), (7) the manual removal of placenta (5 questions), (8) neonatal resuscitation (5 questions), and (9) how to care for a newborn (7 questions).

The postintervention test (*posttest*) questions were the same as the pretest questions. Questions were aligned to the content included in the e-learning program and the curriculum of the included educational videos. The questions were developed by the research team, which included specialists in practice and training in obstetrics and gynecology in Nigeria. Furthermore, consultation with state governments and policy makers occurred during the study planning to ensure that the curriculum of the e-learning program as well as the pre- and posttest questions aligned with the federal government and WHO guidelines for maternal and child health (eg, staff attitudes and provision of respectful maternity care) and that the pre- and posttest questions were clear and easy to understand. Participants who completed the nine VTR modules were automatically prompted via the tablet to take the posttest. Those who had not completed the posttest after 4 weeks of registering for and starting the intervention received fortnightly mobile telephone reminders from the intervention support staff of InStrat to complete the posttest. Multimedia Appendix 1 outlines the questions asked for each topic. Those who had not completed the posttest by the 18th week received weekly text messages from the intervention support staff. No other tests were conducted outside the pre- and postintervention tests in this study. For each user, the system collected data on whether each question was correctly answered. We then calculated our primary outcome as the overall percentage of questions correctly answered in the pre- and posttests. We also created several secondary outcomes based on the percentage of questions correctly answered in both the pre- and posttests, but within each test topic separately. In addition to the pre- and posttest outcome data, we also collected data on FHWs' gender, staff type (CHEW or nurses and midwives), facility type (PHC, comprehensive health centers, health post, or basic health centers), SatCom availability at their facility, facility location (Ondo, Kano, or FCT), and the date of their pre- and posttests, which we used to create a variable measuring the number of days between FHWs' pre- and posttests.

Statistical Analyses

We calculated that a sample size of 324 would provide >80% power to detect an overall increase of 20 percentage points between the pre- and posttest scores, assuming the most conservative overall prescore of 50%, using a two-sided hypothesis test with a significance level of 0.05 and assuming a modest, typical design effect of 1.5, in the absence of any comparable or pilot data, and 10% loss to follow-up. To describe the characteristics of FHWs and health facilities in our study sample, we produced relevant descriptive statistics. To estimate the change in the overall FHW test score results between the pre- and posttests, we first fitted a multilevel linear regression model with the outcome of test score (including both pre- and

posttest scores for every FHW with complete data) and fixed effects for the test period (pre- or posttest), gender (male or female), staff type (CHEW or nurses and midwives [merged due to low sample sizes]), facility type (PHC or comprehensive health centers, or health post and basic health centers [merged due to low sample sizes]), facility SatCom status (yes or no), state (Ondo, FCT, or Kano), and the number of days between FHWs' pre- and posttests. The model also included a random intercept for individual FHWs to account for the repeated outcomes within individuals (ie, the pre- and posttest scores) and a separate random intercept for health facilities to account for any clustering effects at the health facility level. Using the fitted model, we then estimated the overall mean pre- and posttest scores, and the overall posttest minus pretest change in test scores (ie, the estimated change from before to after the intervention), along with the associated 95% CI and *P* values of these means. The means were based on estimated marginal means, also known as least-squares means or adjusted means, as calculated from the fitted model. The estimated marginal means assume a balanced population across all covariates, and when estimating them, we set the only numerical variable in the model (days between pre- and posttest) to its mean value across the sample.

We then estimated test score results within the following mutually exclusive sets of subgroups: (1) male or female FHWs; (2) CHEWs, or nurses and midwives; (3) FHWs in PHCs, or FHWs in comprehensive health centers or FHWs in health posts and basic health centers; (4) FHWs in facilities with SatCom available or FHWs in facilities without the availability of SatCom; and (5) FHWs based in facilities located in Ondo, FCT, or Kano states. To calculate these results, for each set of subgroups, we fitted the same multilevel linear regression model described above, but with an additional term for the interaction between the test period (pre or post) and the relevant categorical variable defining the relevant set of subgroups (eg, gender for male or female FHWs). Using each of these models, we then estimated the pre- and posttest scores and the posttest minus pretest change in the test score for each subgroup, along with the associated 95% CI and *P* values. Within each set of subgroups, we then explored whether the observed changes in test scores between the pre- and posttest periods differed between each subgroup (eg, male vs female FHWs). To do this, we used the same models with interaction terms described above to calculate the differences in estimated test score changes (from pre to post) between the relevant subgroups, taking one subgroup within each set as the reference or comparison group, along with their associated 95% CI and *P* values (again based on estimated marginal means). Finally, we also calculated adjusted overall posttest scores and their 95% CI for each separate topic covered by the test, by repeatedly fitting multilevel linear regression models with outcomes of each topic-specific posttest score, in turn, and independent variables and random intercepts that were the same as described above for the overall primary outcome analysis, excluding a variable for the test period.

For all analyses, we excluded observations (FHWs) if they were missing any outcome or required covariate data (ie, complete case analyses). We calculated CI and *P* values based on *t* statistics using the Kenward–Roger degrees of freedom approximation. We checked for adherence or violation of model assumptions using the standard range of residual and influence plots for multilevel linear regression models, but found no issues. All results were calculated using R version 3.5.2 statistical software (R Foundation for Statistical Computing) [22], with all models fitted using the *lme4* package [23], and all estimated marginal means calculated using the *emmeans* package [24].

Qualitative Data Collection and Analysis

To assess the acceptability and feasibility of using the VTR mobile education intervention, we conducted face-to-face semistructured qualitative interviews with 34 participants in 3 states—12 FHWs, 12 facility managers, and 10 policy makers. Participants were recruited between February 19 and March 9, 2018. Interviews were conducted by 4 medical doctors (KO, AA, OD, and RMY) and a sociologist (DA), who were trained in qualitative interviewing techniques. Only 1 of the 5 data collectors was a female (OD). The research staff provided study information sheets to potential participants to help them understand the objectives of and decide whether to participate in the study. Participants were given at least 24 hours to express interest in participating in the study. Interview guides (Multimedia Appendix 2) were pretested before they were administered to the field. Interviews, which lasted about 30 minutes each, were conducted in a private setting in the workplace of respondents, audio-recorded, transcribed verbatim, and where appropriate, translated into English for analysis. The framework approach was used for analysis, while allowing for the emergence of new themes. The framework analysis involves the stages of familiarization with data, coding (done by the 5 interviewers above), indexing and charting, mapping, and interpretation [25]. The analysis was performed manually.

We drew on the technology acceptance model (TAM) to help explain stakeholders' acceptance and use of VTR Mobile intervention in the workplace environment [26]. The TAM proposes that an individual's acceptability of (ie, intent to use) and use behavior (ie, actual use) of a technology is determined by two variables. These are the perceived usefulness of the technology to enhance job performance, and the perceived ease of use of the technology, that is, the effort needed to learn and use a given technology. An individual's motivation to use an emerging technology is higher if the technology is easy to use. The TAM also proposes that factors such as an individual's understanding of a technology and organizational support measures have positive effects on the perception of usefulness and adoption of technology.

Ethics Approval

Approval for the study was granted by the University of Leeds School of Medicine Research Ethics Committee (MREC16-178) and the Ondo State Government Ministry of Health (AD.4693 Vol. II/109), the Kano State Ministry of Health (MOH/Off/797/T1/350), and the Federal Capital Health Research Ethics Committee (FHREC/2017/01/42/12-05-17).

Results

Overview

We recruited and registered 349 FHWs for this study. However, 2.2% (8/349) of FHWs were doctors and did not complete the pretest. Of the remaining FHWs, 96.1% (328/341) completed both the pre- and posttest, and all had the necessary covariate data (Table 2). All pretests were completed between August 8, 2017, and March 16, 2018, in Ondo state; between August 10, 2017, and February 21, 2018, in Kano state; and between July 10, 2017, and March 15, 2018, in FCT, whereas all posttests were completed between February 23, 2018, and May 21, 2018, in Ondo state; between March 7, 2018, and May 2, 2018, in Kano state; and between March 5, 2018, and April 30, 2018, in FCT. After taking their pretest, FHWs took a mean of 152 days to complete their posttest; however, this varied substantially, ranging from 2 to 279 days (IQR 138.75; Table 3). More specifically, 28.3% (93/328) completed in 4 to 90 days, 17.9% (59/328) completed in 91 to 180 days, and 53.6% (176/328) completed in 181 to 279 days. This result highlights that most FHWs completed their posttest in the final 3 months of the (approximately) 9-month period during which all posttests were completed. Recruited FHWs were primarily female CHEWs based within PHCs and comprehensive health centers, but because comprehensive health centers contain multiple FHWs (mean 12.2), HPs typically contain only 1 or 2 FHWs (mean 1.1), and there were approximately 3 times more HPs than CHCs in the study (Tables 3 and 4).

Less than one-third of all FHWs accessed eHealth materials via SatCom, and SatCom access was available for all health posts and basic health centers in the study, whereas availability was nearly evenly split for PHCs and CHCs (Table 4). Just over half of the FHWs (180/328, 54.9%) were based in facilities in Ondo state, just over a third (115/328, 35.1%) were based in the FCT state, and 10% (33/328) were based in the Kano state (Table 3).

Table 2. Overview of recruitment and the completion of pre- and posttests by frontline health worker staff type (N=349).

	Pretest completion, n (%)	Posttest completion, n (%)
Doctors (n=8)	0 (0)[a]	0 (0)
Nurses and midwives (n=31)	31 (100)	31 (100)
CHEWs[b] (n=310)	297 (95.8)	297 (95.8)
Total (N=349)	328 (93.9)	328 (93.9)

[a]A total of 8 persons were excluded from the analysis due to incomplete tests.

[b]CHEW: community health extension worker.

Table 3. Frontline health worker characteristics (N=328).

Characteristics	Values
Gender, n (%)	
Female	220 (67)
Male	108 (32.9)
Cadre, n (%)	
CHEW[a]	297 (90.5)
Nurses and midwives	31 (9.4)
Type of facility, n (%)	
Primary health center	235 (71.6)
Comprehensive health center	61 (18.5)
Health post	25 (7.6)
Basic health center	7 (2.1)
SatCom[b] available at facility, n (%)	
No	233 (71)
Yes	95 (28.9)
Facility location, n (%)	
Ondo	180 (54.9)
FCT[c]	115 (35.1)
Kano	33 (10)
Days between pretest and posttest, mean (SD)	152.7 (82.1)

[a]CHEW: community health extension worker.

[b]SatCom: satellite communication (including internet).

[C]FCT: Federal Capital Territory.

Table 4. Facility characteristics (N=138).

Facility type	Mode of delivery of eHealth interventions, n (%)		Facility type total, n (%)
	SatCom[a] sites[b]	Non-SatCom sites[c]	
Primary health center	40 (28.9)	59 (42.7)	99 (71.7)
Health post	25 (18.1)	0 (0)	25 (18.1)
Comprehensive health center	4 (2.8)	4 (2.8)	8 (5.7)
Basic health center	6 (4.3)	0 (0)	6 (4.3)

[a]SatCom: satellite communication (including internet).

[b]Total satellite communication sites: 54.3% (75/138).

[c]Total satellite communication sites: 45.6% (63/138).

Overall, FHWs achieved a mean pretest score of 51% (95% CI 48%-54%) and a mean posttest score of 69% (95% CI 66%-72%), and after adjusting for key covariates, this represented an overall mean increase in test score between the pre- and posttest of 17 percentage points (95% CI 15-19; *P*<.001; Table 5). There was an indication that male FHWs' test scores increased slightly less on average than female FHWs (−5 percentage points, 95% CI −9 to 0; *P*=.03), and a much clearer indication that FHWs in Ondo state increased their test scores much more on average than FHWs in Kano state (9 percentage points, 95% CI 3-16; *P*=.005; Table 5). However, there were no clear differences in the observed changes in test scores between different types of FHWs, FHWs in different types of facilities, or FHWs in facilities with and without SatCom availability.

Table 5. Overall subgroup-specific and between-subgroup estimates of frontline health workers' mean pre- and posttest scores and mean pretest to posttest change in test scores[a].

Characteristics	Value, n (%)	Pretest score (95% CI; %)	Posttest score (95% CI; %)	Pre- to posttest scores		Between-subgroup difference in pre- to posttest scores	
				Point change (95% CI; %)	*P* value	Point change (95% CI; %)	*P* value
Overall	328 (100)	51 (48 to 54)	69 (66 to 72)	17 (15 to 19)	<.001	N/A[b]	N/A
Staff sex							
Male	108 (32.9)	52 (48 to 55)	66 (62 to 69)	14 (11 to 18)	<.001	−5 (−9 to 0)	.03
Female	220 (67)	52 (48 to 55)	71 (67 to 74)	19 (16 to 21)	<.001	Reference	N/A
Staff type							
CHEW[c]	297 (90.5)	49 (46 to 52)	66 (63 to 69)	17 (15 to 19)	<.001	−1 (−8 to 6)	.75
Nurses and midwives	31 (9.4)	54 (48 to 59)	72 (67 to 77)	18 (12 to 25)	<.001	Reference	N/A
Facility type							
PHC[d]	235 (71.6)	53 (49 to 57)	70 (66 to 74)	17 (15 to 19)	<.001	2 (−4 to 9)	.50
Comprehensive health center	61 (18.5)	50 (43 to 57)	70 (63 to 78)	21 (16 to 25)	<.001	6 (−2 to 14)	.13
Health posts and basic health centers	32 (9.7)	51 (43 to 59)	66 (58 to 73)	15 (8 to 21)	<.001	Reference	N/A
SatCom[e]							
Yes	95 (28.9)	51 (47 to 55)	67 (63 to 71)	16 (12 to 19)	<.001	−2 (−7 to 2)	.29
No	233 (71)	52 (48 to 56)	70 (67 to 74)	18 (16 to 20)	<.001	Reference	N/A
State							
Ondo	180 (54.8)	53 (48 to 58)	76 (71 to 80)	23 (20 to 25)	<.001	9 (3 to 16)	.005
FCT[f]	115 (35)	56 (51 to 62)	67 (61 to 72)	10 (7 to 13)	<.001	−3 (−10 to 4)	.39
Kano	33 (10)	48 (40 to 56)	61 (53 to 69)	13 (7 to 19)	<.001	Reference	N/A

[a]All values are based on estimated marginal means calculated from multilevel linear regression models. All models have an outcome of frontline health workers' (FHWs') pre- and posttest score values, which were measured as the percentage of correct answers on a 48-question multiple-choice test of FHW knowledge, attitudes, and reported practices with respect to maternal, newborn, and child health. All models include independent variables for FHWs' sex, FHWs' staff type (community health extension workers or nurses and midwives), FHWs' facility type (primary health care, comprehensive health center or health posts and basic health centers), FHWs' facility SatCom status (yes/no), FHWs' facility location (Ondo, Federal Capital Territory, or Kano), and the number of days between FHWs' pre- and posttests. All models also include random intercepts for FHW and facility to account for clustering of pre- and posttest outcome scores within FHWs and facilities. All CIs and *P* values are based on *t* statistics using the Kenward–Roger degrees of freedom approximation. Any FHWs with missing outcome or covariate data were excluded. Refer to the *Methods* section for full details.

[b]N/A: not applicable.

[c]CHEW: community health extension worker.

[d]PHC: primary health care.

[e]SatCom: satellite communication.

[f]FCT: Federal Capital Territory.

Our analysis of the topic-specific test scores showed that, on average, FHWs appeared to do worse on questions about warning signs in pregnancy (topic 2 in Table 6), prevention and management of postpartum hemorrhage (topic 5), and neonatal resuscitation (topic 8) compared with questions on all other topics, but better on questions on respectful maternity care (topic 2) and how to care for a newborn (topic 9) compared with questions on all other topics.

Table 6. Overall frontline health workers' topic-specific posttest scores based on the percentage of correct answers to questions on specific topics on maternal and child health care knowledge and reported attitudes and practices.

Topic number[a] and description	Scores (95% CI)
1. Focused antenatal care	63 (60-65)
2. Respectful maternity care	84 (82-87)
3. Warning signs in pregnancy	51 (49-54)
4. How to use a partograph	67 (65-70)
5. Prevention of PPH[b]	53 (51-56)
6. Management of PPH in a low-resource setting	65 (63-68)
7. Manual removal of placenta	66 (64-69)
8. Neonatal resuscitation	50 (47-52)
9. How to care for a newborn	78 (76-79)

[a]Topic numbers refer to the order in which topics were asked in the test.

[b]PPH: postpartum hemorrhage.

Characteristics of Stakeholders Interviewed

A total of 34 stakeholders were interviewed in 3 states regarding the acceptability, feasibility, and use of the VTR Mobile intervention (Table 7). Approximately 35% (12/34) of respondents (stakeholders) interviewed were FHWs. Another 35% (12/34) of respondents were health facility managers, whereas the remaining 29% (10/34) were policy makers.

Table 7. Characteristics of respondents by stakeholder group interviewed (N=34).

Respondent group	Respondents, n (%)			
	FCT[a,b]	Kano state[c]	Ondo state[d]	Total[e]
FHW[f]	4 (11)	4 (11)	4 (11)	12 (35)
Facility managers	4 (11)	4 (11)	4 (11)	12 (35)
Policymakers	3 (8)	3 (8)	4 (11)	10 (35)

[a]FCT: Federal Capital Territory.

[b]Total respondents in Federal Capital Territory: 32% (11/34).

[c]Total respondents in Kano state: 32% (11/34).

[d]Total respondents in Ondo state: 35% (12/34).

[e]Total respondents interviewed: 100% (34/34).

[f]FHW: frontline health worker.

Qualitative Findings: Factors Affecting the Acceptance and Use of VTR Technology

Findings from IDIs with FHWs, facility managers, and policy makers showed a wide acceptance of VTR mobile technology as an important tool for enhancing the quality of training for health workers and the standard of health care delivery. Stakeholders described the introduction of VTR Mobile as highly beneficial for FHWs in Nigeria and cited the following five drivers of acceptance of tablet-optimized VTR: (1) perceived ease of use of VTR Mobile app and platform; (2) accessibility to tablet computers; (3) convenience of offline access to training content, making videos reliable reference materials; (4) perceived usefulness of training content for improving life-saving skills; and (5) cost-effectiveness of VTR. As respondents often referred to two or three drivers of VTR mobile acceptance in their responses, the quotations outlined supporting drivers (detailed in Table 8) may allude to more than one determinant of acceptability in the same extract.

Despite the barriers outlined above, participants reflected views that the VTR was a promising means for (1) providing high-quality training in a cost-effective way; (2) reinforcing knowledge, enhancing skills, and increasing FHW confidence; (3) improving health care–seeking behavior among women; and (4) reducing maternal and infant mortality and morbidity.

Table 8. Key themes of the acceptability and feasibility of using VTR[a] Mobile with supporting quotes.

Determinants of acceptability and feasibility and description of key findings	Supporting participant quote
Ease of use of VTR Mobile is linked to perceived accessibility	
• Facility managers and FHWs[b] found it easy to use the technology. Respondents highlighted how access to tablet computers in participating health facilities made it easy to use the VTR Mobile App and, by doing so, increased acceptability of the intervention.	• "Having this video and tablet here [in the facility] contributes enough [to making us use it] and the videos give us more guidance on how to manage some minor illnesses that we didn't have more knowledge on them before.... So it is a welcome development." [Facility Manager, Gwagwalada, FCT[c]]
• The feasibility of using tablet computers and the VTR Mobile app was aided by the introductory training provided at the beginning of the study and ongoing support to solve technical problems that arose during the use of devices. The training increased familiarity with and use of digital devices and VTR Mobile app. The link between accessibility to devices and motivation to use the training app in the above quote is supported by a policy maker, who also underlined the prospect of tracking technology for use in monitoring completion and noncompletion of pre- and postintervention tests.	• "The use of VTR has been more regular and on ground. It is regular [because when you have the tablet], you can watch the videos as many times as you want, countless times...you can just open it and watch.... But they [InStrat] have their own assessment too, we also check with them [InStrat] to know how many [participants] are watching the contents. But the assessment you do, they are a kind of mini exam, we will know who are watching and applying what they have learnt from it." [Policy Maker, Gwagwalada, FCT]
Training videos as aide-mémoire for clinical practice	
• Once the videos are downloaded from the VTR Mobile platform onto tablet computers, the training content can be watched repeatedly (offline use) at no additional cost to FHWs and facility managers. The convenience of repeated use of the videos has made them a reliable reference material for FHWs. • It was common for FHWs to gather in groups to watch the videos. They found the video-viewing sessions and their associated vignettes helpful for knowledge exchange, problem-solving, and peer support.	• "I think some staff are beginning to look at this thing [clinical videos on tablets] as an important, err, aspect of work. The video helps us...it makes the work easier. If we are in any difficulty, we just go to that particular video and watch it many times. And now, we know what to do for clients." [Facility manager in Gwagwalada, FCT]
Perceived usefulness of VTR Mobile for improving service provision	
• A key driver of acceptability cited by FHWs was the perceived usefulness of VTR Mobile for improving the quality of health care provision. • Participant narratives demonstrated the utility of computer-optimized videos for increasing knowledge, clarifying areas of confusion, boosting FHW capability to manage birth-related complications and preventing infant mortality. • Although we expected the integration of VTR Mobile into routine health care practice to increase FHW knowledge and confidence to deliver high-quality life-saving care, an unanticipated finding was that access to VTR Mobile empowered FHWs to use the videos to conduct health education sessions for women attending ANC[d]. This unplanned use of VTR Mobile inspired pregnant women to broadcast the availability of new technology at intervention facilities in the community, leading to increased attendance at ANC classes (ie, improved health care–seeking behavior). • Participants described how VTR Mobile also led to improved attitudes of FHWs toward patients, which in turn increased service users' confidence in FHWs, which subsequently increased FHW confidence to provide respectful and high-quality care.	• "You know sometimes when you don't know something [about a topic], and when you continue to receive cases related to that topic, you will be feeling demoralized and doubtful, because you don't know what to do. But when you already have the solution in your hands [as we now do with VTR], then, you will be very confident in what you do, and in the answers, you provide to patients." [FHW in Onikokodiya, Ondo state] • "The last time we used the tablet [to watch videos], it helped us to resuscitate a newborn baby. Before now, when a baby was born, and the baby was not breathing normally, we used to do the mouth-to-mouth [resuscitation] but...our attempts did not work because of this tab on the Ambu bag. This time around, when we had a baby that had trouble breathing, we opened the computer tablet and watched the video, and immediately we picked up our Ambu bag and used it as was shown in the video and the baby began to breath normally again." [Facility manager in Aseigbo, Ondo] • "VTR has really helped a lot because attitude of staff towards patient is one of the most important aspects of healthcare delivery because if you don't show your good conduct, it's like you are driving away your clients. But the more you are closer to patients, the friendlier you are with them, the more they will come to the health facility. They will have more confidence in you, and they will tell you more about themselves and it will help in treatment and other services that you want to provide. Patient-staff relationship is particularly important because that's where confidence will develop. That'll make your work even easier and better because they will not hide anything [from you]. They will tell you the truth and it help you in your diagnosis and even in standard of your work too." [FHW in the FCT]
Cost-effectiveness of VTR	

Determinants of acceptability and feasibility and description of key findings	Supporting participant quote
• What constitutes usefulness, however, may be different across different stakeholder groups depending on whether respondents were health workers or policy makers. Although FHWs and facility managers regarded improved knowledge, skills, and confidence as benefits of VTR, policy makers, in contrast, saw tablet computers loaded with VTR Mobile as a cost-effective intervention for reducing the cost of conventional training of FHWs.	• "In my own assessment, VTR Mobile has been awesome and the experience is encouraging in the sense that health care staff will not need to travel and go anywhere [for training]. Training materials are now available with them [on tablets in health facilities]. While visiting a facility in a very remote area, I heard a testimony from a staff who had a patient with PPH [postpartum hemorrhage] and how she had watched the video clip on the training to be able to manage PPH. She came out [of the labor room] after helping the patient and told us what she did in line with what the video directed her to do. It was an awesome experience...when I heard it, I was happy." [Policymaker in Gwagwalada, FCT]
• Policy makers seemingly found the experience of using clinical videos to improve health care provision in rural areas particularly intriguing, with the potential for mobile technology to improve primary health care. • Policy makers believed that the access to VTR Mobile provided an opportunity to substantially reduce the cost of training while allowing the government to train more FHWs despite several competing priorities. The average cost of training fell by 79.6% to US $509 per year in the project intervention sites compared with US $2489 per year for face-to-face training of CHEWs[e] in Sub-Saharan Africa [27]. The figure of US $509 per year includes the cost of delivering 40-hours of VTR content, supplying hardware, and providing technical support to FHWs.	• "Very well it [VTR] is improving health care positively...I go round like these places [visiting facilities], and the testimonies or reports coming from all sites, particularly the hard-to-reach areas [are remarkable]....These are places where there are no ambulances to convey anybody. Not even motorcycle to transport anybody to the nearest clinic [or hospital]. I am convinced you can strengthen Primary Health care through this project...if we can strengthen Primary Health care [using VTR], other aspects of health care will get this right too." [Policy Maker, Gwagwalada, FCT]
External barriers to the use of VTR	
• Despite the reported positivity toward VTR, three external (structural) barriers affected the adoption and use of VTR: (1) poor internet connection to log into the VTR Mobile platform, (2) poor electricity supply to charge devices, and (3) workload issues that prevented FHWs from completing pre- and posttests. • Regarding internet connectivity and electricity supply, a minority of respondents reported how they temporarily stopped using VTR Mobile due, in part, to the lack of electricity to power the tablets and the lack of internet connectivity (previously enabled through satellite communication technology). Participants were unable to charge devices when rechargeable solar batteries installed at each participating health facility were drained and flat and went outside the facility to charge the device at a nominal cost.	• "Like that of tablet, there was a problem of network, and that of electricity that you are talking about. You know, there is no NEPA [referring to National Electric Power Authority, the former name of an organization governing the use of electricity in Nigeria] supply [ie, mains electricity] to the facility. So if you want to watch the video, you either charge the tablet, [and as there's not mains electricity], or they [the staff] will carry the tablet outside there [to the nearby town] and then I give them money [to pay to charge the device]." [Facility manager in FCT]
• A few health workers reported that the workload at the facility level prevented them from watching all VTR Mobile videos and completing the pre- and posttest promptly. However, it is unclear what proportion of late completers of pre- and posttests referred to in the quantitative section of this paper were affected by clinic workload and tight schedules. Policy makers outlined efforts to address workload issues through providing alternative access to the VTR Mobile app, for example, by supporting FHWs to install the app on personal Android phones to facilitate self-study at home.	• "[For FHWs struggling to complete the tests]...they [InStrat] have made it so easy that they can upload VTR on their phones [of FHWs]. During the pilot study, we encouraged them [FHWs] to get it. Those who have android phones that they can have it on their phones, so that it is not only when they get to the health facility where they have only one tablet that they can do it." [Policymaker in Ondo state]

[a]VTR: video training.

[b]FHW: frontline health worker.

[c]FCT: Federal Capital Territory.

[d]ANC: antenatal care.

[e]CHEW: community health extension worker.

Discussion

Principal Findings

Although digital health approaches have shown promise in improving health care provision in LMICs, they are infrequently implemented at scale [28]. This study focused on understanding the acceptability, feasibility, and potential effectiveness of an e-learning video–based intervention on MNCH transmitted at scale to FHWs in rural areas of Nigeria. We found that following the use of the e-learning intervention, FHWs demonstrated substantial improvements in their scores on a test of their

knowledge, attitudes, and reported practices about MNCH, indicating that the e-learning intervention is potentially effective in improving knowledge, attitudes, and practices in this area. Using the TAM to guide qualitative data analysis, we also identified five determinants of acceptance and three barriers to the use of VTR Mobile for training FHWs.

The five determinants of acceptance highlighted in the findings of this study were (1) the perceived ease of use of technology, (2) the perceived usefulness of clinical videos to enhance job performance, (3) access to tablet computers in the workplace, (4) the convenience of offline access and repeated use of training content, and (5) the perceived cost-effectiveness of VTR for FHW training. The latter three determinants of acceptance in Nigeria help to extend the classic TAM, which often prioritizes the perceived ease of use and usefulness of technologies as principal factors of acceptance. Qualitative data analysis showed how contextual factors such as previous training of FHWs to understand the e-learning technology, organizational support factors such as access to technology in the workplace, and technical support during the e-learning intervention increased the acceptance of and FHW confidence in using VTR to improve service delivery. In addition, our data also revealed how the convenience of offline access to training content at no cost to FHWs combined with the perceived usefulness of VTR to improve FHWs' knowledge and attitudes toward patients seemingly sparked the habitual use of clinical videos as reference material to guide live-saving procedures, which in turn increased FHW confidence to provide high-quality care. Furthermore, insight from Table 8 highlights how improved staff attitudes toward patients stimulated confidence in FHWs, which apparently generated a virtuous circle of increased FHW confidence to provide respectful care leading to service user confidence in FHWs and improved staff-patient relationships that subsequently boosted FHWs' confidence in providing respectful care. Conversely, three barriers that constrained the adoption and use of VTR Mobile in Nigeria were external factors that were the downstream of VTR technology. These structural factors were as follows: first, poor internet connection in a few health facilities served by the 3G mobile networks prevented FHWs from logging into the VTR Mobile platform, thereby limiting access to the training content. It is important to emphasize that accessibility to the internet was not constrained by affordability issues, as tablet computers used by FHWs were loaded with prepaid data plans to ensure seamless access to the e-learning platform. Second, poor electricity supply affected FHWs' ability to charge tablet computers mainly in health facilities located in rural areas and in facilities with empty and uncharged solar batteries. Third, organizational workload issues arising from technology introduction into the workflows of primary health centers limited FHWs' ability to complete the pre- and posttest surveys in some facilities. Taken together, the foregoing five determinants of acceptance and three barriers to the adoption and use of VTR broadly affected the effectiveness of the e-learning intervention in Nigeria.

Comparison With Previous Work

The distinctive features of this study are its identification of evidence of the potential effectiveness and feasibility of deploying digital health approaches for improving the knowledge, attitudes, and practices of FHWs at scale and to explain the possible mechanisms of the impact of VTR on staff performance and health system functions (service delivery and financing) in an LMIC. This study also addresses earlier calls for studies with larger sample sizes, more quantitative methods of evaluation, and exploration of implementation at scale while assessing the effectiveness of digital health approaches to train health care professionals in resource-limited countries [29-32]. The overall findings of our research align with those of previous digital health research in which multimedia content was delivered at scale to FHWs in Nigeria and other LMICs and which found that (1) FHWs had positive attitudes toward technology-enabled learning; (2) previous training and familiarity with technology increased usability; (3) digital health approaches were potentially effective for increasing FHWs' knowledge, attitudes, and care practices; and (4) digital health approaches also empowered workers with skills and confidence in contexts where technology adoption enhanced their performance and supported their work [14,29,30,33]. These findings have implications for developing strategies that ensure adequate orientation and continuing technical support for FHWs to adopt and use digital technology to achieve individual and organizational goals. Furthermore, these findings increase the evidence base underpinning the potential effectiveness of digital provision of training and educational content for health workers, which has previously lacked evidence to inform health workers' performance, skills, and attitudes [4].

Most digital health interventions to support task-shifting and community health worker (CHW) training in LMICs [34] used smartphones and basic feature phones [35]. Only a few e-learning interventions have adopted tablet-based apps [34,36]. However, evaluations of digital health projects to scale up CHW training in Pakistan [36] and India [37] have reported effective training interventions that increase CHW knowledge, motivation, and competence. For example, the *Sangoshthi* project, which scaled up CHW training to benefit more than 900,000 CHWs across India [37] recorded knowledge gains of 16% between pre- and posttest assessments. This is comparable with our findings from Nigeria, which showed a knowledge change of 17% between pre- and posttest scores; however, the *Sangoshthi* project was silent about the effects of its intervention on patient outcomes (micro or individual level), service delivery (meso or organizational level), and policy decisions (macro or wider system). In contrast, our study suggested that delivering video-based training at a scale can positively impact the micro, meso, and macro levels of the health system. The analysis of qualitative data showed that at the micro level, access to tablet devices, training opportunities, and ongoing technical support can increase staff confidence and motivation. At the meso level, better trained and skilled staff at intervention facilities felt empowered to deliver improved services, which manifested as respectful care, better management of complications, and enhanced patient engagement activities during antenatal clinic classes. This suggests that organizational contexts that provide essential equipment (in this case, tablet devices and clinical videos) to support the work of frontline staff can empower them to improve their performance. This further suggests that institutional readiness in human and infrastructural resources is necessary for e-learning; however, it is not always present in

LMICs [38]. At the meso level, we also found that the availability of clinical videos that incorporated vignettes with relevant questions sparked team-level discussions that provided opportunities for problem-solving, knowledge-sharing, and peer support. Finally, at the macro level, we saw how implementing a digital health approach to training can influence strategic policy decisions about which model of training to invest in scarce resources.

It is vital to underline that the macrolevel support was crucial for the benefits and impacts realized at the micro and meso levels of this project. A policy environment that enabled digital health care to function in the three states facilitated the successful adoption and implementation of SatCom and VTR mobile technologies. State-level ICT policies were aligned with the Federal Government of Nigeria's prioritization of ICTs as a strategy for achieving universal access to essential PHC services for mothers and their children [9]. Furthermore, this project is an example of thriving public-private partnerships between the government and local technology companies to develop digital health initiatives in Sub-Saharan Africa with initial financial inputs, technical know-how, and operational efficiency to ensure long-term adoption of digital health care. Such private sector involvement is essential for sustainable implementation if universal health coverage is achieved in Nigeria. Although the approach described in this paper has been successful, it relied on the use of SatCom technology alongside existing 3G mobile networks to overcome telecommunications challenges, as more than half of PHC facilities located in rural areas lacked internet connectivity. This further highlights the need to include the development of critical ICT infrastructure as a strategy to increase the quality of health care delivery in LMICs. Facilitating uninterrupted access to data and networks is a prerequisite to delivering approaches such as VTR Mobile, which have clear benefits for FHWs and recipients of their care. The application of space technologies (eg, SatCom) as used to overcome these challenges in this project required building partnerships with commercial partners. Although there are increasing examples of the application of space technologies such as SatCom technology in global health research, there is a need for improved awareness, training, and collaboration of the research community in such endeavors [39]. We have demonstrated the potential of PHCs to access education resources in the context of this project and suggest this as a priority area for digital health research in the context of LMICs.

Limitations

Although we provide evidence indicating the potential effectiveness of the intervention, our quantitative study design has several important limitations. First, an uncontrolled before-and-after comparison lacks the robustness and comparability of a randomized experimental design [40], which could compare an e-learning intervention to conventional face-to-face training. More specifically, uncontrolled before-and-after studies face threats to internal validity because of several possible biases. These include regression-to-the-mean bias, maturation bias (ie, changes in participants' cognitive abilities due to aging), attrition or loss to follow-up bias, retrospective bias, history bias (sometimes referred to as *secular effects*), and test-retest bias. We do not believe that our study was at risk of any substantial regression-to-the-mean bias because we did not restrict our recruitment of health workers based on any characteristics, but allowed all health workers within a selected facility to participate if they wished. It is also unlikely that our study would have experienced much maturation bias due to the relatively short time between the before-and-after comparisons. We also do not believe that we suffered any substantial attrition bias due to our high level of follow-up (328/341, 96.1%), and we avoided any retrospective bias by using a prospective design. We also believe that the duration of time between completion of pre- and posttest scores reduces the likelihood of history bias, with a limited opportunity for external initiatives to have influenced changes in participant scores. It is likely that our study suffered from test-retest bias, which implies that an unknown amount of the observed increase in test scores is probably due to participants remembering their pretest errors and correcting them, rather than having improved their knowledge via the intervention. Therefore, collectively, these risks must be considered while interpreting the results. Second, we only used pre- and posttest assessments of responses to questions on knowledge, attitudes, and reported practices related to MNCH to assess the effectiveness of the intervention. More objective and rigorous measures of performance and appropriate care behaviors would enhance the robustness of this or future evaluation. Third, although interview questions were developed and refined with key stakeholders and interviews conducted by experienced researchers, interview guides were not pilot-tested before use. Interview guides clarified the intended data to be derived from questions, but pilot testing may have refined the items used and improved the richness of data received from participants. Fourth, evaluation at multiple follow-up periods would provide useful information on the longevity of any intervention effects. Finally, the average cost of training described in this project excludes the cost of the SatCom component of the project, which made the project unsustainable for state governments in Nigeria. We did not evaluate the cost-effectiveness of the intervention as part of this study, which should be explored in future studies. However, the qualitative component did highlight that policy makers perceive tablet computers loaded with VTR Mobile as a means of reducing the costs associated with conventional training of FHWs.

Conclusions

This is the first report of combining SatCom with existing 3G mobile networks to support the VTR of FHWs at scale in LMICs. The study showed a widespread acceptance of VTR among FHWs with five determinants of acceptance in Nigeria: ease of use, perceived usefulness of VTR for improving service delivery, access to tablet computers in the workplace, convenience of offline access to training contents, and cost-effectiveness of VTR. The evaluation also demonstrated the potential effectiveness of the e-learning intervention in improving the knowledge, attitudes, and confidence of rural FHWs. It also identified the possible mechanisms of the impact of this e-learning approach at the micro, meso, and macro levels of the health system. Nonetheless, it raises questions about structural barriers to VTR adoption and use in areas that lack internet connectivity, experience poor electricity supply, and

increased FHW workloads arising from technology introduction. Policy strategies for improving workforce performance should create working environments that provide critical infrastructure to ensure an uninterrupted electricity supply and internet connectivity that support sustained access to reliable training content and enable workers to apply their knowledge and skills to deliver respectful and high-quality health care that promotes UHC.

Acknowledgments

The UK Space Agency (grant IPPC1-30) funded the study. All the views expressed in this publication are of the authors only. The authors also wish to acknowledge the contributions of the staff at the Federal and State Ministry of Health in Ondo State; Kano State; and Federal Capital Territory, Abuja, to the study design.

Authors' Contributions

BE and BO jointly conceived the study. BE, MJA, GOA, BO, and JT developed the manuscript. JPH, BE, and MJA wrote the manuscript, with contributions from BO, GOA, JT, KO, DA, AA, OD, RMY, OO, and TM. All authors read and approved the final version of the manuscript.

Conflicts of Interest

OO is cofounder and CEO of InStrat Global Health Solutions, the company that implemented the VTR technology used in this study. All authors declare no competing interests (JPH, MJA, BO, GOA, KO, DA, AA, OD, RMY, JT, OO, TM, and BE).

Multimedia Appendix 1
Items for pre- and posttest.
[DOCX File , 23 KB - mhealth_v9i9e24182_app1.docx]

Multimedia Appendix 2
Topic guide.
[DOCX File , 33 KB - mhealth_v9i9e24182_app2.docx]

References

1. Orton M, Agarwal S, Muhoza P, Vasudevan L, Vu A. Strengthening delivery of health services using digital devices. Glob Health Sci Pract 2018 Oct 10;6(Suppl 1):S61-S71 [FREE Full text] [doi: 10.9745/GHSP-D-18-00229] [Medline: 30305340]
2. Labrique A, Vasudevan L, Mehl G, Rosskam E, Hyder AA. Digital health and health systems of the future. Glob Health Sci Pract 2018 Oct 10;6(Suppl 1):S1-S4 [FREE Full text] [doi: 10.9745/GHSP-D-18-00342] [Medline: 30305334]
3. mHealth: use of appropriate digital technologies for public health: report by the Director-General. World Health Organization Institutional Repository for Information Sharing. URL: https://apps.who.int/iris/handle/10665/274134 [accessed 2021-08-27]
4. WHO Guideline Recommendations on Digital Interventions for Health System Strengthening. Geneva: World Health Organization; 2019.
5. Seventy-first world health assembly WHA71.7. World Health Organization. URL: https://apps.who.int/gb/ebwha/pdf_files/WHA71/A71_R7-en.pdf [accessed 2021-08-27]
6. Lee S, Cho Y, Kim S. Mapping mHealth (mobile health) and mobile penetrations in sub-Saharan Africa for strategic regional collaboration in mHealth scale-up: an application of exploratory spatial data analysis. Global Health 2017 Aug 22;13(1):63 [FREE Full text] [doi: 10.1186/s12992-017-0286-9] [Medline: 28830540]
7. Lee SH, Nurmatov UB, Nwaru BI, Mukherjee M, Grant L, Pagliari C. Effectiveness of mHealth interventions for maternal, newborn and child health in low- and middle-income countries: systematic review and meta-analysis. J Glob Health 2016 Jun;6(1):010401 [FREE Full text] [doi: 10.7189/jogh.06.010401] [Medline: 26649177]
8. Watterson JL, Walsh J, Madeka I. Using mHealth to improve usage of antenatal care, postnatal care, and immunization: a systematic review of the literature. Biomed Res Int 2015;2015:153402 [FREE Full text] [doi: 10.1155/2015/153402] [Medline: 26380263]
9. Assessing the enabling environment for ICTs for health in Nigeria: a review of policies. United Nations Foundation. 2014. URL: https://ehealth4everyone.com/wp-content/uploads/2016/01/nigeria-policy-report.pdf [accessed 2021-08-26]
10. Saving one million lives project (Federal Ministry of Health). The World Bank. 2017. URL: https://documents1.worldbank.org/curated/en/910491526905973918/pdf/Saving-One-Million-Lives-Project.pdf [accessed 2021-08-27]
11. Wilson K, Gertz B, Arenth B, Salisbury N. The journey to scale: moving together past digital health pilots. Program for Appropriate Technology in Health (PATH). URL: https://path.azureedge.net/media/documents/TS_dhs_journey_to_scale.pdf [accessed 2021-08-27]
12. Trends in maternal mortality: 1990 to 2015. World Health Organization. 2015. URL: https://www.unfpa.org/publications/trends-maternal-mortality-1990-2015 [accessed 2021-08-27]

13. Ebenso B, Allsop MJ, Okusanya B, Akaba G, Tukur J, Okunade K, et al. Impact of using eHealth tools to extend health services to rural areas of Nigeria: protocol for a mixed-method, non-randomised cluster trial. BMJ Open 2018 Oct 18;8(10):e022174 [FREE Full text] [doi: 10.1136/bmjopen-2018-022174] [Medline: 30341123]
14. Labrique AB, Wadhwani C, Williams KA, Lamptey P, Hesp C, Luk R, et al. Best practices in scaling digital health in low and middle income countries. Global Health 2018 Nov 03;14(1):103 [FREE Full text] [doi: 10.1186/s12992-018-0424-z] [Medline: 30390686]
15. National Human Development Report 2018: achieving human development in North East Nigeria. OCHA. 2018. URL: https://reliefweb.int/report/nigeria/national-human-development-report-2018-achieving-human-development-north-east-nigeria [accessed 2021-08-27]
16. Nkwo P, Lawani LO, Ubesie AC, Onodugo VA, Obu HA, Chinawa JM. Poor availability of skilled birth attendants in Nigeria: a case study of enugu state primary health care system. Ann Med Health Sci Res 2015;5(1):20-25 [FREE Full text] [doi: 10.4103/2141-9248.149778] [Medline: 25745571]
17. Training midwives in Somaliland using media. Medical Aid Films. URL: https://www.medicalaidfilms.org [accessed 2021-08-27]
18. Global Health Media. URL: https://globalhealthmedia.org [accessed 2021-08-27]
19. Connecting Frontline Health Workers to resources and each other to expand their knowledge, organize content into courses, and share their learning with the community. ORB. URL: http://www.health-orb.org [accessed 2021-08-27]
20. Mpowering frontline health workers. Partnership for Maternal, Newborn & Child Health. URL: https://www.who.int/pmnch/about/members/database/mpowering/en/ [accessed 2021-09-01]
21. Mobile healthcare solutions. InStrat Global Health Solutions. URL: http://instratghs.com [accessed 2021-08-27]
22. R: A Language and Environment for Statistical Computing. -: R Foundation for Statistical Computing; 2008.
23. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using. J Stat Soft 2015 Oct 7;67(1):-. [doi: 10.18637/jss.v067.i01]
24. emmeans: Estimated Marginal Means, aka Least-Squares Means. Cran Project. 2021. URL: https://cran.r-project.org/web/packages/emmeans/index.html [accessed 2021-08-27]
25. Ritchie J, Spencer L. Qualitative data analysis for applied policy research. In: The Qualitative Researcher's Companion. Thousand Oaks, CA: Sage Publication; 2002:305-329.
26. Portz JD, Bayliss EA, Bull S, Boxer RS, Bekelman DB, Gleason K, et al. Using the technology acceptance model to explore user experience, intent to use, and use behavior of a patient portal among older adults with multiple chronic conditions: descriptive qualitative study. J Med Internet Res 2019 Apr 08;21(4):e11604 [FREE Full text] [doi: 10.2196/11604] [Medline: 30958272]
27. Tani K, Exavery A, Baynes CD, Pemba S, Hingora A, Manzi F, et al. Unit cost analysis of training and deploying paid community health workers in three rural districts of Tanzania. BMC Health Serv Res 2016 Jul 08;16:237 [FREE Full text] [doi: 10.1186/s12913-016-1476-5] [Medline: 27391368]
28. Campbell JI, Aturinda I, Mwesigwa E, Burns B, Santorino D, Haberer JE, et al. The Technology Acceptance Model for Resource-Limited Settings (TAM-RLS): a novel framework for mobile health interventions targeted to low-literacy end-users in resource-limited settings. AIDS Behav 2017 Nov;21(11):3129-3140 [FREE Full text] [doi: 10.1007/s10461-017-1765-y] [Medline: 28421356]
29. O'Donovan J, Bersin A, O'Donovan C. The effectiveness of mobile health (mHealth) technologies to train healthcare professionals in developing countries: a review of the literature. BMJ Innov 2014 Dec 12;1(1):33-36. [doi: 10.1136/bmjinnov-2014-000013]
30. Agarwal S, Perry HB, Long L, Labrique AB. Evidence on feasibility and effective use of mHealth strategies by frontline health workers in developing countries: systematic review. Trop Med Int Health 2015 Aug;20(8):1003-1014 [FREE Full text] [doi: 10.1111/tmi.12525] [Medline: 25881735]
31. Long L, Pariyo G, Kallander K. Digital technologies for health workforce development in low- and middle-income countries: a scoping review. Glob Health Sci Pract 2018 Oct 10;6(Suppl 1):S41-S48 [FREE Full text] [doi: 10.9745/GHSP-D-18-00167] [Medline: 30305338]
32. Barteit S, Jahn A, Banda SS, Bärnighausen T, Bowa A, Chileshe G, et al. E-Learning for medical education in Sub-Saharan Africa and low-resource settings: viewpoint. J Med Internet Res 2019 Jan 09;21(1):e12449 [FREE Full text] [doi: 10.2196/12449] [Medline: 30626565]
33. Abejirinde IO, Ilozumba O, Marchal B, Zweekhorst M, Dieleman M. Mobile health and the performance of maternal health care workers in low- and middle-income countries: a realist review. Int J Care Coord 2018 Sep;21(3):73-86 [FREE Full text] [doi: 10.1177/2053434518779491] [Medline: 30271609]
34. Winters N, Langer L, Nduku P, Robson J, O'Donovan J, Maulik P, et al. Using mobile technologies to support the training of community health workers in low-income and middle-income countries: mapping the evidence. BMJ Glob Health 2019 Jul 30;4(4):e001421 [FREE Full text] [doi: 10.1136/bmjgh-2019-001421] [Medline: 31413872]
35. Ghana telemedicine. Novartis. URL: https://www.novartisfoundation.org/our-work/reimagining-healthcare-through-digital-technology/ghana-telemedicine [accessed 2021-08-27]

36. Rahman A, Akhtar P, Hamdani SU, Atif N, Nazir H, Uddin I, et al. Using technology to scale-up training and supervision of community health workers in the psychosocial management of perinatal depression: a non-inferiority, randomized controlled trial. Glob Ment Health (Camb) 2019 May 16;6:e8 [FREE Full text] [doi: 10.1017/gmh.2019.7] [Medline: 31157115]
37. Yadav D, Singh P, Montague K, Kumar V, Sood D, Balaam M, et al. Sangoshthi: empowering community health workers through peer learning in rural India. In: Proceedings of the 26th International Conference on World Wide Web. 2017 Presented at: WWW '17: Proceedings of the 26th International Conference on World Wide Web; Apr 3-7, 2017; Perth Australia. [doi: 10.1145/3038912.3052624]
38. Frehywot S, Vovides Y, Talib Z, Mikhail N, Ross H, Wohltjen H, et al. E-learning in medical education in resource constrained low- and middle-income countries. Hum Resour Health 2013 Feb 04;11:4 [FREE Full text] [doi: 10.1186/1478-4491-11-4] [Medline: 23379467]
39. Dietrich D, Dekova R, Davy S, Fahrni G, Geissbühler A. Applications of space technologies to global health: scoping review. J Med Internet Res 2018 Jun 27;20(6):e230 [FREE Full text] [doi: 10.2196/jmir.9458] [Medline: 29950289]
40. Ho AM, Phelan R, Mizubuti GB, Murdoch JA, Wickett S, Ho AK, et al. Bias in Before-After Studies: Narrative overview for anesthesiologists. Anesth Analg 2018 May;126(5):1755-1762. [doi: 10.1213/ANE.0000000000002705] [Medline: 29239959]

Abbreviations

CHEW: community health extension worker
CHW: community health worker
FCT: Federal Capital Territory
FHW: frontline health worker
ICT: information and communications technology
IDI: in-depth interview
LMIC: low- and middle-income country
MNCH: maternal, newborn, and child health
PHC: primary health care
SatCom: satellite communication
TAM: technology acceptance model
VTR: video training
WHO: World Health Organization

Edited by R Kukafka, G Eysenbach; submitted 07.09.20; peer-reviewed by KC Wong, A Odutola; comments to author 07.10.20; revised version received 12.11.20; accepted 01.08.21; published 16.09.21.

Please cite as:
Hicks JP, Allsop MJ, Akaba GO, Yalma RM, Dirisu O, Okusanya B, Tukur J, Okunade K, Akeju D, Ajepe A, Okuzu O, Mirzoev T, Ebenso B
Acceptability and Potential Effectiveness of eHealth Tools for Training Primary Health Workers From Nigeria at Scale: Mixed Methods, Uncontrolled Before-and-After Study
JMIR Mhealth Uhealth 2021;9(9):e24182
URL: https://mhealth.jmir.org/2021/9/e24182
doi:10.2196/24182
PMID:34528891

Tutorial

Development of Digital Health Messages for Rural Populations in Tanzania: Multi- and Interdisciplinary Approach

Christine Holst[1], MPhil; Ghislain Maurice Norbert Isabwe[2,3], PhD; Felix Sukums[4], PhD; Helena Ngowi[5], VMD, PhD; Flora Kajuna[5], MA, VMD; Danica Radovanović[6], PhD; Wisam Mansour[6], MA; Elibariki Mwakapeje[7], PhD; Peter Cardellichio[8], PhD; Bernard Ngowi[9,10], MD, PhD; Josef Noll[6,11], PhD; Andrea Sylvia Winkler[1,12], MD, PhD

[1]Centre for Global Health, Department of Community Medicine and Global Health, Institute of Health and Society, University of Oslo, Oslo, Norway

[2]Future Competence International, Kigali, Rwanda

[3]Department of Information and Communication Technology, University of Agder, Grimstad, Norway

[4]Directorate of Information and Communication Technology, Muhimbili University of Health and Allied Sciences, Dar es Salaam, United Republic of Tanzania

[5]Department of Veterinary Medicine and Public Health, Sokoine University of Agriculture, Morogoro, United Republic of Tanzania

[6]Basic Internet Foundation, Kjeller, Norway

[7]Epidemiology and Diseases Control Section, Ministry of Health, Community Development, Gender, Elderly and Children, Dodoma, United Republic of Tanzania

[8]Global Health Media Project, Waitsfield, VT, United States

[9]Muhimbili Medical Research Centre, National Institute for Medical Research, Dar es Salaam, United Republic of Tanzania

[10]Mbeya College of Health and Allied Sciences, University of Dar es Salaam, Mbeya, United Republic of Tanzania

[11]Department of Technology Systems, University of Oslo, Oslo, Norway

[12]Center for Global Health, Department of Neurology, Technical University of Munich, Munich, Germany

Corresponding Author:
Christine Holst, MPhil
Centre for Global Health
Department of Community Medicine and Global Health, Institute of Health and Society
University of Oslo
Postbox 1130
Blindern
Oslo, 0318
Norway
Phone: 47 48234044
Email: christine.holst@medisin.uio.no

Abstract

Background: Health workers have traditionally delivered health promotion and education to rural communities in the Global South in paper leaflet formats or orally. With the rise of digital technologies, health promotion and education can be provided in innovative and more effective formats, which are believed to have a higher impact on disease prevention and treatment.

Objective: The aim of this tutorial is to illustrate how a multi- and interdisciplinary approach can be applied in the design process of digital health messages for use in the Global South.

Methods: The multi- and interdisciplinary team of the Non-discriminating access for Digital Inclusion (DigI) project digitalized and customized available government-approved paper-based health promotion messages into a screen-suitable format. The team worked closely together and used its diverse expertise to develop digital health messages with disease-specific content in Tanzania's national language (Swahili) as well as English. The development process included the following phases: a local needs assessment; identification of government-approved health promotion materials in a nondigital format; identification of key health messages; creation of a practical and engaging story, easy to understand for the general public; drafting of a storyboard for an animated video with review, feedback, and revisions; forward and backward translation; audio recording of the story in both languages; finalization and presentation of the animations; development of relevant questions related to the health messages in each domain; and development of web and mobile apps to access the digital health messages.

Results: Between 2017 and 2019, we developed key health messages, quizzes, and animated health videos to address HIV/AIDS, tuberculosis, Taenia solium cysticercosis and taeniasis, and anthrax, all of which are of public health importance in Tanzania. Feedback from local stakeholders and test users was included in various phases of the process. The 4 videos and other content are available in local information spots on a digital health platform (DigI platform), established by the DigI project, in both Tanzanian Swahili and English.

Conclusions: Our methodological multi- and interdisciplinary approach ensures that the digital health messages for the public are clear, high quality, and align with the government's objectives for health promotion. It also demonstrates the diversity of scientific disciplines required when collaborating on a digital health project. We recommend this approach to be applied to the development of other digital health messages for a wide range of diseases.

International Registered Report Identifier (IRRID): RR2-10.2196/25128

(JMIR Mhealth Uhealth 2021;9(9):e25558) doi:10.2196/25558

KEYWORDS

digital health; eHealth; mHealth; Tanzania; health education; HIV/AIDS; tuberculosis; cysticercosis; tapeworm; anthrax; mobile phone

Introduction

Background

The World Health Organization (WHO) emphasizes the use of digital technologies to enable people to access information, goods, and services to improve their lives [1]. Sustainable Development Goals number 3 (good health and well-being) [2] and number 9 (industries, innovation, and infrastructure) [3] are both linked to digital health. More specifically, this is reflected by strengthening the capacity for early warning, risk reduction, and management of national and global health risks (target 3.D) and significantly increasing access to information and communication technology (target 9.C). Digital health is connected to multiple disciplines and critically depends on the country and the health system context [4]. Multidisciplinary research is about coordinating efforts that bring several disciplines together to provide complementary contributions in the service of a common goal [5] and is essential in the progress of any high-quality cross-cutting research project today. Interdisciplinarity relates to analyzing, synthesizing, and harmonizing the efforts into a common coordinated and coherent entity [6], and has previously been defined as "a mode of research by teams or individuals that integrates information, data, techniques, tools, perspectives, concepts, and/or theories from two or more disciplines or bodies of specialized knowledge to advance fundamental understanding or to solve problems whose solutions are beyond the scope of a single discipline or area of research practice." [7]

It is well known that health promotion and participation in health promotion activities will be revolutionized by digital innovations and that health literacy increases when people begin using digital tools to access information and make informed decisions related to their own health [8]. However, low levels of literacy and digital skills are barriers to accessing digital information, especially in rural areas of low-income countries [9]. Thus, it is important that the design of a digital solution is inclusive and ensures that users develop the skills needed to take advantage of such digital opportunities [10].

Designing digital health messages for a Global South audience can be challenging considering the abovementioned shortfalls in literacy, digital literacy, access to internet, and access to devices. Power structures, poverty, gender inequality, and health inequities at large, including poor infrastructure, complicate the situation for several groups, including those marginalized. However, the future of health education and health promotion participation in sub-Saharan Africa may very well be digital [11]. The importance of early partnering with local stakeholders and users in a co-design process is crucial for success [12]. In the next section, we will present our project exemplified by a case study from Tanzania, before illustrating how we developed content for a global digital health intervention using a multi- and interdisciplinary approach.

The Non-Discriminating Access for Digital Inclusion Project: Digital Health Promotion and Education in Tanzania

In Tanzania, community health workers and health facilities, in addition to nongovernmental organizations, have traditionally provided health promotion and education orally or in printed format (leaflets, billboards, banners, posters, etc) to neighborhoods and communities [13]. The Tanzanian digital landscape is evolving rapidly, and together with significant economic growth [14], we have witnessed increasing penetration of mobile phones in both urban and rural areas and one of the most advanced mobile money markets in sub-Saharan Africa [15]. Tanzania, East Africa's third largest economy with 54.2 million people, had 43.6 million mobile phone subscribers as of June 2018 (80% of the population), compared with 40.1 million a year earlier and in contrast to only 10.2 out of 40.7 million people (25% of the population) in 2007 [16]. However, only 13% of subscribers reported to own a smartphone in 2017 [17]. The Tanzania Digital Health Strategy 2019-2024 seeks to expand the use of digital health technologies to promote healthy behavior through access to health information, education, and communication [18]. The Tanzanian government intends to scale up and intensify the achievements of the previous eHealth Strategy 2013-2018, which implemented various electronic information systems (including the use of television, radio, social media, etc) to provide and promote health education. Although community-based health services focus on health promotion and disease prevention, mobile health and social media can be more widely used to provide quality health

education, information, and communication, enabling communities to adopt healthier behaviors and to increase their health literacy. Impactful health messages are clearly needed in the global fight against infectious diseases such as HIV/AIDS, tuberculosis (TB), and zoonosis (diseases that affect animals and humans) such as Taenia solium cysticercosis and taeniasis (TSCT) and anthrax. The Tanzanian Health Sector Strategic Plan states that the disease control programs for HIV/AIDS and TB have been successful in the areas of early detection and treatment. However, further improvements are needed in the field of prevention [13]. As for TSCT, which is recognized as a neglected tropical disease by the WHO [19], prevention is crucial because the disease is highly endemic in Tanzania [20]. Anthrax is an emerging neglected zoonotic disease, endemic in the African region and present in the north of Tanzania (especially in the Maasai communities) where repeated outbreaks in livestock, wildlife, and humans have made clear the need to strengthen prevention strategies [21].

The Non-discriminating access for Digital Inclusion (DigI) project, in which the development of health messages described in this paper is part, represents a classical multi- and interdisciplinary project funded as an innovation project, with the aim of connecting rural villages to digital information. The DigI project's multi- and interdisciplinary team, comprising 11 partners from 8 countries, includes scientists in the disciplines of humanities, social sciences, formal sciences, and applied sciences. These main disciplines include a variety of subdisciplines: medicine and health, epidemiology, public health, veterinary medicine, electronic engineering, visual and creative arts with graphic and interaction design, human-computer interaction, communication, education or educational technology, information science, internet science, and ethics. Elements from anthropology, human geography, health psychology, and sociology of health and illness have been applied in the various steps of planning and implementation of this project. Through our collaboration of Tanzanian, German, Norwegian, Rwandan, and American partners, our objective is to develop digital health messages by converting printed materials into digital formats such as animated videos (animations), quizzes, graphics, and texts. The majority of the DigI partners either belonged to the information technology (IT) and design task force, or the health task force. The Tanzanian and German project partners with backgrounds in health and medicine provided expertise on HIV/AIDS, TB, TSCT, and anthrax. The team was able to draw on their many years of experience working with each disease in the target areas. The selected diseases in the DigI project are endemic in the chosen geographical areas and thus are important diseases in the Tanzanian public health context. All these diseases are of high priority and require preventive strategies including information dissemination to communities and populations.

A digital health education platform, the DigI platform [22], herein referred to as *the platform* was planned and developed between 2017 and 2019 by the DigI team. It consists of a dashboard (Multimedia Appendix 1) where viewers can navigate between various pages with information about the diseases. The platform features texts, quizzes, graphics, and animated videos that are available for all. The platform with the health messages described in this paper was established as part of an information spot (InfoSpot) in the rural villages of Selela and Esilalei in the Arusha region (for anthrax) and Migoli and Izazi in the Iringa region (for HIV/AIDS, TB, and TSCT) in November 2019. It is possible to access the health messages on the platform, even if the user is not connected to the internet, because the platform is also locally stored on a village server.

In this tutorial, we want to share our contextualization of digital health message creation for a broader audience to illustrate how it is beneficial to take advantage of the various backgrounds in a multinational and multi- and interdisciplinary digital health project. Each step of the process is illustrated with a general takeaway message and lessons learned from our project.

Methods

Definition of Terms

In this paper, the terms health promotion, health education, health communication, health information, health literacy, health message, and digital literacy are used. Here, health education and health promotion are 2 terms that are sometimes used interchangeably [23], whereas health information and health communication are broad terms simply referring to all information and communication related to health. Moreover, health literacy can be explained as a result of successful campaigns of all the aforementioned terms, and health messages are pieces of information designed for health behavior change. In Multimedia Appendix 2 [23-32], we have provided definitions of the terms to offer the reader a more detailed explanation of their differences.

Development of the Digital Content

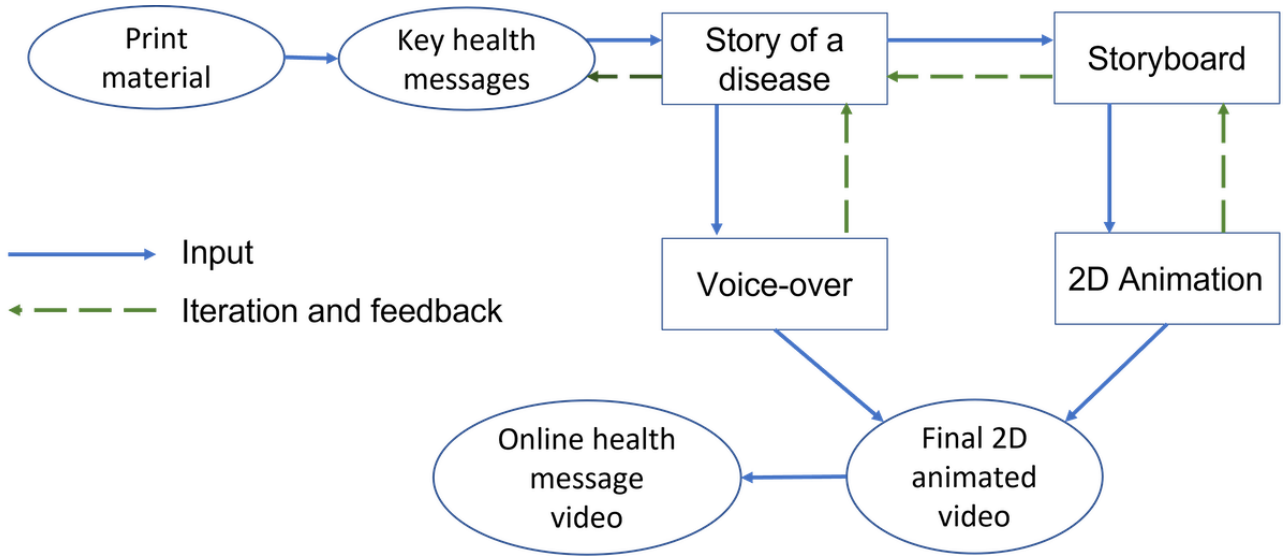
Overview

On the basis of the information in the print materials, the digitization process began by creating a narrative story capturing the key health messages, with elements of storytelling to maintain the *clients'* attention. *Clients* is defined as "members of the public that are potential or current users of health services, including promotion activities" [33]. The stories were adapted to a rural Tanzanian village-life setting, making them valid and authentic for the clients, with an emphasis on keeping the stories simple, truthful, and emotional at times. On the basis of these imaginary stories, storyboards (drawings with captions below each image or scene) were produced to visualize the narrative so that the DigI team could provide more informed feedback. Once the storyboards were approved, the next step consisted of developing animated videos with voiceover narrations in both English and Swahili. The DigI team iterated a series of design-test-evaluate-design cycles, whereby a story was tested and improved accordingly. The digitization process is illustrated in Figure 1.

The exact approach to content creation was established based on the internal discussions. We aimed to meet clients' needs, values, language, and culture and emphasized the rural Tanzanian context throughout the development process.

In the following 10 sections, each phase of our methodological approach is described in detail, with general advice and takeaway messages for each section.

Figure 1. Health message digitization process.



A Local Needs Assessment

A human-centered design process [34] is required when developing apps to meet user requirements. In our project, a survey was conducted at the very beginning to gather users' needs in the project villages, and a participatory and community-driven approach was deliberately preferred to the introduction of the digital health project from outside [35]. Before developing the digital health messages, several site visits were undertaken by health- and IT-researchers in the project. During these visits, local needs and wishes were gathered through discussions with local stakeholders, such as village leaders, health workers, and other community members. During the needs assessment, it became clear to the DigI team that the preferred health education format was animation. The locals explained the difficulties with low literacy in rural societies, suggesting cartoons or films for learning about health. Tanzanian Swahili was obviously a prerequisite, as the vast majority of people in Tanzania do not speak English. Although some ethnic groups have their own languages, most people understand Tanzanian Swahili. Some village stakeholders also expressed the need for other information, such as village news, village meetings, and agriculture information.

Our general advice based on our experience is to spend substantial time exploring the local needs and involving local stakeholders as partners to form the process and shape the end product from the very beginning.

Identification of Government-Approved Health Promotion Materials in a Nondigital Format

Health strategies and guidelines may vary from one country to another, and it is important to pay attention to the efforts that have already been laid down by policy makers, health workers, and researchers at a national level when mapping the local environment. Our video animations represent up-to-date, high-quality, and clear health messages that are based on approved health information materials, as a point of departure for digitization. The messages are taken from leaflets, posters, brochures, banners, guidelines, and strategies, all carefully reviewed for key health messages aimed at a public health audience (Figure 2). We used the Ministry of Health, Community Development, Gender, Elderly and Children–approved health education messages for HIV and TB. HIV health education messages were obtained from the National AIDS Control Programme [36]. In addition, posters and presentations were obtained from the National Tuberculosis and Leprosy Control Programme, whereas the National Strategic Plan V (2015-2020) for Tuberculosis and Leprosy [37] were used as source documents. The health educational material for TSCT was provided by the Cysticercosis Working Group in Eastern and Southern Africa, and the United Republic of Tanzania One Health Strategic Plan 2015-2020 [38] was used for inspiration. The leaflets for the prevention and control of anthrax in humans and animals were jointly prepared by the Ministry of Health, Community Development, Gender, Elderly and Children, Ministry of Livestock and Fisheries, and the Ministry of Natural Resources and Tourism supported by the Food and Agriculture Organization of the United Nations and WHO country offices in Tanzania, using a One Health approach. The hard copies were converted to soft copies and shared with the DigI team for conversion into digital formats.

From this step, we recommend exercising caution and respect toward the national health authorities. National strategies and guidelines may contain important health messages to be conveyed, especially to target populations, because of context-specific reasons, and may differ from international and global guidelines.

Figure 2. Examples of the education material used as a point of departure for digitization (Taenia solium cysticercosis and taeniasis on the left and tuberculosis on the right). Sources: Cysticercosis Working Group in Eastern and Southern Africa and Ministry of Health, Community Development, Gender, Elderly and Children, respectively.

Identification of Key Health Messages

Establish the most important key messages that you want to convey, depending on the health topic and target population. In our project, we dealt with infectious diseases for a rural population in Tanzania, thus dividing our key health messages into the following domains: (1) prevalence of the disease, (2) cause or transmission of the disease, (3) signs or symptoms of the disease, (4) treatment of the disease, and (5) prevention of the disease.

Once the domains were identified, we extracted key health messages from the materials described in the previous section. In the course thereof, short paragraphs, emphasizing the most important messages for each domain, were created. This is transferable to all health topics: to keep it short and simple.

The first domain intends to raise local awareness and includes health messages regarding the disease and its prevalence in the area. The second domain contains information on how diseases spread and infect others. The third domain describes the signs and symptoms of the disease, to help with early detection and management of the disease. Health messages encourage and motivate clients to seek medical advice if symptoms appear. The fourth domain is related to the treatment of diseases. The last and perhaps most important domain includes information on how people can protect themselves and their families and how the disease can be prevented from spreading between individuals and within communities.

This work established the basis for the written key health messages on the platform and in the animated health videos and moreover identified the most important information that the DigI team health task force wanted to provide to clients. The key health messages typically consisted of 50-150 words and were available on the platform. An example of a key health message (from the domain of symptoms in the TB section) is as follows:

> *Most TB patients show the following signs and symptoms: cough for more than 2 weeks, fever for more than 2 weeks, weight loss, night sweats and lymph node enlargement. People with these signs and symptoms should immediately visit a health facility for proper TB diagnosis and treatment.*

Our general advice based on our experience is to rather focus on few key messages at the time and keep the messages short and simple so as to not overload the audience with information.

Creation of a Practical and Engaging Story, Easy to Understand for the General Public

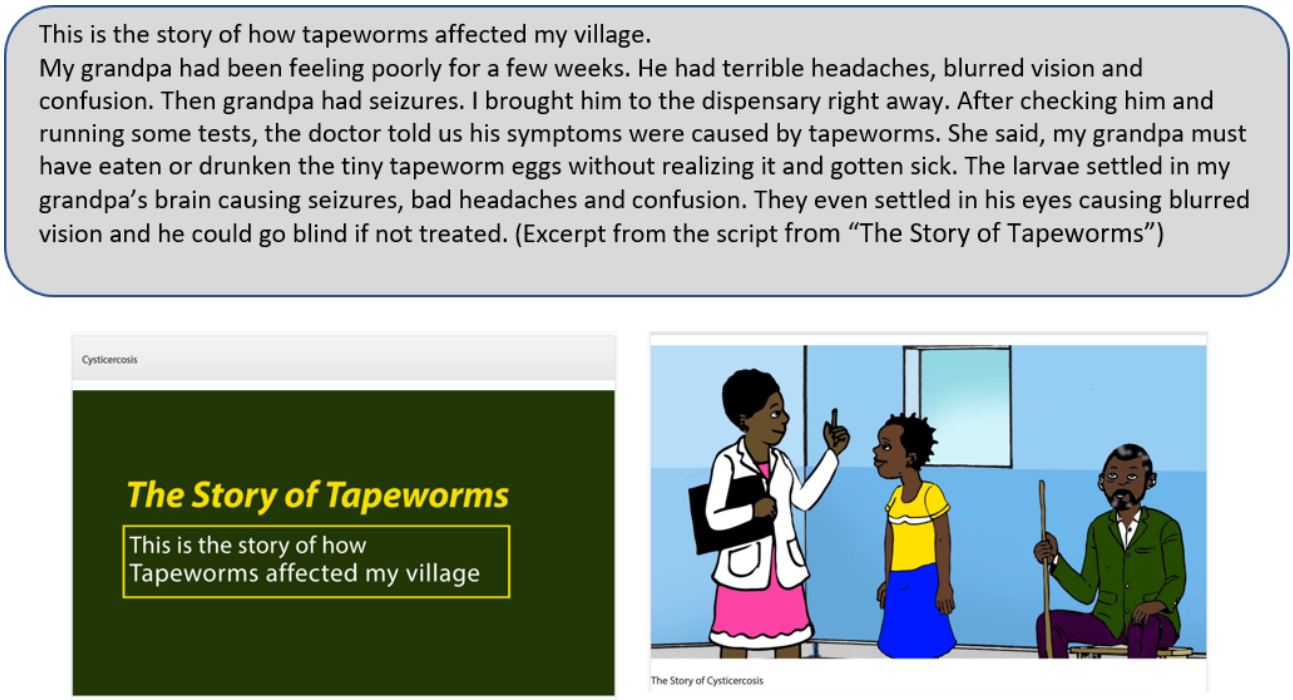
Throughout time, people have learned from stories. Stories can be a powerful way to pass on knowledge and specific messages to other people [39]. In our project, we used storytelling techniques to draft a compelling everyday narrative story that clients could recognize and relate to. All animations were cooperatively developed within the DigI team; however, some partners led the process of writing the story. The draft was discussed by the DigI team. Thereafter, health content was assured by the health task force. After drafting the content of the story, our team worked on creating a word document that contained the script for each story. In our case, the script comprised facts about the disease and a creative mixture of characters, actions, and location settings within the context of a village in rural Tanzania. The scripts were carefully revised by Tanzanian DigI members to ensure local compatibility.

The stories are presented by a narrator and include references to family members as well as the community affected by the diseases (see TSCT example in Figure 3). The narrative includes elements to attract and maintain clients' attention by describing

everyday activities and the realities of life in a rural sub-Saharan community.

Our general advice based on our experience from the story creation is to include a variety of stakeholders to revise the scripts. Specialists and locals can contribute equally, but with different perspectives. Specialists want to ensure that the health content is being presented in an adequate way and that no key messages are left out, whereas locals can point out words and phrases that are difficult to understand. It can be useful to ask questions such as “Are all key messages promoted clearly? Is the language understandable and the story credible?”

Figure 3. Excerpt from “The Story of Tapeworms”.

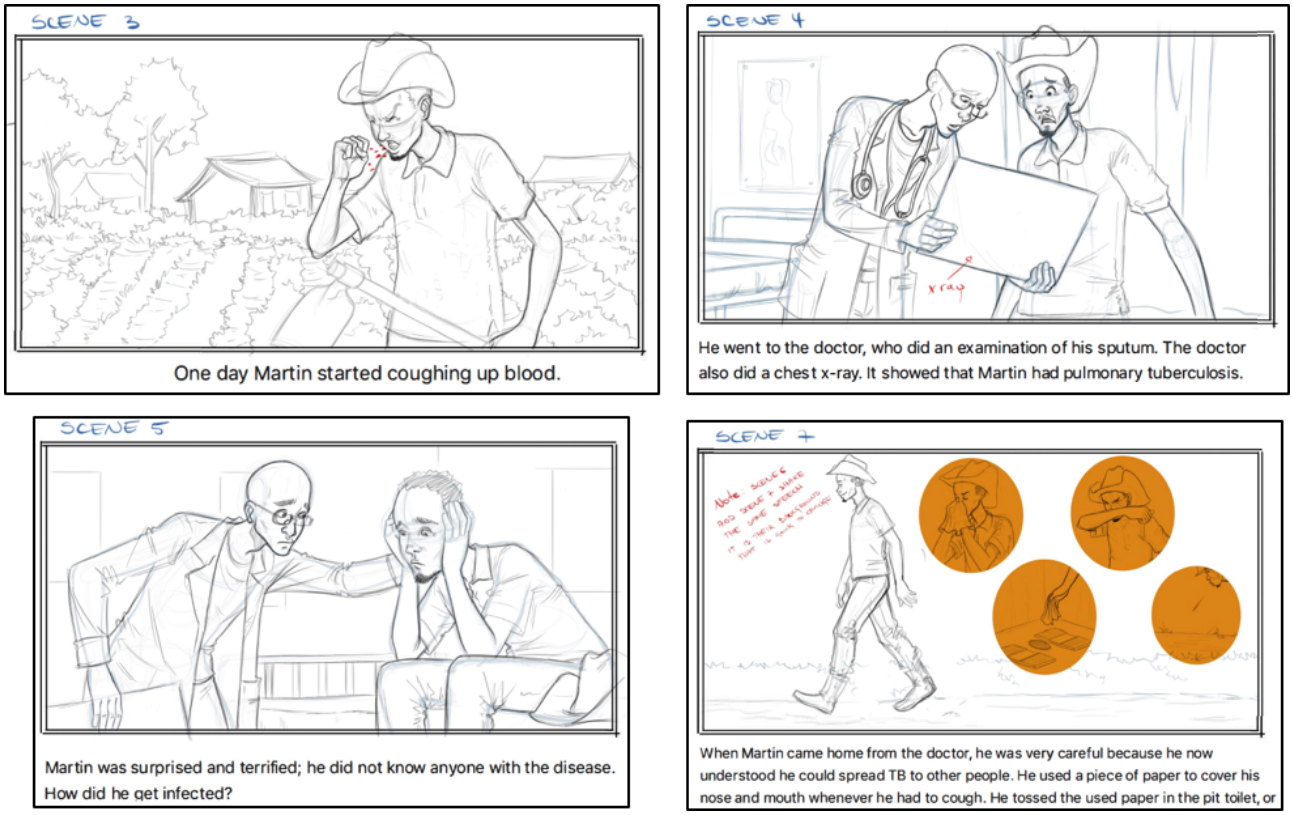


Drafting of a Storyboard for an Animated Video with Review, Feedback, and Revisions

After the script is approved, a storyboard needs to be created if the aim is to create a film or animation. A storyboard is a drawing with text below, illustrating scenes such as farming scenes that may be a common experience for rural clients (Figure 4). It is advisable to share the storyboards with the target population early on to obtain feedback in this design phase. In our project, we planned to use simplified human line drawings but were advised by the local population to incorporate more genuine characters to get the communities more involved when viewing the animations. We also received feedback from the same group to make the characters and environment more similar to those found in rural Tanzania. This included, for example, changing the façade of the health facilities and the clothes of the characters. The core health messages from the written material were left untouched, as these messages were approved, and changing them would require new ethical approval. Following an iterative process of design-test-evaluate-design, an animated video was created based on the original script, revisions, and feedback from the DigI team members and local stakeholders.

Our general advice based on our experience is to establish the storyboards as soon as the script is approved. The storyboards make the animation visualization more realistic and can be effective in identifying scenes of importance for key message uptake. It is beneficial to include target users in this stage to point out recognizability and authenticity, or lack thereof, in different scenes.

Figure 4. Extracts from the storyboard used in the creation of "The Story of Tuberculosis." TB: tuberculosis.



Forward and Backward Translation

Working in 2 languages can be challenging and may complicate the process. Hence, qualified translators with experience in the field are needed to ensure that translation is consistent throughout the various versions of the script. In our project, the health messages and story scripts were created in English and then translated to Swahili for the target audience. The process was as follows: English health messages and story scripts were given to a person to translate them into Swahili. Thereafter, the Swahili versions were given to another person to translate them back into English. Further, the back translation was compared with the original English health messages and story scripts to ensure that the messages were the same. The Swahili messages were also read by the medical doctors who spoke Swahili to ensure that the messages had the same information as the English messages.

Our experience indicated that alterations of the script were done several times, also after translation; thus, it is important to ensure that the final script represents the same key messages in both languages.

Audio Recording of the Story in Both Languages

When the script and translations are final, it is time to record the story. Although it is recommended to use a professional recording studio, the most important factor is to be able to present the audio file without background noise and with a clear and understandable narrator. In our project, we used professional narrators, and voiceover recordings were produced in both English and Swahili. Technically, recording itself was a straightforward process. It was emphasized that the narrators spoke slowly and clearly in an informal tone. The audio file was carefully revised by team members who were Swahili natives and near-natives in English. We recommend paying particular attention to the initial revision process of the audio file. There are cases where a written sentence conveys health messages more effectively than when reading aloud, and to save costs, it may be useful to read the script aloud to team members before recording it.

Finalization and Presentation of the Animation

Once the voice recordings and animation scenes are ready, the animated video can be fully produced and shared with the team for revision and improvement. In our case, all team members carefully reviewed the English version of the draft video. After at least 2 rounds of revisions, the final animated videos were produced and then presented to an audience in rural Tanzania. The *Story of Tapeworms* animation was completed in 2018. The *Story of HIV/AIDS* animation and the *Story of TB* animation were finalized in March 2019. The *Story of Anthrax* animation was completed in May 2019. The animations can be found in Multimedia Appendices 3-10.

Generally, this phase can be costly, but it is worthwhile to address small changes in the final product, which can be of importance for knowledge transfer.

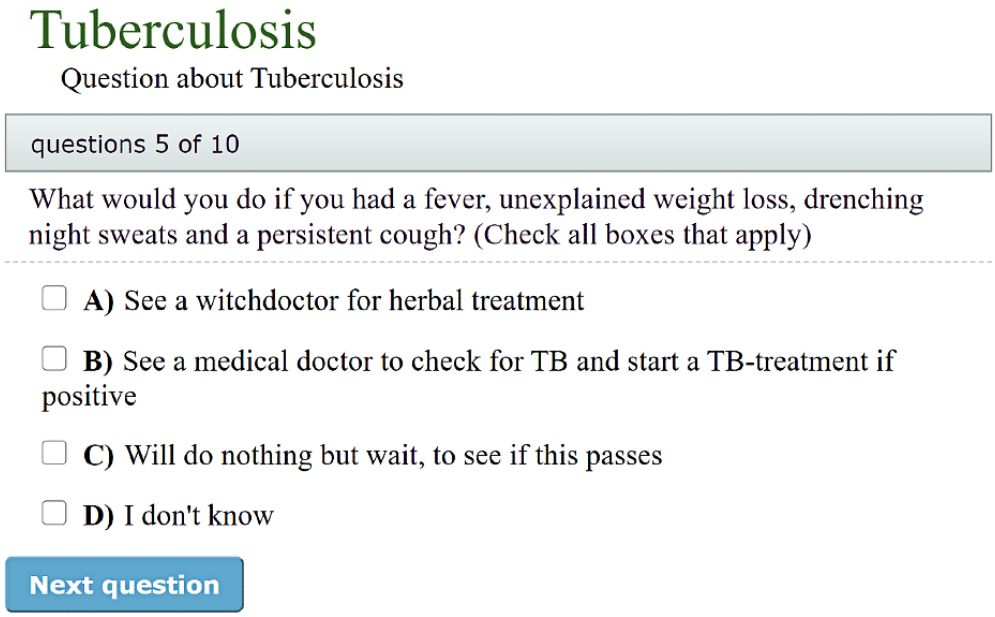
Development of Relevant Questions Related to the Health Messages in Each Domain

To increase knowledge uptake, it may be useful to establish questions from digital health messages, such as animations, and present them to the audience in a quiz format, for example, after the audience has viewed the animation. We derived knowledge questions related to different diseases from the key health messages for each domain. These questions were constructed

for research purposes to test people's knowledge before and after being exposed to our digital health messages in a quantitative study. The same questions were published on the platform in a quiz format so that users could test their own learning. We ensured that the questions did not use medically advanced wording or information. In particular, the health task force assessed their quality as they related to the health content, and the social scientists made the questions easier to understand for the clients. An example of a quiz question is shown in Figure 5.

Our experience is that the questions created should not be too difficult, but should reflect the health messages, and work as reinforcers of the information provided in the animations.

Figure 5. Sample questions presented on the platform. TB: tuberculosis.



Development of Web and Mobile Apps to Access the Digital Health Messages

Access to information can be a bottleneck in disease prevention in the Global South. Animations and other digital formats for health education in the local language are only useful if they are accessible to the target audience. Hence, in the DigI project, it was equally important to develop an easy-to-use digital platform so that people in rural areas could access health messages. Key requirements for the success of the platform included that it would be easy to learn *how to use* it and that it could be used with a basic level of digital literacy (Multimedia Appendix 2). Given that most people are likely to access health messages on a mobile phone, we adopted the *mobile-first* design approach, creating a responsive web application and an Android-based mobile app. The web application presents health messages interactively, whereby a user is presented with an animated video, key health messages, graphics, and a quiz.

As part of the human-centered design process [34], the project team specified the requirements of web and mobile app systems based on inputs from village stakeholders in Migoli and Izazi in the Iringa region. The proposed designs were iteratively improved from a conceptual design to low-fidelity prototypes and high-fidelity interactive prototypes using Adobe XD [40]. The low-fidelity prototypes were simple sketches and wireframes to present early user interface designs for discussions in the DigI team. These helped in gathering user feedback and suggestions on what could be improved. Furthermore, high-fidelity prototypes presented more enriched user interfaces, with interactivity for the test users to experience the interaction with the designed tool. In addition to testing and evaluation by the medical task force and other team members in Tanzania, user-based testing was conducted using the IT task force to uncover potential usability problems. User testing focuses on efficiency, effectiveness, and overall user satisfaction. Using purposive sampling, a total of 12 test users, aged 15-25 years, participated in 2 rounds of usability testing conducted in Kigali (Rwanda) as part of an iterative design process. All test participants had postprimary school education, but none had completed secondary school. The first high-fidelity prototypes were tested by 8 participants (3 women and 5 men) in October 2018, and the second improved versions of the web and mobile app were tested by 4 participants in July 2019. The findings indicated that it was easy to access health information on the platform, view the videos, and read key health messages for different diseases. Test users reported that it was easy to navigate the system and that the content was logically structured. However, some respondents said that "it was hard to find the videos" in the early versions of the platform. Hence, global navigation was redesigned, adding on the top menu a direct link to a page with videos for all diseases. Test participants also suggested making two separate videos for each disease: one video for a story and another with only the health messages. However, after deliberations, the project team decided to keep 1 integrated animation per disease, not too long to keep the audience's attention. Furthermore, test users suggested additional requirements, such as the possibility of self-diagnosis through search functionality on symptoms. This indicated the user's perception of the platform as a health system. Subsequently, we used design constraints to prevent users from attempting to do anything that the system is not intended for,

to make clear that the goal of the platform is to provide health information rather than medical diagnosis. It is in this regard that global navigation includes links to *Health Information*, including videos, text, and quizzes, in addition to a contact form, but there is no search bar. However, given the small screen size of mobile devices, the mobile app includes a search field to help users easily access information on the diseases included in the platform.

The test participants were given a number of tasks to carry out, while a usability evaluator observed their performance to find any problems with the software under testing. Although posttest focus group interviews were conducted to gain insights into users' impressions, a self-reporting web-based survey helped to collect users' opinions on a Likert scale. The resulting web designs, shown in Figure 6, have been deployed on village servers so that clients can access health messages [22] from an InfoSpot free of charge, or on the internet. The viewer can explore video animations, key health messages, and quizzes when clicking on the different diseases. The cholera animation was developed by The Global Health Media Project before the DigI project and was included on the platform as an add-on. The dashboard also provides the user with an option to view more videos.

Given that an increasing number of people have access to low-cost mobile devices, an Android mobile app, shown in Figure 7, was also created to present the same health messages as on the platform. Mobile app users have the option to download all the videos on their devices so that they can view them without an internet connection or being close to the InfoSpot. One of the possible scenarios for using the app is that someone visits an InfoSpot with Wi-Fi, then installs the app and downloads all videos for later viewing together with the family back home.

On the basis of our experience, we recommend digital health content developers to assess the access and context that users have to any digital information source. Digital health education projects are only useful if the target population is able to access health messages either supervised or unsupervised.

Figure 6. The platform user interface. The dashboard on the main health page is shown on the left. On the right, an example of the key health messages and quiz questions within the transmission domain of Anthrax is portrayed.

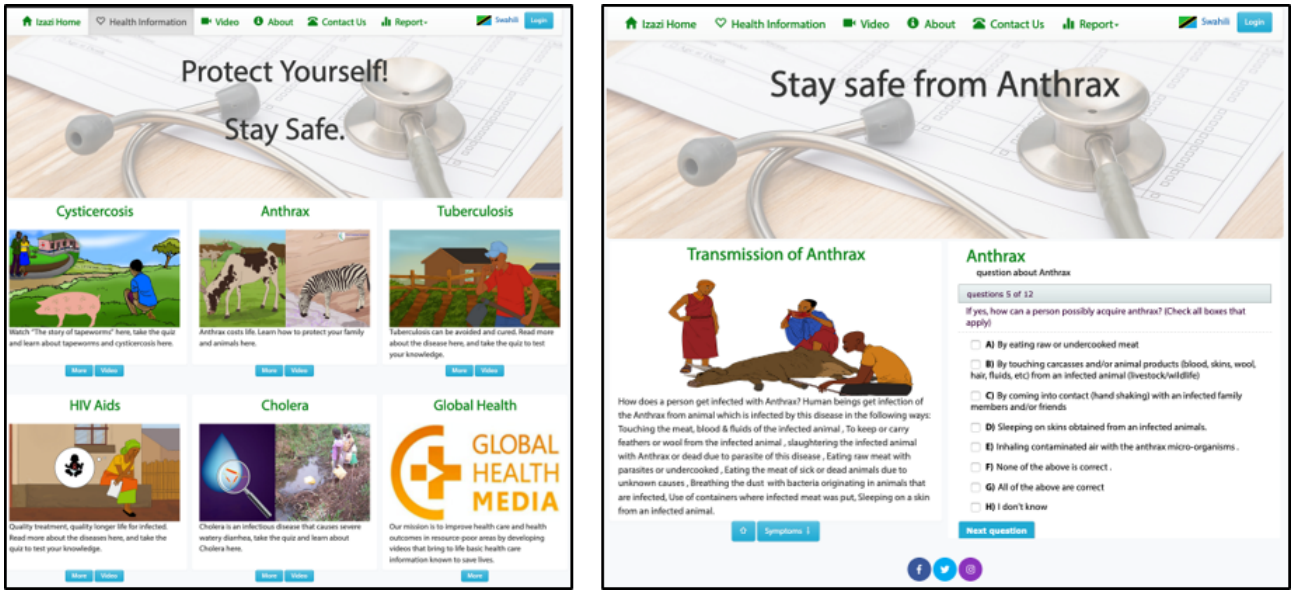
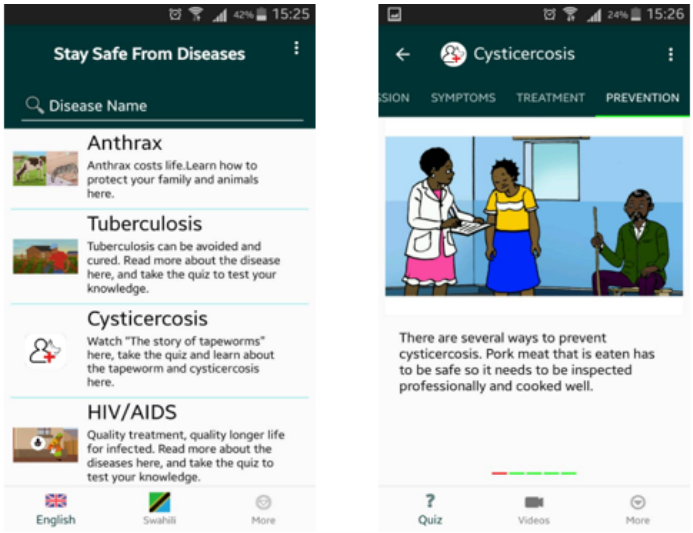


Figure 7. Health Messages mobile app "Linda Afya".

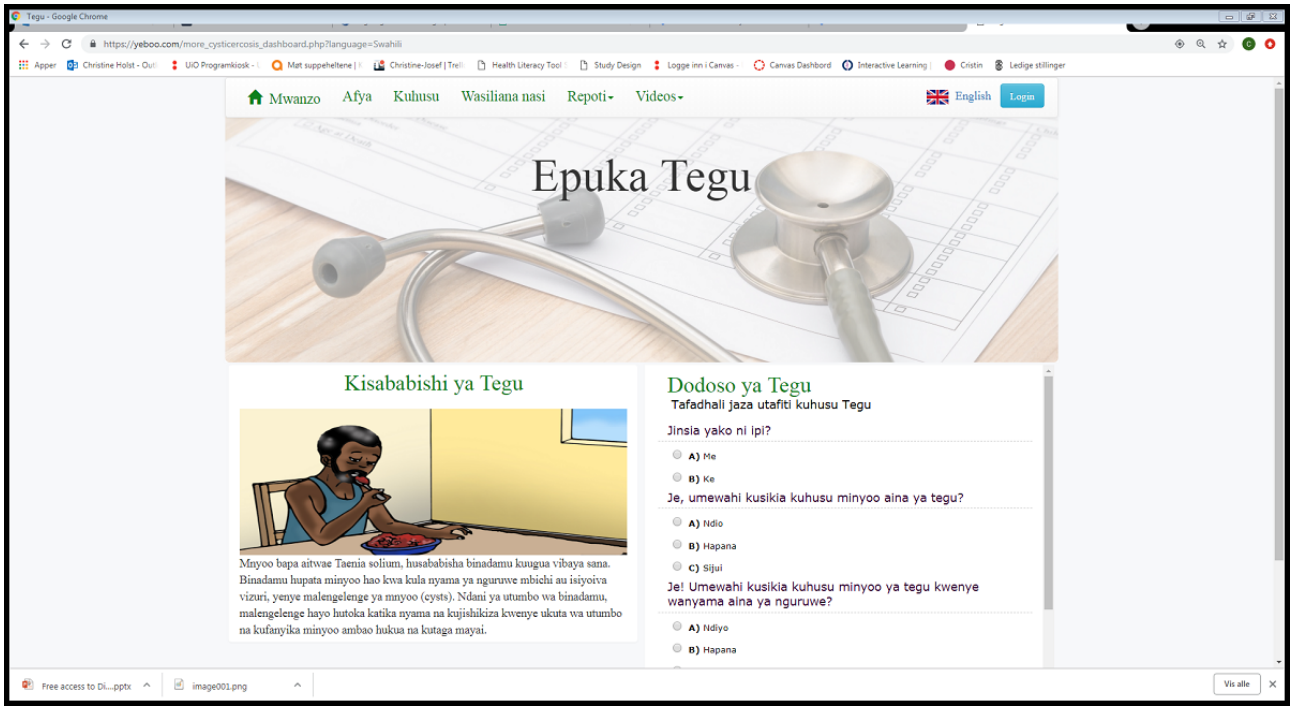


Results

The resulting apps developed through the above-described process can be found via the internet [22] or accessed via Wi-Fi in InfoSpots using locally stored village servers in rural Tanzania. When accessing the platform, the client is introduced to diseases of public health importance in Tanzania (as shown in Figure 6). At any time, the client can change the language from English to Swahili or from Swahili to English. The diseases are HIV/AIDS, TB, TSCT, and anthrax, with all digital information developed as part of the DigI project. In addition, users can find more videos on a variety of health topics, including a cholera animation. The client has 2 choices per disease: *Video* and *More*. When clicking on the *Video* button, an animated video starts immediately. When clicking on the *More* button, the client is taken to a page addressing the specified disease. The client can then navigate through the 5 domains, read about the key health messages, and take the quiz for each disease. Graphics from the animations are used as illustrations for the domains (Figure 8).

In November 2019, seven InfoSpots were available and accessible with the platform, in Esilalei, Selela, Migoli, and Izazi. InfoSpots are still accessible and available for the local population. As of September 3, 2020, the mobile app had been installed by 425 users in 3 months since the launch of the latest version in June 2020. By February 1, 2021, the app store listing had been visited by 1531 people, out of which 1127 were users registered in Kenya. The app can be installed worldwide by searching the app called *Linda Afya* on the Google Play app store.

Figure 8. Swahili Taenia solium cysticercosis and taeniasis health education on the platform.



Discussion

Principal Findings

In this paper, we have provided a tutorial for the various steps in the development process of digital health messages for a rural population in sub-Saharan Africa. The result of our project is a functional health education platform, and the feedback from the local stakeholders and test users was included to improve the platform's usability and impact for the target users. In this section, we present a discussion of the chosen approach used in this project.

A multi- and interdisciplinary approach has become increasingly important in complex research and innovation projects [41]. The advantage of multi- and interdisciplinary collaborations is the ability to draw from each other's knowledge, perspectives, and experiences and channel it toward a joint objective. In this project, the collaboration among team members with medical, social sciences, and IT backgrounds was especially beneficial. Throughout, close collaboration among the local and international health experts, translators, and village stakeholders was important for good results. Furthermore, the storyboard, animated videos, web-based platform, and mobile app could not have been created without input from the visual designers and IT task force. The high quality of digital health messages could not have been achieved without advice from the Tanzanian members of the health task force. International advice on local health matters may differ from the national health advice. Hence, local health researchers were also irreplaceable in identifying the guidelines and written government-approved materials related to each disease. When identifying the key information and creating short, clear messages for the basic health promotion material, the team benefited from members with experience in communication and linguistics. This was important because medical doctors, researchers, and health workers tend to use language with terminology that is difficult for the general public and patients to understand, which can have an exclusive effect [42].

A mixed methods study is currently ongoing within the DigI project to gather evidence for health information uptake and retention based on the above-described animations and access to the platform in 2 villages in rural Tanzania [43]. The data are currently being analyzed, and results are expected late this year. This evaluation will play an important role when revising key messages to become even clearer for the local population. By performing iterations in response to feedback from local communities over time, a human-centered design of global health education is emphasized with its ultimate goal of empowering communities [12].

Regarding the technical aspects of the animation development process, there are various lessons learned that we want to share. We started with the identification of relevant key health messages and then created local stories around those messages, as described above. We had wishes and recommendations from the local community that increased the technical difficulty level but were crucial for the reception in the communities. The many inputs from local stakeholders were acted upon as the DigI team emphasized a co-design process.

Feedback from people in rural areas indicated that the proposed platform should provide information in three main categories: (1) village information, (2) health information, and (3) information on social life and activities. On the basis of this, the design team created templates and user interfaces. In the next stage of the co-design process, people were presented with the interfaces, and they gave positive feedback that the platform was easy to use. However, it seemed difficult to collect the necessary information for publication on the platform, particularly concerning the villages' information and social information. The involved users had limited digital literacy for creating and updating information, for example, using the social media feeds provided by the platform. The research team realized the need to work with a selected group of individuals in each village to increase the usefulness and acceptance of the platform.

The inclusion of digital literacy training programs for the local population is a key factor in the successful implementation of digital health projects. Undoubtedly, the importance of digital literacy is evident when it comes to the use of digital health technologies by people with lower literacy levels, as these technologies could help overcome limitations and surely include new groups in the information society. The DigI team also set up the Key Performance Indicators framework for digital society development and, specifically, to provide a success indicator framework for this project. One of the most relevant success indicators is the level of digital literacy skills, and the health knowledge retention of the participants after the digital skills programs were obtained. Digital literacy goes hand in hand with digital inclusion and social empowerment; thus, it is important that Key Performance Indicators become an integral part of digital literacy initiatives and projects [24].

The drafted stories seemed only to meet the agreement of the full team after several rounds; therefore, the iterative process went on for a long time. In retrospect, the drafting of the storyboards should have been done at an earlier stage within the project, as this process would have provided the group with a good overview of the key health messages and scenes that we wanted to disseminate to the local population. It would have been much easier to imagine the final result if it had been possible to review the different scenes of the animation at an earlier stage.

The translation process of the first animation (TSCT) was performed too early. As the DigI team worked in English and the editing process went on over time, we had to update the Swahili version several times. When the final script was completed in English, the translators had to go through the Swahili version again to ensure that the 2 scripts convey the same story. Consequently, at the stage of the actual recording, the TSCT animation had to be recorded twice because the spoken version sounded different from how it was read. We also had to make adjustments to some wording to avoid misunderstandings. Rerecording was costly and time consuming and could have been avoided. In the following productions, we were more careful with the voiceover recording in the working language and the finalization of the script before translation.

The IT, social science, and health researchers in the DigI project undertook a participatory research approach and spent substantial time partnering with the local population to identify the preferences for digital health message design and knowledge gaps. The WHO has recommended a similar approach of using the existing Ministry of Health documentation and the engagement of different stakeholders [44]. Members of the community, including health workers, school teachers, village leaders, and community leaders, are also key stakeholders to be involved through discussions and interviews to prepare appropriate and relevant health messages and other information. It is essential to engage communities through representatives, including community leaders, to cocreate digital content but also to reach expert consensus on health messages after receiving feedback from the communities. Another important aspect to be considered is user adoption of the messages, whereby health messages are expected to be disseminated to, and accessed by, all or a large proportion of the targeted audience via the platform.

A study that introduced a remote measurement technology platform reported that during the introduction of the remote measurement technology, a multi-stakeholder approach, including patients and research clinicians, provided knowledge about varying requirements for the design and development of the platform [45]. The involvement of patients empowered them to understand the value of the system and to provide their views and needs, thus facilitating the development of a user-adapted platform. Furthermore, the need has been emphasized for developers of digital health information to use numerous innovative strategies to meet both the needs and expectations of targeted audiences [46]. During consultative meetings in the early phases of our project, village leaders indicated the need to use the platform to disseminate news and information to villagers. This included schedules for village meetings, agricultural information on the market, and prices of their products, which were initially not foreseen by the experts. This once again indicates that it is paramount to involve the right people at the right time in an iterative manner to co-design information to be included in an intervention [47]. This multi-

and interdisciplinary approach allows for continuous improvement to meet the needs of the targeted audiences.

Conclusions and Future Plans

The digital health messages, video animations, quizzes, and web applications described in this paper were created through an integrated approach based on various scientific disciplines in addition to engagement of and input from village stakeholders and test users, thus taking advantage of the network and expertise of the DigI team members.

We believe that the methodological approach described in this paper, referring to the digitization of approved printed health education materials, should be promoted. The production process resulted in high-quality educational material that can be used in different forms and in different settings, such as stationary health messages on an internet platform or animated videos telling the story of four diseases in their local context in rural Tanzania. Potentially, health animations developed through the approach described in this paper could also be used in national knowledge portals, as the health messages they contain have been fully approved.

Acknowledgments

The DigI project was funded by The Norwegian Ministry of Foreign Affairs, The Norwegian Agency for Development Cooperation and Norwegian Research Council. The research assistants Ernest Nyoni, Sarah Swai, Bernadetha Tungu, Joan Kalugendo, Pudensiana Hilary, and Getrude Maganya gave constructive feedback during the Morogoro meeting. Merete Taksdal from LHL International provided feedback on TB content.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Screenshot of the platform.
[PNG File , 587 KB - mhealth_v9i9e25558_app1.png]

Multimedia Appendix 2
Definition of terms.
[DOCX File , 22 KB - mhealth_v9i9e25558_app2.docx]

Multimedia Appendix 3
HIV/AIDS health animated video (English).
[MP4 File (MP4 Video), 58029 KB - mhealth_v9i9e25558_app3.mp4]

Multimedia Appendix 4
Tuberculosis health animated video (English).
[MP4 File (MP4 Video), 28571 KB - mhealth_v9i9e25558_app4.mp4]

Multimedia Appendix 5
Taenia solium cysticercosis and taeniasis health animated video (English).
[MOV File , 78062 KB - mhealth_v9i9e25558_app5.mov]

Multimedia Appendix 6
Anthrax health animated video (English).
[MOV File , 38758 KB - mhealth_v9i9e25558_app6.mov]

Multimedia Appendix 7
HIV/AIDS health animated video (Swahili).
[MP4 File (MP4 Video), 56923 KB - mhealth_v9i9e25558_app7.mp4]

Multimedia Appendix 8
Tuberculosis health animated video (Swahili).
[MP4 File (MP4 Video), 36238 KB - mhealth_v9i9e25558_app8.mp4]

Multimedia Appendix 9

Taenia solium cysticercosis and taeniasis health animated video (Swahili).
[MOV File , 43603 KB - mhealth_v9i9e25558_app9.mov]

Multimedia Appendix 10
Anthrax health animated video (Swahili).
[MP4 File (MP4 Video), 60353 KB - mhealth_v9i9e25558_app10.mp4]

References

1. Seventy-first World Health Assembly. Draft thirteenth general programme of work, 2019–2023. World Health Organization. 2018. URL: https://www.un.org/sustainabledevelopment/health [accessed 2021-05-18]
2. Goal 3: Ensure healthy lives and promote well-being for all at all ages. United Nations Sustainable Development Goals. URL: https://www.un.org/sustainabledevelopment/health [accessed 2021-05-18]
3. Goal 9: Build resilient infrastructure, promote sustainable industrialization and foster innovation. United Nations Sustainable Development Goals. URL: https://www.un.org/sustainabledevelopment/infrastructure-industrialization/ [accessed 2021-05-18]
4. WHO guideline: Recommendations on digital interventions for health system strengthening. World Health Organization. URL: https://apps.who.int/iris/bitstream/handle/10665/311941/9789241550505-eng.pdf?ua=1 [accessed 2021-05-18]
5. Fiore SM, Salas E. Problems and possibilities: Strategically pursuing a science of learning in distributed environments. In: Fiore SM, Salas E, editors. Toward a Science of Distributed Learning. Washington, D. C: American Psychological Association; 2007:237-264.
6. Choi BC, Pak AW. Multidisciplinarity, interdisciplinarity and transdisciplinarity in health research, services, education and policy: 1. Definitions, objectives, and evidence of effectiveness. Clin Invest Med 2006 Dec;29(6):351-364. [Medline: 17330451]
7. National Academy of Sciences, National Academy of Engineering, Institute of Medicine. Facilitating Interdisciplinary Research. Washington, D. C: The National Academies Press; 2005:2.
8. World Health Organization. Promoting Health in the Sdgs: Report on the 9th Global Conference for Health Promotion, Shanghai, China, 21–24 November 2016: All for Health, Health for All. Geneva: World Health Organization; 2017. URL: https://www.who.int/publications/i/item/promoting-health-in-the-sdgs [accessed 2021-09-13]
9. GSM Association. The State of Mobile Internet Connectivity 2020. URL: https://www.gsma.com/r/wp-content/uploads/2020/09/GSMA-State-of-Mobile-Internet-Connectivity-Report-2020.pdf [accessed 2021-05-18]
10. Zelezny-Green R, Vosloo S, Conole G. Digital inclusion for low-skilled and low-literate people: a landscape review. United Nations Educational, Scientific and Cultural Organization. 2018. URL: https://unesdoc.unesco.org/ark:/48223/pf0000261791 [accessed 2021-05-18]
11. Holst C, Sukums F, Radovanovic D, Ngowi B, Noll J, Winkler A. Sub-Saharan Africa-the new breeding ground for global digital health. Lancet Digital Health 2020 Apr;2(4):160-162 [FREE Full text] [doi: 10.1016/S2589-7500(20)30027-3] [Medline: 33328076]
12. Adam M, McMahon SA, Prober C, Bärnighausen T. Human-centered design of video-based health education: an iterative, collaborative, community-based approach. J Med Internet Res 2019 Jan 30;21(1):e12128 [FREE Full text] [doi: 10.2196/12128] [Medline: 30698531]
13. Health Sector Strategic Plan July 2015 – June 2020 (HSSP IV). The United Republic of Tanzania. Ministry of Health and Social Welfare. 2015. URL: https://tinyurl.com/46mz24h4 [accessed 2021-05-18]
14. Tanzania. The World Bank. URL: https://www.worldbank.org/en/country/tanzania/overview#1 [accessed 2021-05-18]
15. GSM Association. Digital Transformation in Tanzania - The Role of Mobile Technology and Impact on Development Goals. 2019. URL: https://www.gsmaintelligence.com/research/?file=783bb9b0ab8e6e53361607a838d25dcb&download [accessed 2021-05-18]
16. Tanzania in figures 2018. The United Republic of Tanzania. National Bureau of Statistics. URL: https://www.nbs.go.tz/nbs/takwimu/references/Tanzania_in_Figures_2018.pdf [accessed 2021-05-18]
17. Internet connectivity seen as having positive impact on life in Sub-Saharan Africa. Pew Research Center. 2018. URL: https://www.pewresearch.org/global/2018/10/09/majorities-in-sub-saharan-africa-own-mobile-phones-but-smartphone-adoption-is-modest/ [accessed 2021-05-18]
18. Digital Health Strategy, July 2019 - June 2024. The United Republic of Tanzania. Ministry of Health, Community Development, Gender, Elderly and Children. URL: http://moh.go.tz/en/strategic-plans?download=395:digital-health-strategy [accessed 2021-05-18]
19. Taeniasis/cysticercosis. World Health Organization. URL: https://www.who.int/news-room/fact-sheets/detail/taeniasis-cysticercosis [accessed 2021-05-18]
20. Ngowi HA, Winkler AS, Braae UC, Mdegela RH, Mkupasi EM, Kabululu ML, et al. Taenia solium taeniosis and cysticercosis literature in Tanzania provides research evidence justification for control: A systematic scoping review. PLoS One 2019;14(6):e0217420 [FREE Full text] [doi: 10.1371/journal.pone.0217420] [Medline: 31166983]

21. Mwakapeje ER, Høgset S, Fyumagwa R, Nonga HE, Mdegela RH, Skjerve E. Anthrax outbreaks in the humans - livestock and wildlife interface areas of Northern Tanzania: a retrospective record review 2006-2016. BMC Public Health 2018 Jan 05;18(1):106 [FREE Full text] [doi: 10.1186/s12889-017-5007-z] [Medline: 29304765]
22. Digi platform. Non-Discriminating Access for Digital Inclusion. URL: https://izazi.yeboo.com/health_info_dashboard.php [accessed 2021-05-18]
23. Kumar S, Preetha G. Health promotion: an effective tool for global health. Indian J Community Med 2012 Jan;37(1):5-12 [FREE Full text] [doi: 10.4103/0970-0218.94009] [Medline: 22529532]
24. Radovanović D, Holst C, Belur S, Srivastava R, Houngbonon G, Le Quentrec E, et al. Digital literacy key performance indicators for sustainable development. Soc Inclusion 2020 May 14;8(2):151-167 [FREE Full text] [doi: 10.17645/si.v8i2.2587]
25. The Ottawa Charter for Health Promotion. World Health Organization. 1986. URL: https://www.who.int/teams/health-promotion/enhanced-wellbeing/first-global-conference [accessed 2021-08-26]
26. Nutbeam D. Health promotion glossary. Health Prom Int 1998 Jan 01;13(4):349-364. [doi: 10.1093/heapro/13.4.349]
27. Nutbeam D. Health literacy as a public goal: a challenge for contemporary health education and communication strategies into the 21st century. Health Prom Int 2000;15(3):259-267. [doi: 10.1093/heapro/15.3.259]
28. Cerqueira MT, Coe GA. Communication, Education and Participation: a Framework and Guide to Action. Washington, D.C: Pan American Health Organization; 1996.
29. What is Health Information? The American Health Information Management Association (AHIMA). URL: https://www.ahima.org/certification-careers/certifications-overview/career-tools/career-pages/health-information-101/ [accessed 2021-08-26]
30. Dodson S, Good S, Osborne RH. Health literacy toolkit for lowand middle-income countries: a series of information sheets to empower communities and strengthen health systems. New Delhi, India: World Health Organization, Regional Office for South-East Asia; 2015.
31. Morrison FP, Kukafka R, Johnson SB. Analyzing the structure and content of public health messages. AMIA Annu Symp Proc 2005:540-544 [FREE Full text] [Medline: 16779098]
32. World Health Organization. Education for Health: A Manual on Health Education in Primary Health Care. Geneva: World Health Organization; 1988:1-261.
33. Classification of digital health interventions v1.0. A shared language to describe the uses of digital technology for health. World Health Organization. Geneva: World Health Organization; 2018. URL: http://apps.who.int/iris/bitstream/handle/10665/260480/WHO-RHR-18.06-eng.pdf [accessed 2021-05-18]
34. Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems. International Standardization Organization (ISO). 2019. URL: https://www.iso.org/obp/ui/#iso:std:iso:9241:-210:ed-2:v1:en [accessed 2021-05-18]
35. Cornwall A, Jewkes R. What is participatory research? Soc Sci Med 1995 Dec;41(12):1667-1676. [doi: 10.1016/0277-9536(95)00127-s]
36. National Aids Control Programme. National Guidelines for the Management of HIV and AIDS. The United Republic of Tanzania. Ministry of Health, Community Development, Gender, Elderly and Children. 2017. URL: https://tinyurl.com/r9emcp34 [accessed 2021-05-18]
37. The National Strategic Plan V (2015-2020) for Tuberculosis and Leprosy. The United Republic of Tanzania. Ministry of Health and Social Welfare. 2015. URL: https://ntlp.go.tz/site/assets/files/1074/national_strategic_plan_2015_2020.pdf [accessed 2021-05-18]
38. One Health Strategic Plan 2015 - 2020. The United Republic of Tanzania. URL: https://tinyurl.com/48pssd5u [accessed 2021-05-18]
39. Schank RC, Berman TR. The pervasive role of stories in knowledge and action. In: Green MC, Strange JJ, Brock TC, editors. Narrative Impact: Social and Cognitive Foundations. New York: Lawrence Erlbaum Associates Publishers; 2002:287-313.
40. Adobe XD. URL: https://www.adobe.com/products/xd.html [accessed 2021-05-18]
41. Parmar V. A multidisciplinary approach to ICT development. Information Technologies and International Development 2009;5(4):89-96 [FREE Full text]
42. Koch-Weser S, Dejong W, Rudd RE. Medical word use in clinical encounters. Health Expect 2009 Dec;12(4):371-382 [FREE Full text] [doi: 10.1111/j.1369-7625.2009.00555.x] [Medline: 19709316]
43. Holst C, Sukums F, Ngowi B, Diep LM, Kebede TA, Noll J, et al. Digital health intervention to increase health knowledge related to diseases of high public health concern in Iringa, Tanzania: protocol for a mixed methods study. JMIR Res Protoc 2021 Apr 22;10(4):e25128 [FREE Full text] [doi: 10.2196/25128] [Medline: 33885369]
44. World Health Organization. Monitoring and Evaluating Digital Health Interventions: a Practical Guide to Conducting Research and Assessment. Geneva. Licence: CC BY-NC-SA 3.0 IGO: World Health Organization; 2016. URL: https://apps.who.int/iris/bitstream/handle/10665/252183/9789241511766-eng.pdf [accessed 2021-05-18]

45. Craven MP, Andrews JA, Lang AR, Simblett SK, Bruce S, Thorpe S, RADAR-CNS Consortium. Informing the development of a digital health platform through universal points of care: qualitative survey study. JMIR Form Res 2020 Nov 26;4(11):e22756 [FREE Full text] [doi: 10.2196/22756] [Medline: 33242009]
46. Comulada WS, Step M, Fletcher JB, Tanner AE, Dowshen NL, Arayasirikul S, Special Projects Of National Significance Social Media Initiative Study Group. Predictors of internet health information-seeking behaviors among young adults living with HIV across the United States: longitudinal observational study. J Med Internet Res 2020 Nov 02;22(11):e18309 [FREE Full text] [doi: 10.2196/18309] [Medline: 33136057]
47. World Health Organization. Youth-Centred Digital Health Interventions: A Framework for Planning, Developing and Implementing Solutions With and for Young People. Geneva. Licence: CC BY-NC-SA 3.0 IGO: World Health Organization; 2020. URL: https://www.who.int/publications/i/item/9789240011717 [accessed 2021-05-18]

Abbreviations

DigI: Non-discriminating access for Digital Inclusion
InfoSpot: information spot
IT: information technology
TB: tuberculosis
TSCT: Taenia solium cysticercosis and taeniasis
WHO: World Health Organization

Edited by L Buis; submitted 06.11.20; peer-reviewed by A Ugargol, J Dol, L McCann, D Pförringer; comments to author 15.12.20; revised version received 09.02.21; accepted 17.06.21; published 22.09.21.

Please cite as:
Holst C, Isabwe GMN, Sukums F, Ngowi H, Kajuna F, Radovanović D, Mansour W, Mwakapeje E, Cardellichio P, Ngowi B, Noll J, Winkler AS
Development of Digital Health Messages for Rural Populations in Tanzania: Multi- and Interdisciplinary Approach
JMIR Mhealth Uhealth 2021;9(9):e25558
URL: https://mhealth.jmir.org/2021/9/e25558
doi:10.2196/25558
PMID:34550081

Original Paper

Pediatric Weight Management Through mHealth Compared to Face-to-Face Care: Cost Analysis of a Randomized Control Trial

Louise Tully[1], MSc; Jan Sorensen[2], MA, MSc; Grace O'Malley[1,3], PhD

[1]Obesity Research and Care Group, Division of Population Health Sciences, School of Physiotherapy, Royal College of Surgeons in Ireland, University of Medicine and Health Sciences, Dublin, Ireland

[2]Healthcare Outcomes Research Centre, Royal College of Surgeons in Ireland, University of Medicine and Health Sciences, Dublin, Ireland

[3]W82GO Child and Adolescent Weight Management Service, Children's Health Ireland at Temple Street, Dublin, Ireland

Corresponding Author:
Louise Tully, MSc
Obesity Research and Care Group, Division of Population Health Sciences
School of Physiotherapy
Royal College of Surgeons in Ireland, University of Medicine and Health Sciences
Beaux Lane House, Level -1
Mercer St
Dublin, D02 DH60
Ireland
Phone: 353 014022100
Email: louisetully@rcsi.com

Abstract

Background: Mobile health (mHealth) may improve pediatric weight management capacity and the geographical reach of services, and overcome barriers to attending physical appointments using ubiquitous devices such as smartphones and tablets. This field remains an emerging research area with some evidence of its effectiveness; however, there is a scarcity of literature describing economic evaluations of mHealth interventions.

Objective: We aimed to assess the economic viability of using an mHealth approach as an alternative to standard multidisciplinary care by evaluating the direct costs incurred within treatment arms during a noninferiority randomized controlled trial (RCT).

Methods: A digitally delivered (via a smartphone app) maintenance phase of a pediatric weight management program was developed iteratively with patients and families using evidence-based approaches. We undertook a microcosting exercise and budget impact analysis to assess the costs of delivery from the perspective of the publicly funded health care system. Resource use was analyzed alongside the RCT, and we estimated the costs associated with the staff time and resources for service delivery per participant.

Results: In total, 109 adolescents participated in the trial, and 84 participants completed the trial (25 withdrew from the trial). We estimated the mean direct cost per adolescent attending usual care at €142 (SD 23.7), whereas the cost per adolescent in the mHealth group was €722 (SD 221.1), with variations depending on the number of weeks of treatment completion. The conversion rate for the reference year 2013 was $1=€0.7525. The costs incurred for those who withdrew from the study ranged from €35 to €681, depending on the point of dropout and study arm. The main driver of the costs in the mHealth arm was the need for health professional monitoring and support for patients on a weekly basis. The budget impact for offering the mHealth intervention to all newly referred patients in a 1-year period was estimated at €59,046 using the assessed approach.

Conclusions: This mHealth approach was substantially more expensive than usual care, although modifications to the intervention may offer opportunities to reduce the mHealth costs. The need for monitoring and support from health care professionals (HCPs) was not eliminated using this delivery model. Further research is needed to explore the cost-effectiveness and economic impact on families and from a wider societal perspective.

Trial Registration: ClinicalTrials.gov NCT01804855; https://clinicaltrials.gov/ct2/show/NCT01804855

(JMIR Mhealth Uhealth 2021;9(9):e31621) doi:10.2196/31621

KEYWORDS

childhood obesity; pediatric weight management; economic evaluation; digital health; telemedicine; mHealth

Introduction

Digital Delivery of Pediatric Weight Management

Mobile health (mHealth), a subcategory of telemedicine whereby clinical care is provided via mobile devices, for weight management in pediatric populations with clinical obesity is an emerging field [1]. For excess adiposity in childhood, family-orientated multidisciplinary weight management, consisting of nutrition and physical activity support with integration of evidence-based behavior-change techniques, is recommended as the cornerstone of treatment [2-4]. There is evidence that telemedicine interventions can support self-management of nutrition and physical activity in children and adolescents [5]; however, there is a scarcity of studies focusing on the economic evaluations of such interventions, particularly for mHealth interventions developed to incorporate evidence-based approaches [1,5,6].

The use of mHealth may improve capacity in terms of delivering health care with a wider geographical reach and may overcome barriers to attending physical appointments experienced by some families using ubiquitous devices such as smartphones and tablets. During the COVID-19 pandemic, technology facilitated alternative modes of delivery for weight management services, which, for many people, meant avoiding long periods of time without professional weight management support [7]. In the long term, such interventions could also expand capacity to areas where long waiting lists, geographical constraints, and staff shortages impose barriers to accessing care.

Study Rationale

To inform decisions about implementation of mHealth, it is necessary to demonstrate "value for money" in addition to clinical effectiveness for novel treatments and health technologies [8]. Previous studies on mHealth applications for self-management in adult populations have shown promise for potential cost savings [9,10]. However, economic evaluations of telemedicine interventions present methodological challenges in ensuring that the true costs of digital services and their reach are captured in a systematic way that is comparable to face-to-face care [11]. This challenge is compounded by efforts to simultaneously account for the rapid evolution of technology and its effect on resource use and availability; in contrast, the research process (including trial design, implementation, analysis, dissemination, and policy implications) can take many years.

A Tier 3 accredited center of excellence (European Association for the Study of Obesity Centre for Obesity Management) [12] consisting of a multidisciplinary weight management service (the W82GO service) is available for children and adolescents with obesity in Children's Health Ireland at Temple Street, an urban tertiary care pediatric hospital in the Republic of Ireland [13]. Clinical appointments are either delivered as part of group or one-to-one interventions depending on the needs of the child or adolescent and the preferences of the family. A pilot randomized controlled trial (RCT) [13] tested the clinical effectiveness of a bespoke evidence-based mHealth platform (Android app, clinical portal, and backend database) as an alternative to usual care for the maintenance phase of weight management (3 face-to-face booster sessions over 46 weeks) using a noninferiority design with adolescents. The change in the BMI standardized deviation score (BMI-SDS) was assessed as the primary outcome. The platform was designed using a participatory approach with the intended end users (adolescents with obesity and their parents) [14,15]. The findings of the RCT suggested that substituting face-to-face maintenance care with the mHealth intervention did not adversely affect the change in the primary outcome (BMI-SDS) of the overall treatment. Although study attrition was substantial and was similar to other pediatric trials [16], there was insufficient power to statistically confirm noninferiority [17]. As a result of this, and in addition to high levels of missing secondary outcome data including health-related quality-of-life data, a full cost-effectiveness analysis was not possible despite conducting a clinical trial.

Study Aim

We aimed to assess the direct costs of delivering the mHealth intervention to participants in the trial relative to usual care participants to inform future designs of mHealth trials to assess effectiveness and cost-effectiveness within this population as well as contribute to the evidence base for the economic viability of integrating mHealth into pediatric weight management services in future.

Methods

Design

The pilot noninferiority RCT was approved by the ethics committee of Children's Health Ireland at Temple Street (reference number 11–033; ClinicalTrials.gov trial registration: NCT01804855). We undertook a microcosting analysis to assess and compare the costs of treatment groups participating in the RCT for pediatric weight management, namely usual care versus mHealth delivered using the "Reactivate" system. We also carried out a budget impact analysis for a 12-month period.

Sample Size and Recruitment

The null hypothesis in the trial protocol was that the mHealth intervention would have a positive effect on change in the BMI-SDS but that this change will be inferior to that observed in usual care. Based on a reduction of 0.21 in the BMI-SDS at 12 months, an SD of 0.24 in the usual care group, and a noninferiority limit of 0.12, the sample size at 80% power was calculated to be 50 per group or 100 in total. To allow for expected attrition, the target recruitment sample size was 134 [13]. Eligible trial participants were recruited from the W82GO Child and Adolescent Weight Management Service, which is the only dedicated Tier 3 service for children and adolescents with obesity in the Republic of Ireland. All new adolescent referrals made to the service by a pediatrician were screened against the inclusion/exclusion criteria. Those eligible were invited to participate in the study following the consideration of the study by their parents and upon receipt of parental consent and adolescent assent forms.

In total, 109 adolescent participants with clinical obesity (40 boys, 69 girls) were recruited through the W82GO service and received phase 1 of the treatment face to face before being randomized to receive the maintenance phase (phase 2) of

treatment either through usual care (three additional face-to-face booster sessions with the multidisciplinary team either through one-to-one sessions or group sessions) or remotely via the mHealth app (Reactivate) [13].

Data Collection

Participant data including trial group data, whether they commenced one-to-one or group treatment, the number of sessions attended, and records of treatment completion or withdrawal stages, were collected during the trial and used for this analysis to ascertain variations in costs per patient. Cost data were obtained from multiple sources. For face-to-face maintenance sessions, we used a time-driven activity-based microcosting method [18] to capture the direct costs associated with the face-to-face time of health care professionals (HCPs) with patients. We also included administrative time associated with appointment preparation. We interviewed personnel to map workflow processes associated with usual care to accurately assess the unit costs of program appointments and dropout/nonattendance costs. A record of the trial costs was maintained by the principal investigator, and it included invoices received for contracted mHealth service delivery, the related expenses, and the time allocated for checking in, monitoring, and processing participants. During baseline data collection, parents/carers were asked to provide details of their annual income, current occupation, the make and model of their car (if any), mode of transport, and distance traveled to attend hospital appointments.

Cost Analysis

We carried out our cost analysis based on the detailed unit costs for providing care to both study groups from the perspective of the publicly funded health care system. We undertook the cost analysis under pragmatic "real-world" conditions and their cost implications (ie, estimates of implementing the intervention outside of a research trial) [19], as the trial costs included additional expenses that would not represent the cost of telemedicine if provided as part of usual care (eg, provision of smartphones and mobile data packages to trial participants). We calculated the cost of staff time according to local guidance [20-22], adjusting for pay-related social insurance, pension contributions, annual leave, and overheads. Salaries were calculated using the midpoints from the salary scales for the trial period [23]. For the cost comparison assessment, we also included equipment frequently used for clinical appointments. The unit costs and breakdown of these are shown in Multimedia Appendix 1. Variations in the costs allocated to individuals were based on their treatment group, the completion status, and the number of weeks/sessions completed.

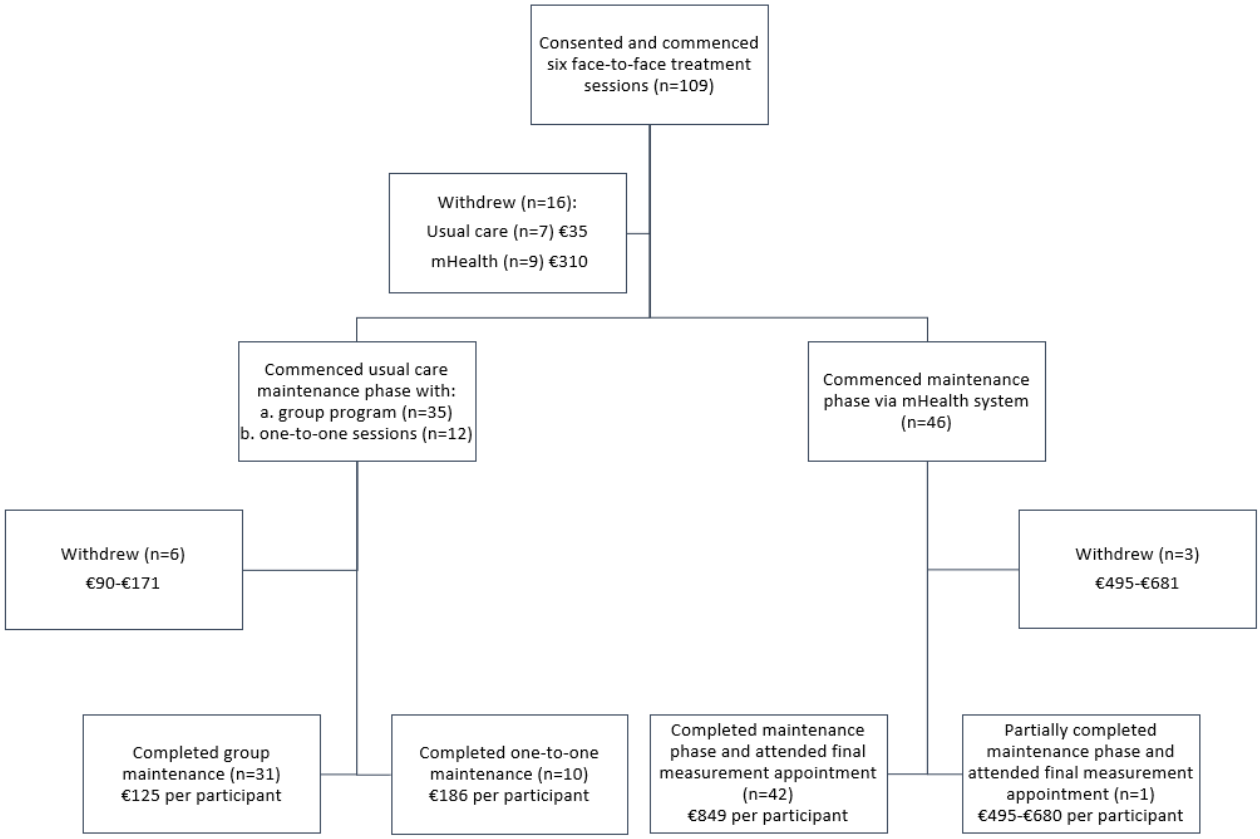
We also undertook a budget impact analysis to assess the cost of providing the mHealth intervention to all eligible adolescents (new referrals) over a 12-month period. In the sensitivity analysis, we evaluated cost assumptions by changing the base case parameters, such as the annual cost of software maintenance, equipment, and variations in the time spent by HCPs in monitoring and supporting adolescents in the mHealth arm. We also examined the impact on the cost per adolescent by changing the optimum treatment cohort size by varying the annual number of users.

We assessed the costs incurred by families based on prospectively collected trial data, but these were not included in the main cost comparison owing to high levels of missing and incomplete data. Therefore, this study considered only the 12-month costs incurred by the publicly funded health care system.

Results

In total, 109 adolescents and their families provided consent for participation in the trial; only 84 participants completed the trial as 25 adolescents withdrew from the study (13 from the usual care group and 12 from the mHealth group) after allocation, as shown in Figure 1.

Figure 1. Trial allocation and completion among participants with base case cost estimates. mHealth: mobile health.



Health Care System Perspective Costs

The conversion rate for the reference year 2013 was $1=€0.7525. We estimated the mean direct cost per adolescent who completed one-to-one usual care in the maintenance phase of treatment at €186 for all three sessions, as shown in Table 1. For an adolescent who participated in group maintenance sessions, the cost was estimated at €125 (assuming a maximum capacity of 15 families per group). Withdrawal or partial completion costs ranged from €35 to €171 per adolescent for one-to-one sessions depending on the number of sessions missed; withdrawal costs for those in the group treatment were estimated at €153, as their place in the group was lost and could not be filled by another patient. For adolescents who were randomized to use the mHealth system and who completed the program, the mean cost per adolescent was estimated at €849 (based on the intention-to-treat cost divided over all the adolescents allocated to the mHealth arm; n=55). Withdrawal from or partial completion of the mHealth intervention was estimated to cost €310 to €680, depending on when the participant dropped out (see Multimedia Appendix 1).

Table 1. Cost per adolescent by treatment group.

Treatment group	Estimated direct cost per participant, mean (SD)
Usual care (one-to-one program)	€176.58 (22.41)
Usual care (group program)	€132.52 (12.18)
mHealth[a]	€722.36 (221.07)

[a]mHealth: mobile health.

Accounting for partial completion and attrition costs, the mean cost incurred for those in the usual care arm was €142 (SD 23.7) (group participants: mean €133, SD 12.2; one-to-one participants: mean €177, SD 22.4). The mean cost for those randomized to use mHealth was estimated to be €722 (SD 221).

The costs for the design and development of the mHealth service (website and app domain name registration and hosting, videography, iconography, device updates for firmware, app development, maintenance costs, and cloud hosting) were independent of the number of users. The sensitivity analysis showed that the main driver of costs for the mHealth group was the HCP time spent managing the mHealth service arm of the trial (platform administration, individualized care plans, providing feedback, troubleshooting, checking in). This was estimated to be approximately 12 hours per adolescent over 46 weeks (approximately 15 minutes per adolescent per week) during the trial. Sensitivity analysis showed that this would need to be reduced to 1.5 hours (2 minutes per adolescent per week), with the number of users increased to 160 before the cost per person would match that of one-to-one, in-person care (€186 per participant). Further, we tested our assumptions around the

estimated costs of software maintenance and data storage costs per annum through increasing these by 10%, and this increased the cost per adolescent for the mHealth arm (n=55) by €10, which became negligible once extrapolated to large numbers of users and had a negligible impact on the cost comparison with usual care.

Budget Impact

Using the cost per adolescent who completed the mHealth intervention and considering the capacity of the weight management service to be 120 new patient referrals per year, we estimated the budget impact of offering the maintenance phase of treatment to all eligible adolescents (BMI ≥98th centile) face to face instead of using mHealth, from the perspective of the health care system. Offering phase 2 of the face-to-face treatment to each eligible adolescent with obesity using the base case has a direct cost of approximately €19,074, whereas the mHealth service would cost €78,120 (excluding app development costs). As such, the direct budget impact of replacing face-to-face maintenance treatment and offering the mHealth intervention to all eligible adolescents in one year would be €59,046, without accounting for potential cost and time savings to be gained by offering mHealth care only.

Family Perspective

Of the families who took part in the trial, 65% (71/109) provided details about their travel, work, and school arrangements for attending clinical appointments. Further, 17% (19/109) of the families used public transport, at a mean cost of €7 per hospital visit (range €1-€41), whereas 33% (36/109) families drove an average of 17.7 km (range 2-64 km) to their in-person appointments, costing approximately €11 each way (based on a previous study estimating the cost as €0.62 per kilometer including running costs and depreciation [24] plus a €3.10 hourly parking fee). In addition, 4 out of 109 families (4%) took a taxi, with a mean cost of €12 each way (range €5-€15). Using the data provided, the mean cost of travel to and from appointments per adolescent was €18 per visit (€54 for the full face-to-face maintenance phase).

Furthermore, 27% (29/109) of the adolescents had missed school for their appointment on the day of clinic, with an average of 3 hours missed (ranging from 20 minutes to the full school day). As for parents, 21% (22/109) reported that they needed to take time off from work to attend their child's appointment. Among these 22 parents, 7 needed a full day off and 11 required closer to half a day off; the others did not provide details. Of those who required time off, 18 parents reported their annual income, with 7 earning less than €15,000 per annum and 3 others earning less than €25,000 per annum. In addition, 4 parents earned more than €40,000 per annum and 4 did not report their income. The mean daily salary (adjusted to the whole time equivalent) per parent who provided details of their income was €122 (median €100).

Discussion

Principal Findings

This study assessed the treatment costs based on trial data from a pragmatic noninferiority pilot RCT. There was a 23% attrition rate for the trial (25/109); however, this is broadly in line with pediatric RCTs [25] and weight management interventions in general [16], where dropouts are common owing to the intensive nature of these interventions.

The results show that this mHealth intervention, developed using evidence-based approaches, is associated with higher health care costs than face-to-face pediatric weight management. The design of the trial was such that all adolescents attended face-to-face treatment before randomization to either the digital or face-to-face maintenance phase; therefore, this partially digital intervention arm incurred appointment and mHealth costs. The sensitivity analysis results demonstrated that if rolled out to a larger number of users, the main driver of the costs for the mHealth arm is the staff cost related to HCP monitoring and support on a weekly basis. If the mHealth service were to be automated, it could be to reduce these costs; however, further studies would be required to explore the clinical impact of delivering the mHealth service to this clinical population with inputs from less-experienced clinical staff or via increased automation and the associated ethical considerations.

Comparison With Prior Work

Our finding that staff costs are the most sensitive drivers of the overall cost has been shown in economic evaluations of mHealth in other fields, including care after pregnancy termination [26]. We also found that substantial costs were incurred by the families, but we were unable to fully explore the costs from the perspective of parents and families owing to incomplete data collection. However, this is an important consideration for future research, as assessing ways to reduce inequalities that may be exacerbated by the burden of attending face-to-face appointments is crucial. It is important to explore ways to collect cost data from families, which does not substantially add to the burden of participating in research.

Previous studies have demonstrated that families who live further from clinics, or for whom travel to in-person appointments is more burdensome or complex, tend to view telemedicine more favorably [27]. Despite this, most published economic evaluations of telemedicine consider the perspectives of only the health care service/provider, as shown for cardiovascular disease management [28], obesity prevention [6], and eHealth more broadly [29]. It is important for researchers to assess delivery costs for future evaluations of digitally delivered pediatric weight management to build an evidence base for this population with unique care needs [30]. It is also vital that economic evaluations adopt a societal perspective to capture costs apart from direct health system costs. This has been recommended for mHealth in caring for the elderly as well [31]. The financial strain on parents/caregivers is a documented barrier for pediatric chronic disease management [32] and access to childhood obesity treatment [33], particularly in rural communities [34]. Nelson and colleagues also reported that a telerehabilitation program was not cost-effective for patients recovering from hip replacements but that the reduced burden on patients and caregivers was notable [35]. Clinical pediatric populations are comparable in that patient and caregiver time as well as travel are required for appointments. However, digital interventions

may also incur further costs for families, some of which we did not capture, such as internet and phone bills that were covered within the research budget but would add to the burden on families. This highlights the need for further cost studies incorporating a wider perspective that is not limited to the health care provider.

In general, the economic evidence for mHealth is mixed [36]. It is an emerging field, and much of the work to date has evaluated mHealth for health promotion or in the self-management of chronic conditions to prevent the need for health service usage. Such studies are not directly comparable with this trial that evaluated an evidence-based adolescent obesity intervention requiring consistent appointment attendance with an obesity intervention delivered via an mHealth platform as a remote alternative. The body of work associated with the development and testing of the Reactivate mHealth system [13-15,17] has provided novel evidence for the feasibility of using mHealth for pediatric weight management with transparent accounts of the limitations identified, including those relating to the collection of cost data, which will be valuable for informing the design of future robust trials with this vulnerable and complex population.

A recent scoping review [37] on the use of eHealth in diabetes care highlighted critical issues such as staff training, monitoring, technological infrastructure support and maintenance, and how these differ by setting and intervention. Although the mHealth intervention that we evaluated for the maintenance phase of the treatment did not prove economically viable in its prototype form, our results point to design and development aspects where amendments may produce cost savings. The need for 15 minutes of HCP time per participant per week may be a modifiable intervention component. Our study assessed the costs based on the time spent by a senior registered pediatric physiotherapist; however, the option of a more junior staff member managing the mHealth intervention may be feasible, or there may be scope for automating some of the tasks, such as feedback on engagement with the app.

Input from families and HCPs involved in the service could further help identify the acceptability of such modifications. Further, the option of offering both phases of treatment via the mHealth intervention may present a more economically attractive alternative to face-to-face treatment although their clinical effectiveness is unknown. Exploring how this option might suit patients with less complex obesity and fewer complications or comorbidities may also yield evidence for its appropriateness. The acceptability of receiving only remote care for adolescent obesity is also unknown; however, when the mHealth trial was being designed, most families specified a preference for some face-to-face care. This was considered during the design of the pilot RCT. More recently, during the COVID-19 pandemic, up to 40% of families refused virtual appointments from the Child and Adolescent Weight Management Service and preferred to wait longer for face-to-face care. Notwithstanding the preferences of families who are already engaged in treatment, there may be scope to increase access to care through using the mHealth platform with families whose access to evidence-based obesity treatment is limited (eg, children and adolescents who live in rural areas, those who may age out of eligibility for pediatric health care, or those who have no local pediatric obesity treatment services). It may also be possible to achieve cost savings by providing earlier access to treatment via the mHealth platform to adolescents in the community setting and negate the need to join a waiting list for a Tier 3 obesity service. Earlier interventions can reduce or prevent obesity-related complications; given the promising preliminary data on the clinical effectiveness of the mHealth system [17], offering such care to adolescents may mitigate the health effects of obesity at a crucial time during their development.

Limitations

This study had several strengths and limitations. Assessing the costs incurred by both treatment arms alongside a pragmatic pilot RCT was an important strength of the study, as it reflected the actual costs of delivery in a real-world clinical setting and allowed assumptions that were underpinned by clinical experiences. The microcosting analysis also enabled detailed and accurate direct costing for usual care within the pediatric weight management service. However, the study did not meet the target recruitment number within the available time period, and coupled with the attrition rate, this led to insufficient power for demonstrating statistically significant noninferiority. In addition, low response rates for health-related quality-of-life measures used contributed to the decision of undertaking only a direct cost comparison. As a result, our cost analysis does not provide a full economic evaluation. Further, although it was the only treatment center available nationwide, we acknowledge the limited external validity of our findings given the recruitment through a single center for obesity management. Cost was also not a prespecified outcome for this trial and this study was undertaken as an exploratory analysis after completion of the trial.

Nonetheless, it is important to provide transparent accounts of studies undertaken to assess mHealth interventions with this clinical population, for whom no previous cost studies have been undertaken. It is especially pertinent to document data to describe the economic viability of mHealth, which is often presumed to be a cost-saving alternative to traditional care [38] given the emphasis on digital interventions within the European digital health strategy [39] compounded by the shift in processes resulting from the COVID-19 pandemic.

In addition, access to treatment for obesity is severely limited in Ireland and elsewhere with only approximately 20% of primary care providers reporting sufficient capacity to offer treatment [40]. Therefore, developing and evaluating mHealth interventions for obesity is a high priority for health services. This preliminary research will allow for improved processes and designs aiming to maximize resources while maintaining clinical effectiveness and acceptability among users.

Conclusions

Childhood obesity remains a leading concern in public health and health services, and the lifetime societal costs have been shown to be substantial [41]. It is important for researchers and practitioners to find new ways to improve the reach and effectiveness of treatments to ensure equitable care. The

analyzed digital approach, implemented for the maintenance phase of weight management, was over four times more expensive to deliver than face-to-face maintenance sessions in a pilot RCT. When implemented outside a clinical trial, this cost is likely to reduce owing to the economics of scale and lower costs associated with technology usage. Our results highlight the importance of conducting further research to explore the cost-effectiveness of evidence-informed mHealth interventions in treating chronic diseases such as obesity across multiple centers.

Acknowledgments

This study was funded by the RCSI University of Medicine and Health Sciences - StAR program (grant 2151) and carried out as part of the Health Research Board (HRB) SPHeRE training program (SPHeRE/2013/1). The randomized controlled trial on which this study was based was funded by the HRB (HFP/2011/54) and the Children's Fund for Health & National Children's Research Centre of Ireland (PAC11-58). The funders had no role in the design of this study, including the collection, analyses, or interpretation of data, or in the preparation of the manuscript. We are grateful to Professor Amanda Burls for her supervision and guidance during the design and development of the Reactivate trial and the planning of this cost analysis. The authors wish to acknowledge all the staff at CHI Temple Street, especially the W82GO team who facilitated data collection for this study, as well as the study participants and families for their time.

Conflicts of Interest

The senior author (GO'M) led the design and ongoing development of the Reactivate system.

Multimedia Appendix 1
Development of unit costs.
[DOCX File , 30 KB - mhealth_v9i9e31621_app1.docx]

Multimedia Appendix 2
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1239 KB - mhealth_v9i9e31621_app2.pdf]

References

1. Tully L, Burls A, Sorensen J, El-Moslemany R, O'Malley G. Mobile health for pediatric weight management: systematic scoping review. JMIR Mhealth Uhealth 2020 Jun;8(6):e16214 [FREE Full text] [doi: 10.2196/16214] [Medline: 32490849]
2. Ells L, Mead E, Atkinson G, Corpeleijn E, Roberts K, Viner R, et al. Surgery for the treatment of obesity in children and adolescents. Cochrane Database Syst Rev 2015 Jun(6):CD011740. [doi: 10.1002/14651858.CD011740] [Medline: 26104326]
3. Mead E, Atkinson G, Richter B, Metzendorf MI, Baur L, Finer N, et al. Drug interventions for the treatment of obesity in children and adolescents. Cochrane Database Syst Rev 2016 Nov;11(11):CD012436 [FREE Full text] [doi: 10.1002/14651858.CD012436] [Medline: 27899001]
4. Mead E, Brown T, Rees K, Azevedo LB, Whittaker V, Jones D, et al. Diet, physical activity and behavioural interventions for the treatment of overweight or obese children from the age of 6 to 11 years. Cochrane Database Syst Rev 2017 Dec;6(6):CD012651. [doi: 10.1002/14651858.CD012651] [Medline: 28639319]
5. Rose T, Barker M, Maria JC, Morrison L, Lawrence W, Strömmer S, et al. A systematic review of digital interventions for improving the diet and physical activity behaviors of adolescents. J Adolesc Health 2017 Dec;61(6):669-677 [FREE Full text] [doi: 10.1016/j.jadohealth.2017.05.024] [Medline: 28822682]
6. Brown V, Tran H, Downing KL, Hesketh KD, Moodie M. A systematic review of economic evaluations of web-based or telephone-delivered interventions for preventing overweight and obesity and/or improving obesity-related behaviors. Obes Rev 2021 Jul;22(7):e13227. [doi: 10.1111/obr.13227] [Medline: 33763956]
7. Woo Baidal JA, Chang J, Hulse E, Turetsky R, Parkinson K, Rausch JC. Zooming towards a telehealth solution for vulnerable children with obesity during COVID-19. Obesity (Silver Spring) 2020 Apr;28(7):1184-1186. [doi: 10.1002/oby.22860] [Medline: 32352650]
8. Drummond M, Sculpher M, Claxton K, Stoddart G, Torrance G. Methods for the Economic Evaluation of Health Care Programmes. Fourth edition. United Kingdom: Oxford University Press; 2015:2.
9. Luxton DD, Hansen RN, Stanfill K. Mobile app self-care versus in-office care for stress reduction: a cost minimization analysis. J Telemed Telecare 2014 Dec;20(8):431-435. [doi: 10.1177/1357633X14555616] [Medline: 25316037]
10. Armstrong KA, Semple JL, Coyte PC. Replacing ambulatory surgical follow-up visits with mobile app home monitoring: modeling cost-effective scenarios. J Med Internet Res 2014 Sep;16(9):e213 [FREE Full text] [doi: 10.2196/jmir.3528] [Medline: 25245774]

11. McNamee P, Murray E, Kelly MP, Bojke L, Chilcott J, Fischer A, et al. Designing and undertaking a health economics study of digital health interventions. Am J Prev Med 2016 Nov;51(5):852-860. [doi: 10.1016/j.amepre.2016.05.007] [Medline: 27745685]
12. EASO COM applications Teddington, EASO Secretariat; 2021 online. European Association for the Study of Obesity. URL: https://easo.org/coms-2/easo-com-applications/ [accessed 2021-07-26]
13. O'Malley G, Clarke M, Burls A, Murphy S, Murphy N, Perry IJ. A smartphone intervention for adolescent obesity: study protocol for a randomised controlled non-inferiority trial. Trials 2014 Jan;15:43 [FREE Full text] [doi: 10.1186/1745-6215-15-43] [Medline: 24485327]
14. O'Malley G, Dowdall G, Burls A, Perry IJ, Curran N. Exploring the usability of a mobile app for adolescent obesity management. JMIR Mhealth Uhealth 2014 Jun;2(2):e29 [FREE Full text] [doi: 10.2196/mhealth.3262] [Medline: 25098237]
15. O'Malley GC. Childhood obesity treatment: integrating mobile health technology into a paediatric obesity service. Cork Open Research Archive. Cork: University College Cork; 2015. URL: https://cora.ucc.ie/handle/10468/2772 [accessed 2021-07-26]
16. Warschburger P, Kröller K. Loss to follow-up in a randomized controlled trial study for pediatric weight management (EPOC). BMC Pediatr 2016 Nov;16(1):184 [FREE Full text] [doi: 10.1186/s12887-016-0727-2] [Medline: 27842526]
17. O'Malley G, Perry IJ, Brinkley A, Sheridan N, Dowdall G, Murphy S, et al. Exploring the clinical effectiveness of a mobile health intervention for adolescent weight management. In: Obesity Reviews. 2020 Presented at: Joint European and International Congress on Obesity (ECO-ICO) 2020; September 2020; Online URL: https://doi.org/10.1111/obr.13115
18. Quigley E, O'Donnell S, Doyle G. Time-driven activity-based costing: a step-by-step guide to collecting patient-level health care costs. In: Research Methods Cases: Medicine and Health. London: SAGE; 2020.
19. How to use a cost effectiveness analysis to evaluate your digital health product. Public Health England. London URL: https://www.gov.uk/guidance/cost-effectiveness-analysis-health-economic-studies?utm_source=cc8def40-25cf-44da-be0f-771c4a285085&utm_medium=email&utm_campaign=govuk-notifications&utm_content=weekly [accessed 2021-07-26]
20. Guidelines for the economic evaluation of health technologies in Ireland. Health Information and Quality Authority. Dublin: Health Information and Quality Authority; 2020. URL: https://www.hiqa.ie/sites/default/files/2020-09/HTA-Economic-Guidelines-2020.pdf [accessed 2021-08-30]
21. PRSI Class A rates. Department of Employment Affairs and Social Protection. Dublin: Government of Ireland URL: https://www.gov.ie/en/publication/ffa563-prsi-class-a-rates/ [accessed 2021-07-26]
22. HR Circular 034/2019: standardisation of annual leave health and social care professionals Dublin.: Health Service Executive; 2019. URL: https://www.hse.ie/eng/staff/resources/hr-circulars/hr-circular-034-2019-re-annual-leave-for-health-social-care-professionals.pdf [accessed 2021-07-26]
23. Payscales for HSE staff. Health Service Executive. 2013. URL: https://healthservice.hse.ie/filelibrary/staff/november-2013-consolidated-salary-scales.pdf [accessed 2021-07-26]
24. Bösch PM, Becker F, Becker H, Axhausen KW. Cost-based analysis of autonomous mobility services. Transport Policy 2018 May;64:76-91. [doi: 10.1016/j.tranpol.2017.09.005]
25. Karlson CW, Rapoff MA. Attrition in randomized controlled trials for pediatric chronic conditions. J Pediatr Psychol 2009 Aug;34(7):782-793. [doi: 10.1093/jpepsy/jsn122] [Medline: 19064607]
26. Hill J, McGinn J, Cairns J, Free C, Smith C. A mobile phone-based support intervention to increase use of postabortion family planning in Cambodia: cost-effectiveness evaluation. JMIR Mhealth Uhealth 2020 Feb;8(2):e16276 [FREE Full text] [doi: 10.2196/16276] [Medline: 32130166]
27. Tully L, Case L, Arthurs N, Sorensen J, Marcin JP, O'Malley G. Barriers and facilitators for implementing paediatric telemedicine: rapid review of user perspectives. Front Pediatr 2021;9:630365 [FREE Full text] [doi: 10.3389/fped.2021.630365] [Medline: 33816401]
28. Farabi H, Rezapour A, Jahangiri R, Jafari A, Rashki Kemmak A, Nikjoo S. Economic evaluation of the utilization of telemedicine for patients with cardiovascular disease: a systematic review. Heart Fail Rev 2020 Nov;25(6):1063-1075. [doi: 10.1007/s10741-019-09864-4] [Medline: 31760593]
29. Sanyal C, Stolee P, Juzwishin D, Husereau D. Economic evaluations of eHealth technologies: a systematic review. PLoS One 2018;13(6):e0198112 [FREE Full text] [doi: 10.1371/journal.pone.0198112] [Medline: 29897921]
30. Moorman EL, Koskela-Staples NC, Mathai BB, Fedele DA, Janicke DM. Pediatric obesity treatment via telehealth: current evidence and future directions. Curr Obes Rep 2021 Jul;10(3):1-14. [doi: 10.1007/s13679-021-00446-w] [Medline: 34302603]
31. Jiang X, Yao J, You JH. Telemonitoring versus usual care for elderly patients with heart failure discharged from the hospital in the United States: cost-effectiveness analysis. JMIR Mhealth Uhealth 2020 Jul;8(7):e17846 [FREE Full text] [doi: 10.2196/17846] [Medline: 32407288]
32. Javalkar K, Rak E, Phillips A, Haberman C, Ferris M, Van Tilburg M. Predictors of caregiver burden among mothers of children with chronic conditions. Children (Basel) 2017 May;4(5):39 [FREE Full text] [doi: 10.3390/children4050039] [Medline: 28509853]
33. Kulik NL, Thomas EM, Iovan S, McKeough M, Kendzierski S, Leatherwood S. Access to primary care child weight management programs: urban parent barriers and facilitators to participation. J Child Health Care 2017 Dec;21(4):509-521. [doi: 10.1177/1367493517728401] [Medline: 29110523]

34. Lim CS, Janicke DM. Barriers related to delivering pediatric weight management interventions to children and families from rural communities. Children's Health Care 2013 Jul;42(3):214-230. [doi: 10.1080/02739615.2013.816596]
35. Nelson M, Russell T, Crossley K, Bourke M, McPhail S. Cost-effectiveness of telerehabilitation versus traditional care after total hip replacement: a trial-based economic evaluation. J Telemed Telecare 2021 Jul;27(6):359-366. [doi: 10.1177/1357633X19869796] [Medline: 31530065]
36. Iribarren SJ, Cato K, Falzon L, Stone PW. What is the economic evidence for mHealth? a systematic review of economic evaluations of mHealth solutions. PLoS One 2017 Feb;12(2):e0170581 [FREE Full text] [doi: 10.1371/journal.pone.0170581] [Medline: 28152012]
37. Fisher L, Glasgow RE, Huebschmann A. A scoping review and general user's guide for facilitating the successful use of eHealth programs for diabetes in clinical care. Diabetes Technol Ther 2021 Feb;23(2):133-145. [doi: 10.1089/dia.2020.0383] [Medline: 32865431]
38. Sadeghian R, Gooch R, Phelps L, Preud'homme D. The feasibility and satisfaction of using telemedicine to provide tertiary paediatric obesity care. J Int Soc Telemed eHealth 2018 Sep;6(1):1-9 [FREE Full text]
39. A Europe fit for the digital age. European Commission Directorate-General for Communication. 2020. URL: https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age_en [accessed 2021-07-26]
40. Kuhle S, Doucette R, Piccinini-Vallis H, Kirk SFL. Successful childhood obesity management in primary care in Canada: what are the odds? PeerJ 2015;3:e1327 [FREE Full text] [doi: 10.7717/peerj.1327] [Medline: 26623175]
41. What are the estimated costs of childhood overweight and obesity on the island of Ireland? Safefood. Cork; 2017. URL: https://www.safefood.net/research-reports/estimated-costs-childhood-obesity [accessed 2021-07-26]

Abbreviations

HCP: health care professional
mHealth: mobile health
RCT: randomized controlled trial
SDS: standardized deviation score

Edited by G Eysenbach; submitted 28.06.21; peer-reviewed by B Nievas Soriano; comments to author 22.07.21; revised version received 26.07.21; accepted 01.08.21; published 14.09.21.

Please cite as:
Tully L, Sorensen J, O'Malley G
Pediatric Weight Management Through mHealth Compared to Face-to-Face Care: Cost Analysis of a Randomized Control Trial
JMIR Mhealth Uhealth 2021;9(9):e31621
URL: https://mhealth.jmir.org/2021/9/e31621
doi:10.2196/31621
PMID:34519665

Original Paper

An mHealth-Based Intervention for Adolescents With Type 1 Diabetes and Their Parents: Pilot Feasibility and Efficacy Single-Arm Study

Bree Holtz[1], PhD; Katharine M Mitchell[1], MA; Amanda J Holmstrom[2], PhD; Shelia R Cotten[3], PhD; Julie K Dunneback[4], MSN, APRN, BC, CPNO, CDE; Jose Jimenez-Vega[5], MD; Deborah A Ellis[6], PhD; Michael A Wood[7], MD

[1]Department of Advertising and Public Relations, Michigan State University, East Lansing, MI, United States
[2]Department of Communication, Michigan State University, East Lansing, MI, United States
[3]Office of Research Development, Clemson University, Clemson, SC, United States
[4]Sparrow Health System, Lansing, MI, United States
[5]DeVos Children's Hospital, Grand Rapids, MI, United States
[6]Department of Family Medicine, Wayne State University, Detroit, MI, United States
[7]Pediatric Endocrinology, University of Michigan Medical School, Ann Arbor, MI, United States

Corresponding Author:
Bree Holtz, PhD
Department of Advertising and Public Relations
Michigan State University
404 Wilson Road
Room 309
East Lansing, MI, 48824
United States
Phone: 1 5178844537
Email: bholtz@msu.edu

Abstract

Background: Type 1 diabetes (T1D) affects more than 165,000 individuals younger than 20 years in the United States of America. The transition from parent management to parent-child team management, with the child taking on increased levels of self-care, can be stressful and is associated with a deterioration in self-management behaviors. Therefore, a mobile app intervention, *MyT1DHero,* was designed to facilitate diabetes-specific positive parent-adolescent communication and improve diabetes-related outcomes. The *MyT1DHero* intervention links an adolescent with T1D and their parent through 2 separate app interfaces and is designed to promote positive communication regarding T1D management.

Objective: The aim of this pilot study was to determine (1) the initial efficacy of the *MyT1DHero* intervention in improving diabetes outcomes in adolescents, specifically the hemoglobin A_{1c} (HbA_{1c}) levels, diabetes care adherence, and quality of life, and (2) the adolescents' overall satisfaction with this intervention.

Methods: This pilot study included 30 adolescent-parent pairs who used the *MyT1DHero* app in a 12-week single-arm clinical trial. Participants were recruited from the local pediatric endocrinology subspecialty clinic via snowball sampling. HbA_{1c} levels, diabetes care adherence, quality of life, family conflict, and satisfaction levels were measured and analyzed using paired sample two-sided *t* tests and linear regression analyses.

Results: The final analysis included 25 families. The mean age of the adolescents was 12.28 (SD 1.62) years. Half of the participants (13/25) reported a diabetes diagnosis of less than 5 years. After 12 weeks of the intervention, diabetes care adherence significantly improved (before the study: mean 3.87 [SD 0.59]; after the study: mean 4.19 [SD 0.65]; t_{21}=–2.52, P=.02, d=0.52) as did quality of life (before the study: mean 4.02 [SD 0.84]; after the study: mean 4.27 [SD 0.73]; t_{24}=2.48, P=.01, d=0.32). HbA_{1c} levels (before the study: mean 8.94 [SD 1.46]; after the study: mean 8.87 [SD 1.29]; t_{24}=0.67, P=.51, d=0.04) and family conflict (before the study: mean 2.45 [SD 0.55]; after the study: mean 2.61 [SD 0.45]; t_{23}=0.55, P=.14, d=0.32) changed in the

hypothesized direction, but the change was not significant. However, higher use of the mobile app was associated with more improvement in HbA_{1c} levels ($F_{1,20}$=9.74, P<.005; R^2=0.33). Overall, the adolescents were satisfied with the app intervention.

Conclusions: In a 12-week pilot study of the mobile app intervention designed to facilitate parent-adolescent communication for improving diabetes outcomes, significant benefits were demonstrated in self-care adherence and quality of life. A randomized controlled trial with a longer intervention is needed to replicate these findings and to determine the stability of the intervention effects.

Trial Registration: ClinicalTrials.gov NCT03436628; https://clinicaltrials.gov/ct2/show/NCT03436628

(JMIR Mhealth Uhealth 2021;9(9):e23916) doi:10.2196/23916

KEYWORDS

mobile health (mHealth); adolescents; type 1 diabetes; mobile phone; parent-adolescent; chronic disease; feasibility; diabetes management

Introduction

Background

Type 1 diabetes (T1D) affects more than 165,000 individuals younger than 20 years in the United States [1,2]. If the child is young at diagnosis, parents initially take over the responsibility for the management of T1D owing to the complexity of the disease and its management to prevent acute and long-term health complications. However, during adolescence, the child begins to take on more responsibility for diabetes self-management. While they are learning the skills needed for self-management, it is vital that parents remain engaged in the process, but they should also prepare their adolescents for independence [3,4].

Transition of T1D Care

The transition from parent management to a parent-adolescent team management—with the adolescent taking on increased responsibilities of self-care and parents learning to relinquish control—can be stressful and is associated with a deterioration in diabetes self-management adherence behaviors [5]. Diabetes management requires monitoring the blood glucose levels multiple times per day, counting the carbohydrates for meals and snacks, calculating and administering insulin doses by injection or pump, and adjusting the insulin levels or food depending on the glucose readings during a physical activity or illness [6-9]. The transition period, which begins when the child is an adolescent, generally lasts until early adulthood. However, during adolescence, there is often a decrease in the frequency of blood glucose monitoring, an increase in hemoglobin A_{1c} (HbA_{1c}) levels, and an increased risk for hospitalization associated with diabetic ketoacidosis. Such deterioration in management can result in poorer glycemic control and potentially lifelong complications [10]. Therefore, it is important to develop interventions to assist in the transition from parent to parent-adolescent team management in order to improve the health-related outcomes for the adolescent.

Parent and Adolescent Communication

Communication between parents and adolescents is critical at this juncture but is often fraught with difficulty and conflict that is further aggravated by adolescents' deteriorating adherence to their diabetes management and metabolic control [5,11,12]. Additionally, parents often report feeling worried about their child's health and feel compelled to check on them frequently [13]. In turn, many adolescents feel that their parents are frequently nagging them about their diabetes care [14]. Reducing parent-adolescent conflict around diabetes management and creating an open and trusting relationship during this tumultuous time is imperative, as a review of the literature found that parent-adolescent diabetes-related conflict is associated with poorer diabetic outcomes [15]. Studies have demonstrated that increased positive reciprocity, problem-solving, and positive communication between parents and adolescents, including optimistic communication, parental social support, and shared decision-making, are associated with positive T1D outcomes [16-19]. Reducing conflict and improving parent engagement during this transition may be possible though the use of technology (ie, mobile phones) to improve communication around diabetes care, thereby improving diabetes outcomes.

Mobile Health Apps

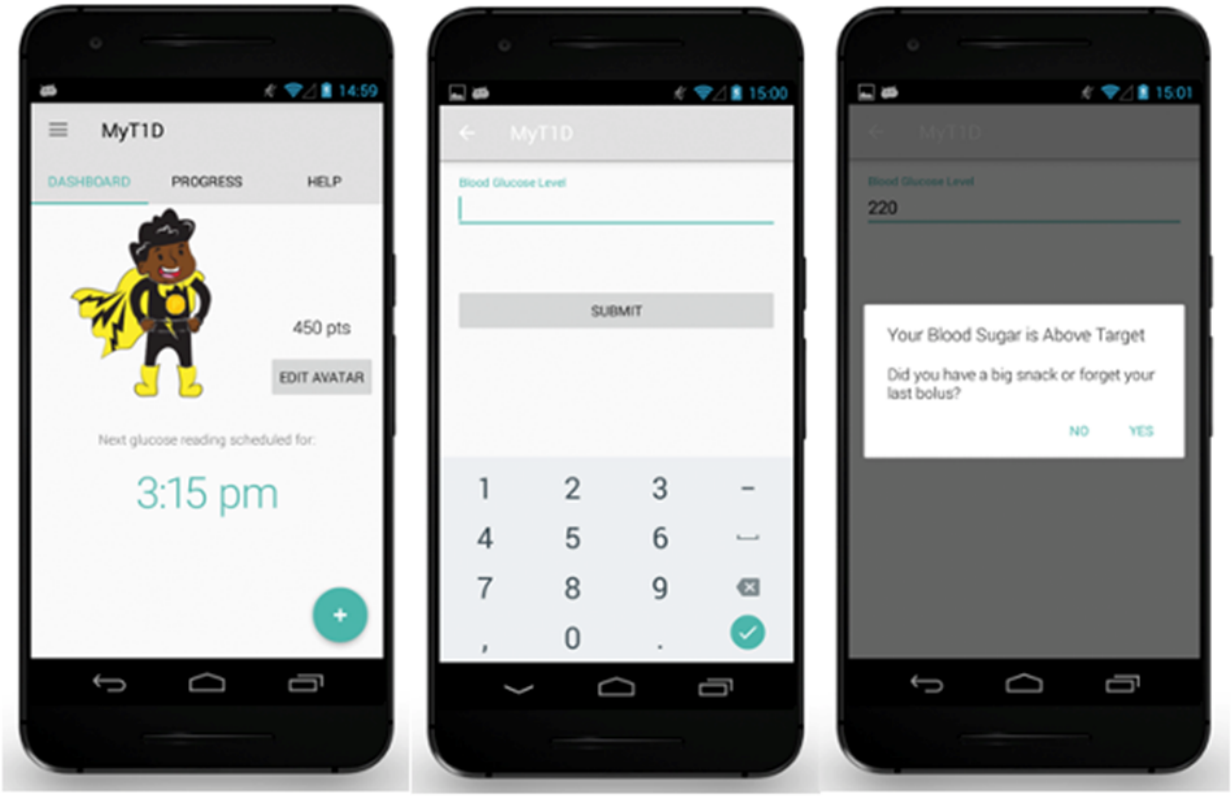
Mobile health (mHealth) apps have been rated as highly desirable by adolescents with T1D [20]. mHealth enables the use of more engaging strategies such as gamification and customization that have been shown to improve the frequency of blood glucose monitoring in adolescents with T1D [21], and in some cases, improve HbA_{1c} levels when compared to standard care [22,23]. Therefore, an mHealth intervention, *MyT1DHero*, was developed to help support and improve parent-adolescent communication related to diabetes management. The focus of *MyT1DHero*'s is to improve positive communication regarding T1D in an effort to reduce "blood sugar nagging" by parents [24]. These improvements in family communication were, in turn, expected to reduce family conflict, improve adolescents' adherence to blood glucose monitoring, and thus improve HbA_{1c} levels and quality of life.

MyT1DHero Intervention

The *MyT1DHero* intervention links the adolescent with T1D and their parent through 2 separate app interfaces—one for the adolescent and one for the parent. This app is designed to promote positive communication regarding T1D management. In order to frame the communication in that way, we asked that parents and adolescents establish a blood glucose testing schedule on the app, which will create time-specific reminders for the adolescent to test and enter their blood glucose readings. The app intervention also provides adolescents with tips on how

to address out-of-range blood glucose values. For example, when an adolescent enters a low blood glucose reading, the app provides a list of possible snacks (both standard and family-entered). When the child successfully measures and enters the blood glucose values into the app, points are awarded that the adolescents can use to "purchase" accessories for their "hero" avatar on the app. These accessories include capes, boots, logos, masks, and different hair colors and styles. The parents receive notifications about each blood glucose value their child enters. Additionally, parents and adolescents may send each other preprogrammed messages within the app that act as conversation prompts to facilitate positive communication about management, which also earns the adolescent points. In addition to the main functions of the intervention described above, there are also links to videos of other adolescents with T1D telling their stories and providing affirming messages. The reinforcement of diabetes behaviors has been found to help in supporting the adolescents to be more independent, while still keeping the parent aware. See Figure 1 for the screenshots of the app. This intervention is currently only available on an Android platform. Those without a phone or with a phone that was not compatible with *MyT1DHero* were provided a mobile phone for this study. The literature on the development and usability testing of *MyT1DHero* can be found elsewhere [14,25]. The objective of this pilot test was to determine (1) the initial efficacy of the *MyT1DHero* intervention in improving diabetes outcomes, specifically the HbA_{1c} levels, diabetes care adherence, and quality of life, and (2) the adolescents' overall satisfaction with the *MyT1DHero* intervention.

Figure 1. Screenshots of the MyT1DHero app.



Methods

Study Participants

Eligible adolescents were 10-15 years of age, had a T1D diagnosis for at least 6 months, with HbA_{1c} levels ≥7.0%, fluent in English, had a parent/guardian willing to participate, and allowed to use a mobile phone for this study. Participants were recruited from the local pediatric endocrinology subspecialty clinic through the Southeast Michigan JDRF (formerly Juvenile Diabetes Research Foundation) groups, as well as Facebook, Twitter, word-of-mouth, and snowball sampling. Recruitment to this study took place from September 2017 to December 2017. Thirty adolescent-parent pairs were enrolled in this study. This study was approved by the institutional review board at the Michigan State University and Sparrow Health System.

Study Design

This pilot study was a single-arm preclinical/postclinical trial without a control group. The intervention was tested over a 12-week period. At the initial visit, informed consent (parent) and assent (adolescent) were obtained. Additionally, venipuncture for determining baseline HbA_{1c} levels was performed at a laboratory (specifically for this study), and pretest survey measures were collected from both the adolescents and the parents.

Measures

The surveys included questions from the Diabetes Behavior Rating Scale [26], which assesses how well the adolescent is managing diabetes care, and the survey was given to both the adolescent and the parent (see Textbox 1 for example questions, α=.64-.65; all reliabilities reported are based on this study's

sample). This scale has been tested in this population as a measure of adherence several times and includes daily and long-term adherence items [26,27]. This scale has 5 response categories ranging from 1 (Never) to 5 (Always). A higher score indicates a greater level of adherence.

Textbox 1. Survey for the assessment of diabetes care.

Examples of questions in the survey for adolescents

How often…

- was the amount of insulin that your doctor prescribed (including adjustments for diet or blood glucose level) actually taken?
- were your blood sugar numbers written in your log, diary, or chart?
- is insulin correctly adjusted for meals you eat away from home (eg, at restaurants, parties)?
- are clinic or doctor's appointments kept?
- is your doctor/nurse called for changes in insulin dose if you get frequent "high" or "low" blood sugar levels?
- is your doctor/nurse called if you have severe diabetic symptoms that you cannot correct (eg, drinking a lot, needing fast sugar a lot)?

Examples of questions in the survey for parents

How often…

- was the amount of insulin that your child's doctor prescribed (including adjustments for diet or blood glucose level) actually taken?
- were your child's blood sugar numbers written in their log, diary, or chart?
- is insulin correctly adjusted for your child's meals that they eat away from home (eg, at restaurants, parties)?
- are clinic or doctor's appointments kept?
- is your child's doctor/nurse called for changes in insulin dose if you get frequent "high" or "low" blood sugar levels?
- is your child's doctor/nurse called if you have severe diabetic symptoms that you cannot correct (eg, drinking a lot, needing fast sugar a lot)?

Quality of life [28] was self-reported by the adolescent using the PedsQL generic scale (α=.96). This 23-item scale contained 5 response categories ranging from 1 (Never) to 5 (Almost Always); a lower score indicates higher perception of quality of life.

Family conflict was measured using the Revised Diabetes Family Conflict Scale [29] given to both the adolescent and parent. Reliability ranged from α=.94 to α=.95. This 12-item scale has 3 response categories ranging from 1 (Almost Always) to 3 (Never), with a higher score indicating a lower level of conflict.

Satisfaction with the app was measured at the end of the study by using 8 questions from the Poststudy System Usability Questionnaire [30] (α=.84-.85). Satisfaction was measured by asking participants about the different aspects of the app, for example, "I found the *MyT1DHero* app easy to use." Response options ranged from 1 (Strongly Agree) to 7 (Strongly Disagree), with a lower score indicating higher satisfaction.

Usage was measured in multiple ways. For the number of days used, a day of usage was counted if the child entered the app. The number of days used was grouped as low, medium, and high usage. The low category indicated entering the readings in the app once or less per day for the month, the medium category indicated twice or less per day for the month, and the high category indicated at least 3 times per day. The "high" category indicates proper use of the app as we asked for at least 4 blood sugar entries per day. In addition, the number of blood sugar entries per day and the number of messages sent per day were also considered, and these data were stored by the server. Finally, changes in app usage across the 90-day study period were measured.

Procedures

At enrollment, participants received a detailed explanation of the *MyT1DHero* app intervention. The parents and adolescents were asked to use the app for 12 weeks, with specific instructions for adolescents to enter their blood glucose readings at least 4 times per day. At the conclusion of the intervention, a repeat venipuncture for HbA_{1c} levels was performed at a laboratory, and posttest surveys were conducted (either online or in-person). Participants were compensated US $50 (US $25 for parents and US $25 for adolescents) after completing the tasks at the initial enrollment meeting. Participants received an additional US $50 (US $25 for parents and US $25 for adolescents) after completing the repeat venipuncture. Finally, an additional US $100 (US $50 for parents and US $50 for adolescents) was given to the participants after completing the posttest surveys and returning phones, if applicable.

Statistical Analysis

Statistical analyses were performed using SPSS (v.25, IBM Corp). Descriptive statistics were conducted to summarize the sample characteristics and study variables. Paired-sample two-sided *t* tests were conducted to determine if there were statistical differences in HbA_{1c} levels and the other study outcomes from preintervention to postintervention. Additionally, linear regression analysis was conducted to evaluate the association between the adolescents' usage of the app (high usage, 70-90 days; low usage, 69 days or less) and change in HbA_{1c} levels (pre-HbA_{1c} subtracted from post-HbA_{1c}). We also

tested the correlations between the adolescent and parent measures, as well as between the age of the adolescent and the satisfaction of the intervention.

Results

Participant Characteristics

From September 2017 to December 2017, 30 families were screened and enrolled in this study. The final analysis included 25 families (5 never completed posttest assessments, Figure 2). The majority of the adolescents were Whites (22/25, 88%). The mean age of the adolescents was 12.28 (SD 1.62) years and the majority of the parents was in the age range of 35-44 years (n=10) (see Table 1). Comparing the participants who dropped out, there were no differences in the demographics, except in the length of diagnosis. All the participants who dropped out (n=5) reported that they had been diagnosed for 5 or more years. Of those who completed the study, 12 reported that they had been diagnosed for 5 or more years, 11 had been diagnosed for 1 to 5 years, and 2 had been diagnosed that year (t_{22}=1.73, P<.001).

Figure 2. Participant recruitment flowchart.

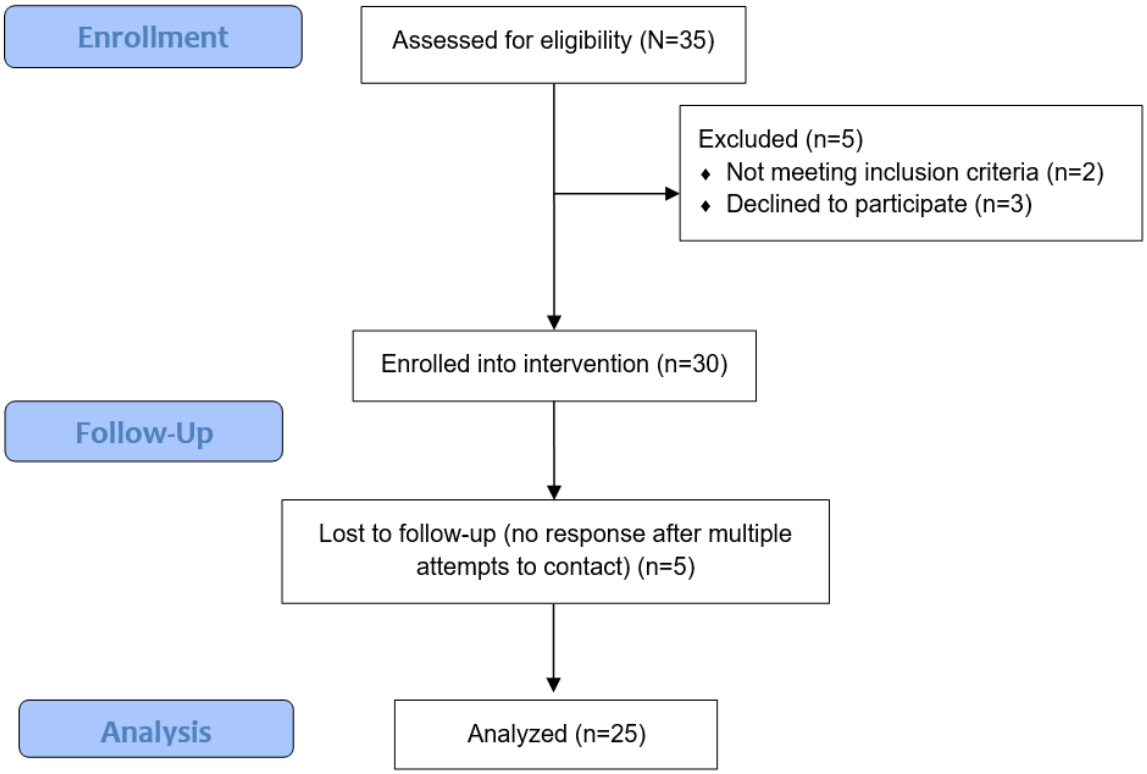


Table 1. Demographics of the participants and outcomes (N=25).

Demographics	Values, n (%)
Adolescent age (years)	
10	4 (16)
11	4 (16)
12	7 (28)
13	5 (20)
14	1 (4)
15	4 (16)
Adolescent gender	
Female	12 (48)
Male	13 (52)
Length of diagnosis	
6 months to less than 1 year	2 (8)
1 year to less than 5 years	11 (44)
5 years or more	12 (48)
Ethnicity/race	
White	22 (88)
Black or African American	2 (8)
Hispanic or Latino	1 (4)
Parent relationship to adolescent	
Biological mother	22 (88)
Biological father	1 (4)
Adoptive mother	1 (4)
Adoptive father	1 (4)
Parent age (years)	
25-34	1 (4)
35-44	10 (40)
45-54	8 (32)
55-64	3 (12)
Income (US $)	
Less than $25,000	2 (8)
$25,000-$49,999	4 (16)
$50,000-$74,999	5 (20)
$75,000-$99,999	4 (16)
$100,000 or more	10 (40)
Insurance	
Employer provided	18 (72)
Medicaid	2 (8)
Medicare	1 (4)
Self-insured, non-Medicare, or Medicaid	1 (4)
Other	3 (12)
Treatment[a]	
Pump	21 (84)

Demographics	Values, n (%)
Continuous glucose monitor	12 (45)
Multiple daily injections	4 (16)

[a]The participants in this category were on multiple treatments, which is reflected in the percentages in the subcategories.

Study Outcomes

HbA_{1c} levels

After 12 weeks, there were no significant changes (Table 2) in the HbA_{1c} levels (before the study: mean 8.94 [SD 1.46]; after the study: mean 8.87 [SD 1.29]; t_{24}=0.67, P=.51, d=0.04).

Table 2. Results of the outcome variables.

Group, outcome variables	Pretest values, mean (SD)	Posttest values, mean (SD)	*P* value
Adolescents			
Hemoglobin A_{1c} levels	8.94 (1.46)	8.87 (1.29)	.51
Diabetes behavior rating scale (1=Never to 5=Always)	3.87 (0.59)	4.19 (0.65)	.02
Quality of life scale (1=Never to 5=Almost Always)	4.02 (0.84)	4.27 (0.73)	.01
Revised diabetes family conflict scale (1=Almost Always to 3=Never)	2.45 (0.55)	2.61 (0.45)	.14
Parents			
Diabetes behavior rating scale (1=Never to 5=Always)	4.04 (0.50)	4.34 (0.47)	.02
Revised diabetes family conflict scale (1=Almost Always to 3=Never)	2.47 (0.78)	2.51 (0.85)	.78

Diabetes Care Adherence and Quality of Life

Significant improvements were found in the diabetes behavior and quality of life of the adolescents. Posttest measurements (mean 4.19 [SD 0.65]) of diabetes behavior demonstrated improvement compared to the pretest measurements (mean 3.87 [SD 0.59]; t_{21}=–2.52, P=.02, d=0.52). The parents' perceptions of diabetes behavior also improved from before the study (mean 4.04 [SD 0.50]) to after the study (mean 4.34 [SD 0.47]; t_{21}=–2.58, P=.02, d=0.62). Additionally, the scores of the child and the parent Diabetes Behavior Rating Scale were significantly correlated (r_{24}=0.46, P=.04). The quality of life significantly improved after the intervention for the adolescents (mean 4.27 [SD 0.73]) compared to that before the intervention (mean 4.02 [SD 0.84]; t_{24}=2.48, P=.01, d=0.32). Although there was an improvement in the family conflict for the adolescents from before the intervention (mean 2.45 [SD 0.55]) to after the intervention (mean 2.61 [SD 0.45]; t_{23}=0.55, P=.14, d=0.32), the improvement was not significant. The results of the parents' perception of conflict did not change significantly from before the intervention (mean 2.47 [SD 0.78]) to after the intervention (mean 2.51 [SD 0.85]; t_{21}=–0.28, P=.78, d=0.28). The scores of the child and the parent Revised Diabetes Family Conflict Scale were significantly correlated (r_{22}=0.47, P=.03).

Satisfaction With the Intervention

Satisfaction with the intervention was measured through 8 questions. Overall, both the adolescents (mean 2.19 [SD 0.94]) and the parents (mean 2.26 [SD 1.29]) rated the intervention as satisfactory. The results of Pearson correlation indicated that there was no association between the age of the child and the satisfaction level (r_{24}=–0.02, P=.92). Each satisfaction item with the descriptive statistics can be found in Table 3.

Table 3. Satisfaction of the participants with the MyT1DHero app measured using the poststudy system usability questionnaire.[a]

Variables (scored on a scale of 1-7)	Adolescents		Parents	
	Mean (SD)	Range	Mean (SD)	Range
Overall, I'm satisfied with how easy it is to use this app.	2.70 (1.59)	1-7	2.10 (1.62)	1-7
Learning to use this app was easy.	1.74 (0.94)	1-5	2.00 (1.77)	1-7
The characters on the screen were easy to see.	1.63 (0.79)	1-3	1.48 (0.68)	1-3
The app is understandable.	2.15 (1.13)	1-5	1.97 (1.60)	1-7
It was easy to become skilled using this app.	2.11 (1.25)	1-6	2.00 (1.55)	1-7
I found the app easy to use.	2.30 (1.68)	1-7	1.87 (1.59)	1-7
I think using this app is a good idea.	2.93 (1.94)	1-7	2.19 (1.76)	1-7
Proper type 1 diabetes terms were used throughout the app.	2.04 (1.13)	1-5	2.45 (1.79)	1-7
All satisfaction items	2.19 (0.94)	1-7	2.56 (1.29)	1-7

[a]Response categories: 1=strongly agree, 4=neutral, 7=strongly disagree.

App Usage

The average number of days the app was used was 63 (SD 27.70) days. The majority of the participants had usage in the medium or high use categories (21/25, 84%). In addition, participants entered their blood sugar readings an average of less than 3 times per day (mean 2.70 [SD 2.01]) and sent less than 1 message on average per day (mean 0.20 [SD 0.24]) or an average of approximately 23 messages over the course of the entire study period (mean 22.60 [SD 22.38]). Over the 90-day study period, daily use varied. At the start of the study (days 1-5), participants used the app the most times per day (mean 24.54 [SD 9.36]). In the following 60 days, participants used the app on average 7 times per day (mean 7.46 [SD 2.36]). The average use per day then fell further during the final 25 days of the study (mean 4.45 [SD 1.39]). However, linear regression analysis indicated a significant effect between the *overall* use of the app (high/medium/low) and improvement in HbA_{1c} levels ($F_{1,20}$=9.74, P<.005; R^2=0.33), demonstrating that the adolescents who used *MyT1DHero* more had greater improvements in their HbA_{1c} levels. We then conducted a separate linear regression to examine if there were any specific functions that predicted the change in HbA_{1c} levels ($F_{5,7}$=–5.681, P=.21; R^2=0.80). The data demonstrated that the more the adolescents logged their blood glucose readings and the more messages they sent to their parents, the more their HbA_{1c} levels improved (Table 4).

Table 4. Linear regression of the types of usage predicting the change in the hemoglobin A_{1c} levels.

Function	Unstandardized β	SE	β	*t (df)*	*P* value
Constant	–1.356	0.332	N/A[a]	–4.089 (24)	.005
Glucose logging (child only)	.004	0.001	1.098	4.798 (24)	.002
Messaging (child)	–.069	0.020	–1.168	–3.492 (24)	.01
Messaging (parent)	.042	0.014	.945	3.024 (24)	.02
Glucose-specific messaging (child)	.066	0.044	.648	1.494 (24)	.18
Glucose-specific messaging (parent)	–.013	0.012	–.470	–1.091 (24)	.31

[a]N/A: not applicable.

Further, participants who used a continuous glucose monitor (CGM) were asked to enter their calibration blood glucose values 1-2 times per day but were allowed to use their CGM number for the other entries. Importantly, there were no significant differences between CGM users and non-CGM users in the usage of the app by category (t_{24}=0.561, P=.58) nor in the days used (t_{24}=0.092, P=.93).

Discussion

Overview of This Study

The purpose of this study was to conduct a preliminary evaluation of the efficacy of the *MyT1DHero* mobile app intervention. The goal of the *MyT1DHero* intervention is to facilitate positive diabetes-related communication between adolescents with T1D and their parents to improve diabetes outcomes and help in the transition of diabetes management responsibilities. As the adolescent transitions to more self-care, they need to perform these management behaviors independently, such as monitoring blood glucose levels and taking the appropriate actions if these values are out of range.

Adherence Improvements in Diabetes Care

The results from this preliminary evaluation demonstrate that diabetes care adherence as reported by both the adolescents and the parents and the quality of life as reported by adolescents

improved after using the *MyT1DHero* app. Demonstrating improvement in diabetes management behaviors is critical for this population, especially in this age group, as these behaviors need to be learned and integrated into everyday life for the rest of the adolescents' life. Studies have demonstrated that perceptions of quality of life significantly affect adherence to medical recommendations and overall health outcomes [30-33]. These findings suggest that participating in this type of intervention could improve primary clinical outcomes such as glycemic control; however, additional studies with longer follow-up periods are needed to see if improvements in diabetes care would be manifested in improved blood glucose levels.

HbA_{1c} Improvements

Although there was no overall significant effect on the HbA_{1c} levels, the results of our study demonstrate that those who used the app more had more improvements in their HbA_{1c} levels. This indicates the importance of creating an engaging intervention that is tailored to individual families in order to encourage long-term engagement [34]. Moreover, improvements in diabetes behavior and quality of life through long-term use of the *MyT1DHero* app intervention could also positively affect clinical diabetes outcomes; however, further research is needed.

Adolescent App Satisfaction

This study shows promise for interventions delivered via mobile phones. Past research has shown that mobile phones and apps are perceived by adolescents as a more acceptable way to communicate with a parent, particularly in social contexts in which adolescents may be reluctant to engage in diabetes management, for example, when they are with their peers [35]. Furthermore, while our sample spans 2 developmental stages, this study was conducted to understand the acceptability of the intervention. However, there was no relationship between age and satisfaction of the app intervention, indicating that more qualitative work should be done in this area to determine the acceptability and appropriateness for different ages.

Communication Between Parents and Adolescents

Future research should include not only assessments of the intervention's effect on negative communication patterns between parents and adolescents but also the positive interactions. For example, although the use of the intervention did not result in a significant drop in family conflict, it is possible that positive communication behaviors such as self-disclosure, expressions of affection, and social support did increase throughout the length of the intervention. Further, our data demonstrate that when parents initiate the messaging that is perceived to be negative for the child, it can be considered as "nagging" by the parents. However, when the child reached out, there was more positive impact [24]. Future studies should also measure different types of communication—both positive and negative.

Limitations of This Study

This research has the following limitations. The small sample size was fairly homogenous. The sample reported relatively low levels of baseline family conflict, which may also have limited the study's ability to show significant reductions in conflict at the 12-week follow-up. This study used self-report measures, which should be noted as a limitation. Providing a mobile phone might have also affected the results. We had to provide phones to participants who had an Apple phone; therefore, these participants ended up carrying 2 phones. Additionally, although we asked families to provide us with the blood glucose trends via a meter download at baseline and at study end, the majority of the families failed to do this. Therefore, an objective measure of blood glucose checking was not available for use in the analyses. Finally, we cannot know for sure if the adolescents entered their blood glucose values accurately and honestly in the app. Further studies are necessary to determine the intervention efficacy in promoting positive communication around T1D management.

Conclusion

The strengths of this study demonstrate a positive impact on adolescents' diabetes behaviors and quality of life through the use of an app focused on fostering positive communication with their parents. Mobile phone apps provide a promising avenue for improving parental involvement in the transition of diabetes management from children to adolescents. Using a family-based app intervention provides parents access to their child's health care treatment and allows adolescents to easily ask for help even when the family is not together. More work regarding the use of similar app interventions is needed to further demonstrate their positive and long-term effectiveness for adolescents with T1D. This study provides preliminary support for the feasibility and efficacy of the *MyT1DHero* intervention for improving health outcomes in adolescents with T1D. mHealth interventions developed to improve the communication between adolescents and their parents during the transition to adolescent self-care could be a promising method to improve health outcomes in this population.

Acknowledgments

This work was supported by the American Diabetes Association (#1-16-ICTS-045). All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

References

1. Pettitt DJ, Talton J, Dabelea D, Divers J, Imperatore G, Lawrence JM, SEARCH for Diabetes in Youth Study Group. Prevalence of diabetes in U.S. youth in 2009: the SEARCH for diabetes in youth study. Diabetes Care 2014 Mar;37(2):402-408 [FREE Full text] [doi: 10.2337/dc13-1838] [Medline: 24041677]
2. Majumder E, Cogen FR, Monaghan M. Self-Management Strategies in Emerging Adults With Type 1 Diabetes. J Pediatr Health Care 2017;31(1):29-36 [FREE Full text] [doi: 10.1016/j.pedhc.2016.01.003] [Medline: 26861574]
3. Nordfeldt S, Ängarne-Lindberg T, Nordwall M, Krevers B. Parents of adolescents with type 1 diabetes--their views on information and communication needs and internet use. A qualitative study. PLoS One 2013;8(4):e62096 [FREE Full text] [doi: 10.1371/journal.pone.0062096] [Medline: 23626772]
4. Palmer DL, Berg CA, Butler J, Fortenberry K, Murray M, Lindsay R, et al. Mothers', fathers', and children's perceptions of parental diabetes responsibility in adolescence: examining the roles of age, pubertal status, and efficacy. J Pediatr Psychol 2009 Mar;34(2):195-204 [FREE Full text] [doi: 10.1093/jpepsy/jsn073] [Medline: 18632787]
5. Ellis DA, Podolski C, Frey M, Naar-King S, Wang B, Moltz K. The role of parental monitoring in adolescent health outcomes: impact on regimen adherence in youth with type 1 diabetes. J Pediatr Psychol 2007 Oct;32(8):907-917. [doi: 10.1093/jpepsy/jsm009] [Medline: 17426045]
6. Downing J, Gleeson HK, Clayton PE, Davis JRE, Wales JK, Callery P. Transition in endocrinology: the challenge of maintaining continuity. Clin Endocrinol (Oxf) 2013 Jan;78(1):29-35. [doi: 10.1111/j.1365-2265.2012.04473.x] [Medline: 22734661]
7. Herbert L, Owen V, Pascarella L, Streisand R. Text message interventions for children and adolescents with type 1 diabetes: a systematic review. Diabetes Technol Ther 2013 May;15(5):362-370. [doi: 10.1089/dia.2012.0291] [Medline: 23550554]
8. Timms N, Lowes L. Autonomy or non-compliance in adolescent diabetes? Br J Nurs 1999;8(12):794-7, 800. [doi: 10.12968/bjon.1999.8.12.6574] [Medline: 10670296]
9. Weissberg-Benchell J, Wolpert H, Anderson BJ. Transitioning from pediatric to adult care: a new approach to the post-adolescent young person with type 1 diabetes. Diabetes Care 2007 Oct;30(10):2441-2446. [doi: 10.2337/dc07-1249] [Medline: 17666466]
10. Morrison S, Dashiff C, Abdullatif H, Moreland E. Parental separation anxiety and diabetes self-management of older adolescents: a pilot study. Pediatric Nursing 2012;38(2):88-95 ISSN: 0097-9805 [FREE Full text]
11. Gustafsson PA, Cederblad M, Ludvigsson J, Lundin B. Family interaction and metabolic balance in juvenile diabetes mellitus. A prospective study. Diabetes Res Clin Pract 1987 Nov;4(1):7-14. [doi: 10.1016/s0168-8227(87)80027-x] [Medline: 3691301]
12. Hauser ST, Jacobson AM, Lavori P, Wolfsdorf JI, Herskowitz RD, Milley JE, et al. Adherence among children and adolescents with insulin-dependent diabetes mellitus over a four-year longitudinal follow-up: II. Immediate and long-term linkages with the family milieu. J Pediatr Psychol 1990 Aug;15(4):527-542. [doi: 10.1093/jpepsy/15.4.527] [Medline: 2258799]
13. Holtz BE, Mitchell KM, Nuttall AK, Cotten SR, Hershey DD, Dunneback JK, et al. Using User-Feedback to Develop a Website: MyT1DHope, for Parents of Children with T1D. Health Commun 2020 Mar;35(3):281-288. [doi: 10.1080/10410236.2018.1560579] [Medline: 30592228]
14. Holtz BE, Murray KM, Hershey DD, Dunneback JK, Cotten SR, Holmstrom AJ, et al. Developing a Patient-Centered mHealth App: A Tool for Adolescents With Type 1 Diabetes and Their Parents. JMIR Mhealth Uhealth 2017 May 19;5(4):e53 [FREE Full text] [doi: 10.2196/mhealth.6654] [Medline: 28428167]
15. Tsiouli E, Alexopoulos EC, Stefanaki C, Darviri C, Chrousos GP. Effects of diabetes-related family stress on glycemic control in young patients with type 1 diabetes: Systematic review. Can Fam Physician 2013 Mar;59(2):143-149 [FREE Full text] [Medline: 23418238]
16. Wysocki T, Greco P, Harris MA, Bubb J, White NH. Behavior therapy for families of adolescents with diabetes: maintenance of treatment effects. Diabetes Care 2001 Mar;24(3):441-446. [doi: 10.2337/diacare.24.3.441] [Medline: 11289465]
17. Hanson CL, Henggeler SW, Burghen GA. Social competence and parental support as mediators of the link between stress and metabolic control in adolescents with insulin-dependent diabetes mellitus. Journal of Consulting and Clinical Psychology 1987;55(4):529-533. [doi: 10.1037/0022-006x.55.4.529]
18. Wysocki T, Lochrie A, Antal H, Buckloh LM. Youth and parent knowledge and communication about major complications of type 1 diabetes: associations with diabetes outcomes. Diabetes Care 2011 Aug;34(8):1701-1705 [FREE Full text] [doi: 10.2337/dc11-0577] [Medline: 21659641]
19. Katz ML, Volkening LK, Butler DA, Anderson BJ, Laffel LM. Family-based psychoeducation and Care Ambassador intervention to improve glycemic control in youth with type 1 diabetes: a randomized trial. Pediatr Diabetes 2014 Mar;15(2):142-150 [FREE Full text] [doi: 10.1111/pedi.12065] [Medline: 23914987]
20. Markowitz JT, Cousineau T, Franko DL, Schultz AT, Trant M, Rodgers R, et al. Text messaging intervention for teens and young adults with diabetes. J Diabetes Sci Technol 2014 Oct;8(5):1029-1034 [FREE Full text] [doi: 10.1177/1932296814540130] [Medline: 25172879]
21. Cafazzo JA, Casselman M, Hamming N, Katzman DK, Palmert MR. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. J Med Internet Res 2012 May 08;14(3):e70 [FREE Full text] [doi: 10.2196/jmir.2058] [Medline: 22564332]

22. Kitsiou S, Paré G, Jaana M, Gerber B. Effectiveness of mHealth interventions for patients with diabetes: An overview of systematic reviews. PLoS One 2017;12(3):e0173160 [FREE Full text] [doi: 10.1371/journal.pone.0173160] [Medline: 28249025]
23. Holtz B, Lauckner C. Diabetes management via mobile phones: a systematic review. Telemed J E Health 2012 May;18(3):175-184. [doi: 10.1089/tmj.2011.0119] [Medline: 22356525]
24. Holtz BE, Mitchell KM, Holmstrom AJ, Cotten SR, Hershey DD, Dunneback JK, et al. Teen and parental perspectives regarding transition of care in type 1 diabetes. Children and Youth Services Review 2020 Mar;110:104800. [doi: 10.1016/j.childyouth.2020.104800]
25. Holtz BE, Murray KM, Hershey DD, Richman J, Dunneback JK, Vyas A, et al. The design and development of MyT1DHero: A mobile app for adolescents with type 1 diabetes and their parents. J Telemed Telecare 2017 Dec 11;25(3):172-180. [doi: 10.1177/1357633x17745470]
26. Iannotti RJ, Nansel TR, Schneider S, Haynie DL, Simons-Morton B, Sobel DO, et al. Assessing regimen adherence of adolescents with type 1 diabetes. Diabetes Care 2006 Oct 26;29(10):2263-2267. [doi: 10.2337/dc06-0685] [Medline: 17003304]
27. Hood KK, Peterson CM, Rohan JM, Drotar D. Association between adherence and glycemic control in pediatric type 1 diabetes: a meta-analysis. Pediatrics 2009 Dec;124(6):e1171-e1179. [doi: 10.1542/peds.2009-0207] [Medline: 19884476]
28. Varni JW, Burwinkle TM, Jacobs JR, Gottschalk M, Kaufman F, Jones KL. The PedsQL in type 1 and type 2 diabetes: reliability and validity of the Pediatric Quality of Life Inventory Generic Core Scales and type 1 Diabetes Module. Diabetes Care 2003 Mar;26(3):631-637. [doi: 10.2337/diacare.26.3.631] [Medline: 12610013]
29. Hood KK, Butler DA, Anderson BJ, Laffel LM. Updated and revised Diabetes Family Conflict Scale. Diabetes Care 2007 Jul;30(7):1764-1769 [FREE Full text] [doi: 10.2337/dc06-2358] [Medline: 17372149]
30. Lewis JR. IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use. International Journal of Human-Computer Interaction 1995 Jan;7(1):57-78. [doi: 10.1080/10447319509526110]
31. Graue M, Wentzel-Larsen T, Hanestad BR, Søvik O. Health-related quality of life and metabolic control in adolescents with diabetes: the role of parental care, control, and involvement. J Pediatr Nurs 2005 Oct;20(5):373-382. [doi: 10.1016/j.pedn.2005.08.005] [Medline: 16182097]
32. Nansel T, Weisberg-Benchell J, Wysocki T, Laffel L, Anderson B, Steering Committee of the Family Management of Diabetes Study. Quality of life in children with Type 1 diabetes: a comparison of general and diabetes-specific measures and support for a unitary diabetes quality-of-life construct. Diabet Med 2008 Dec;25(11):1316-1323 [FREE Full text] [doi: 10.1111/j.1464-5491.2008.02574.x] [Medline: 19046222]
33. Nardi L, Zucchini S, D'Alberton F, Salardi S, Maltoni G, Bisacchi N, et al. Quality of life, psychological adjustment and metabolic control in youths with type 1 diabetes: a study with self- and parent-report questionnaires. Pediatr Diabetes 2008 Oct;9(5):496-503. [doi: 10.1111/j.1399-5448.2008.00414.x] [Medline: 18507786]
34. Mitchell KM, Holtz BE, McCarroll A. Patient-Centered Methods for Designing and Developing Health Information Communication Technologies: A Systematic Review. Telemed J E Health 2019 Nov;25(11):1012-1021. [doi: 10.1089/tmj.2018.0236] [Medline: 30576258]
35. Davie R, Panting C, Charlton T. Mobile phone ownership and usage among pre-adolescents. Telematics and Informatics 2004 Nov;21(4):359-373. [doi: 10.1016/j.tele.2004.04.001]

Abbreviations

CGM: continuous glucose monitor
HbA_{1c}: hemoglobin A_{1c}
mHealth: mobile health
T1D: type 1 diabetes

Edited by L Buis; submitted 27.08.20; peer-reviewed by SHM Guo; comments to author 02.11.20; revised version received 16.12.20; accepted 23.08.21; published 14.09.21.

Please cite as:
Holtz B, Mitchell KM, Holmstrom AJ, Cotten SR, Dunneback JK, Jimenez-Vega J, Ellis DA, Wood MA
An mHealth-Based Intervention for Adolescents With Type 1 Diabetes and Their Parents: Pilot Feasibility and Efficacy Single-Arm Study
JMIR Mhealth Uhealth 2021;9(9):e23916
URL: https://mhealth.jmir.org/2021/9/e23916
doi:10.2196/23916
PMID:34519670

Original Paper

Using an mHealth App (iGAM) to Reduce Gingivitis Remotely (Part 2): Prospective Observational Study

Guy Tobias[1], DMD, MPH, PhD; Assaf B Spanier[2], PhD

[1]Department of Community Dentistry, Faculty of Dental Medicine, The Hebrew University-Hadassah School of Dental Medicine, Jerusalem, Israel
[2]Department of Software Engineering, Azrieli College of Engineering, Jerusalem, Israel

Corresponding Author:
Guy Tobias, DMD, MPH, PhD
Department of Community Dentistry
Faculty of Dental Medicine
The Hebrew University-Hadassah School of Dental Medicine
Ein Kerem
Jerusalem, 91120
Israel
Phone: 972 52 705 2333
Email: guy.tobias@mail.huji.ac.il

Related Articles:

Companion article: https://mhealth.jmir.org/2020/8/e19433/

Companion article: https://formative.jmir.org/2021/10/e26125

Abstract

Background: Gingivitis is a nonpainful, inflammatory condition that can be managed at home. Left untreated, gingivitis can lead to tooth loss. Periodic dental examinations are important for early diagnosis and treatment of gum diseases. To contain the spread of the coronavirus, governments, including in Israel, have restricted movements of their citizens which might have caused routine dental checkups to be postponed.

Objective: This study aimed to examine the ability of a mobile health app, iGAM, to reduce gingivitis, and to determine the most effective interval between photograph submissions.

Methods: A prospective observational cohort study with 160 unpaid participants divided into 2 equal groups using the iGAM app was performed. The intervention group photographed their gums weekly for 8 weeks. The wait-list control group photographed their gums at the time of recruitment and 8 weeks later. After photo submission, the participants received the same message “we recommended that you read the information in the app regarding oral hygiene habits.” A single-blinded researcher examined all the images and scored them according to the Modified Gingival Index (MGI).

Results: The average age of the intervention group was 26.77 (SD 7.43) and 28.53 (SD 10.44) for the wait-list control group. Most participants were male (intervention group: 56/75,74.7%; wait-list control group: 34/51, 66.7%) and described themselves as “secular”; most were “single” non-smokers (intervention group: 56/75, 74.7%; wait-list control group: 40/51, 78.4%), and did not take medications (intervention group: 64/75, 85.3%; wait-list control group: 40/51, 78.4%). A total of 126 subjects completed the study. A statistically significant difference (P<.001) was found in the dependent variable (MGI). Improvements in gingival health were noted over time, and the average gingivitis scores were significantly lower in the intervention group (mean 1.16, SD 1.18) than in the wait-list control group (mean 2.16, SD 1.49) after 8 weeks. Those with more recent dental visits had a lower MGI (P=.04). No association was found between knowledge and behavior. Most participants were familiar with the recommendations for maintaining oral health, yet they only performed some of them.

Conclusions: A dental selfie taken once a week using an mobile health app (iGAM) reduced the signs of gingivitis and promoted oral health. Selfies taken less frequently yielded poorer results. During the current pandemic, where social distancing recommendations may be causing people to avoid dental clinics, this app can remotely promote gum health.

(JMIR Mhealth Uhealth 2021;9(9):e24955) doi:10.2196/24955

KEYWORDS

mHealth; public health; oral health promotion; gum health; COVID-19

Introduction

COVID-19 and Dental Implications

One and a half years have passed since the COVID-19 pandemic broke out in Wuhan, China [1]. On March 19, 2019, the Israeli Prime Minister declared a national state of emergency and various restrictions were enforced. “Normal” life resumed with ongoing rules of wearing masks, hand hygiene, and keeping 2 meters between people. Dentistry is a field with close contact between patients and the clinical team, with a high risk of transmission of infections [2]. The aim of these restrictions is to balance the public health issues, which include decreasing or preventing the spread of COVID-19, with the need for minimizing dental and oral pain during this period.

Dental Background

The diseases of the oral cavity can be divided according to the type of affected tissue: soft or hard tissue disease. Bacteria in the dental plaque are the main cause of these diseases [3].

Hard Tissue Disease

Caries or tooth decay [4] affects over 90% of the world’s population. This is caused by bacterial plaque on tooth surfaces that convert the sugars from food into acids which then cause minerals to leech out of the outer layer of the tooth (enamel) and dissolve it. The bacteria then penetrate the deeper layers of the teeth, namely the dentine and eventually the pulp. Tooth decay, also known as caries, is a progressive disease that can be painful and cause systemic infection and thus cannot be ignored.

Soft Tissue Disease

Gingivitis [5] is a nonpainful, reversible inflammatory condition characterized by swelling and redness, along with occasional bleeding when brushing and bad breath. This reversible stage can be treated with home remedies including strict oral hygiene, routine dental visits, and cleanings by a dental hygienist. Untreated gingivitis can progress into periodontitis, where irreversible damage to the tooth-supporting tissues can lead to tooth mobility and tooth loss.

Periodic dental examinations are important for the early diagnosis and treatment of dental problems [6], and the issue of pain figures prominently in the decision to go to the dental clinic [7-9]. Between visits, dentists do not know what changes are occurring in their patients’ mouths and patients usually do not notice changes in the oral cavity until they experience pain. During this period of lockdown and avoidance of routine dental checkups, we speculate that many people are experiencing a decline in their oral health.

eHealth and mobile health

eHealth [10] is the use of information and communication technologies for health. Mobile health (mHealth) is a component of eHealth [11] and is defined by the Global Observatory for eHealth as “medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants, and other wireless devices” [12].

There are currently over 318,000 mHealth apps, with hundreds of new apps added daily. mHealth apps can monitor health conditions and alert the patient or attending physician about deterioration [13,14].

In August 2020, we published a paper describing the development of an mHealth app in the field of dentistry named iGAM [15]. The paper presented here is the second part in the series and presents the results of an observational study conducted to examine the ability of a cellular app to reduce gingivitis. The third part will include the results of a mixed methods study that investigated the acceptance of this type of mHealth app. Using computerized technologies in dentistry can assist in remotely monitoring patients during times of social restrictions (eg, COVID-19). A study conducted in Italy by Giudice et al [16] in 2020 found that teledentistry minimizes contact and therefore decreases COVID-19 dissemination. Another study noted the urgent need to incorporate mHealth apps into routine dental practice as a complementary tool to manage with those patient needs that arise due to social distancing [17].

To the best of our knowledge, there are currently no mHealth apps that monitor gingival health. The aim of this observational study was to examine the ability of an mHealth app, iGAM, to reduce gingivitis without researcher intervention and to determine the best interval between photograph submissions.

Methods

Design and Development

This prospective observational study was conducted between September 2019 and May 2020 at the Department of Community Dentistry Faculty of Dental Medicine, The Hebrew University-Hadassah School of Dental Medicine. The protocol was approved by the Hadassah research ethics committee (IRB 0212-18-HMO), and written informed consent was obtained from all participants. There was no payment for participation. Advertisements were posted in academic institutions and hospitals in Jerusalem for 1 month (August 2019). Potential participants met with the primary investigator (GT), and, after a thorough explanation about the study, they could download the app and fill in the informed consent form. The first participant was randomly assigned by flipping a coin to either the intervention group or wait-list control group, and from that point onward, the app was assigned the groups such that sequential enrollees were assigned to different groups. The intervention group photographed their gums once a week for 8 weeks. The wait-list control group photographed their gums at the time of recruitment for the study and at the end of the study 2 months later. During the first login, the users set a password known only to them and were then asked to provide anonymous personal information. The data were then stored on a Google

encrypted server, Firebase. Each participant was given a randomly generated user number.

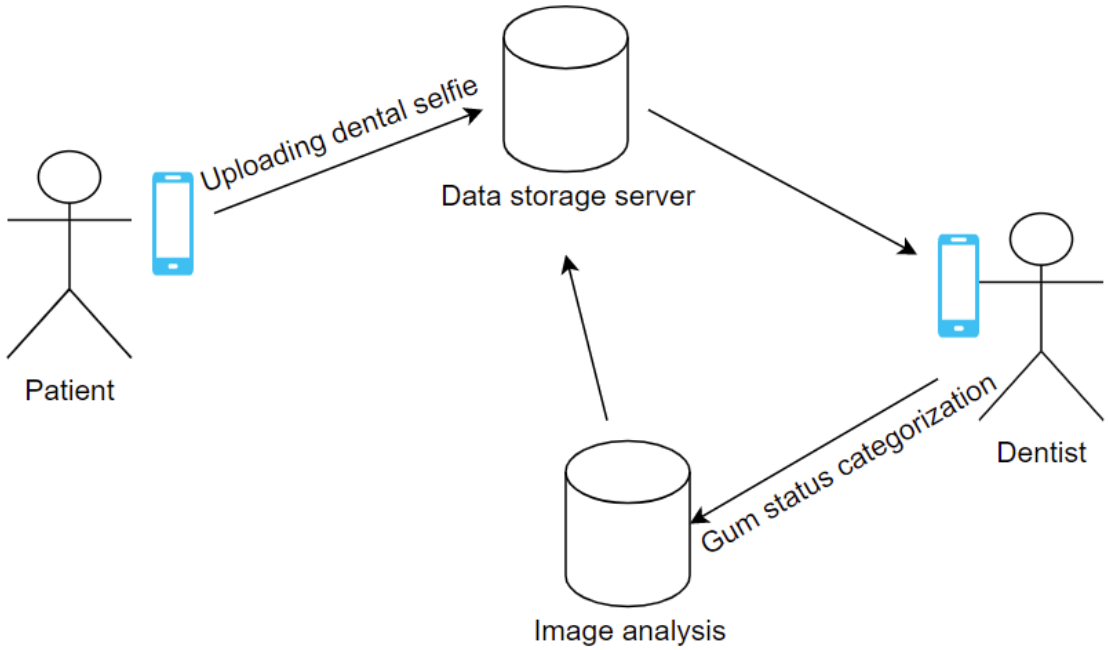
This study used the Modified Gingival Index (MGI) [18] that was introduced in the mid-1980s and was found to be reliable in the visual diagnosis of gingivitis and for assessment of the prevalence and severity of gingivitis without the use of any dental instruments (ie, this index can be used in a home setting). This index focuses on the gingival unit, which includes 3 parts: (1) the free gingiva, the part of the gums around the tooth not attached to the tooth surface; (2) the attached gingiva, the firm part of the gums that is tightly attached to underlying tissues of the tooth; and (3) the alveolar mucosa, the soft tissue between the lips and the gums. The index has a scale from 0 to 4, which describes the relative status of the gums as follows: 0=absence of inflammation; 1=mild inflammation—slight change in color, and small area of altered texture but not the entire marginal or papillary gingival unit; 2=mild inflammation—as above but involving the entire marginal or papillary gingival unit; 3=moderate inflammation—glazing, redness, edema, and/or hypertrophy of the marginal gingival unit; and 4=severe inflammation—marked redness, edema and/or hypertrophy of the marginal or papillary gingival unit, spontaneous bleeding, congestion, or ulceration.

After registration, the user was transferred to the part of the app where the photos of the gums are taken. Upon submission of the photograph, the researcher (AS) received an SMS text message, “User (number) submitted a photograph.” The user received an SMS text message, “The photograph was submitted,” with the submission date and the number of remaining submissions required. Participants received a comprehensive dental kit: toothpaste, toothbrush, mouthwash, dental floss, and toothpicks. In addition, the participants received a reusable mouth opener. The participants were taught how to use the app and take photos of their own gums while using the mouth opener. In addition, participants were instructed to photograph their gums during the day in front of a mirror using white lighting with a white wall in the background. After each photo submission, the participants received the following message “we recommended that you read the information in the app regarding oral hygiene habits.”

All tutorials and instructions were performed by a single researcher (GT) who was blinded to the treatment group and who examined the images and scored them according to the MGI index; the results were not reported to the participants. Figure 1 summarizes the process.

Interactions with the participants were via text messages from within the app. There was only 1 training session, and participants who asked questions were referred to the training documents in the app.

Figure 1. Interaction between participant (patient), app admin (researcher, dentist), and data storage server (Firebase).



Inclusion Criteria

The inclusion criteria for participants included 3 points: (1) over the age of 18 years, (2) could read and understand Hebrew, and (3) in possession of a smartphone with an Android operating system

The iGAM app is currently available on the Google Play Store in Hebrew. The app’s development was described in a recently published article.

iGAM features

The iGAM mHealth app has 3 components: (1) a self-completion questionnaire that deals with knowledge and attitudes toward oral hygiene habits; (2) text accompanied by illustrations regarding brushing techniques as well as short articles about the importance of maintaining oral health in general and during pregnancy, the implications of smoking on gum health, and the connection between gum diseases and general health; and (3) the ability to photograph one’s own gums using the rear camera

of the smartphone. The first volunteer was allocated to the intervention group, the second volunteer to the wait-list control group, and so on.

Statistical Methods

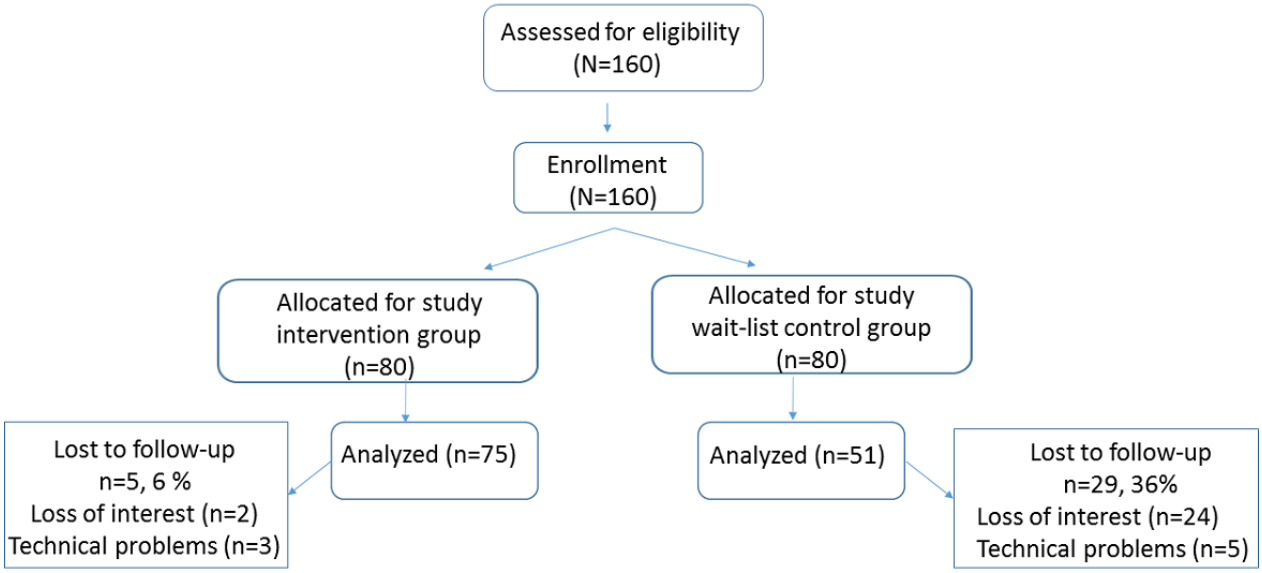
Baseline characteristics are presented as mean and SD for continuous variables and as frequencies and percentages for categorical variables. Specific statistical tests are detailed in the results section. The data were analyzed with SPSS statistics software version 27.0 (IBM Corporation), and significance levels were set at a *P* value of .05.

Results

Demographic and Descriptive Statistics

In total, 175 candidates responded to the recruitment advertisement. The study included 160 participants who met the inclusion criteria and were divided into 2 equal groups. Participants in the intervention group took weekly photographs of their gums for 8 weeks. Those in the wait-list control group were asked to photograph their gums at 2 time points: at the beginning of the study and 8 weeks later. In all, 126 participants successfully completed the tasks: 75 (59.5%) from the intervention group and 51 (40.5%) from the wait-list control group. Figure 2 shows the distribution of the study participants.

Figure 2. Distribution of study participants.



The age range of the intervention group was 19 to 65 years with an average of 26.77 (SD 7.43), while that in wait-list control group was from 20 to 66 years with an average of 28.53 (SD 10.44).

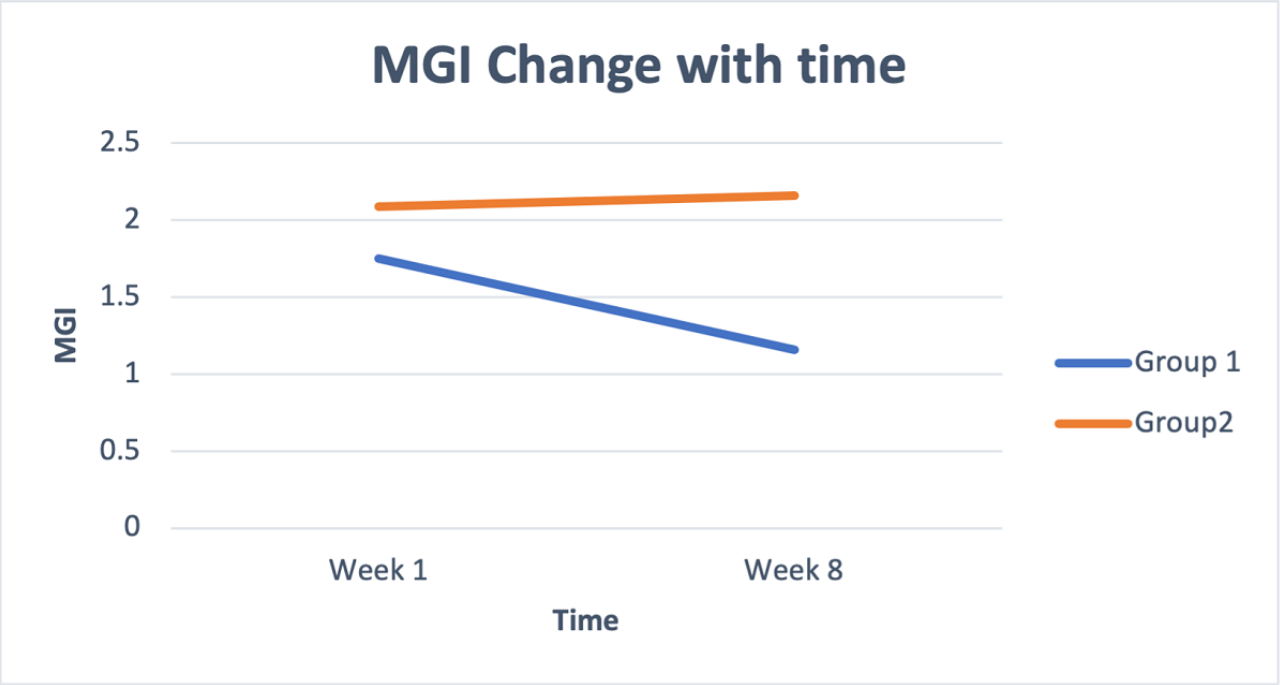
In the comparison of the intervention group and the wait-list control group, the majority were male (intervention group: 56/75, 74.7%; wait-list control group: 34/51, 66.7%), native to Israel (intervention group: 73/75, 97.3%; wait-list control group: 50/51, 98%), and Jewish (intervention group: 75/75, 100%; wait-list control group: 49/51, 96.1%); the common term regarding the level of religiosity in both groups was “secular” (intervention group: 21/75, 28%; wait-list control group: 15/51, 29.4%). In addition, the most common value in terms of marital status was “single” (intervention group: 38/75, 50.7%; wait-list control group: 24/51, 47.1%). The majority of the participants in the wait-list control group were employees (24/51, 47.1%), while in intervention group, the most common term from an occupational point of view was “student” (31/75, 41.3%). Most of the study population did not smoke (intervention group: 56/75, 74.7%; wait-list control group: 40/51, 78.4%) and did not take any medications (intervention group: 64/75, 85.3%; wait-list control group: 40/51, 78.4%; see Table 1).

Table 1. Characteristics of the study population.

Sociodemographic characteristics	Intervention group, n (%) (N=75)	Wait-list control group, n (%) (N=51)
Gender		
Male	56 (75)	34 (67)
Female	19 (25)	17 (33)
Country of Birth		
Israel	73 (97)	50 (98)
Other	2 (3)	1 (2)
Nationality		
Jewish	75 (100)	49 (96)
Other	0/0 (0)	2 (4)
Religiosity		
Secular	21 (28)	15 (29)
Traditional	14 (19)	10 (20)
Orthodox	15 (20)	10 (20)
Ultra-orthodox	17 (23)	7 (14)
Refused to answer	8 (11)	9 (18)
Family Status		
Married	27 (36)	14 (27)
Living with a partner	7 (9)	6 (12)
Divorced	2 (3)	5 (10)
Widow	1 (1)	1 (2)
Single	38 (51)	24 (47)
Refused to answer	0/0 (0)	1 (2)
Occupation		
Salaried employee	28 (37)	24 (47)
Self-employed	13 (17)	7 (14)
Unemployed	2 (3)	0 (0)
Student	31 (41)	19 (37)
Retiree	0 (0)	1 (2)
Soldier	1 (1)	0 (0)
Smoker		
Yes	19 (25)	11 (22)
No	56 (75)	40 (78)
Taking medication		
Yes	11 (15)	11 (22)
No	64 (85)	40 (78)

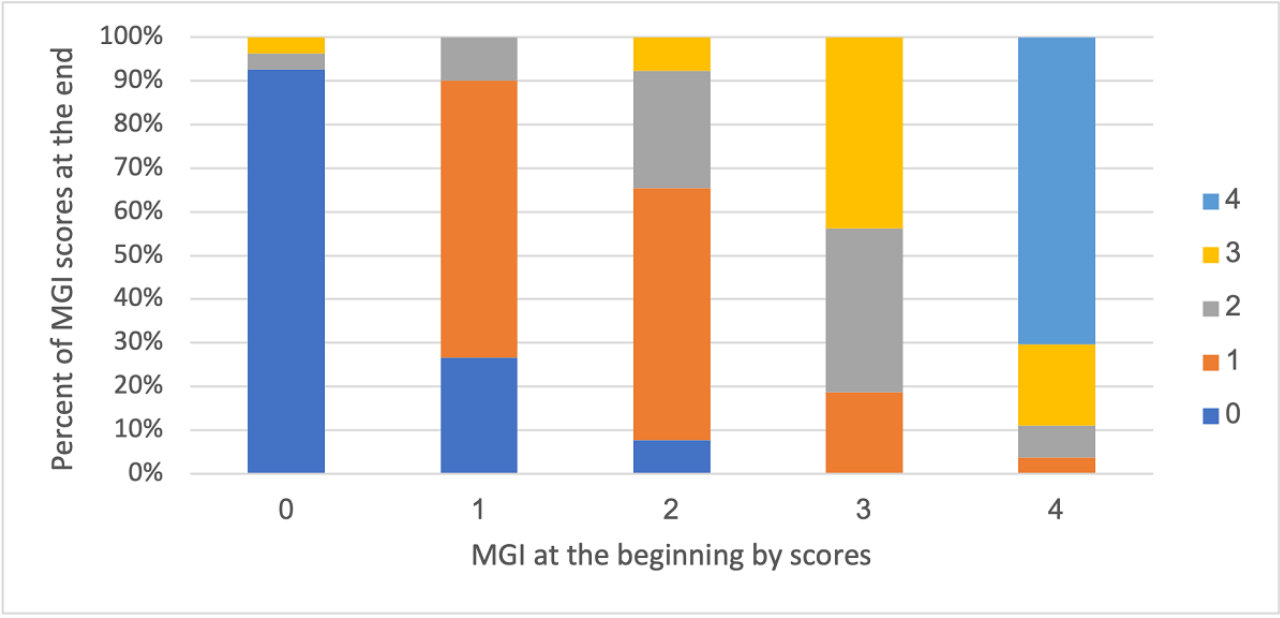
Results of independent *t* tests showed no significant differences between the intervention group (mean 1.75, SD 1.36) and the wait-list control group (mean 2.09, SD 1.55) regarding gingivitis at the start of the study (t_{124}= –1.346; *P*=. 18). In contrast, there were significantly lower gingivitis scores in the intervention group (mean 1.16, SD 1.18) compared to the wait-list control group (mean 2.16, SD 1.49; $t_{91.035}$=–3.998; *P*<.001) after 8 weeks. The intervention group showed improvement, whereas there was a decline in gingival health in the wait-list control group although the changes in wait-list control group were not statistically significant (Figure 3).

Figure 3. Changes in MGI over the study period. MGI: Modified Gingival Index.



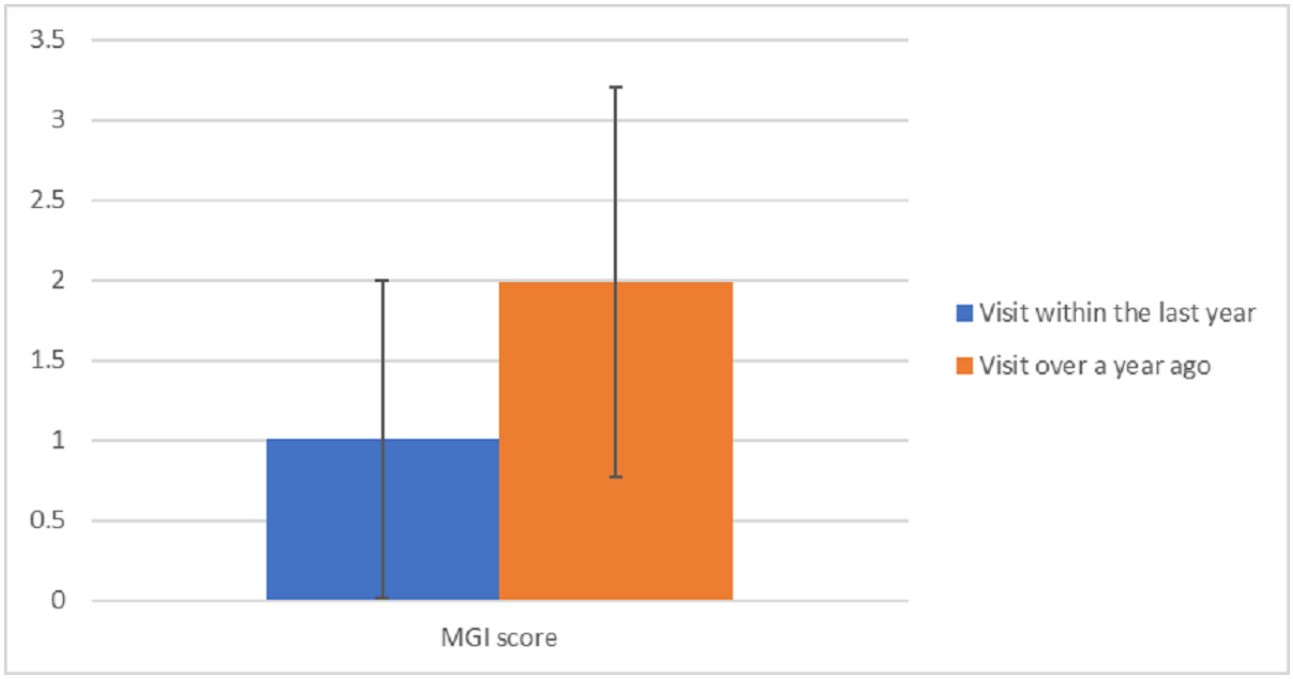
Repeated measures analysis of the data showed a significant effect in the intervention group ($F_{7,68}$=6.672; P<.001). A statistically significant difference was found in the dependent variable (MGI) in a linearly negative manner ($F_{1,74}$=45.054; P<.001); as time passed, the gum condition improved (Figure 3). Of those starting with an MGI=0, 90% (19/21) finished with an MGI=0. Of those starting with an MGI=4, 73% (8/11) finished with an MGI=4 at week 8, and the remaining 27% improved (3/11). Of those starting at an MGI=1, 61% (14/23) finished with an MGI=1, 13% (3/23) with an MGI=2 (ie, worse results), and 26% with an MGI=0 (6/23). Of those starting with an MGI=2, 27% (3/11) remained at an MGI=2 at the end of the study and 64% (7/11) improved (MGI=1: 6/11, 55%; MGI=0: 1/11, 9%). Of those starting with an MGI=3, 44% (4/9) remained at MGI=3, and 56% (5/9) improved (MGI=2: 3/9, 33%; MGI=1: 2/9, 22%; see Figure 4).

Figure 4. Distribution of MGI scores at the beginning (x-axis) and percentages at the end (y-axis). MGI: Modified Gingival Index.



Spearman correlation coefficient showed a statistically significant positive association between MGI score and latest visit to the dentist (Spearman correlation coefficient=186; P=.04). When the latest visit to the dentist was more recent, the chance of lower MGI was higher. An independent samples t test revealed that participants who visited the dentist within the year (mean 0.23, SD 0.76) had significantly less severe gingivitis than did those who had visited more than a year prior (mean 0.46, SD 0.79; t_{124}=–1.669; P=.049; see Figure 5).

Figure 5. MGI score and timing of last dental visit. MGI: Modified Gingival Index.



Most participants brushed their teeth twice a day, with 76% (57/75) in the intervention group and 72.5% (37/51) in the wait-list control group answering correctly regarding recommended brushing frequency (intervention group: 66/75, 86.7%; wait-list control group: 46/51, 90.2%). Most reported they rarely used floss (intervention group: 53/75, 70.7%; wait-list control group 43/51, 66.7%) or mouthwash (intervention group: 44/74, 58.7%; wait-list control group 25/51, 49%). The most common value in terms of visits to the dentist in the wait-list control group was “in the last six months” (22/51, 43.1%), whereas in the intervention group, the most common values were “between six months and a year” (24/75, 32%) and “over two years” (24/75, 32%). Most participants knew that the recommended frequency dental visits are “every six months” (intervention group: 43/75, 57; wait-list control group: 39/51, 76.5%). The most common reason for the visit was “treatment” (intervention group: 37/75, 49.3%; wait-list control group: 15/51, 29.4%). In the intervention group, the most common value in terms of a last visit to the dental hygienist was “between one and two years” (23/75, 30.7%), and in the wait-list control group the most common value was “in the last six months” (20/51, 39.2%). Interestingly, most of the participants knew that the recommended frequency of visits to the hygienist is “every six months” (intervention group: 41/75, 54.7%; wait-list control group: 35/51, 68.6%; see Table 2 for details).

Table 2. Oral health knowledge and habits.

Question	Intervention control group, n (%) (N=75)	Wait-list control group, n (%) (N=51)
How often do you brush your teeth?		
Once a day	17 (22.7)	12 (23.5)
Twice a day	(76) 57	37 (72.5)
More than twice a day	0 (0)	0 (0)
Rarely/irregularly	1 (1.3)	2 (3.9)
As far as you know, what is the recommended frequency for daily brushing?		
1	8 (10.7)	4 (7.8)
2	65 (86.7)	46 (90.2)
More than 2	1 (1.3)	0 (0)
Do not know	1 (1.3)	1 (2)
How often do you floss?		
Once a day	9 (12)	7 (13.7)
Twice a day	3 (4)	0 (0)
Two to three times a week	6 (8)	4 (7.8)
Once a week	1 (1.3)	4 (7.8)
Two to three times a month	0 (0)	1 (2)
Rarely/irregularly	53 (70.7)	34 (66.7)
Never	3 (4)	1 (2)
How often do you use mouthwash?		
Once a day	7 (9.3)	5 (9.8)
Two to three times a week	2 (2.7)	3 (5.9)
Once a week	1 (1.3)	3 (5.9)
Two to three times a month	4 (5.3)	4 (7.8)
Rarely/irregularly	44 (58.7)	25 (49.0)
Never	17 (22.7)	11 (21.6)
When was the last time you visited the dentist?		
In the last six months	19 (25.3)	22 (43.1)
Between six months and a year	24 (32)	12 (23.5)
Between one and two years	7 (9.3)	10 (19.6)
Over two years	24 (32)	6 (11.8)
Never	1 (1.3)	1 (2)
What was the reason for visiting the dental clinic?		
Routine dental examination	11(14.6)	10 (19.6)
Treatment by a hygienist	11(14.6)	10 (19.6)
Treatment such as restoration, tooth extraction, and root canal treatment	37 (49.3)	15 (29.4)
Treatment such as bridge, dentures, orthodontics, and dental implants	12 (16)	13 (25.5)
Gingival treatment	2 (2.7)	3 (5.9)
Treatment after an accident or fall	2 (2.7)	0 (0)
What is the recommended frequency of dental visits?		
Every six months	43 (57.3)	39 (76.5)
Every year	23 (30.7)	10 (19.6)

Question	Intervention control group, n (%) (N=75)	Wait-list control group, n (%) (N=51)
Every two years	4 (5.3)	0 (0)
Do not know	5 (6.7)	2 (3.9)
When was the last time you visited a hygienist?		
In the last six months	17 (22.7)	20 (39.2)
Between six months and a year	17 (22.7)	9 (17.6)
Between one and two years	23 (30.7)	8 (15.7)
Over two years	12 (16)	13 (25.5)
I have never visited	6 (8)	1 (2)
What is the recommended frequency of visiting the hygienist?		
Every six months	41 (54.7)	35 (68.6)
Every year	22 (29.3)	12 (23.5)
Every two years	5 (6.7)	2 (3.9)
Do not know	7 (9.3)	2 (3.9)

Discussion

This novel observational study used an mHealth app (iGAM) to examine whether dental selfies reduce gingivitis as determined by the MGI score and investigated the interval between photograph submissions needed to enable an improvement.

COVID-19 has spread rapidly across the world since December 2019, and, in order to reduce infection, social distancing has been recommended [19,20]. Since people rarely go to the dentist if they are asymptomatic [21], we assumed that gingivitis will be neglected and worsen, highlighting a need for remote monitoring in dentistry. mHealth apps are technological tools that can remotely maintain patient care. Indeed, during the COVID-19 period, the number of mHealth apps available and the number of users have increased [22-24].

This observational study demonstrated that a dental mHealth app iGAM can reduce gingivitis. At the beginning of the study, all participants filled in a questionnaire regarding oral hygiene habits and knowledge of oral health maintenance recommendations. Most of the participants were in their third decade of life, men, secular, single, employees and students, healthy, and nonsmokers. These characteristics are consistent with other studies on the use of mobile medical apps [25-28].

We found that the participants with more recent dental visits had significantly better gum health and found that weekly use of the app not only raised awareness of gum health but actually led to its improvement. Studies have shown that the more medical awareness a person has, the greater tendency to use mHealth apps [29-31].

Those taking weekly pictures had reduced gingivitis scores, whereas the other group showed no improvements, and their gum condition slightly deteriorated (the decline in the second group was not statistically significant). Similarly, other observational studies using mHealth apps in which the researcher does not provide any feedback have shown that the apps improve health (eg, reducing asthma attacks) [32-34].

We found a significantly positive correlation between the gum condition at the beginning of the study and at the end of the study in participants with excellent (MGI=0) or terrible (MGI=4) gums. For the remaining participants with scores not at the limits of the index, gingival health improved (ie, MGI scores were lower). There are several possible reasons for these interesting and unique findings.

First, regarding the limits of the MGI, we assume that for an individual with optimal oral health (MGI=0) who knows how to maintain their mouth, no improvement is possible, and therefore they do not need to use the app. Conversely, an individual with the worst score (MGI=4) is probably aware of their poor condition because they have bleeding, swollen gums, and frequent bad breath, and have done nothing about it, so an app is unlikely to change their behavior.

Second, participants with MGI scores between 1 and 3 demonstrated a statistically significant improvement in gum health. Studies have found that most of the population has scores in this range [35-37], with the occasional appearance of signs of inflammation. Therefore, the app assisted in promoting and maintaining gum health over time.

In the examination of knowledge and oral hygiene habits, no association was found between knowledge and behavior. Furthermore, although most participants were familiar with the accepted guidelines for maintaining oral health, they only performed some of them. Most of the participants said they brush twice daily, but rarely use brushing aids. Furthermore, the main reason for visiting the dentist was pain, even though most participants knew that the recommendation is to go to the dentist twice a year for an examination.

More participants in the wait-list control group failed to complete the requirements of the study, even though they had less to do, as they took 2 photographs compared to 8. It seems that having to perform a task every 8 weeks was more challenging than doing so each week. It cannot be argued that participants forgot the date of the photo, as the app sent reminders to each participant on the day they needed to take the

photograph. Studies have found that apps with regular activities are more likely to be used [38].

Our study was limited in 7 main aspects. (1) A causal relationship cannot be established from an observational study; however, many investigations on mHealth apps use this methodology, and predict clinical values from their outcomes. (2) We only enrolled participants fluent in Hebrew, which might have increased the risk for selection bias. The app should be translated into the other official languages of Israel (ie, English and Arabic) in order to gather data and make more generalizable findings. (3) Only Android users were enrolled, which might have increased the risk for socioeconomic bias. For example, a survey [39] among 139,000 US smartphone users showed that Android users have less education than do iPhone users. (4) The advertisement for the study clearly invited participants aged 18 years and above, yet the mean age of the study participant was ~27; this possible selection bias might have been due to the advertisements being posted in academic institutions and university hospitals. A larger-scale study with participants from different age groups would yield more generalizable findings. (5) The 8-week time period for self-monitoring was selected for 2 reasons: gingivitis heals within 10 to 14 days with proper oral hygiene, so improvement is more immediately evident; and most studies on gingivitis are for this period, and thus a longer study might have showed more significant differences between the groups. (6) The questionnaire used was developed exclusively for this study, which limits comparison with questionnaires with established validity. (7) The high dropout rates, especially in the wait-list control group, should not be taken for granted. The impact of the high dropout rate, particularly of the wait-list control group, on the results needs further investigation using qualitative research methods to understand the acceptance and usability of the iGAM mHealth app among the dropouts and those completing the study. We will publish this data in the third paper in this series.

This observational study found that a dental selfie taken once a week using an mHealth app (iGAM) reduces the signs of gingivitis and promotes oral health. In this period of the COVID-19 pandemic where social distancing recommendations may be causing people to avoid dental clinics, this app can remotely promote gum health.

Conflicts of Interest

Authors TG and ABS were involved in the development of iGAM app.

References

1. Archived: WHO timeline - COVID-19. World Health Organization. URL: https://www.who.int/news-room/detail/27-04-2020-who-timeline---covid-19 [accessed 2020-09-01]
2. Harrel SK, Molinari J. Aerosols and splatter in dentistry: a brief review of the literature and infection control implications. J Am Dent Assoc 2004 Apr;135(4):429-437 [FREE Full text] [doi: 10.14219/jada.archive.2004.0207] [Medline: 15127864]
3. Bratthall D, Petersen P, Stjernswärd J. Oral and craniofacial diseases and disorders. In: Jamison DT, Breman JG, Measham AR, editors. Disease Control Priorities in Developing Countries. 2nd edition. Washington (DC): The International Bank for Reconstruction and Development / The World Bank; 2006.
4. Selwitz RH, Ismail AI, Pitts NB. Dental caries. Lancet 2007 Jan 06;369(9555):51-59. [doi: 10.1016/S0140-6736(07)60031-2] [Medline: 17208642]
5. Page RC. Gingivitis. J Clin Periodontol 1986 May;13(5):345-359. [doi: 10.1111/j.1600-051x.1986.tb01471.x] [Medline: 3522644]
6. Hom JM, Lee JY, Silverman J, Casamassimo PS. State Medicaid early and periodic screening, diagnosis, and treatment guidelines: adherence to professionally recommended best oral health practices. J Am Dent Assoc 2013 Mar;144(3):297-305 [FREE Full text] [doi: 10.14219/jada.archive.2013.0117] [Medline: 23449906]
7. Karibe H, Koeda M, Aoyagi-Naka K, Kato Y, Tateno A, Suzuki H, et al. Differences in the perception of dental sounds: a preliminary study. Patient Prefer Adherence 2019;13:1051-1056 [FREE Full text] [doi: 10.2147/PPA.S204465] [Medline: 31308637]
8. Pohjola V, Rekola A, Kunttu K, Virtanen JI. Association between dental fear and oral health habits and treatment need among University students in Finland: a national study. BMC Oral Health 2016 Mar 27;16:26 [FREE Full text] [doi: 10.1186/s12903-016-0179-y] [Medline: 26922793]
9. Armfield JM, Slade GD, Spencer AJ. Dental fear and adult oral health in Australia. Community Dent Oral Epidemiol 2009 Jun;37(3):220-230. [doi: 10.1111/j.1600-0528.2009.00468.x] [Medline: 19508270]
10. Eysenbach G. What is e-health? J Med Internet Res 2001 Jun;3(2):E20. [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
11. Silva BMC, Rodrigues JJPC, de LTDI, López-Coronado M, Saleem K. Mobile-health: A review of current state in 2015. J Biomed Inform 2015 Aug;56:265-272. [doi: 10.1016/j.jbi.2015.06.003] [Medline: 26071682]
12. mHealth. New horizons through mobile technologies. World Health Organization. URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2020-09-01]
13. Number of mHealth apps available in the Apple App Store from 1st quarter 2015 to 1st quarter 2021. Statista. URL: https://www.statista.com/statistics/779910/health-apps-available-ios-worldwide/ [accessed 2020-08-15]
14. Number of mHealth apps available in the Google Play Store from 1st quarter 2015 to 1st quarter 2021. Statista. URL: https://www.statista.com/statistics/779919/health-apps-available-google-play-worldwide/ [accessed 2020-08-15]

15. Tobias G, Spanier AB. Developing a mobile app (iGAM) to promote gingival health by professional monitoring of dental selfies: user-centered design approach. JMIR Mhealth Uhealth 2020 Aug 14;8(8):e19433 [FREE Full text] [doi: 10.2196/19433] [Medline: 32795985]
16. Giudice A, Barone S, Muraca D, Averta F, Diodati F, Antonelli A, et al. Can teledentistry improve the monitoring of patients during the Covid-19 dissemination? A descriptive pilot study. Int J Environ Res Public Health 2020 May 13;17(10):1-9 [FREE Full text] [doi: 10.3390/ijerph17103399] [Medline: 32414126]
17. Ghai S. Teledentistry during COVID-19 pandemic. Diabetes Metab Syndr 2020;14(5):933-935 [FREE Full text] [doi: 10.1016/j.dsx.2020.06.029] [Medline: 32593116]
18. Lobene RR, Weatherford T, Ross NM, Lamm RA, Menaker L. A modified gingival index for use in clinical trials. Clin Prev Dent 1986;8(1):3-6. [Medline: 3485495]
19. Lewnard JA, Lo NC. Scientific and ethical basis for social-distancing interventions against COVID-19. Lancet Infect Dis 2020 Jun;20(6):631-633 [FREE Full text] [doi: 10.1016/S1473-3099(20)30190-0] [Medline: 32213329]
20. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, COVID-19 Systematic Urgent Review Group Effort (SURGE) study authors. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. Lancet 2020 Jun 01:1973-1987 [FREE Full text] [doi: 10.1016/S0140-6736(20)31142-9] [Medline: 32497510]
21. Shahrabani S. Factors affecting oral examinations and dental treatments among older adults in Israel. Isr J Health Policy Res 2019 May 10;8(1):43-52 [FREE Full text] [doi: 10.1186/s13584-019-0312-x] [Medline: 31077257]
22. Ohannessian R, Duong TA, Odone A. Global telemedicine implementation and integration within health systems to fight the COVID-19 pandemic: a call to action. JMIR Public Health Surveill 2020 Apr 02;6(2):e18810-e18810 [FREE Full text] [doi: 10.2196/18810] [Medline: 32238336]
23. Bokolo Anthony Jnr. Use of telemedicine and virtual care for remote treatment in response to the COVID-19 pandemic. J Med Syst 2020 Jun 15;44(7):132-141 [FREE Full text] [doi: 10.1007/s10916-020-01596-5] [Medline: 32542571]
24. Wosik J, Fudim M, Cameron B, Gellad ZF, Cho A, Phinney D, et al. Telehealth transformation: COVID-19 and the rise of virtual care. J Am Med Inform Assoc 2020 Apr 20:957-962 [FREE Full text] [doi: 10.1093/jamia/ocaa067] [Medline: 32311034]
25. Hung M, Jen W. The adoption of mobile health management services: an empirical study. J Med Syst 2012 Jun;36(3):1381-1388. [doi: 10.1007/s10916-010-9600-2] [Medline: 20878452]
26. Zhang X, Guo X, Lai K, Guo F, Li C. Understanding gender differences in m-health adoption: a modified theory of reasoned action model. Telemed J E Health 2014 Jan;20(1):39-46. [doi: 10.1089/tmj.2013.0092] [Medline: 24161004]
27. Bhuyan SS, Lu N, Chandak A, Kim H, Wyant D, Bhatt J, et al. Use of mobile health applications for health-seeking behavior among US adults. J Med Syst 2016 Jun;40(6):153. [doi: 10.1007/s10916-016-0492-7] [Medline: 27147516]
28. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov 04;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
29. Rai A, Chen L, Pye J, Baird A. Understanding determinants of consumer mobile health usage intentions, assimilation, and channel preferences. J Med Internet Res 2013 Aug 02;15(8):e149 [FREE Full text] [doi: 10.2196/jmir.2635] [Medline: 23912839]
30. Symul L, Wac K, Hillard P, Salathé M. Assessment of menstrual health status and evolution through mobile apps for fertility awareness. NPJ Digit Med 2019;2:64 [FREE Full text] [doi: 10.1038/s41746-019-0139-4] [Medline: 31341953]
31. Australian Research Centre for Population Oral Health, The University of Adelaide, South Australia. Periodontal diseases in the Australian adult population. Aust Dent J 2009 Dec;54(4):390-393 [FREE Full text] [doi: 10.1111/j.1834-7819.2009.01167.x] [Medline: 20415940]
32. Jake-Schoffman DE, Silfee VJ, Waring ME, Boudreaux ED, Sadasivam RS, Mullen SP, et al. Methods for evaluating the content, usability, and efficacy of commercial mobile health apps. JMIR Mhealth Uhealth 2017 Dec 18;5(12):e190 [FREE Full text] [doi: 10.2196/mhealth.8758] [Medline: 29254914]
33. Harada N, Harada S, Ito J, Atsuta R, Hori S, Takahashi K. Mobile health app for Japanese adult patients with asthma: clinical observational study. J Med Internet Res 2020 Aug 14;22(8):e19006 [FREE Full text] [doi: 10.2196/19006] [Medline: 32795993]
34. Chan YY, Wang P, Rogers L, Tignor N, Zweig M, Hershman SG, et al. The Asthma Mobile Health Study, a large-scale clinical observational study using ResearchKit. Nat Biotechnol 2017 Mar 13:354-362 [FREE Full text] [doi: 10.1038/nbt.3826] [Medline: 28288104]
35. Carvajal P, Gómez M, Gomes S, Costa R, Toledo A, Solanes F, et al. Prevalence, severity, and risk indicators of gingival inflammation in a multi-center study on South American adults: a cross sectional study. J Appl Oral Sci 2016 Oct;24(5):524-534 [FREE Full text] [doi: 10.1590/1678-775720160178] [Medline: 27812624]
36. Elías-Boneta AR, Encarnación A, Rivas-Tumanyan S, Berríos-Ouslán BC, García-Godoy B, Murillo M, et al. Prevalence of gingivitis in a group of 35- to 70-year-olds residing in Puerto Rico. P R Health Sci J 2017 Sep;36(3):140-145. [Medline: 28915302]

37. Funieru C, Klinger A, Băicu C, Funieru E, Dumitriu HT, Dumitriu A. Epidemiology of gingivitis in schoolchildren in Bucharest, Romania: a cross-sectional study. J Periodontal Res 2017 Apr;52(2):225-232. [doi: 10.1111/jre.12385] [Medline: 27093564]
38. Robbins R, Krebs P, Jagannathan R, Jean-Louis G, Duncan DT. Health app use among us mobile phone users: analysis of trends by chronic disease status. JMIR Mhealth Uhealth 2017 Dec 19;5(12):e197 [FREE Full text] [doi: 10.2196/mhealth.7832] [Medline: 29258981]
39. Android users poorer, shorter, unhealthier, less educated, far less charitable than Apple iPhone users. MacDailyNews. URL: https://macdailynews.com/2013/11/13/android-users-poorer-shorter-unhealthier-less-educated-far-less-charitable-than-apple-iphone-users/ [accessed 2020-09-01]

Abbreviations

mHealth: mobile health
MGI: Modified Gingival Index

Edited by G Eysenbach, L Buis; submitted 12.10.20; peer-reviewed by G Capocasale, J Zhang; comments to author 03.11.20; revised version received 28.11.20; accepted 02.08.21; published 16.09.21.

Please cite as:
Tobias G, Spanier AB
Using an mHealth App (iGAM) to Reduce Gingivitis Remotely (Part 2): Prospective Observational Study
JMIR Mhealth Uhealth 2021;9(9):e24955
URL: https://mhealth.jmir.org/2021/9/e24955
doi:10.2196/24955
PMID:34528897

Original Paper

Barriers and Facilitators Associated With App-Based Treatment for Female Urinary Incontinence: Mixed Methods Evaluation

Nienke J Wessels[1*], MD; Anne M M Loohuis[1*], MD; Henk van der Worp[1], PhD; Linde Abbenhuis[1], MD; Janny Dekker[1], PhD; Marjolein Y Berger[1], PhD; Julia E W C van Gemert-Pijnen[2], PhD; Marco H Blanker[1], MD, PhD

[1]Department of General Practice and Elderly Care Medicine, University Medical Centre Groningen, University of Groningen, Groningen, Netherlands
[2]Center for eHealth Research and Disease Management, Faculty of Behavioural Sciences, University of Twente, Enschede, Netherlands
[*]these authors contributed equally

Corresponding Author:
Nienke J Wessels, MD
Department of General Practice and Elderly Care Medicine
University Medical Centre Groningen
University of Groningen
FA21, PO Box 196
Groningen,
Netherlands
Phone: 31 05 03616746
Email: n.j.wessels@umcg.nl

Abstract

Background: App-based treatment for urinary incontinence is a proven effective and cost-effective alternative to care as usual, but successful implementation requires that we identify and address the barriers and facilitators associated with app use.

Objective: The goal of the research was to explore the factors influencing app-based treatment for urinary incontinence and identify which barriers or facilitators are associated with treatment success or failure.

Methods: We used a sequential explanatory mixed methods design to connect the results of a randomized controlled trial with data from semistructured interviews. This previous RCT had shown the noninferiority of app-based treatment compared with care as usual for urinary incontinence over 4 months. Participants who reported success or failure with app-based treatment, as measured by the change in International Consultation on Incontinence Modular Questionnaire Urinary Incontinence Short Form symptom score, were selected for telephone interview by purposive sampling (n=17). This study reports mainly on the qualitative component of our mixed methods study. Qualitative analyses were conducted in two ways. First, we analyzed the qualitative data of all interviewed participants and discussed the relationships between the main themes. Second, the experiences between the success (n=9) and failure group (n=8) were compared and contrasted to explore factors that were positively or negatively associated with the quantitative effect of app-based treatment. These factors were then interpreted as barriers to and facilitators of successful app-based treatment.

Results: Four interrelated themes were identified as affecting the app based treatment effect: adherence, personal factors, app factors, and awareness. Qualitative analyses of the relationships between the themes showed that adherence-related factors directly influenced treatment effect in both a positive and negative matter. In turn, adherence was also positively and negatively influenced by the other 3 themes. Additionally, awareness was positively influenced by the treatment effect. Within these themes, several factors were identified that acted as barriers (eg, unrealistic expectation of time investment and interfering personal circumstances), facilitators (eg, strict integration of exercises and prior pelvic floor muscle therapy), or both (eg, personality traits and increased awareness of symptoms).

Conclusions: This study shows that the effect of app-based treatment for urinary incontinence is mainly influenced by adherence, which in turn is affected by personal factors, app-based factors, and awareness. The identified factors could function as both facilitators and barriers depending on the user and interaction with other themes. Insight into these facilitators and barriers could lead to improved implementation and increased treatment effectiveness by targeting women most likely to benefit and through further development of the app.

International Registered Report Identifier (IRRID): RR2-10.1002/nau.23507

(JMIR Mhealth Uhealth 2021;9(9):e25878) doi:10.2196/25878

KEYWORDS

mHealth; female; mixed methods; primary health care; urinary incontinence

Introduction

The use of mobile health (mHealth) for urinary incontinence can be an effective and cost-effective alternative to care as usual [1-3]. Although implementation must now proceed for us to realize these benefits for patients and caregivers, successful uptake of the app requires that we identify and address the barriers and facilitators associated with this treatment modality [4]. The complexity of mHealth interventions with multiple interacting components calls for a thorough evaluation of the connection between patient experience, adherence, and effectiveness but is often lacking [5,6]. Previous qualitative studies of women suffering from urinary incontinence have identified factors that could affect app- or internet-based treatment for urinary incontinence by exploring their expectations and experiences [7-10]. Women expected that internet-based treatment would be more accessible, more flexible, and improve treatment adherence, but they expressed concern about the lack of contact with a caregiver [7,8]. Only two studies focused specifically on user experiences with mHealth for urinary incontinence over periods of 6 weeks to 3 months [9,10]. Women commented on several positive and negative effects: support via reminders, insecurity of the treatment result, and increased awareness of symptoms [9,10]. However, these experiences with internet- or app-based treatment for urinary incontinence were never assessed in relation to quantitative treatment success or failure. It is important to explore if a relation exists between the factors identified in qualitative research and the actual success or failure of the intervention [11]. This could reveal strategies for tailoring the app, increasing its effects, or targeting women most likely to benefit.

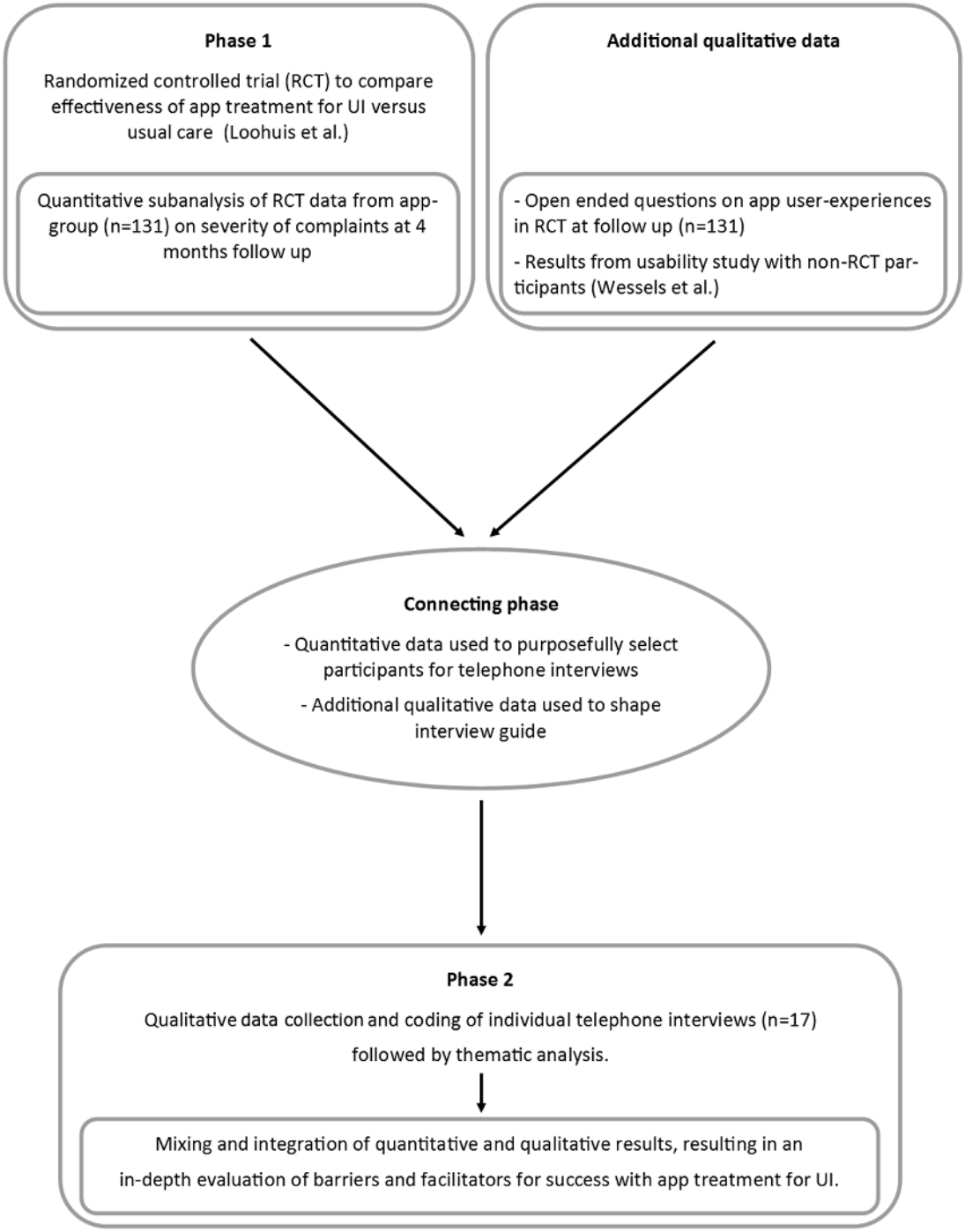
In this study, we aimed to explore the factors influencing app-based treatment for urinary incontinence and identify which barriers or facilitators are associated with treatment success or failure.

Methods

Study Design

We conducted a mixed methods study with a sequential explanatory design that built on a previously reported quantitative randomized controlled trial (RCT) by integrating the results of a qualitative analysis of interviews. The qualitative phase reported in this manuscript follows from a quantitative phase that was reported elsewhere [3,12] (Multimedia Appendix 1) and links both phases in a connecting phase (Figure 1). The original RCT showed the noninferiority of app-based treatment for urinary incontinence (containing a step-by-step self-management program based on Dutch general practitioner and international guidelines [13,14]) compared with care as usual after 4 months [3], and for this study we used the quantitative outcomes of the URinControl RCT to select participants for telephone interview by purposive sampling. The qualitative results from the interviews were expected to refine and explain the quantitative results by exploring participants' views in more depth [11,15,16]. The Research Ethics Committee (no. M17.207954) and Medical Ethical Review Board (METc-no.: 2014/574) of the University Medical Center Groningen, Netherlands, approved the study. All participants gave written informed consent.

Figure 1. Description of sequential explanatory mixed methods study to explore barriers and facilitators for success with app treatment for urinary incontinence.



Participants

We used purposive sampling to select women for interview from the app-based treatment group according to the change in symptom severity measured by the International Consultation on Incontinence Modular Questionnaire Urinary Incontinence Short Form (ICIQ-UI-SF) at 4 months [17]. We ranked the change in ICIQ-UI-SF score from the largest increase to the largest decrease in symptoms and invited participants by working inward from these extremes. In this way, we created two groups: a treatment success group and a treatment failure group. We approached women who had completed the 12-month follow-up requirement to avoid influencing the ongoing RCT. Women were invited by telephone, after which an appointment was made for a telephone interview.

Data Collection

The semistructured interview guide contained several broad themes to ensure that all relevant topics were covered in the interviews. These were selected based on a literature review and the results of a study of the experiences of URinControl app users not included in the RCT [10]. We also reviewed the answers to open-ended questions regarding the experiences of all users in the app group to help further shape the interview guide (see connecting stage in Figure 1).

A female medical master's student (LA) who had no prior relationship with the participants conducted telephone interviews in April and May 2019. She is experienced in performing in-person interviews and prepared for the current task by conducting an extensive literature review on this subject. Additionally, we held a pilot interview and regular peer debriefings to evaluate the quality of the interviews. The interviewer encouraged participants to elaborate on their experiences and asked them to raise any subject they felt relevant that had not yet been covered. Interviews where audio recorded and transcribed verbatim.

Data Analysis

Qualitative analysis was driven by an inductive approach, allowing new patterns and categories to emerge from the raw data. Interview transcripts were coded separately by two researchers (NW, LA) using Atlas.ti (version 8.4, Atlas.ti Scientific Software Development GmbH), and the codes and emerging categories were compared and checked for consensus. Additionally, we regularly discussed broader themes emerging from the categories within the research group and compared the raw data to ensure that the themes covered all aspects. Interviews were conducted until saturation (no new categories emerged in 3 consecutive interviews). Analysis then proceeded in 2 stages. First, we focused on the coded data of all interviewed participants and discussed the relationships between the main themes. Second, we integrated the quantitative and qualitative data by comparing and contrasting the experiences between the success and failure groups, describing the between-group differences in subthemes and the relations between main themes. Additionally, between-group differences in subthemes were checked by frequency counts. Multimedia Appendix 2 and 3 provide a more detailed description of the qualitative analysis and the coding tree.

The descriptive analysis of participant characteristics was conducted with SPSS (version 26, IBM Corp). Reporting was in accordance with the Consolidated Criteria for Reporting Qualitative Research (COREQ) [18].

Results

Participant Selection and Characteristics

The change in urinary incontinence severity measured with the ICIQ-UI-SF in the 102 women with complete follow-up at 4 months ranged from –8 to +3 points (mean 2.2, SD 2.56; Figure 2). As mentioned, women were invited by working inward from the largest increase and largest decrease in the change in ICIQ-UI-SF score. Three women had not completed 12 months of follow-up and were not invited, and 5 women declined the invitation to participate. Data saturation was reached after the 17th interview. The interviewed women were aged 35 to 78 years and had suffered from urinary incontinence for between 3 months and 20 years (Table 1).

Figure 2. Overview of the interview participants (n=17) with respect to the total randomized controlled trial app group (n=102). International Consultation on Incontinence Modular Questionnaire Urinary Incontinence Short Form change scores: negative scores indicate symptom improvement (success) and positive scores indicate symptoms increasing (failure).

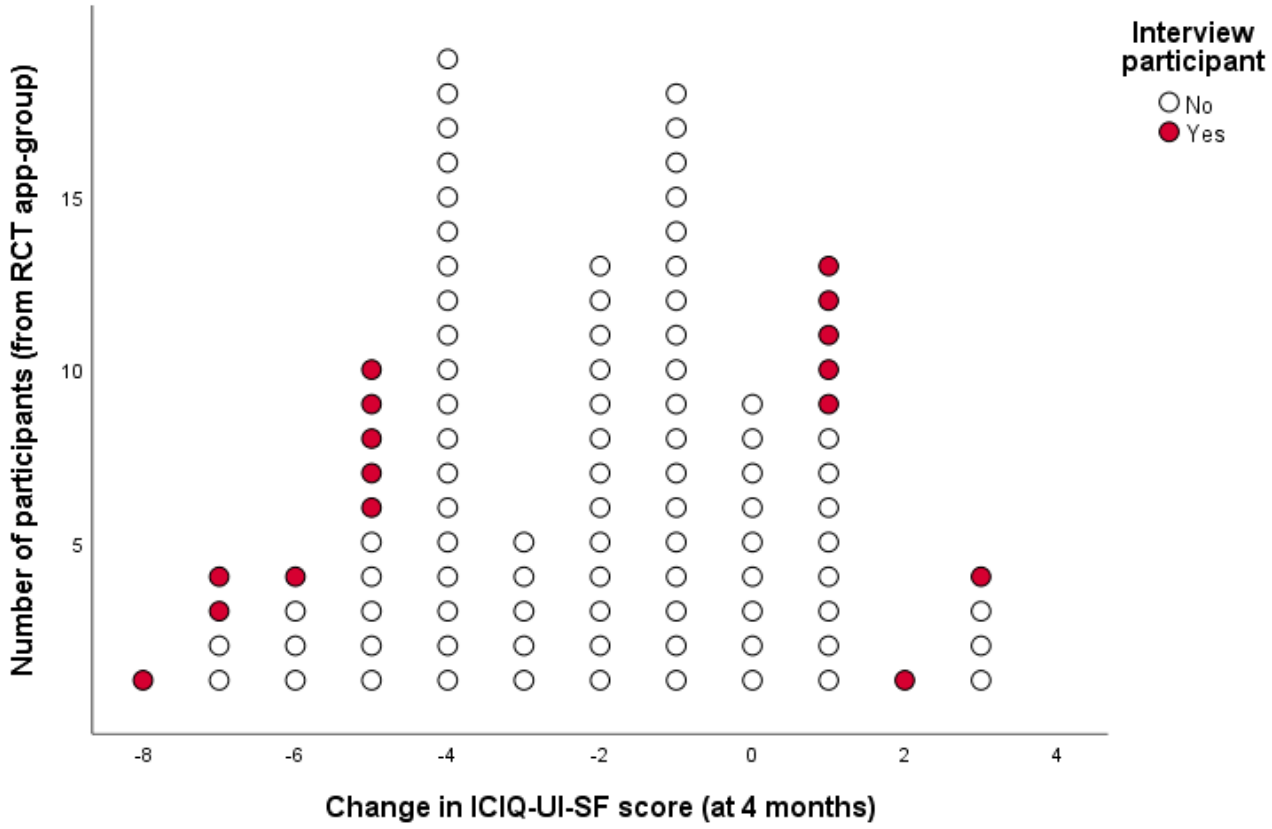


Table 1. Characteristics of the interview participants[a].

Participants			UI[b] outcomes		UI at baseline				Relevant experience	
#	Age (years)	Level of education[c]	Severity[d] change	PGI-I[e]	Severity score	Impact[f] score	Type	Duration (years)	Previous PFMT[g]	Smartphone/tablet[h] (years)
Treatment success										
1	65	Higher	–6	6	10	33	Stress	4	No	8
2	67	Higher	–8	6	16	59	Urge	5	No	2
3	54	Lower	–7	7	12	42	Stress	20	Yes	2
4	61	Lower	–5	6	10	37	Stress	20	Yes	5
5	46	Higher	–5	6	7	32	Stress	2	No	8
6	48	Higher	–7	5	13	36	Stress	6	No	15
7	71	Higher	–5	5	14	51	Urge	15	No	8
8	78	Lower	–5	6	10	37	Stress	20	Yes	1
9	44	Lower	–5	5	11	32	Urge	15	No	—[i]
Treatment failure										
10	54	Lower	2	6	6	32	Stress	5	No	5
11	65	Lower	3	4	5	23	Stress	10	No	6
12	48	Lower	1	6	12	51	Stress	12	Yes	10
13	48	Higher	1	5	4	25	Urge	16	No	10
14	43	Higher	1	4	5	27	Urge	0.25	No	3
15	63	Higher	1	4	4	32	Urge	3	No	7
16	42	Higher	1	6	7	26	Stress	0.42	No	6
17	35	Lower	1	5	9	27	Urge	20	Yes	—

[a]Women using app-based treatment purposefully sampled based on change of urinary incontinence severity (ICIQ-UI-SF score) after 4 months. All measures were self-reported and recorded at baseline except for the ICIQ-UI-SF change score and the PGI-I, which were recorded at 4-month follow-up.

[b]UI: urinary incontinence.

[c]Lower: primary or secondary education; higher: tertiary education or higher.

[d]Severity was based on International Consultation on Incontinence Modular Questionnaire Urinary Incontinence Short Form; range 0-21, higher score means worse incontinence.

[e]PGI-I: Patient Global Impression of Improvement, Likert scale ranging from 0 (very much worse) to 7 (very much better), with 4 reflecting no change.

[f]Impact based on International Consultation on Incontinence Modular Questionnaire Lower Urinary Tract Symptoms Quality of Life; range 19-67, higher score reflects larger impact of urinary incontinence on quality of life.

[g]PFMT: pelvic floor muscle therapy.

[h]Years in possession of device.

[i]Not applicable.

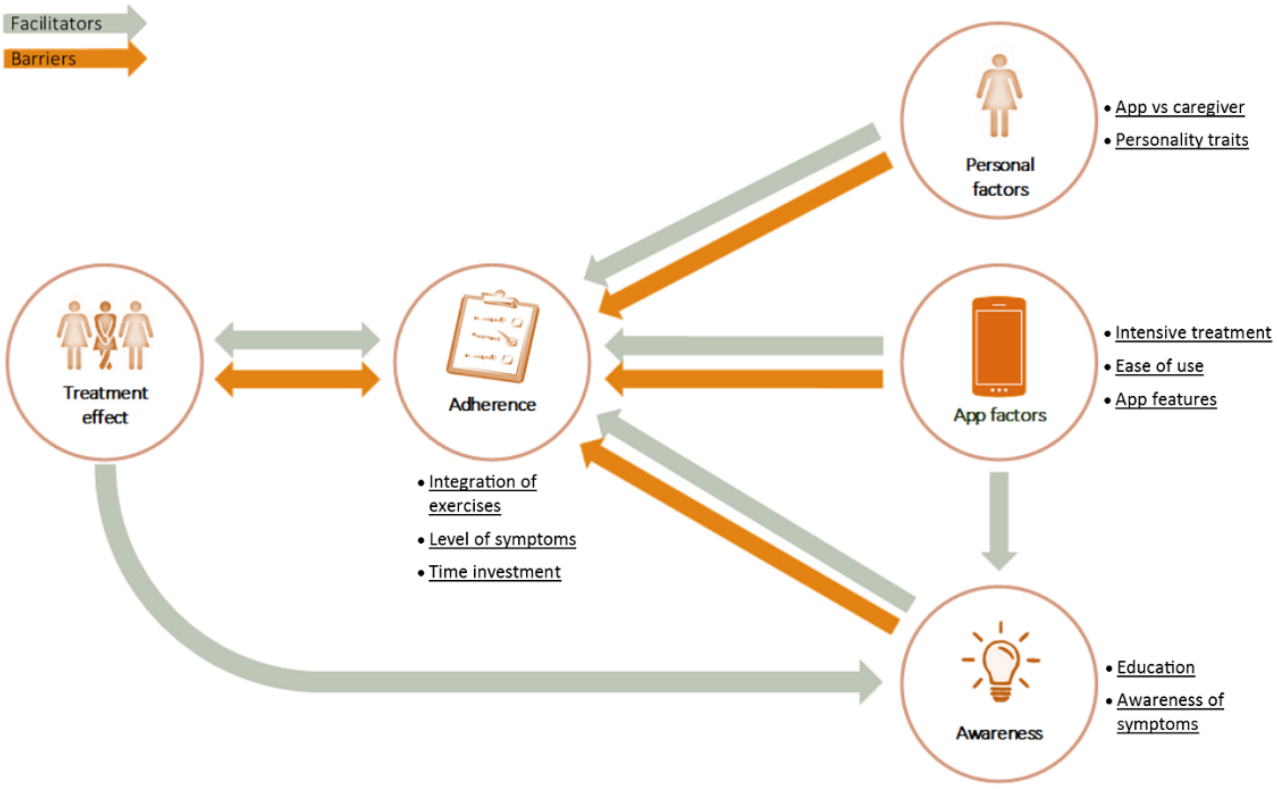
Overall, 9 women experienced treatment success and 8 experienced treatment failure (Figure 2). The change in ICIQ-UI-SF score for women from the success group ranged from –8 to –5 points (median –5.5); for the failure group, it ranged from 1 to 3 (median 1.5). Patients from the success group seemed to have worse urinary incontinence–specific measures and higher age at baseline. Educational level and relevant experiences seemed comparable between groups. None of the patients from the failure group experienced a worsening of symptoms (Patient Global Impression of Improvement [PGI-I] <4).

Semistructured Interviews

Main Themes

We identified adherence, personal factors, app factors, and awareness as the main themes related to overall treatment effect. Discussion of the relationships between the themes resulted in a cross-thematic network (Figure 3). Factors in the adherence theme directly influenced app-based treatment effects as a barrier and facilitator. Adherence was further influenced by factors in the personal factors, app factors, and awareness themes (barriers and facilitators). Finally, awareness was facilitated by the treatment effect and by app factors.

Figure 3. Cross-thematic network of interrelated themes resulting from the qualitative analysis of telephone interviews (n=17). Subthemes show the barriers or facilitators for successful app treatment.



There were no differences between the success and failure groups in the main themes or relationship directions, but there were differences between those groups in the subthemes and in the strength of the relationships between the main themes. The frequency counts for quotes showed between-group differences in subthemes with clear patterns that matched those found in the interviews (Table 2).

Table 2. Themes and subthemes by treatment success and failure[a].

Theme and subtheme	Groups	
	Success group	Failure group
Personal factors		
App versus caregiver	__[b]	—
Prior pelvic floor muscle therapy	3	—
Being independent of care provider	3	4
Insecure about correctly performing exercises	2	1
Lowering shame barriers	2	—
Personality traits	—	—
Positive (eg, go-getter, disciplined)	2	—
Negative (eg, slacking off)	2	4
App factors		
Intensive treatment	3	2
Ease of use	—	—
Devices	—	—
Tablet	3	3
Smartphone	6	6
Complex user interface	3	2
Lessons (exercise levels)	—	—
Useful	7	5
Not useful	1	1
App features	—	—
Reminders	—	—
Useful	4	3
Not useful	6	2
Timing inconvenient	3	4
Graphs	—	—
Useful	3	1
Not useful	5	7
Awareness		
Education	5	3
Awareness of symptoms	—	—
Positive	9	5
Negative	—	2
Adherence		
Integration of exercises	4	6
Level of symptoms	—	—
Recurrence of symptoms (positive)	5	1
Improvement of symptoms (positive)	1	—
Time investment (negative)	—	2
Personal circumstances (negative)	3	5

[a]Numbers are representative of how many participants mentioned the subtheme throughout the interviews.

[b]Not applicable.

Adherence

General Findings

The adherence theme covered factors affecting the level to which participants felt they adhered to treatment advice (ie, app use and performing exercises). Women in both groups felt that their adherence was directly related to the treatment effect, and it was evident that increasing and decreasing symptoms each affected their motivation to adhere to treatment. Several subthemes emerged.

Integration of Exercises

The intensive treatment and frequent reminders provided by the app helped women perform exercises at set times. This enabled them to establish exercise routines that suited their schedules and contributed to the overall treatment effect women experienced. Women in the failure group tended to describe less strict exercise regimes than women in the success group.

> *After a while you'll get into a certain rhythm. I did it on my way to work.... those are the moments you remember, so you get a regularity to it. That is pretty nice.* [P6, success]
>
> *Or when I'm in the car, I have nothing to do, and I'm bored; during a long drive, for example. They are the kind of exercises you can do everywhere with no one taking notice.* [P10, failure]

Level of Symptoms

Women in both groups mentioned that symptom severity influenced adherence, both positively and negatively. One woman in the failure group stated that her low symptom level meant that she lacked the motivation to persevere with the exercises, resulting in minimal treatment effect. Women in the success group more often stated that both symptom improvement and symptom recurrence after a period of less adherence motivated them to start again, hereby enhancing their training results.

> *No, I think that my complaints need to be more severe for that [increased adherence]. Now I just think "I'll use a pantyliner and I'll be done with it.*" [P15, failure]
>
> *…That's why I keep doing it. Because I stopped for about three weeks, because I had surgery on my foot and was hospitalized for it. But after that I could notice that I hadn't done it.* [P8, success]

Time Investment

Some women in the failure group felt that the treatment program was much more time consuming than expected, which markedly decreased their adherence and limited their treatment effect. A lack of time due to personal circumstances such as illness, family reasons, or life events also negatively influenced adherence in both groups, but this was mentioned more by women in the failure group.

Personal Factors

This theme covers personality traits and attitudes toward app-based treatment in comparison with treatment by a care provider.

App Versus Caregiver

Women in both groups valued the concept of 24-hour treatment availability and liked being independent of a care provider. This enabled them to be in control of their own treatment and combine it with their busy and irregular lifestyles, which made it easier for them to adhere to the treatment.

> *Great [opinion about being randomized in app group], because in all honesty, I wasn't looking forward to it at all. I thought, then I'll have to go a general practitioner and make another appointment again. But with an app, you are the one in control, which is much easier.* [P15, failure]

A few women mentioned they occasionally wondered if they performed the exercises correctly, and although none had consulted a care provider, they stated that they might do so in the future. For some women in the success group, preference for the app arose from having experienced insufficient results from prior physical therapy for incontinence. Additionally, one woman stated that she preferred the app because she felt a major barrier when talking about her symptoms.

> *I thought it would be very convenient to try the app, because this [urinary incontinence] is not something I would easily consult my general practitioner for. … It's just not something people talk about.* [P1, success]

Personality Traits

Personal characteristics were frequently mentioned as barriers or facilitators of success. Women in the success group mainly declared it was a matter of just doing it and being a bit of a go-getter to continue with the exercises on a regular basis. Conversely, women in the failure group tended to focus on negative traits and described knowing themselves as sloppy and not being able to persevere, which negatively impacted their adherence.

> *It's just a matter of carrying on, and you'll start getting results.* [P3, success]
>
> *That is the same as going to the dentist and thinking, maybe I should brush my teeth thoroughly for a change [laughs]. ... It's not so much the app having to change, I think it [low adherence] is something engrained in humans.* [P12, failure]

App Factors

Subthemes related to app factors (ie, experiences with different features) included the intensity and extensiveness of treatment, ease of use of the app, and features within the app.

Intensive Treatment

Women in both groups appreciated the intensive and extensive treatment program offered by the app, indicating that they felt this was something a caregiver could not provide.

> *...I think you are more dedicated to it, especially at the start. When you go to the physical therapist, you get some exercises, You go home, and you do those. But with the app, you just do it every day.* [P4, success]

Ease of Use

Most women installed the app on their smart phone because this device was always at hand, which made it easier to adhere to the intensive treatment. Others preferred a tablet because of the larger display or because they did not possess a smartphone or know how to operate one. Most women in both groups found the app easy to use and appreciated the clear instruction provided in the lessons. However, a few women stated that the app's user interface was overly complex, taking too long to identify where to start and how to get an overview of the content. This negatively impacted their motivation to use the app.

App Features

Most women tried the app's reminder function, but it yielded mixed feelings regarding the effect on their treatment adherence. Despite being able to set 3 reminders per day, many women in both groups found the timings inconvenient or did not want to receive a reminder when they were with other people. Also, some women were unaware that the app provided this function. Overall, despite many women appreciating the inclusion of a reminder function, a slightly larger cohort (mainly from the success group) stated that they ultimately stopped using this feature.

> *No, I felt those [reminders] were actually only annoying because I already had my own vision of when I was generally going to practice.* [P12, failure]

> *I was planning on doing them only when I was by myself. I did not want to receive a reminder when I was out somewhere.* [P1, success]

Some women stated that the graphs provided insights about progression and made them more aware of their symptoms, but only a few participants used and appreciated this function. When used, the function did give a sense of being on the right path and encouraged perseverance. One participant from each group stated they thought it would have been more motivational to have a graphical display showing symptom changes. However, many women, mostly from the failure group, found that the graphs were difficult to interpret or that they added little. One woman declared that she found looking at the graphs to be too confrontational.

> *I would leave that out; when you've practiced and all the statistics. If you skipped that for a day, you'll start feeling guilty.* [P16, failure]

Awareness

General Findings

There was increased awareness in several domains. Awareness increased concerning knowledge of the disease (education) and awareness of symptoms. This increased awareness could act as either a facilitator of or a barrier to adherence. Furthermore, awareness increased directly with both app factors (reminders) and treatment effects (symptom improvement or recurrence).

Education

Women in both groups found the information provided by the app useful. Many had thought urinary incontinence was a part of life they had to accept. Knowledge about the possible effect of conservative therapies enhanced their motivation to carry on with the exercise program. Others stated that they felt less alone dealing with urinary incontinence knowing that other women experienced the same symptoms.

Awareness of Symptoms

The intensive treatment resulted in increased awareness of the impact of symptoms and of coping strategies used. In general, women in both groups appreciated this aspect. In the failure group, women stated that they liked knowing what to do to improve their symptoms. In the success group, women tended to report putting this knowledge into action, stating several key benefits: they felt more confidence in their treatment and the treatment helped them to make lifestyle changes and lessened the sense of taboo when talking with other women about symptoms.

> *Well, I certainly know what to do now to get results. I know that I have to do it for months, then it will work. That I understand.* [P11, failure]

> *It gave me some reassurance in the sense of "that should be doable." And that already made it easier to postpone toilet visits.* [P7, success]

Conversely, a few women in the failure group stated that they did not like the increased focus on themselves and their problems, which made them less motivated to use the app. One even wondered if this had led to her symptoms increasing.

Discussion

Principal Findings

Our findings provide new insights into the barriers and facilitators associated with successful app-based treatment for urinary incontinence. Additionally, this study contributes to the growing understanding of barriers or facilitating factors influencing mHealth use and ways to overcome or improve them [6].

Our results principally show that the effect of each explored factor results from whether there is treatment success or failure. Moreover, the views of patients concerning adherence to app use and performing the recommended exercises were key. Comparison between the success and failure group revealed several factors that facilitated treatment success, namely strict integration of exercises, previous experience of face-to-face pelvic floor muscle therapy (PFMT) with insufficient effect, and being a so-called go-getter; by contrast, we identified the barriers as being unrealistic expectations of time investment, interfering personal circumstances, and being unable to persevere. Of note, however, the graphs and reminder functions did not have the expected facilitating effect and, indeed, sometimes acted as a barrier. It was interesting that the general increased awareness after treatment and the awareness of symptom change positively and negatively affected adherence and treatment effectiveness. We believe these facilitators and

barriers can be used to improve outcomes with app-based therapy.

Strengths and Limitations

This is the first study using a sequential explanatory design to assess the facilitators of and barriers to app-based treatment for urinary incontinence. We consider the mixing and integration of qualitative and quantitative data throughout the study to be an important strength, helping to improve the quality of our conclusions [12]. This approach produces a whole that is greater than the sum of the individual qualitative and quantitative parts [19]. We also selected high- and low-performing cases to explore the contrast between treatment success and failure [11], which enabled us to identify facilitators and barriers associated with the desired treatment effect. Other strengths of our design were the use of previously collected qualitative data to build the interview guide, the reevaluation of themes within each outcome group, and the use of quote frequency counts.

Despite these notable strengths, however, there were some important limitations. For example, there was no member check due to logistic difficulties, and 8 women were unavailable for interview, potentially affecting the identified themes. Additionally, it should be noted that the exploratory nature of this type of (qualitative) research allows for hypothesis generation, not hypothesis testing. Therefore, when interpreting the results, one should keep in mind that this research is not able and does not seek to predict treatment effect. This research rather explores the factors influencing treatment from a qualitative participant perspective and relates these to the quantitative treatment effects.

Participant selection for the interviews was based on follow-up outcomes at 4 months, which we anticipated would reflect the optimum treatment effect. However, interviews were postponed until after the 12-month follow-up to limit interference with the trial. Although this extension allowed us to explore facilitators and barriers in both the short- and long-term, it could have introduced recall bias in the women's experiences and perception of factors influencing effectiveness in the first 4 months.

There was also some inconsistency with the concepts of failure and success. Among the women with a deterioration in urinary incontinence severity on the ICIQ-UI-SF at 4 months, none perceived a worsening on the PGI-I at that time and none reported treatment failure in the interviews after 12 months. Recall bias could explain the inconsistency between the ICIQ-UI-SF at 4 months and the interview after 12 months but not the difference between the PGI-I and the ICIQ-UI-SF, both of which were measured at 4 months. Thus, it may be that these differences indicate that the perception of improvement reflects not only the change in urinary incontinence symptoms but also better coping strategies or decreased shame due to increased knowledge.

Comparison With Existing Literature

Previous studies have included women with no experience with eHealth for urinary incontinence, using eHealth for urinary incontinence for 6 weeks to 3 months, and with no case selection based on treatment effect [7-10]. In this study, we explored the experiences of women using the app for 12 months who had showed a clear worsening or improvement of symptoms. Although we identified similar main themes, we could also further explore the relationship between those themes and the treatment effect. Consistent with existing research, women from both of our study groups expressed positive views about the availability, flexibility, privacy, and education provided by eHealth for urinary incontinence [7-10].

Insecurity About Exercise Performance

Women's feelings in our study were mixed with regard to insecurities about the correct performance of exercises. Firet et al [7] described that women had experienced their pelvic floor muscles to be difficult to contract correctly during face-to-face PFMT. Elsewhere, Asklund et al [9] reported that a lack of reassurance created insecurity when women thought contractions were good enough but were left wondering if personal instruction could lead to improvements. In our study, women expressed these insecurities in both the success and failure groups, but despite being instructed to consult a health care professional if they needed, none sought further advice. This suggests that the presence of insecurity about treatment is not a differentiating factor for treatment success or failure. Instead, treatment failure in women with insecurities may have reflected other barriers (eg, not being able to persevere or having interfering personal circumstances) or different coping strategies, with insecurities and doubts keeping them from consulting a caregiver.

Awareness of urinary incontinence symptoms and treatment options acted as both a facilitator and a barrier for women in our study, whereas in other studies, increased awareness was mainly described positively [8,9].

Increased Awareness: Positive Effects

Positive effects found in our study were an increased awareness of symptoms and treatment options, which lessened the sense of taboo around the topic and encouraged women to change their lifestyles. Additionally we confirmed that awareness of symptom recurrence after a period of lower adherence stimulated motivation, as Asklund et al [9] also described.

Increased Awareness: Negative Effects

Negative effects were related to a negative focus on symptoms and a decrease in adherence to treatment. The increased negative focus on symptoms in some women acted as a barrier as it kept them from continuing app use, which was also reported by Wessels et al [10]. Also, for some women, awareness of symptom improvement during treatment led to decreased motivation to adherence to the treatment.

App Features

Additionally, it was notable that reminders did not facilitate treatment success and the graph function was deemed too confrontational or unhelpful, contrasting with our expectation that these would positively affect motivation and adherence [9,10]. This may be related to the long 12-month follow-up period. For example, the facilitating effect of reminders may have been small or only present early on, potentially being lost due to recall bias. The sense that the graphs were confrontational

may have appeared over time in response to a lack of treatment effect, but this may also have resulted because the graphs only monitored lack of adherence, rather than progress or change in urinary incontinence symptoms. Asklund et al [9] showed the same statistics in their graph use but did not report this confrontational effect.

Implications for Research and Practice

The findings of our study can be used to increase the effect of app-based treatment by targeting women who are most likely to benefit and showing how we can better tailor app-based treatments. When the app is made available to the wider public, it will be important to inform potential users about the various factors that can influence the treatment effect. When care providers discuss the use of app-based treatment for a patient with urinary incontinence, our findings indicate it is crucial they consider personality traits (eg, highly self-motivated), expectations of time investment, and previous experiences with regular PFMT. We can tailor the app-based treatment to increase the treatment effect by modifying the graph and reminder functions. Graphs could be an optional tool that are simplified to emphasize urinary incontinence symptom progression rather than lack of adherence. To reduce the perceived intrusion of the reminder function, this could be revised to a daily to-do list with no preset times. Finally, future research could be focused on further examining the characteristics of women in whom app-based treatment failed because this might be a distinct group with similar personality traits. This knowledge could help health care professionals provide the necessary support for patients to achieve treatment success.

In conclusion, this study shows that the effect of app-based treatment for urinary incontinence is mainly influenced by adherence, which in turn is affected by personal factors, app-based factors, and awareness. However, it was notable that the identified factors could function as both facilitators and barriers depending on the user and the interaction with other factors. Insight into these facilitators and barriers can be used to increase the treatment effect of app-based treatment for urinary incontinence by ensuring that we target women most likely to benefit. Introducing some minor changes to the graph and reminder functions could improve the usability of our app.

Acknowledgments

We thank the participants for their time and effort, Dr Robert Sykes [20] for providing editorial services, and Jan Eise Kuipers for translating the citations. This work was supported by a grant from ZonMw, the Dutch Organization for Health Research and Development (project number: 837001508) and subfunded by a grant from the PW Boer Foundation. The study won the 2016 Professor Huygen Award for best study proposal in general practice, which resulted in additional funding. The funders had no role in data collection, data analysis, the decision to publish, or manuscript preparation. All authors had full access to all the study data and take responsibility for the integrity of the data and accuracy of the data analysis.

Conflicts of Interest

NJW, AMML, HvdW, LA, JD, MYB, and MHB are affiliated with the Department of General Practice and Elderly Care Medicine at the University Medical Centre Groningen, which owns the intellectual property of the URinControl app. The URinControl app is provided to all women in the Netherlands without costs.

Multimedia Appendix 1
Design summary.
[DOCX File , 17 KB - mhealth_v9i9e25878_app1.docx]

Multimedia Appendix 2
Supplemental methods.
[DOCX File , 17 KB - mhealth_v9i9e25878_app2.docx]

Multimedia Appendix 3
Coding tree including major topics and subthemes.
[DOCX File , 49 KB - mhealth_v9i9e25878_app3.docx]

References

1. Asklund I, Nyström E, Sjöström M, Umefjord G, Stenlund H, Samuelsson E. Mobile app for treatment of stress urinary incontinence: a randomized controlled trial. Neurourol Urodyn 2016 Sep 09;36(5):1369-1376. [doi: 10.1002/nau.23116] [Medline: 27611958]
2. Sjöström M, Lindholm L, Samuelsson E. Mobile app for treatment of stress urinary incontinence: a cost-effectiveness analysis. J Med Internet Res 2017 May 08;19(5):e154 [FREE Full text] [doi: 10.2196/jmir.7383] [Medline: 28483745]
3. Loohuis AMM, Wessels NJ, Dekker JH, van Merode NAM, Slieker-Ten Hove MCP, Kollen BJ, et al. App-based treatment in primary care for urinary incontinence: a pragmatic, randomized controlled trial. Ann Fam Med 2021;19(2):102-109 [FREE Full text] [doi: 10.1370/afm.2585] [Medline: 33685871]

4. Schreiweis B, Pobiruchin M, Strotbaum V, Suleder J, Wiesner M, Bergh B. Barriers and facilitators to the implementation of ehealth services: systematic literature analysis. J Med Internet Res 2019 Nov 22;21(11):e14197 [FREE Full text] [doi: 10.2196/14197] [Medline: 31755869]
5. Greenhalgh T, Abimbola S. The NASSS framework: a synthesis of multiple theories of technology implementation. Stud Health Technol Inform 2019 Jul 30;263:193-204. [doi: 10.3233/SHTI190123] [Medline: 31411163]
6. Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]
7. Firet L, Teunissen D, Verhoeks C, Lagro-Janssen A. Expectations regarding eHealth among women with stress urinary incontinence. Int Urogynecol J 2019 Nov;30(11):1955-1963. [doi: 10.1007/s00192-018-3849-2] [Medline: 30594948]
8. Björk A, Sjöström M, Johansson EE, Samuelsson E, Umefjord G. Women's experiences of internet-based or postal treatment for stress urinary incontinence. Qual Health Res 2014 Apr;24(4):484-493. [doi: 10.1177/1049732314524486] [Medline: 24598777]
9. Asklund I, Samuelsson E, Hamberg K, Umefjord G, Sjöström M. User experience of an app-based treatment for stress urinary incontinence: qualitative interview study. J Med Internet Res 2019 Mar 14;21(3):e11296 [FREE Full text] [doi: 10.2196/11296] [Medline: 30869644]
10. Wessels NJ, Hulshof L, Loohuis AMM, van Gemert-Pijnen L, Jellema P, van der Worp H, et al. User experiences and preferences regarding an app for the treatment of urinary incontinence in adult women: qualitative study. JMIR Mhealth Uhealth 2020 Jul 12;8(6):e17114 [FREE Full text] [doi: 10.2196/17114] [Medline: 32530431]
11. Morgan D. Follow-up qualitative extensions to core quantitative research projects. In: Integrating Qualitative and Quantitative Methods: A Pragmatic Approach. Thousand Oaks: Sage Publications; 2014:62-67.
12. Loohuis AMM, Wessels NJ, Jellema P, Vermeulen KM, Slieker-Ten Hove MC, van Gemert-Pijnen JEWC, et al. The impact of a mobile application-based treatment for urinary incontinence in adult women: design of a mixed-methods randomized controlled trial in a primary care setting. Neurourol Urodyn 2018 Sep;37(7):2167-2176. [doi: 10.1002/nau.23507] [Medline: 29392749]
13. Damen-van Beek Z, Wiersma T. Practice guideline: urinary incontinence in women. Dutch College of General Practitioners. URL: https://richtlijnen.nhg.org/files/pdf/43_Incontinentie%20voor%20urine%20bij%20vrouwen_juli-2015.pdf [accessed 2020-07-20]
14. Dumoulin C, Hunter KF, Moore K, Bradley CS, Burgio KL, Hagen S, et al. Conservative management for female urinary incontinence and pelvic organ prolapse review 2013: summary of the 5th International Consultation on Incontinence. Neurourol Urodyn 2016 Jan;35(1):15-20. [doi: 10.1002/nau.22677] [Medline: 25400065]
15. Ivankova NV, Creswell JW, Stick S. Using mixed-methods sequential explanatory design: from theory to practice. Field Methods 2016 Jul 21;18(1):3-20. [doi: 10.1177/1525822X05282260]
16. Creswell J, Plano Clark VL, Gutmann ML, Hansen WE. Advanced mixed methods research designs. In: Tashakkori A, Teddlie C, editors. Handbook of Mixed Methods in Social and Behavioral Research. Thousand Oaks: Sage Publications; 2003:209-240.
17. Avery K, Donovan J, Peters TJ, Shaw C, Gotoh M, Abrams P. ICIQ: a brief and robust measure for evaluating the symptoms and impact of urinary incontinence. Neurourol Urodyn 2004;23(4):322-330. [doi: 10.1002/nau.20041] [Medline: 15227649]
18. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
19. Fetters MD, Freshwater D. The 1 + 1 = 3 integration challenge. J Mixed Methods Res 2015 Apr 09;9(2):115-117. [doi: 10.1177/1558689815581222]
20. Doctored. URL: https://www.doctored.org.uk/ [accessed 2021-08-09]

Abbreviations

COREQ: Consolidated Criteria for Reporting Qualitative Research
ICIQ-UI-SF: International Consultation on Incontinence Modular Questionnaire Urinary Incontinence Short Form
mHealth: mobile health
PFMT: pelvic floor muscle therapy
PGI-I: Patient Global Impression of Improvement
RCT: randomized controlled trial

Edited by L Buis; submitted 19.11.20; peer-reviewed by A Tretiakov, D Parry; comments to author 15.01.21; revised version received 24.02.21; accepted 18.05.21; published 17.09.21.

Please cite as:
Wessels NJ, Loohuis AMM, van der Worp H, Abbenhuis L, Dekker J, Berger MY, van Gemert-Pijnen JEWC, Blanker MH
Barriers and Facilitators Associated With App-Based Treatment for Female Urinary Incontinence: Mixed Methods Evaluation
JMIR Mhealth Uhealth 2021;9(9):e25878
URL: https://mhealth.jmir.org/2021/9/e25878
doi:10.2196/25878
PMID:34533466

Original Paper

Effects of an Artificial Intelligence–Assisted Health Program on Workers With Neck/Shoulder Pain/Stiffness and Low Back Pain: Randomized Controlled Trial

Tomomi Anan[1,2], MD; Shigeyuki Kajiki[1,2,3], MD, PhD; Hiroyuki Oka[4], MD, PhD; Tomoko Fujii[4], MD, PhD; Kayo Kawamata[4], PT; Koji Mori[1], MD, PhD; Ko Matsudaira[4], MD, PhD

[1]Department of Occupational Health Practice and Management, Institute of Industrial Ecological Sciences, University of Occupational and Environmental Health, Japan, Kitakyushu, Japan

[2]Cenxus Occupational Physicians' Firm, Tokyo, Japan

[3]Advanced Occupational Health Research and Consulting Inc, Tokyo, Japan

[4]Department of Medical Research and Management for Musculoskeletal Pain, 22nd Century Medical and Research Center, Faculty of Medicine, The University of Tokyo, Tokyo, Japan

Corresponding Author:
Shigeyuki Kajiki, MD, PhD
Department of Occupational Health Practice and Management
Institute of Industrial Ecological Sciences
University of Occupational and Environmental Health, Japan
1-1 Iseigaoka, Yahatanishi-ku
Kitakyushu, 807-8555
Japan
Phone: 81 93 691 7523
Fax: 81 93 692 4590
Email: skajiki@med.uoeh-u.ac.jp

Abstract

Background: Musculoskeletal symptoms such as neck and shoulder pain/stiffness and low back pain are common health problems in the working population. They are the leading causes of presenteeism (employees being physically present at work but unable to be fully engaged). Recently, digital interventions have begun to be used to manage health but their effectiveness has not yet been fully verified, and adherence to such programs is always a problem.

Objective: This study aimed to evaluate the improvements in musculoskeletal symptoms in workers with neck/shoulder stiffness/pain and low back pain after the use of an exercise-based artificial intelligence (AI)–assisted interactive health promotion system that operates through a mobile messaging app (the AI-assisted health program). We expected that this program would support participants' adherence to exercises.

Methods: We conducted a two-armed, randomized, controlled, and unblinded trial in workers with either neck/shoulder stiffness/pain or low back pain or both. We recruited participants with these symptoms through email notifications. The intervention group received the AI-assisted health program, in which the chatbot sent messages to users with the exercise instructions at a fixed time every day through the smartphone's chatting app (LINE) for 12 weeks. The program was fully automated. The control group continued with their usual care routines. We assessed the subjective severity of the neck and shoulder pain/stiffness and low back pain of the participants by using a scoring scale of 1 to 5 for both the intervention group and the control group at baseline and after 12 weeks of intervention by using a web-based form. We used a logistic regression model to calculate the odds ratios (ORs) of the intervention group to achieve to reduce pain scores with those of the control group, and the ORs of the subjective assessment of the improvement of the symptoms compared to the intervention and control groups, which were performed using Stata software (version 16, StataCorp LLC).

Results: We analyzed 48 participants in the intervention group and 46 participants in the control group. The adherence rate was 92% (44/48) during the intervention. The participants in the intervention group showed significant improvements in the severity of the neck/shoulder pain/stiffness and low back pain compared to those in the control group (OR 6.36, 95% CI 2.57-15.73; $P<.001$). Based on the subjective assessment of the improvement of the pain/stiffness at 12 weeks, 36 (75%) out of 48 participants

in the intervention group and 3 (7%) out of 46 participants in the control group showed improvements (improved, slightly improved) (OR 43.00, 95% CI 11.25-164.28; *P*<.001).

Conclusions: This study shows that the short exercises provided by the AI-assisted health program improved both neck/shoulder pain/stiffness and low back pain in 12 weeks. Further studies are needed to identify the elements contributing to the successful outcome of the AI-assisted health program.

Trial Registration: University hospital Medical Information Network-Clinical Trials Registry (UMIN-CTR) 000033894; https://upload.umin.ac.jp/cgi-open-bin/ctr_e/ctr_view.cgi?recptno=R000038307.

(JMIR Mhealth Uhealth 2021;9(9):e27535) doi:10.2196/27535

KEYWORDS

neck pain; shoulder pain; shoulder stiffness; low back pain; musculoskeletal symptoms; digital intervention; mobile app; mHealth; eHealth; digital health; mobile phone

Introduction

Musculoskeletal symptoms of the neck and shoulder and low back pain are common health problems in the working population [1,2]. In Japan, shoulder stiffness and low back pain are the most common somatic symptoms. In men, low back pain is the most common, followed by stiff shoulders; in women, the order is reversed—stiff shoulders rank first, while low back pain ranks second [3]. Health problems in employees incur substantial costs in terms of medical care expenditure and poor work productivity, which is marked by absenteeism and presenteeism [4,5]. Absenteeism is defined as health-related absence from work, and presenteeism is the condition where employees are physically present at work but are unable to fully engage themselves [4,5]. A recent survey reported that neck and shoulder stiffness followed by low back pain were the leading causes of presenteeism in Japanese workers [5].

The causes of most musculoskeletal symptoms are not clearly known. However, primary neck/shoulder pain/stiffness and nonspecific low back pain are related to poor posture and psychological stress. Neck/shoulder stiffness can be classified into primary or secondary complaints. The causes of secondary neck/shoulder stiffness or pain include cervical spine diseases, glenohumeral joint diseases, cardiovascular diseases, eye fatigue, and temporomandibular disorder [6]. The causes of primary neck/shoulder stiffness or pain are not well known. However, hemodynamics in the trapezius muscle is expected to be involved [7,8]. Besides, head-down posture, psychological factors [9], and physical inactivity [10] are related to chronic neck and shoulder pain.

Low back pain is mainly classified into 2 categories: specific and nonspecific. Specific low back pain occurs when the symptoms are caused by a specific pathophysiological mechanism such as lumbar disk herniation, infection, osteoporosis, rheumatoid arthritis, fracture, or tumor [11]. It accounts for only about 10% of all low back pain cases [11]. In contrast, about 90% of patients with low back pain experience nonspecific low back pain, where the symptoms do not have a clear and specific cause [11].

Several studies have reported methods to minimize the discomfort from such symptoms—a combination of exercises and psychological approaches seems effective for patients with musculoskeletal symptoms. Moderate-to-strong evidence suggests that exercise therapy is effective in relieving pain and improving function in musculoskeletal disease [12-18]. However, there seems to be little evidence on the types of exercises and programs that are effective in relieving musculoskeletal pain [12,14,15,19]. There are also several psychological treatments or therapies for musculoskeletal symptoms [12]. In a study on patients with chronic low back pain, both groups (one that received only exercise therapy and another that received a combination of cognitive behavioral therapy and exercise therapy) showed improvements in pain intensity and quality of life compared to baseline [20]. Patients with chronic pain who received acceptance and commitment therapy, a third-generation cognitive behavioral therapy based on mindfulness, experienced minor improvements in pain intensity and the degree of depressive symptoms and dysfunction and moderate improvements in the degree of anxiety and disability caused by pain compared to patients in the normal treatment group and control group [21].

One problem with exercise therapy is the low level of adherence to the prescribed exercises. Two systematic reviews reported that up to 70% of participants did not adhere to the prescribed exercises [22,23]. The lower the adherence to exercise, the lesser effective is the therapy [24]. Therefore, adherence is a critical factor that determines the outcomes of the intervention process [25]. The type of exercise does not seem to influence adherence [25], but support from experts seems important [26,27]. Factors that enhance the adherence to exercise programs include the attractiveness of the programs, feedback from and interaction with experts, evaluation of patients' performance, and the feeling of being supported by experts [24]. Supervised exercises, review sessions, and visual or audio aids are also effective [25]. Additional support such as phone calls, email reminders, and text messages promotes the engagement of digital interventions [28-33]. In Japan, the medical system has not been able to provide sufficient services for such conditions. Some enthusiastic medical professionals, preventive medical services, and occupational health services provide care for people with functional impairments such as musculoskeletal symptoms. In most other situations, patients need to look for ways to improve their symptoms on their own.

Currently, digital health programs using smartphones, tablets, and computers are relatively inexpensive and are widely accepted, especially by the young and the middle-aged. Three systematic reviews involving musculoskeletal symptoms that

included a plethora of studies with digital interventions, for example, mobile phone apps, websites, and web-based software have been performed [28,34,35]. About half of these showed that digital interventions were effective in reducing low back pain [28,30-32,36,37]. There are limited studies on digital interventions for stiff neck/shoulders; an intervention for shoulder stiffness using a software program that promoted regular breaks did not show a significant effect on the severity and frequency of the symptoms [38]. Besides, 65% of smartphone apps that perform self-monitoring and self-managing of chronic pain have been developed without the involvement of experts and proper supporting evidence [39]. Conversely, it has also been suggested that digital interventions may have possibilities to improve adherence in the target population [26]. Some studies including digital health interventions have shown enhanced self-management and adherence to medications in patients with asthma, chronic obstructive pulmonary disease, hypertension, and diabetes [40-43].

The aim of this study was to improve neck/shoulder pain/stiffness and low back pain of workers who experienced those symptoms by continuing to do exercises that included stretching and mindfulness. As a measure of encouraging them to continue the exercises, we provided them with secaide Ver.0.9 [44], an artificial intelligence (AI)–assisted interactive health promotion system using a mobile messaging app (the AI-assisted health program). We hypothesized that this digital intervention would support participants to continue exercising and enhance their adherence to the exercises, resulting in greater improvement of symptoms. To the best our knowledge, this is the first study to use a chatbot as a health care support measure through a messaging app to improve the musculoskeletal symptoms of workers.

We implemented a randomized controlled trial (RCT) among workers with either neck/shoulder pain/stiffness or low back pain or both in a company setting to evaluate the improvement of musculoskeletal symptoms by using an AI-assisted interactive health promotion system through a mobile messaging app to habituate exercise.

Methods

Study Design

This study was a two-armed, randomized, controlled, and unblinded trial on workers with either neck/shoulder pain/stiffness or low back pain or both. We set the intervention and control groups, and the participants of the intervention group used the AI-assisted health program for 12 weeks and those of the control group continued their regular exercise routine at their workplace. We assessed the subjective severity of the neck and shoulder pain/stiffness and low back pain in both the intervention and control groups at baseline and 12 weeks immediately after the intervention. We provided an explanatory document regarding the study to the applicants and obtained informed consent. This trial was conducted with the approval of the ethics committee of the University of Tokyo Hospital (ID 12035) and the ethical review of the target company. This trial was registered at University hospital Medical Information Network-Clinical Trials Registry (ID 000033894). This RCT was performed based on CONSORT-EHEALTH (Consolidated Standards of Reporting Trials-eHealth) guidelines.

Study Population

We conducted this study in a company that develops, designs, and manufactures precision electronic components and that has approximately 2200 employees. Some employees were manufacturing engineers who managed and supervised the manufacturing process of the precision electronic components, but most employees were white-collar employees who were engaged in the design, development, or clerical work of the product. The occupational health staff of the company recruited employees with remarkable musculoskeletal symptoms, either or both neck/shoulder stiffness/pain and low back pain. They chose participants based on data from the periodic health survey conducted by the target company. The company performs the survey once a year to check the physical and mental health conditions of the employees. The employees used a specific URL to sign into the health check system and answer specific questions, according to the instructions provided. We targeted those who answered “frequently” or “almost always” in either of the following questions in the survey:

1. How often do you experience neck/shoulder stiffness?

(1) Almost never, (2) Occasionally, (3) Frequently, (4) Almost always

2. How often do you have low back pain?

(1) Almost never, (2) Occasionally, (3) Frequently, (4) Almost always

They recruited applicants by email notifications between September 3 and 14, 2018. The inclusion criteria were as follows: employees aged 20-64 years, who had their own cell phones and the apps could be installed on the phones, and who understood the purpose and agreed to the publishing of the contents and results of the study. The exclusion criteria were as follows: employees who disagreed with the study, were pregnant or may have been pregnant, had cardiopulmonary diseases, participated in other clinical trials, and had any other obvious disabilities or exercise restrictions.

Randomization

A total of 121 employees applied for the study. After we confirmed that the applicants met all the inclusion criteria and did not violate any of the exclusion criteria, all 121 participants were randomized to 2 groups by generating random numbers on a computer and stratified by 10-year age groups and separated by the age group with a 1:1 allocation ratio.

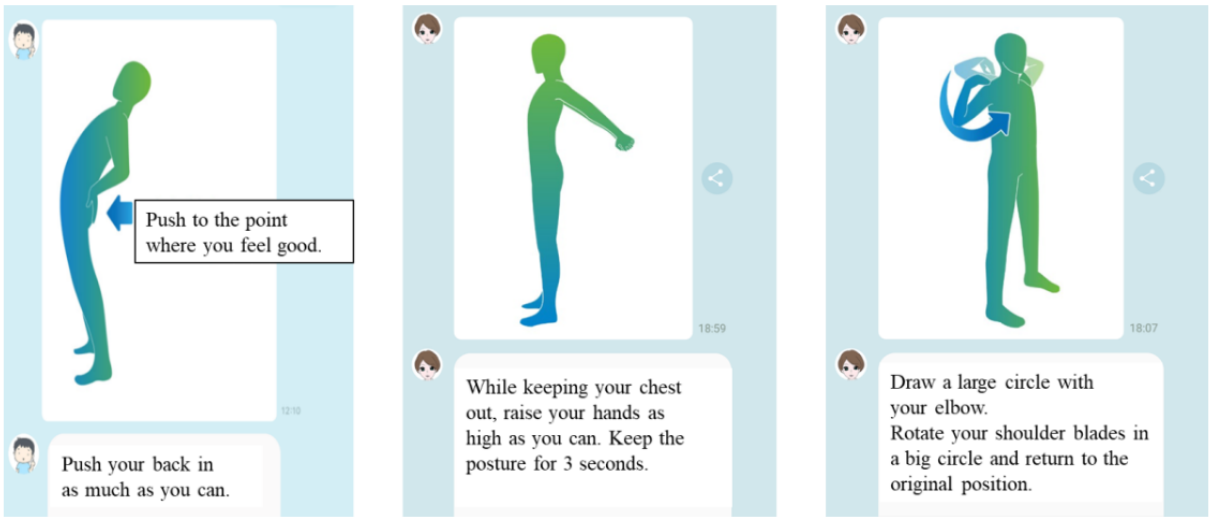
Intervention

We held an initial training session for the applicants to explain the purpose of the study and obtained informed consent. We also explained to the participants of the intervention group how to install and use the AI-assisted health program on their cell phones to inform them a Quick Response code and a passcode to access the program on September 26 and 27, 2018. After that, they started to use the program. The control group continued with their regular exercise routine, which included exercising for about 3 minutes during the break time provided by the

company every day to prevent stiff shoulders and back pain. We also allowed the control group to use the AI-assisted health program after the 12-week intervention.

The AI-assisted chatbot was programmed to send the users messages with the exercise instructions and some tips on what they can do in their daily lives to improve those symptoms. The messages were sent every day at a fixed time through LINE. The notification time could be changed by the users to a time convenient for them. The participants could finish their exercise within 1 minute each day. The program is interactive and the participants can respond to the messages by using a simple selection list; the chatbot offers them tailored replies depending on their responses. The exercise provided by the program had 3 components: stretching [45-47], maintaining good posture [48,49], and mindfulness [21] (Figure 1). When the participants interrupted the exercise, the chatbot motivated them to continue exercising. The program we used in this study was named secaide Ver.0.9, which was patented and has been created and developed since June 2017 by Travoss Co, Ltd and an orthopedist who is specialized in musculoskeletal disorders. Until we performed this interventional study, this program had not been used or evaluated previously. During the intervention period, we notified participants that occupational health staff at the company would respond to inquiries about changes in their physical conditions and the company that provided the program would reply to technical questions about accessing the program, but we did not provide any human support such as specific advice on how to perform the exercises or recommendation to continue the exercise.

Figure 1. Examples of exercises with instructions from the artificial intelligence–assisted health program.



Adherence to the Program

We counted the number of participants who continued to access and reply the chatbot's messages at least once every 3 days and excluded the participants who did not access and reply for 3 weeks in a row. Then, we calculated the adherence rate by dividing the abovementioned count by the number of participants at the start of the intervention.

Outcomes

In this study, we set 2 types of outcomes. The first was a subjective assessment of the degree of pain on a scale of 1 to 5; this included subjective ratings of the neck/shoulder stiffness/pain and low back pain at baseline and after 12 weeks. A score of 4 or more was defined as severe pain. The second was a subjective assessment of whether there was an improvement. The participants were asked to subjectively rate whether their pain had improved after 12 weeks; they chose from the following options: improved, slightly improved, unchanged, slightly worse, and worse. Those who responded that their pain had improved or slightly improved were defined as the group that showed subjective improvement. All participants answered the questions through a web-based form.

Statistical Analysis

A linear regression analysis was used to compare the intervention and control groups for change in the subjective pain scores after the program. The odds ratios (ORs) of the intervention group to achieve a subjective pain score of less than 3 compared with those of the control group was estimated using the logistic regression model. In addition, the OR of the subjective improvement of the symptoms compared to the intervention and control groups was estimated using the logistic regression model. All analyses were performed using Stata software (version 16, StataCorp LLC).

Results

Study Population

Figure 2 shows the CONSORT flow diagram. A total of 121 employees applied for this study. All the employees were engineers who were engaged in developing or designing precision devices or clerical workers, which meant that they spent most of their working hours doing sedentary work. After we confirmed the eligibility of 121 applicants who wished to participate in the study, we randomly assigned them to an intervention group (n=61) and control group (n=60), because we had to notify the schedule of the initial session during their

working hours to the participants. Unfortunately, 13 and 14 applicants allotted in the intervention group and control group, respectively, could not participate in the session or answer the baseline survey. Therefore, the intervention started with 48 and 46 participants in the intervention and control groups, respectively. We could follow up 48 and 42 participants in the intervention and control groups, respectively. Table 1 shows the baseline characteristics of the participants who answered the survey at baseline.

Figure 2. CONSORT (Consolidated Standards of Reporting Trials) flow diagram.

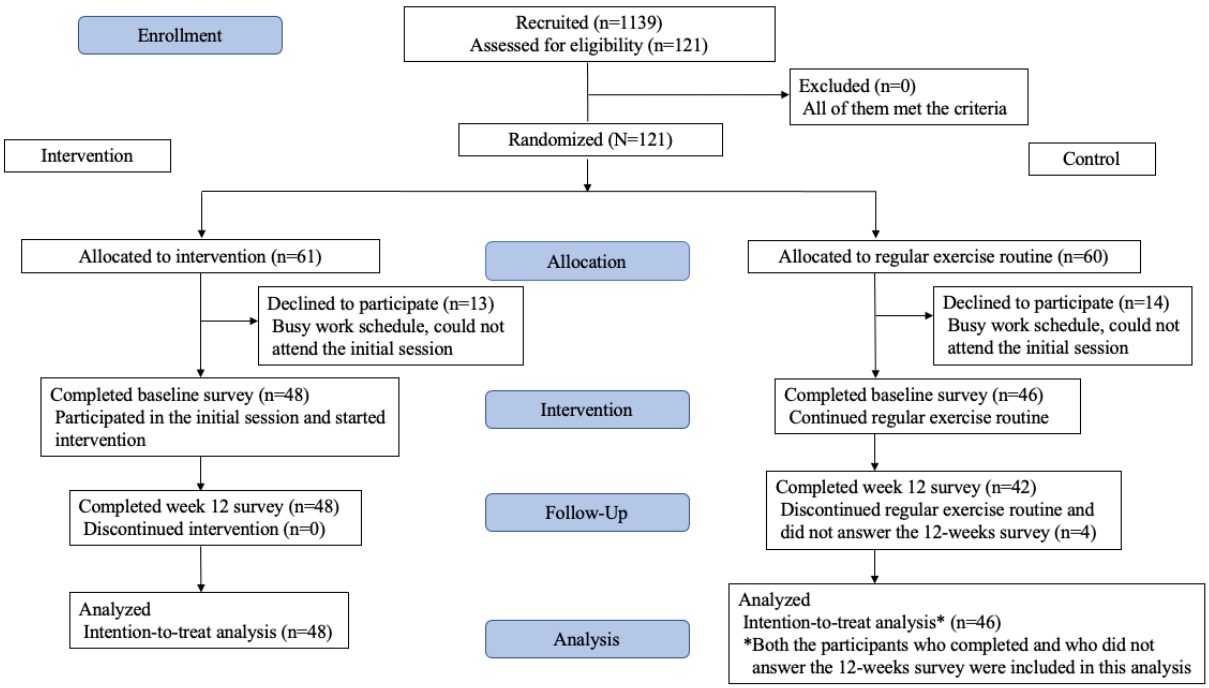


Table 1. Baseline characteristics of the participants.

Characteristics	Intervention group (n=48)	Control group (n=46)
Age (years), mean (SD)	41.8 (8.7)	42.4 (8.0)
Sex, n (%)		
Women	9 (19)	13 (28)
Men	39 (81)	33 (72)
Severe neck/shoulder stiffness/pain,[a] n (%)	34 (71)	31 (67)
Severe low back pain,[a] n (%)	11 (23)	18 (39)
Pain level of neck/shoulder stiffness or low back pain, n (%)		
3[b]	10 (21)	9 (20)
4[b]	24 (50)	19 (41)
5[b]	13 (27)	16 (35)

[a]A score of 4 or more indicates severe pain.

[b]Subjective ratings of the degree of pain on a scale of 1 to 5.

Adherence to the Program

Of the 48 participants in the intervention group, 47 started the AI-assisted health program and 44 continued the exercise for the entire intervention period. The adherence rate was 92% (44/48) in this study.

Outcomes

Table 2 shows the results of the outcomes. At 12 weeks, the average pain level of the neck/shoulder stiffness/pain or low back pain was 3.0 (SD 1.1) in the intervention group and 4.0 (SD 0.8) in the control group; the difference was statistically significant (P<.001). Each pain level (scores 1 to 5) also showed a statistically significant difference (P<.001) between the intervention and control groups. We analyzed the outcomes by

dividing them into 2 categories depending on whether symptoms were severe (a score of 4 or more) or not (a score of 3 or less) at 12 weeks. There was a significant difference between the 2 groups (*P*<.001). In the intervention group, the proportion of participants who had severe symptoms decreased from 77% (37/48) at baseline to 33% (16/48) after the intervention. Conversely, in the control group, the proportion of participants who had severe symptoms decreased from 76% (33/46) to 67% (31/46). Furthermore, regarding the subjective assessment of whether there was an improvement, 36 (75%) out of 48 participants in the intervention group and 3 (7%) out of 46 participants in the control group showed improvements (improved, slightly improved), and there was a statistically significant difference (*P*<.001) between the 2 groups.

Table 2. Scores of the outcomes after the intervention.

Outcome	Intervention group (n=48)	Control group (n=46)	*P* value
Pain level of the neck/shoulder pain/stiffness or low back pain after the intervention (12 weeks), mean (SD)	3.0 (1.1)	4.0 (0.8)	<.001
Pain level of the neck/shoulder stiffness/pain and low back pain after the intervention (12 weeks), n (%)			<.001
1[a]	2 (4)	0 (0)	
2[a]	19 (40)	0 (0)	
3[a]	11 (23)	15 (33)	
4[a]	10 (21)	16 (35)	
5[a]	6 (13)	15 (33)	
Presence of severe pain, according to subjective pain scores, Yes (pain score,[a] 4-5), n (%)	16 (33)	31 (67)	<.001
Achievement of subjective symptom improvement, Yes (improved, slightly improved), n (%)	36 (75)	3 (7)	<.001

[a]Subjective rating of the degree of pain on a scale of 1 to 5.

We performed logistic regression analyses to evaluate the differences in symptoms (neck/shoulder, low back) between the baseline and at 12 weeks after adjusting for age and sex. The difference in the worst pain scores of neck/shoulder pain/stiffness and low back pain between baseline and 12 weeks was –1.12 (95% CI –1.53 to –0.70; *P*<.001) (Table 3). We also examined the OR of the outcomes (Table 3). The participants in the intervention group showed significant improvements in the severity of the neck/shoulder pain/stiffness and low back pain compared to those in the control group (OR 6.36, 95% CI 2.57-15.73; *P*<.001). The OR of the subjective improvement in symptoms at 12 weeks was 43.00 (95% CI 11.25-164.28; *P*<.001).

Table 3. The outcomes after the intervention.

Outcome, group	Odds ratio (95% CI)	*P* value
Difference in the worst pain scores[a] between baseline and 12 weeks		<.001
Control	Reference	
Intervention	–1.12[b] (–1.53 to –0.70)	
Absence of severe[c] pain according to subjective pain scores		<.001
Control	Reference	
Intervention	6.36[d] (2.57 to 15.73)	
Achievement of subjective symptom improvement		<.001
Control	Reference	
Intervention	43.00[d] (11.25 to 164.28)	

[a]Subjective rating of the degree of pain on a scale of 1 to 5.

[b]The estimate represents a regression coefficient.

[c]A score of 4 or more indicates severe pain.

[d]The estimate represents an odds ratio from a logistic regression model.

Discussion

Overview

This RCT showed that the 12-week use of an AI-assisted health program that provides 1 short exercise routine per day significantly improved (1) the subjective symptoms of both neck and shoulder pain/stiffness and low back pain after 12 weeks compared to those at baseline and (2) the subjective assessment of improvement following the 12-week intervention in the intervention group compared to that in the control group. In this study, the intervention group showed a high adherence of 92% (44/48) for the whole intervention period as we hypothesized that this digital intervention would support participants to continue exercising and enhance their adherence to the exercises, resulting in greater improvement in symptoms. High adherence is associated with improvement of the symptoms [22,26,27]. The results seemed to be largely attributed to the high adherence of the participants. Previous reports have shown that adherence to home-based exercise was at most 50%-70% [24,26].

This program was designed to improve exercise adherence. This program has been mainly implemented with both instructions and reminder functions. The chatbot sent a message with the exercise instructions and a corresponding illustration at a fixed time every single day, which also functioned as a reminder feature that sent instructions at a fixed time each day. Some studies have shown that social support is essential for enhancing adherence to home-based exercise [26,50]. Concretely, feedback from experts to patients, interactions between patients and caregivers, and supervision from experts improve adherence [24,51]. Digital interventions were more effective with human support [52]. In addition, further support such as phone calls, email reminders, and text messages was used to promote engagement in digital interventions [28-33]. Email reminders increased exercise adherence [53,54]. Interaction through the chatbot app of this program may have also given participants a sense of support. Although the participants were not supervised or provided feedback by the experts, the service provided both good and bad examples of how the exercise was supposed to be performed. Besides, they were motivated from time to time to continue with the exercise every day. Therefore, the participants might have felt supervised. In that respect, the program also had a monitoring function. The positive influence of these functions was consistent with the fact that the roles of human caregivers were essential for patients to continue with prescribed exercises. The beneficial features of this program were that we used a chat app, LINE, and the daily exercise time was the shortest used to date.

LINE is the most commonly used chat app in Japan and is similar to Facebook messenger, WhatsApp in the United States and Europe, and WeChat in China. This app allows users to send messages with emoji and stickers to individuals or groups and make voice or video calls. It is used mainly in Japan, Taiwan, and parts of Southeast Asia, which has approximately 86 million users (2020) [55]. Many users look at the app once they receive a message automatically. Therefore, when the participants received a message from the chatbot, they checked the messages immediately, which led them to continue the program. The high rate of adherence to the program in this intervention may also be attributed to the simple and extremely short exercise offered in this program compared to that offered in previous studies. In this study, daily exercises could be performed within 1 minute per day, and their effects were evident within a short time. Exercising for an average duration of 10 minutes per day significantly improved low back pain [15]. Exercises performed for 10-15 minutes each day 3-5 times a week were effective for stiff shoulders [16,18]. The duration of use of a mobile app to individually manage chronic neck and back pain was recommended for 20-30 minutes per day in another study [56].

We performed this intervention for only 12 weeks, which is considered a reasonable period of time in terms of the effects on the musculoskeletal symptoms, adherence to the program, and follow-up of the participants. Long-term improvement in musculoskeletal symptoms can be expected with continued use of the program, but the adherence to the program could decline in the population. If the exercise is used only during the intervention period, the symptoms may gradually worsen again unless the patients continue the same exercises by themselves even after the end of the interventional period. According to several systematic reviews [22,28,34], the duration of the intervention ranges from 3 weeks to 12 months, but the relationship between the duration of the intervention and the outcome is not clear [28]. Although a few studies have mentioned the long-term effects of exercises for musculoskeletal symptoms, the exercises for low back pain combined with the program aimed at keeping motivation showed significant improvements in pain and function at 5-year follow-ups and about 60% adherence for both the intervention and the control groups [57].

The target population for this study was workers at the same company. They were usually engaged in jobs that required the use of computers and were accustomed to using smartphones on a daily basis. In other words, they had high computer/internet literacy. The program using LINE was highly compatible and it was easy to continue for them. Since this intervention was conducted as a part of the company's health promotion activities, they were dedicated to adhere to and continue the program even though there was no requirement to do so. Hence, it is understandable that the results may differ depending on the intervention population such as community residents, older adults, and outpatients.

A secondary effect of introducing mobile health and eHealth into health promotion activities is that people can easily access and utilize evidence-based health information for their own health care and, in other words, they can improve their own health literacy [58]. In the case of workers, mobile health can be used as an opportunity for those who have difficulty paying attention to their own health management owing to their busy schedules or other reasons. We are not sure which elements of this intervention, including the exercises themselves, were clearly effective. Although we used a specific program in this study, apps or internet services with similar functions such as sending easy and simple instructions every day at a fixed time may be effective in relieving musculoskeletal symptoms.

Limitations

There are several limitations in this study. First, no information was collected on the possible causes and diagnosis of neck/shoulder pain/stiffness and low back pain. The status of treatment such as the use of analgesics was also unknown. The improvement in pain observed in the intervention group may be due not only to the intervention but also to the changes in the treatment and occupational factors. However, since this study is an RCT, we believe that such effects are not necessarily crucial. Second, 13 participants in the intervention group and 14 in the control group dropped out after being randomly allotted before the start of the program. Hence, we could not even obtain their baseline data. Since we performed this intervention as a part of the health promotion activities at workplace during working hours, we needed to provide them the schedule for the introduction of the briefing session of the study in advance to have them participate in it. This matter forced us to randomize the participants as soon as they had enrolled as participants and had checked their eligibility to participate in this study, which resulted in dropouts prior to participation in the intervention. Further, some individuals could not participate the program because it did not align with their schedules; such withdrawal of participation was inevitable. Third, the definition of adherence differs in literature and is often not clearly defined. Although some studies have relied on participants' self-reporting [26], we could estimate the adherence by the access history of the program, as this study was a digital intervention. However, we could only know if they had accessed the program but we could not know if they had done the exercises. It is important to note that the adherence rate is different from the rate of exercising. However, we believe that many of the participants were able to do the exercises because the results showed improvement in the musculoskeletal symptoms. Fourth, this study was unblinded because of the characteristics of the intervention. The influence of the Hawthorne effect should be considered. Although we also provided the AI-assisted health program to the control group after completing the intervention for the intervention group and we did not reveal to the intervention group that they were the intervention group, it cannot be denied that the population of the intervention group could infer that they had been selected and should be able to improve because they were doing the exercises every day. Finally, regarding the outcome, we only adopted the subjective symptoms of neck/shoulder pain/stiffness and low back pain. The subjective symptoms are highly variable depending on their condition. We recruited the target papulation as those who had the symptoms above a certain level in the health survey in the past, but some participants had weak subjective symptoms in the baseline survey. We did not exclude such participants in this study.

Conclusions

This RCT showed that an intervention with simple and short exercises provided by an AI-assisted health program via participants' mobile phones' text messaging app for 12 weeks improved both neck and shoulder stiffness and low back pain. Digital health programs could help busy workers continue with their exercise routines easily without the need for frequent and direct contact with medical professionals. Further studies are needed to identify the elements of the AI-assisted health program that worked.

Acknowledgments

We deeply thank Professor Yoshihisa Fujino, MD, MPH, PhD at University of Occupational and Environmental Health, Japan for his cooperation in data analyses. We would like to thank Editage (www.editage.com) for English language editing. This study was supported by a grant from the Ministry of Health, Labor, and Welfare (H30-rodo-ippan-008) and the 2018 Japan Full Happ Survey Research (Public Interest Incorporated Foundation Japan Small and Medium Enterprise Welfare Business Foundation). The funding agencies had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' Contributions

SK, HO, TF, and TA designed the whole study. SK and TA managed the intervention. K Matsudaira and KK provided the intervention program. TA prepared the manuscript. K Mori supervised the study process.

Conflicts of Interest

The authors declare the following potential conflicts of interest. Ko Matsudaira is a shareholder/advisor of Trunk Solution Co, Ltd, who has a patent for the artificial intelligence (AI)-assisted health program used in this study and received the following support: a research grant from the Ministry of Health, Labor and Welfare for the submitted work; grant support from Sumitomo Dainippon Pharma Co, Ltd, and Okamura Corporation; grant support, including lecture fees, from Ayumi Pharmaceutical Corporation, Nippon Zoki Pharmaceutical Co, Ltd, Ono Pharmaceutical Co, Ltd, Eli Lilly Japan KK, Astellas Pharma Inc, TOTO Ltd, and Eisai Co, Ltd; lecture fees from Pfizer Japan Inc, Hisamitsu Pharmaceutical Co, Inc, Janssen Pharmaceutical KK, Kaken Pharmaceutical Co, Ltd, and Teijin Pharma Limited; lecture fees and advisory fees from Shionogi & Co, Ltd, MTG Co, Ltd, Sompo Holdings, Inc, NuVasive Japan, Murata Manufacturing Co, Ltd; grants and personal fees from The Association for Preventive Medicine of Japan, Shionogi & Co, Ltd, Nippon Zoki Pharmaceutical Co, Ltd, Ono Pharmaceutical Co, Ltd, Ayumi Pharmaceutical Corporation; grants from DeNA, Murata Manufacturing Co, Ltd, Chugai Pharmaceutical Co, Ltd, Sompo Holdings, Inc, MS&AD InterRisk Research & Consulting, Inc, NuVasive Japan, Medical Data Scientist, and Medical AI Device Development Organization, Inotech Co, Ltd; and personal fees from Eli Lilly Japan KK, Pfizer Japan Inc, and Hisamitsu Pharmaceutical Co,

Inc, outside of the submitted work. Hiroyuki Oka reports personal fees received from Ayumi Pharmaceutical Corporation, Nippon Zoki Pharmaceutical Co, Ltd, Ono Pharmaceutical Co, Ltd, Sompo Holdings, Inc, NuVasive Japan, Inc, Promotion of Practical Use of AI Medical Diagnosis Support Equipment, MS&AD InterRisk Research & Consulting, Inc, Inotech Corporation, Chugai Pharmaceutical Co, Ltd, The Association for Preventive Medicine of Japan, Shionogi & Co, Ltd, MTG Co, Ltd, and grants from Pfizer Inc outside of the submitted work.

Multimedia Appendix 1
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 1245 KB - mhealth_v9i9e27535_app1.pdf]

References

1. Coggon D, Ntani G, Palmer KT, Felli VE, Harari R, Barrero LH, et al. The CUPID (Cultural and Psychosocial Influences on Disability) study: methods of data collection and characteristics of study sample. PLoS One 2012;7(7):e39820 [FREE Full text] [doi: 10.1371/journal.pone.0039820] [Medline: 22792189]
2. Sarquis LMM, Coggon D, Ntani G, Walker-Bone K, Palmer KT, Felli VE, et al. Classification of neck/shoulder pain in epidemiological research: a comparison of personal and occupational characteristics, disability, and prognosis among 12,195 workers from 18 countries. Pain 2016 May;157(5):1028-1036 [FREE Full text] [doi: 10.1097/j.pain.0000000000000477] [Medline: 26761390]
3. Comprehensive survey of living condition in 2015. Ministry of Health, Labour and Welfare, Japan. URL: https://www.mhlw.go.jp/toukei/saikin/hw/k-tyosa/k-tyosa16/dl/16.pdf [accessed 2021-01-16]
4. Boles M, Pelletier B, Lynch W. The relationship between health risks and work productivity. J Occup Environ Med 2004 Jul;46(7):737-745. [doi: 10.1097/01.jom.0000131830.45744.97] [Medline: 15247814]
5. Nagata T, Mori K, Ohtani M, Nagata M, Kajiki S, Fujino Y, et al. Total Health-Related Costs Due to Absenteeism, Presenteeism, and Medical and Pharmaceutical Expenses in Japanese Employers. Journal of Occupational and Environmental Medicine 2018;60(5):e273-e280. [doi: 10.1097/jom.0000000000001291]
6. Fujii T, Matsudaira K, Yoshimura N, Hirai M, Tanaka S. Associations between neck and shoulder discomfort (Katakori) and job demand, job control, and worksite support. Modern Rheumatology 2014 Jan 14;23(6):1198-1204. [doi: 10.3109/s10165-012-0824-5]
7. Strøm V, Røe C, Knardahl S. Work-induced pain, trapezius blood flux, and muscle activity in workers with chronic shoulder and neck pain. Pain 2009 Jul;144(1-2):147-155. [doi: 10.1016/j.pain.2009.04.002] [Medline: 19410368]
8. Shiro Y, Arai YP, Matsubara T, Isogai S, Ushida T. Effect of muscle load tasks with maximal isometric contractions on oxygenation of the trapezius muscle and sympathetic nervous activity in females with chronic neck and shoulder pain. BMC Musculoskelet Disord 2012 Aug 13;13:146 [FREE Full text] [doi: 10.1186/1471-2474-13-146] [Medline: 22889146]
9. Eriksen W. Linking work factors to neck myalgia: the nitric oxide/oxygen ratio hypothesis. Med Hypotheses 2004;62(5):721-726. [doi: 10.1016/j.mehy.2003.12.015] [Medline: 15082096]
10. Elfering A, Müller U, Rolli Salathé C, Tamcan O, Mannion AF. Pessimistic back beliefs and lack of exercise: a longitudinal risk study in relation to shoulder, neck, and back pain. Psychol Health Med 2015;20(7):767-780. [doi: 10.1080/13548506.2015.1017824] [Medline: 25726859]
11. Koes BW, van Tulder MW, Thomas S. Diagnosis and treatment of low back pain. BMJ 2006 Jun 17;332(7555):1430-1434 [FREE Full text] [doi: 10.1136/bmj.332.7555.1430] [Medline: 16777886]
12. Babatunde OO, Jordan JL, Van der Windt DA, Hill JC, Foster NE, Protheroe J. Effective treatment options for musculoskeletal pain in primary care: A systematic overview of current evidence. PLoS One 2017;12(6):e0178621 [FREE Full text] [doi: 10.1371/journal.pone.0178621] [Medline: 28640822]
13. Green S, Buchbinder R, Hetrick S. Physiotherapy interventions for shoulder pain. Cochrane Database Syst Rev 2003(2):CD004258. [doi: 10.1002/14651858.CD004258] [Medline: 12804509]
14. Hayden J, van Tulder MW, Malmivaara A, Koes B. Exercise therapy for treatment of non-specific low back pain. Cochrane Database Syst Rev 2005 Jul 20(3):CD000335. [doi: 10.1002/14651858.CD000335.pub2] [Medline: 16034851]
15. Bell JA, Burnett A. Exercise for the primary, secondary and tertiary prevention of low back pain in the workplace: a systematic review. J Occup Rehabil 2009 Mar;19(1):8-24. [doi: 10.1007/s10926-009-9164-5] [Medline: 19219537]
16. Shariat A, Cleland JA, Danaee M, Kargarfard M, Sangelaji B, Tamrin SBM. Effects of stretching exercise training and ergonomic modifications on musculoskeletal discomforts of office workers: a randomized controlled trial. Braz J Phys Ther 2018;22(2):144-153 [FREE Full text] [doi: 10.1016/j.bjpt.2017.09.003] [Medline: 28939263]
17. Tunwattanapong P, Kongkasuwan R, Kuptniratsaikul V. The effectiveness of a neck and shoulder stretching exercise program among office workers with neck pain: a randomized controlled trial. Clin Rehabil 2016 Jan;30(1):64-72. [doi: 10.1177/0269215515575747] [Medline: 25780258]
18. Ylinen J, Kautiainen H, Wirén K, Häkkinen A. Stretching exercises vs manual therapy in treatment of chronic neck pain: a randomized, controlled cross-over trial. J Rehabil Med 2007 Mar;39(2):126-132 [FREE Full text] [doi: 10.2340/16501977-0015] [Medline: 17351694]

19. Wang X, Zheng J, Yu Z, Bi X, Lou S, Liu J, et al. A meta-analysis of core stability exercise versus general exercise for chronic low back pain. PLoS One 2012;7(12):e52082 [FREE Full text] [doi: 10.1371/journal.pone.0052082] [Medline: 23284879]
20. Khan M, Akhter S, Soomro RR, Ali SS. The effectiveness of Cognitive Behavioral Therapy (CBT) with general exercises versus general exercises alone in the management of chronic low back pain. Pak J Pharm Sci 2014 Jul;27(4 Suppl):1113-1116. [Medline: 25016276]
21. Veehof MM, Trompetter HR, Bohlmeijer ET, Schreurs KMG. Acceptance- and mindfulness-based interventions for the treatment of chronic pain: a meta-analytic review. Cogn Behav Ther 2016;45(1):5-31. [doi: 10.1080/16506073.2015.1098724] [Medline: 26818413]
22. Beinart NA, Goodchild CE, Weinman JA, Ayis S, Godfrey EL. Individual and intervention-related factors associated with adherence to home exercise in chronic low back pain: a systematic review. Spine J 2013 Dec;13(12):1940-1950. [doi: 10.1016/j.spinee.2013.08.027] [Medline: 24169445]
23. Essery R, Geraghty AWA, Kirby S, Yardley L. Predictors of adherence to home-based physical therapies: a systematic review. Disabil Rehabil 2017 Mar;39(6):519-534. [doi: 10.3109/09638288.2016.1153160] [Medline: 27097761]
24. Palazzo C, Klinger E, Dorner V, Kadri A, Thierry O, Boumenir Y, et al. Barriers to home-based exercise program adherence with chronic low back pain: Patient expectations regarding new technologies. Ann Phys Rehabil Med 2016 Apr;59(2):107-113 [FREE Full text] [doi: 10.1016/j.rehab.2016.01.009] [Medline: 27050664]
25. Jordan J, Holden M, Mason E, Foster N. Interventions to improve adherence to exercise for chronic musculoskeletal pain in adults. Cochrane Database Syst Rev 2010 Jan 20(1):CD005956 [FREE Full text] [doi: 10.1002/14651858.CD005956.pub2] [Medline: 20091582]
26. Argent R, Daly A, Caulfield B. Patient Involvement With Home-Based Exercise Programs: Can Connected Health Interventions Influence Adherence? JMIR Mhealth Uhealth 2018 Mar 01;6(3):e47 [FREE Full text] [doi: 10.2196/mhealth.8518] [Medline: 29496655]
27. Adherence to long-term therapies: evidence for action. World Health Organization. URL: http://apps.who.int/iris/bitstream/handle/10665/42682/9241545992.pdf;jsessionid=F1949BE01A3BF6EBD7C0F5B32849F837?sequence=1 [accessed 2021-01-16]
28. Hewitt S, Sephton R, Yeowell G. The Effectiveness of Digital Health Interventions in the Management of Musculoskeletal Conditions: Systematic Literature Review. J Med Internet Res 2020 Jun 05;22(6):e15617 [FREE Full text] [doi: 10.2196/15617] [Medline: 32501277]
29. Buhrman M, Nilsson-Ihrfeldt E, Jannert M, Ström L, Andersson G. Guided internet-based cognitive behavioural treatment for chronic back pain reduces pain catastrophizing: a randomized controlled trial. J Rehabil Med 2011 May;43(6):500-505 [FREE Full text] [doi: 10.2340/16501977-0805] [Medline: 21533329]
30. del Pozo-Cruz B, Parraca J, del Pozo-Cruz J, Adsuar J, Hill J, Gusi N. An occupational, internet-based intervention to prevent chronicity in subacute lower back pain: a randomised controlled trial. J Rehabil Med 2012 Jun;44(7):581-587 [FREE Full text] [doi: 10.2340/16501977-0988] [Medline: 22674240]
31. Del Pozo-Cruz B, Adsuar JC, Parraca J, Del Pozo-Cruz J, Moreno A, Gusi N. A web-based intervention to improve and prevent low back pain among office workers: a randomized controlled trial. J Orthop Sports Phys Ther 2012 Oct;42(10):831-841. [doi: 10.2519/jospt.2012.3980] [Medline: 22951407]
32. Irvine AB, Russell H, Manocchia M, Mino DE, Cox Glassen T, Morgan R, et al. Mobile-Web app to self-manage low back pain: randomized controlled trial. J Med Internet Res 2015 Jan 02;17(1):e1 [FREE Full text] [doi: 10.2196/jmir.3130] [Medline: 25565416]
33. Petrozzi MJ, Leaver A, Ferreira PH, Rubinstein SM, Jones MK, Mackey MG. Addition of MoodGYM to physical treatments for chronic low back pain: A randomized controlled trial. Chiropr Man Therap 2019;27:54 [FREE Full text] [doi: 10.1186/s12998-019-0277-4] [Medline: 31673330]
34. Nicholl BI, Sandal LF, Stochkendahl MJ, McCallum M, Suresh N, Vasseljen O, et al. Digital Support Interventions for the Self-Management of Low Back Pain: A Systematic Review. J Med Internet Res 2017 May 21;19(5):e179 [FREE Full text] [doi: 10.2196/jmir.7290] [Medline: 28550009]
35. Garg S, Garg D, Turin TC, Chowdhury MFU. Web-Based Interventions for Chronic Back Pain: A Systematic Review. J Med Internet Res 2016 Jul 26;18(7):e139 [FREE Full text] [doi: 10.2196/jmir.4932] [Medline: 27460413]
36. Shebib R, Bailey JF, Smittenaar P, Perez DA, Mecklenburg G, Hunter S. Randomized controlled trial of a 12-week digital care program in improving low back pain. NPJ Digit Med 2019;2:1 [FREE Full text] [doi: 10.1038/s41746-018-0076-7] [Medline: 31304351]
37. Toelle TR, Utpadel-Fischler DA, Haas K, Priebe JA. App-based multidisciplinary back pain treatment versus combined physiotherapy plus online education: a randomized controlled trial. NPJ Digit Med 2019;2:34 [FREE Full text] [doi: 10.1038/s41746-019-0109-x] [Medline: 31304380]
38. van den Heuvel SG, de Looze MP, Hildebrandt VH, Thé KH. Effects of software programs stimulating regular breaks and exercises on work-related neck and upper-limb disorders. Scand J Work Environ Health 2003 Apr;29(2):106-116 [FREE Full text] [doi: 10.5271/sjweh.712] [Medline: 12718496]

39. Wallace LS, Dhingra LK. A systematic review of smartphone applications for chronic pain available for download in the United States. J Opioid Manag 2014;10(1):63-68. [doi: 10.5055/jom.2014.0193] [Medline: 24604571]
40. Conway CM, Kelechi TJ. Digital Health for Medication Adherence in Adult Diabetes or Hypertension: An Integrative Review. JMIR Diabetes 2017 Aug 16;2(2):e20 [FREE Full text] [doi: 10.2196/diabetes.8030] [Medline: 30291093]
41. Patel N, Jones P, Adamson V, Spiteri M, Kinmond K. Chronic Obstructive Pulmonary Disease Patients' Experiences of an Enhanced Self-Management Model of Care. Qual Health Res 2016 Mar;26(4):568-577. [doi: 10.1177/1049732315573013] [Medline: 25711841]
42. Morrison D, Mair FS, Chaudhuri R, McGee-Lennon M, Thomas M, Thomson NC, et al. Details of development of the resource for adults with asthma in the RAISIN (randomized trial of an asthma internet self-management intervention) study. BMC Med Inform Decis Mak 2015 Jul 28;15:57 [FREE Full text] [doi: 10.1186/s12911-015-0177-z] [Medline: 26215651]
43. Wakefield BJ, Holman JE, Ray A, Scherubel M, Adams MR, Hillis SL, et al. Effectiveness of home telehealth in comorbid diabetes and hypertension: a randomized, controlled trial. Telemed J E Health 2011 May;17(4):254-261. [doi: 10.1089/tmj.2010.0176] [Medline: 21476945]
44. Stiff shoulders, back pain, SNS remote presenteeism countermeasures. Secaide.me. URL: https://www.secaide.me/ [accessed 2021-01-16]
45. Matsudaira K, Hiroe M, Kikkawa M, Sawada T, Suzuki M, Isomura T, et al. Can standing back extension exercise improve or prevent low back pain in Japanese care workers? Journal of Manual & Manipulative Therapy 2015 Jan 04;23(4):205-209. [doi: 10.1179/2042618614y.0000000100]
46. Tonosu J, Matsudaira K, Oka H, Okazaki H, Oshio T, Hanaoka I, et al. A population approach to analyze the effectiveness of a back extension exercise "One Stretch" in patients with low back pain: A replication study. J Orthop Sci 2016 Jul;21(4):414-418. [doi: 10.1016/j.jos.2016.03.002] [Medline: 27053155]
47. Oka H, Nomura T, Asada F, Takano K, Nitta Y, Uchima Y, et al. The effect of the 'One Stretch' exercise on the improvement of low back pain in Japanese nurses: A large-scale, randomized, controlled trial. Mod Rheumatol 2019 Sep;29(5):861-866. [doi: 10.1080/14397595.2018.1514998] [Medline: 30130991]
48. Jinnouchi H, Matsudaira K, Kitamura A, Kakihana H, Oka H, Hayama-Terada M, et al. Effects of Low-Dose Therapist-Led Self-Exercise Education on the Management of Chronic Low Back Pain: Protocol for a Community-Based, Randomized, 6-Month Parallel-Group Study. Spine Surg Relat Res 2019;3(4):377-384 [FREE Full text] [doi: 10.22603/ssrr.2019-0005] [Medline: 31768459]
49. Jinnouchi H, Matsudaira K, Kitamura A, Kakihana H, Oka H, Hayama-Terada M, CIRCS Investigators. Effects of brief self-exercise education on the management of chronic low back pain: A community-based, randomized, parallel-group pragmatic trial. Mod Rheumatol 2021 Jul;31(4):890-898. [doi: 10.1080/14397595.2020.1823603] [Medline: 32930621]
50. Brewer BW, Cornelius AE, Van Raalte JL, Petitpas AJ, Sklar JH, Pohlman MH, et al. Age-Related Differences in Predictors of Adherence to Rehabilitation After Anterior Cruciate Ligament Reconstruction. J Athl Train 2003 Jun;38(2):158-162 [FREE Full text] [Medline: 12937528]
51. Friedrich M, Cermak T, Maderbacher P. The effect of brochure use versus therapist teaching on patients performing therapeutic exercise and on changes in impairment status. Phys Ther 1996;76(10):1082-1088. [doi: 10.1093/ptj/76.10.1082]
52. Mohr DC, Cuijpers P, Lehman K. Supportive accountability: a model for providing human support to enhance adherence to eHealth interventions. J Med Internet Res 2011 Mar 10;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
53. Yap TL, Davis LS, Gates DM, Hemmings AB, Pan W. The Effect of Tailored E-mails in the Workplace. AAOHN Journal 2009 Jul 01;57(7):267-273. [doi: 10.1177/216507990905700704]
54. Yap TL, Davis LS, Gates DM, Hemmings AB, Pan W. The Effect of Tailored E-mails in the Workplace. AAOHN Journal 2009 Aug 01;57(8):313-319. [doi: 10.1177/216507990905700805]
55. LINE revenue and usage statistics (2020). Business of Apps. URL: https://www.businessofapps.com/data/line-statistics/ [accessed 2021-04-28]
56. Lo WLA, Lei D, Li L, Huang DF, Tong K. The Perceived Benefits of an Artificial Intelligence-Embedded Mobile App Implementing Evidence-Based Guidelines for the Self-Management of Chronic Neck and Back Pain: Observational Study. JMIR Mhealth Uhealth 2018 Nov 26;6(11):e198 [FREE Full text] [doi: 10.2196/mhealth.8127] [Medline: 30478019]
57. Friedrich M, Gittler G, Arendasy M, Friedrich KM. Long-term effect of a combined exercise and motivational program on the level of disability of patients with chronic low back pain. Spine (Phila Pa 1976) 2005 May 01;30(9):995-1000. [doi: 10.1097/01.brs.0000160844.71551.af] [Medline: 15864148]
58. Lotto M, Strieder AP, Ayala Aguirre PE, Oliveira TM, Andrade Moreira Machado MA, Rios D, et al. Parental-oriented educational mobile messages to aid in the control of early childhood caries in low socioeconomic children: A randomized controlled trial. J Dent 2020 Oct;101:103456 [FREE Full text] [doi: 10.1016/j.jdent.2020.103456] [Medline: 32827598]

Abbreviations

AI: artificial intelligence
CONSORT: Consolidated Standards of Reporting Trials

OR: odds ratio
RCT: randomized controlled trial

Edited by L Buis; submitted 27.01.21; peer-reviewed by S Kanamori, L Sandal, M Lotto; comments to author 19.03.21; revised version received 14.05.21; accepted 30.07.21; published 24.09.21.

Please cite as:
Anan T, Kajiki S, Oka H, Fujii T, Kawamata K, Mori K, Matsudaira K
Effects of an Artificial Intelligence–Assisted Health Program on Workers With Neck/Shoulder Pain/Stiffness and Low Back Pain: Randomized Controlled Trial
JMIR Mhealth Uhealth 2021;9(9):e27535
URL: https://mhealth.jmir.org/2021/9/e27535
doi:10.2196/27535
PMID:34559054

Original Paper

mHealth Interventions for Lifestyle and Risk Factor Modification in Coronary Heart Disease: Randomized Controlled Trial

Jang-Whan Bae[1*], MD, PhD; Seoung-Il Woo[2*], MD, PhD; Joongyub Lee[3], MD, PhD; Sang-Don Park[2], MD, PhD; Sung Woo Kwon[2], MD; Seong Huan Choi[2], MD; Gwang-Seok Yoon[2], MD; Mi-Sook Kim[4], PhD; Seung-Sik Hwang[5], MD, PhD; Won Kyung Lee[6], MD, PhD

[1]Division of Cardiology, Department of Internal Medicine, Chungbuk National University Hospital, Chungbuk National University College of Medicine, Cheongju, Republic of Korea

[2]Department of Cardiology, School of Medicine, Inha University Hospital, Inha University, Incheon, Republic of Korea

[3]Department of Preventive Medicine, Seoul National University College of Medicine, Seoul, Republic of Korea

[4]Medical Research Collaborating Center, Seoul National University Hospital, Seoul, Republic of Korea

[5]Department of Public Health Sciences, Graduate School of Public Health, Seoul National University, Seoul, Republic of Korea

[6]Department of Prevention and Management, School of Medicine, Inha University Hospital, Inha University, Incheon, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Won Kyung Lee, MD, PhD
Department of Prevention and Management, School of Medicine
Inha University Hospital
Inha University
27 Inhang-Ro, Jung-Gu
Incheon
Republic of Korea
Phone: 82 10 6360 7965
Email: bluewhale65@gmail.com

Abstract

Background: Self-management of lifestyle and cardiovascular disease risk factors is challenging in older patients with coronary heart disease (CHD). SMS text messaging could be a potential support tool for self-management and the most affordable and accessible method through a mobile phone. High-quality evidence had been lacking, and previous studies evaluated the effects of SMS text messaging on the subjective measures of short-term outcomes. Recently, a large-sized randomized controlled trial in Australia reported promising findings on the objective measures upon 6-month follow-up. However, an examination of the effectiveness of such interventions in an Asian population with unique demographic characteristics would be worthwhile.

Objective: This study examined the effectiveness of a 1-way SMS text messaging program to modify the lifestyle and cardiovascular disease risk factors of patients who underwent the first percutaneous coronary intervention (PCI).

Methods: A parallel, single-blinded, 1:1 random allocation clinical trial was conducted with 879 patients treated through PCI. They were recruited during hospital admission from April 2017 to May 2020 at 2 university hospitals in the Republic of Korea. In addition to standard care, the intervention group received access to a supporting website and 4 SMS text messages per week for 6 months regarding a healthy diet, physical activity, smoking cessation, and cardiovascular health. Random allocation upon study enrollment and SMS text messaging after hospital discharge were performed automatically using a computer program. The coprimary outcomes were low-density-lipoprotein cholesterol (LDL-C), systolic blood pressure (SBP), and BMI. The secondary outcomes were change in lifestyle and adherence to the recommended health behaviors.

Results: Of the eligible population, 440 and 439 patients who underwent PCI were assigned to the intervention and control groups, respectively. The 1-way SMS text messaging program significantly enhanced physical activity (P=.02), healthy diet (P<.01), and medication adherence (P<.04) among patients with CHD. Hence, more people were likely to control their cardiovascular disease risk factors per the recommendations. The intervention group was more likely to control all 5 risk factors by 62% (relative risk 1.62, 95% CI 1.05-2.50) per the recommendations. On the other hand, physiological measures of the primary outcomes, including LDL-C levels, SBP, and BMI, were not significant. Most participants found the SMS text messaging program useful and helpful in motivating lifestyle changes.

Conclusions: Lifestyle-focused SMS text messages were effective in the self-management of a healthy diet, exercise, and medication adherence, but their influence on the physiological measures was not significant. One-way SMS text messages can be used as an affordable adjuvant method for lifestyle modification to help prevent the recurrence of cardiovascular disease.

Trial Registration: Clinical Research Information Service (CRiS) KCT0005087; https://cris.nih.go.kr/cris/search/detailSearch.do/19282

(JMIR Mhealth Uhealth 2021;9(9):e29928) doi:10.2196/29928

KEYWORDS

coronary heart disease; prevention; lifestyle modification; mobile health; text message; mHealth

Introduction

A healthy lifestyle, risk factor modification, and medication adherence are vital for preventing mortality and recurrent events in individuals with coronary heart disease (CHD). Previous studies have revealed the preventive effect of lifestyle modifications including smoking cessation, exercise, healthy diet, and exercise on mortality among patients with CHD. A Cochrane systematic review demonstrated that exercise-based cardiac rehabilitation reduced the all-cause mortality by 26% [1]. Another systematic review also reported that smoking cessation could reduce the risk of death and myocardial infarction in patients with CHD by 30% [2]. In the 2017 factsheet, the World Health Organization reported that approximately 75% of recurrent vascular events might be prevented when the patients were adherent to medications, such as aspirin, β-blockers, angiotensin-converting enzyme inhibitors (ACEis), and statin, and practices such as smoking cessation [3].

On the other hand, self-management of the cardiovascular disease risk factors is often challenging. A previous study reported that approximately one-third of patients with acute coronary syndrome persisted in smoking or were not adherent to the recommendations for diet or exercise [4]. Therefore, the development of tools to enhance self-management and evaluation of their effectiveness is crucial. Texting using a mobile phone has been suggested as a potential tool because it is affordable and accessible to older people with CHD [5]. Early research suggested the possible benefit of texting interventions on lifestyle modification [6]. On the other hand, a Cochrane review pointed out that evidence was not strong for such interventions because of the small number of participants and the risk of bias [7]. Moreover, previous studies were limited to the short-term consequence of texting interventions and the subjective measures on the outcomes [7-9]. Recently, Chow et al [10] reported encouraging findings in a large-scale RCT that a text message program could change the objective measure of cardiovascular disease risk factors and the self-reported lifestyle upon 6-month follow-up. On the other hand, it is unclear if such interventions would be effective in an Asian population with a different culture and lifestyle.

Therefore, this study examined whether an SMS text messaging program could enhance self-management of lifestyle and risk factor modification with objective and subjective measures in a randomized controlled trial.

Methods

Study Design

This study is a parallel, single-blind, randomized controlled trial that enrolled 879 patients with a 1:1 allocation. The care provider and outcome evaluator were blinded to the assignment. Demographic information was obtained using a questionnaire at baseline. The following objective measures of the coprimary outcomes were obtained at baseline and 6 months after enrollment: low-density-lipoprotein cholesterol (LDL-C) levels, systolic blood pressure (SBP), and BMI. Subjective measures of the secondary outcomes included a self-report of physical activity (PA), diet, and medication adherence in the questionnaire.

The institutional review boards of the Inha University Hospital (IRB number 2017-03-008-001) and the Chungbuk National University Hospital (IRB number 2017-05-016) approved the study protocol, and informed consent was provided by patients who participated in this study. The protocol of this trial was registered retrospectively at the Clinical Research Information Service of the Republic of Korea (registration number KCT0005087).

Participants

Participants were eligible if they were diagnosed with CHD and underwent percutaneous coronary intervention (PCI) for the first time; those younger than 18 years were excluded. Initially, acute myocardial infarction was targeted as the inclusion criterion, but the criteria were extended to CHD, including angina pectoris, treated through PCI after the participating hospitals were confirmed at 1 month of recruitment. A cardiologist diagnosed the participants with CHD after coronary angiography and PCI during hospitalization. The patients were excluded if they had no mobile phone or difficulty reading SMS text messages. The demographic characteristics were evaluated during hospitalization. The income level was obtained through subjective assessment of household income. The participants selected 1 of the following five levels: low, middle-low, middle, middle-high, and high household income.

Participants were recruited after face-to-face assessment during hospital admission at 2 tertiary and university teaching hospitals in Chungcheongbuk-Do and Incheon, Republic of Korea. The study areas, Chungcheongbuk-Do and Incheon, had a population of 1,590,372 and 2,922,121 individuals, respectively, in 2020. Both hospitals had regional cardiocerebrovascular centers

(RCCVCs) established by the Ministry of Health and Welfare to prevent and treat cardiovascular disease [11].

Intervention

Access to a supporting website and SMS text messages regarding lifestyle modification were provided for 6 months to the intervention group. The contents of the SMS text messages were based on the Tobacco, Exercise, and Diet Messages (TEXTME) trial and the Australian Heart Foundation Healthy Living Guidelines [10,12]. The cardiologists, nurses, clinical nutritionists, and preventive medicine experts reviewed the text messages in the TEXTME trial and modified them considering the Asian diet and culture. The SMS text messages related to smoking cessation, diet, physical activity, and general cardiovascular health, including medication adherence. The number of messages was 25 and 27 for the category of smoking cessation and physical activity, respectively. The diet category consisted of 27 and 99 messages for vegetarian and nonvegetarian individuals, respectively. General cardiovascular health categories A and B included 24 messages for all participants and nonsmokers. Category A included 24 messages for heart health and medication adherence, and category B included 24 mixed messages regarding diet, physical activity, general heart health, and information on passive smoking. Category B was designed to prevent nonsmokers from receiving duplicate messages from category A. The message-sending program delivered semipersonalized text messages considering the smoking status and diet pattern of the participants—vegetarian or not—with their names.

Participants in the intervention group received 4 messages per week for 24 weeks in addition to standard care. The message-management program selected a message randomly from each of the following four categories for a week: smoking cessation, physical activity, diet, and general cardiovascular health A for smokers; and physical activity, diet, and general cardiovascular health A and B for nonsmokers. The messages were sent on 4 of 5 randomly selected weekdays and at randomly selected times (9 AM, 12 PM, 3 PM, and 5 PM) of the day. Two text messenger systems were used: the default text messenger of the mobile phone and a commercial messenger (Kakaotalk), which is the most popular messaging app in the Republic of Korea. Text message delivery was ensured by programming the system to send the message to the commercial messenger first and then to the default messenger if the text message could not be delivered. The text messaging program re-sent the message via the default text messenger. Every participant in the intervention group received 96 messages in the 6-month prevention program. The algorithm for message management was developed in accordance with the prespecified rule on selecting the categories and text messages and the frequency and timing of texting. The participants were not meant to reply to the messages because they were designed as a 1-way SMS text messaging program. The participants were told not to respond to the messages and were informed that the messages were managed using a computer program. On the other hand, the study personnel explained to all the participants that they could request to stop the SMS text messaging program through the caller's phone number when they wanted to withdraw.

The supporting websites were the homepages of Chungcheongbuk-Do and Incheon RCCVC [13,14]. They provided information on cardiovascular disease, including a healthy lifestyle and disease management. Nine common action plans were selected as a preventive lifestyle of cardiovascular disease by 14 RCCVCs: quit smoking, avoid heavy drinking, reduce salt intake, exercise regularly, maintain an ideal body weight, reduce stress in daily life, take a regular health examination, adhere to medical treatment, and call an emergency response system if there are symptoms of cardiovascular disease. These were reflected in the contents of the leaflets, booklets, infographics, and video clips on the supporting websites. The study participants in the intervention group received the website link on their mobile phones. Although the website was introduced and recommended to the intervention group, a visit to the website was not required or checked.

Outcomes

The coprimary outcome of the study was the LDL-C level, SBP, and BMI at 6 months, relative to baseline levels. Fasting lipid levels, systolic and diastolic blood pressure with the heart rate, and BMI were measured in accordance with international standardized procedures at baseline and at 6 months. Blood pressure and heart rate were measured using electronic devices (Exhomax plus HBP-1000, HuBDIC Healthcare Co). Three resting measurements in sitting position with digital recordings were taken, each at least 5 minutes apart, with the mean of the last 2 readings used for analyses. The body weight and height were determined using an automatic standardized scale (GBF-500, TransTek) and electronic height rod (BSM330, InBody) with an accuracy of 100 g and 1 mm, respectively. BMI was calculated by dividing the weight in kilograms by the square of the height in meters. Regarding obesity, a BMI of 25 kg/m^2 was used as a cut-off level in accordance with the Asia Pacific guidelines of obesity of the World Health Organization [15]. The lipid levels (LDL-C, high-density-lipoprotein cholesterol, and triglycerides) of the participants were measured from the fasting blood samples and are reported as mg/dL.

The secondary outcomes were healthy lifestyle at the 6-month follow-up: smoking cessation, PA, fruit and vegetable intake up to ≥2 times/day, and medication adherence. At the 6-month visit to the cardiology clinic, smoking status, PA, diet, and medication adherence were acquired using a questionnaire. PA was assessed using a shortened Korean version of the International Physical Activity Questionnaire (IPAQ), a commonly used tool to assess PA. The total hours per week for walking, moderate PA, and vigorous PA were computed using metabolic equivalent (MET) values [16]. As a secondary outcome, PA was reported as the median METs-min/week and the proportion of inactive PA. PA was categorized into active and inactive on the basis of the recommendations of the total level: ≥600 and <600 METs-min/week.

Past and current smoking statuses were assessed through a self-report, including the duration and quantity of smoking and the duration of smoking cessation. The proportion of current smokers among the participants in the intervention and control groups upon 6-month follow-up was compared to that at baseline.

The frequency of fruit and vegetable intake over the last 7 days was assessed in the questionnaire: "How many days or times in the last 7 days did you consume fruit or vegetables?" The participants could select 1 of seven answers: 0 times, 1-2 times/week, 3-4 times/week, 5-6 times/week, 1 time/day, 2 times/day, ≥3 times/day. They were reclassified into ≥2 times/day and <2 times/day in the analysis. Medication adherence was evaluated using the Modified Morisky Scale (MMS) and the number of days the medication was taken per month [17]. Medication adherence was defined as good if the participants took their medication as instructed on 25 or more days in the last month, corresponding to more than 80%. A 6-item MMS was used to assess medication-taking behavior: 2 questions were added to the validated 4-item Morisky Scale to explain the persistence of therapy [18]. The score ranged from 0 to 3 for every 2 domains (knowledge and motivation); the total score ranged from 0 to 6. As a secondary outcome, medication adherence was reported using the proportion of good medication adherence, and the median MMS score was also presented.

Moreover, the secondary outcomes included the proportion achieving the guideline levels of 5 modifiable risk factors: LDL-C levels of <70 mg/dL, blood pressure of <140/90 mmHg, 30 minutes of moderate exercise up to ≥5 days/week, smoking cessation, and BMI of <25 kg/m^2 [10,19]. The effects of the SMS text messaging program on the risk factor modification were evaluated by dividing the participants into those who met all 5 guideline levels and those who did not meet ≤4. Similarly, participants were also classified with the cut-off levels of any 4 and 3 guidelines achieved: ≥4 guideline level and <4 guideline level; ≥3 guideline level and <3 guideline level.

Program Evaluation

In this study, the intervention was a 1-way SMS text messaging program, including a supporting website, and the study participants were asked to not reply to the messages. The program was evaluated by administering a questionnaire to the intervention participants to assess the perceived utility, acceptability, and influence of the intervention on behavioral changes. The questionnaire also included questions on the reading, saving, and sharing of text messages, but it was not validated (Multimedia Appendix 1). These were given to participants at the outpatient clinic after the final outcome assessment, and 349 participants conducted a self-report. The participants reported the level agreement to the statements related to the usefulness and acceptability of the text messages using a 5-point Likert scale: 1=strongly disagree, 2=disagree, 3=neither agree nor disagree, 4=agree, and 5=strongly agree. For analysis, "agree" and "strongly agree" were reclassified as "agree," while all the other categories were classified as "disagree."

Sample Size

A sample size of 880 individuals was calculated considering a 15% loss to follow-up, 90% power (2-tailed; significance level=5%) to detect a difference in the 3 primary outcomes between the 2 groups: 10 mg/dL in LDL-C levels, 5 mmHg in SBP, and 1.2 kg/m^2 in BMI [10]. For the coprimary outcomes, no significance level adjustment was made to the sample size calculations to account for multiple comparisons.

Randomization

After a pilot study, a computerized randomization program was developed for a random 1:1 allocation whose sequence was generated in a block size of 8. A web-based interface was developed for computerized randomization and an SMS text message–sending program. When the study personnel entered the participant information in the secure web interface, the participants were assigned randomly to the intervention or control group. If they were assigned to the intervention and discharged from the hospital, the computer program automatically sent an SMS text message for 6 months after hospital discharge. The study personnel could not ask the participants about their allocation. The study participants were asked to not reveal their allocation at the 6-month follow-up visit to maintain blinding of the study personnel. Except for SMS text messages, both groups received guideline-directed standard care for CHD: treatment including medication, regular follow-up at the outpatient clinic, and education on cardiovascular health and risk factors provided by nurses specialized in patient education at RCCVCs.

Statistical Analysis

Statistical analysis was conducted by an independent biostatistician and an epidemiologist not involved in the study. All evaluations of the intervention were performed on the principle of an intention to treat. Subgroup analyses were specified in the statistical analysis plan by age, sex, smoking status at baseline, recruiting hospital, LDL-C tertile at baseline, and CHD category: acute myocardial infarction and angina pectoris. No interim analysis was planned or conducted.

Continuous variables at baseline are presented as mean (SD) values if they were distributed normally, while the median (IQR) values are used to describe the nonnormally distributed variables. Categorical variables are presented as n (%) values. The baseline characteristics between the intervention and the control groups were compared using an independent samples *t* test and chi-square test. The primary analysis was an analysis of covariance (ANCOVA) and robust Poisson regression to estimate the relative risk with the baseline values of the analyzed parameters, where continuous and binary outcomes were evaluated, respectively [20]. A randomization-based nonparametric ANCOVA was used when the continuous outcomes were not normally distributed.

Regarding the coprimary outcomes, the results were not statistically significant. An adjustment to account for multiple comparisons, such as a Bonferroni correction, was not needed to reduce the overall false-positive rate.

The analyses were conducted using SAS Enterprise Guide (version 7.4, SAS Institute). All statistical tests were 2-tailed, and P<.05 was considered significant.

Data availability

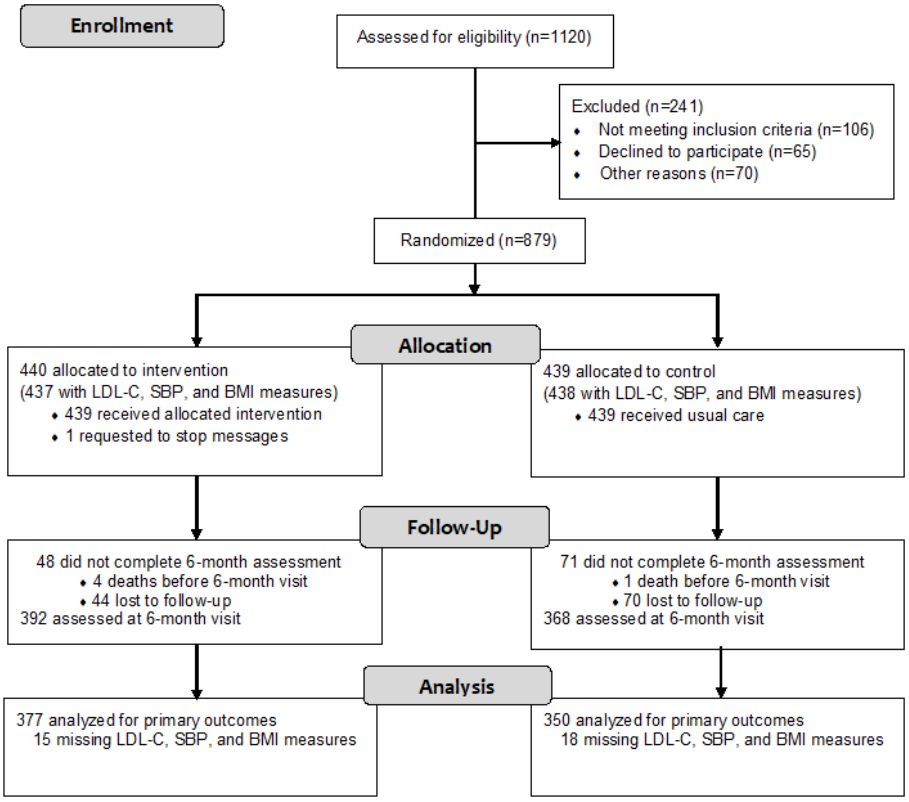
Data are available from the authors upon reasonable request.

Results

Results Overview

Of the 1120 patients screened for eligibility, 879 who were admitted with CHD and underwent PCI from April 2017 to May 2020 were enrolled in this study (Figure 1). In total, 241 patients were excluded: 28 patients who did not have a mobile phone, 78 patients who had difficulty reading SMS text messages, 65 patients who declined to participate in the study, and 70 patients who declined participation for other reasons, including cases of in-hospital mortality. Of the 879 eligible patients, 440 and 439 patients were assigned randomly to the intervention and control groups, respectively. In total, 48 patients were lost to follow-up, including 4 patients who died before the 6-month follow-up in the intervention group, while 71 patients in the control group were lost to follow-up, including 1 death. Recruitment was concluded when the study sample size was achieved, and the follow-up period was from October 2017 to November 2020.

Figure 1. Schematic representation of the randomized controlled trial. LDL-C: low-density-lipoprotein cholesterol, SBP: systolic blood pressure.



Baseline Characteristics of the Study Participants

Table 1 lists the baseline characteristics of the study participants. The mean age of the enrolled participants was 60.4 (SD 10.5) years, and 732 (83.3%) participants were male. The mean LDL-C level, SBP, and BMI were 110.3 mg/dL, 125.3 mmHg, and 24.9 kg/m^2, respectively. The baseline characteristics were similar in the intervention and control groups except for the ACEis/angiotensin receptor blockers (ARBs); the control group was more likely to take ACEis or ARBs before hospital admission.

Table 1. Baseline characteristics of the study participants (N=879).

Characteristics	Intervention (n=440)	Control (n=439)	Total
Demographics			
Age (years), mean (SD)	60.1 (10.6)	60.7 (10.4)	60.4 (10.5)
Males, n (%)	368 (83.6)	364 (82.9)	732 (83.3)
Participating institution, n (%)			
Inha university hospital	323 (73.4)	323 (73.6)	646 (73.5)
Chungbuk university hospital	117 (26.6)	116 (26.4)	233 (26.5)
Education level, n (%)			
Elementary school or less	51 (11.7)	48 (11.0)	99 (11.3)
Middle school	55 (12.6)	55 (12.6)	110 (12.6)
High school	205 (46.9)	207 (47.4)	412 (47.1)
University	102 (23.3)	99 (22.7)	201 (23.0)
Graduate school or higher	24 (5.5)	28 (6.4)	52 (6.0)
Income, n (%)			
Low	57 (13.0)	49 (11.2)	106 (12.1)
Middle-low	58 (13.3)	67 (15.3)	125 (14.3)
Middle	251 (57.4)	246 (56.3)	497 (56.9)
Middle-high	56 (12.8)	49 (11.2)	105 (12.0)
High	15 (3.4)	26 (6.0)	41 (4.7)
Clinical data			
Acute myocardial infarction, n (%)	219 (50.1)	212 (49.3)	431 (49.7)
BMI (kg/m^2), mean (SD)	25.0 (3.4)	24.9 (3.1)	24.9 (3.2)
Waist circumference (cm), median (IQR)	90 (85-96)	90 (85-95)	90 (85-96)
Hip circumference (cm), mean (SD)	94.2 (10.3)	94.0 (11.2)	94.1 (10.7)
Total cholesterol (mg/dL), median (IQR)	174 (145-209)	171 (145-201)	172 (145-205)
Low-density-lipoprotein cholesterol (mg/dL), mean (SD)	110.8 (41.2)	109.8 (38.1)	110.3 (39.7)
High-density-lipoprotein cholesterol (mg/dL), median (IQR)	43 (37-51)	43 (36-50)	43 (37-51)
Triglycerides (mg/dL), median (IQR)	124 (89-186)	125 (90-181)	125 (90-185)
Systolic blood pressure (mmHg), mean (SD)	124.4 (18.9)	126.1 (19.8)	125.3 (19.4)
Diastolic blood pressure (mmHg), mean (SD)	74.0 (12.4)	74.7 (13.0)	74.3 (12.7)
Heart rate (beats/min), mean (SD)	75.3 (12.3)	76.5 (12.8)	75.9 (12.6)
Risk factor level: low-density-lipoprotein cholesterol level ≥ 70 mg/dL, n (%)	358 (81.9)	372 (84.9)	730 (83.4)
Blood pressure, n (%)			
Systolic > 140 mmHg	90 (20.6)	75 (17.2)	165 (18.9)
Diastolic > 90 mmHg	43 (9.8)	33 (7.6)	76 (8.7)
BMI ≥ 25 kg/m^2, n (%)	200 (45.8)	205 (46.8)	405 (46.3)
Physical activity (metabolic equivalents-min/week), median (IQR)	0 (0-720)	0 (0-924)	0 (0-777)
Inactive (<600 metabolic equivalents-min/week), n (%)	312 (71.4)	293 (67.1)	605 (69.2)
Smoking status, n (%)			
Current	189 (43.3)	184 (42.1)	373 (42.7)
Former	145 (33.2)	146 (33.4)	291 (33.3)
Diabetes, n (%)	127 (29.0)	129 (29.5)	256 (29.2)

Characteristics	Intervention (n=440)	Control (n=439)	Total
Hypertension, n (%)	212 (48.4)	212 (48.4)	424 (48.4)
Achieving guideline levels, n (%)			
Low-density-lipoprotein cholesterol level < 70 mg/dL	79 (18.1)	66 (15.1)	145 (16.6)
Blood pressure < 140/90 mmHg	330 (75.5)	306 (69.9)	636 (72.7)
Exercising regularly[a]	125 (28.6)	144 (33.0)	269 (30.8)
Nonsmokers	248 (56.8)	253 (57.9)	501 (57.3)
BMI < 25 kg/m^2	237 (54.23)	233 (53.20)	470 (53.71)
All 5 key guideline levels	8 (1.8)	8 (1.8)	16 (1.8)
Achieving 4 of 5 key guideline levels	55 (12.6)	59 (13.5)	114 (13.0)
Medications before admission, n (%)			
Angiotensin-converting enzyme inhibitor/angiotensin receptor blocker	130 (29.7)	158 (36.1)	288 (32.9)
Aspirin	103 (23.5)	118 (26.9)	221 (25.2)
β-Blocker	48 (11.0)	44 (10.1)	92 (10.5)
Statin	117 (26.7)	122 (27.9)	239 (27.3)
All 4 medications	10 (2.3)	16 (3.7)	26 (3.0)

[a]Regular exercise involves more than 30 minutes of moderate exercise performed ≥5 days/week.

Study Outcomes

The SMS text messaging program yielded modest improvement in the primary outcomes, whereas it improved the PA, diet, and medication adherence significantly (Table 2). Although the LDL-C level tended to decrease in the intervention group, the differences between the 2 groups were not significant for any of the 3 outcomes: LDL-C levels, SBP, and BMI. On the other hand, PA was significantly higher in the intervention group than in the control group, by 220 (95% CI 36-404) min/week. Participants in the intervention group were more likely to eat fruit or vegetables frequently (≥2 times a day) and adhere to their medications (≥80%) compared to those in the control group.

Table 2. Primary and secondary outcomes upon 6-month follow-up.

Parameters	Values		Mean difference (95% CI)	*P* value
	Intervention (n=377)	Control (n=350)		
Primary outcomes				
Low-density-lipoprotein cholesterol level (mg/dL), mean (95% CI)	73.9 (70.8 to 77.0)	77.4 (74.2 to 80.6)	–3.6 (–8.0 to 0.9)	.12
Systolic blood pressure (mmHg), mean (95 CI)	126.6 (125.0 to 128.2)	128.4 (126.7 to 130.0)	–1.8 (–4.1 to 0.5)	.13
BMI (kg/m^2), mean (95% CI)	25.0 (24.9 to 25.2)	25.1 (25.0 to 25.3)	–0.09 (–0.3 to 0.1)	.41
Secondary outcomes[a]				
Current smokers, n (%)	79 (20.6)[b]	81 (23.0)[b]	0.98 (0.76 to 1.3)[c]	.86
Physical activity (metabolic equivalents-min/week), median (IQR)	693 (0 to 1386)	384 (0 to 1162)	220 (36 to 404)	.02
Inactive (<600 metabolic equivalents-min/week), n (%)	173 (45.2)[b]	208 (59.1)[b]	0.75 (0.66, 0.86)[c]	<.001
Fruit/vegetable intake ≥ 2 times/day[d], n (%)	89 (23.2)[b]	52 (14.8)[b]	1.5 (1.1 to 2.1)[c]	.006
Medication adherence > 80%, n (%)	376 (98.2)[b]	324 (92.1)[b]	1.1 (1.0 to 1.1)[c]	<.001
Modified Morisky Scale score[e], median (IQR)	5 (5 to 5)	5 (5 to 5)	0.07 (–0.03 to 0.16)	.19

[a]Regarding the secondary outcome, 383 and 352 participants in the intervention and control groups, respectively, responded to the follow-up survey.

[b]Unadjusted proportion of lifestyle in the control and intervention groups upon 6-month follow-up.

[c]Adjusted relative risk between the control and intervention groups.

[d]Denotes the frequency of fruit and vegetables consumed in the prior 7 days up to ≥2 times/day.

[e]The number of participants who responded to the Modified Morisky Scale upon 6-month follow-up were 380 and 352 in the intervention and control groups, respectively.

Guideline Levels Achieved

The participants in the intervention group were more likely to achieve a healthy lifestyle by following the recommended guidelines (Table 3). The participants were categorized as those who met all 5 guideline levels and those who did not meet ≤4. The intervention group was more likely to control all 5 risk factors than the control group (relative risk [RR] 1.62, 95% CI 1.05-2.50). If the outcomes were set as meeting ≥4 guidelines, the participants in the intervention group were more likely to achieve the guideline level on ≥4 risk factors than those in the control group (RR 1.27, 95% CI 1.03-1.56). On the other hand, there was no significant difference in each risk factor between the intervention and control groups, except PA.

Table 3. Achieving guideline levels of risk factors upon 6-month follow-up.

Parameters	Intervention (n=378), n (%)	Control (n=350), n (%)	Relative risk (95% CI)	*P* value
Low-density-lipoprotein cholesterol level < 70 mg/dL	192 (50.8)	160 (45.7)	1.10 (0.95-1.28)	.19
Blood pressure < 140/90 mmHg	297 (78.6)	264/350 (75.4)	1.04 (0.96-1.12)	.38
Exercising regularly[a]	210[b] (54.8)	144[b] (40.9)	1.37 (1.17-1.60)	<.001
Nonsmoker	304[b] (79.4)	271[b] (77.0)	1.02 (0.95-1.08)	.64
BMI < 25 kg/m^2	196 (51.9)	169 (48.3)	1.02 (0.91-1.16)	.71
Key guideline levels				
Achieving all 5	48 (12.7)	27 (7.7)	1.62 (1.05-2.50)	.03
Achieving ≥ 4	135 (35.7)	97 (27.8)	1.27 (1.03-1.56)	.02
Achieving ≥ 3	278 (73.5)	229 (65.6)	1.11 (1.01-1.22)	.03

[a]Regular exercise involves more than 30 minutes of moderate exercise up to ≥5 days/week.

[b]Regarding physical activity and nonsmoking, 383 and 352 participants in the intervention and control groups, respectively, responded to the follow-up survey.

Subgroup Analysis

Figures 2 and 3 show the results of subgroup analysis on the primary and secondary outcomes. Regarding the objective measures of the primary outcomes, the difference between the intervention and control groups was not significant (Figure 2). On the other hand, with 3 secondary outcomes showing a significant difference between the intervention and control groups, the SMS text messages were more likely to be effective among males, young adults, current smokers, or patients with acute myocardial infarction (Figure 3).

Figure 2. Subgroup analysis of the primary outcomes. LDL-C: low-density-lipoprotein cholesterol.

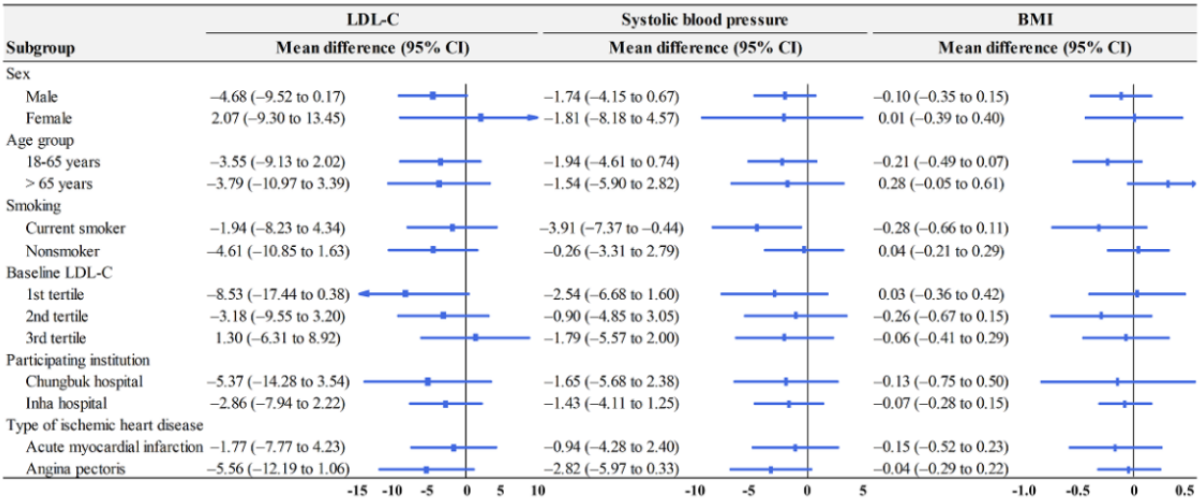
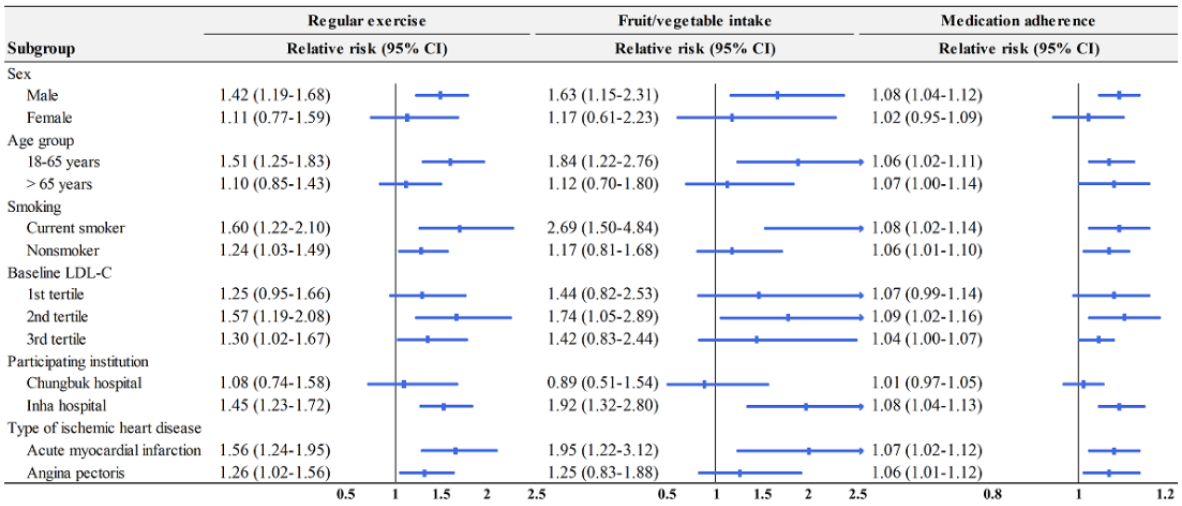


Figure 3. Subgroup analysis of the secondary outcomes: physical activity, fruit/vegetable intake, and medication adherence. LDL-C: low-density-lipoprotein cholesterol.



Program Evaluation

Among 440 participants in the intervention group, 349 completed the questionnaire on the utility and acceptability of the SMS text messaging program (Table 4). Most of the participants responded that the SMS text messages of the program were helpful (82.0%), easy to understand (94.6%), and a good motivation for changing their lifestyle (78.2%). Overall, they reported satisfaction with the frequency, time, and duration of the program. During the 6-month period, only 1 participant opted out of the SMS text messaging program. The cost of the SMS text messaging program was US $2.1 per person for regular SMS text messaging and US $0.5 per person using the commercial messenger app. The numbers of SMS text messages delivered through the default text messenger and the commercial text messenger were 20,658 (49%) and 21,502 (51%), respectively, from among 42,160 messages sent to the intervention group.

Table 4. Utility and perceived acceptability of the SMS text messaging intervention program by the participants (n=349).

Characteristics	Participants, n (%)
Usefulness and understanding	
Found messages useful	286 (82.0)
Messages were easy to understand	330 (94.6)
Influence on motivation and behavior change	
Messages motivated change	273 (78.2)
Diet was more healthy owing to the messages	219 (62.8)
Exercise increased owing to the messages	213 (61.0)
Messages reminded to take medication	184 (52.7)
Message saving and sharing	
Read at least 80% of messages	318 (91.1)
Saved messages	161 (46.1)
Shared messages with family, friends, or clinicians	126 (36.1)
Appropriate message characteristics	
Number of messages per week	291 (83.4)
Program length (6 months)	297 (85.1)
Time of the day when messages were received	303 (86.8)

Discussion

Principal Findings

The 1-way SMS text messaging program and a supporting website enhanced physical activity and encouraged a healthy diet and medication adherence among patients who underwent PCI. Therefore, more people were likely to follow the lifestyle and risk factor modification as recommended. On the other hand, the intervention could not induce a significant decrease in each objective measure of risk factors: LDL-C levels, SBP, and BMI. Most participants found the SMS text messaging program a helpful motivation to change their lifestyle.

Comparison With Prior Work

In this study, the SMS text messages contributed to lifestyle modification, but its impact on the physiological measures of risk factors may not be as much as those on lifestyle modification. These findings were inconsistent with those of the TEXTME study [10]—the clinical trial benchmarked by this study. The TEXTME study [10] reported significant improvement in all 3 objective measures. Regarding the objective measures, the mean differences between the intervention and control groups in this study were smaller than those in the TEXTME trial (LDL-C levels: 3.6 vs 5 mg/dL, SBP: 1.8 vs 7.6 mmHg, and BMI: 0.1 vs 1.3 kg/m^2). This may be owing to different profiles of the lifestyle and risk factors of the study participants between the 2 trials; the participants in this study were older Asian people with a lower BMI and SBP but a more sedentary life than those in the TEXTME trial [10]. Moreover, at our study sites, nurses provided education and counseling about cardiovascular disease to all the patients during hospital admission, regardless of the allocation. This may reduce the gap between the intervention and control groups. For example, the decreases in LDL-C levels were more significant in both groups in this study than in the TEXTME study [10] (73.9 and 77.4 mg/dL in the control and the intervention groups, respectively, from a baseline of 110 mg/dL in this study, and 79 and 84 mg/dL, respectively, from a baseline of 101 mg/dL in the TEXTME study). Similarly, the control group of this study was more likely to quit smoking than that of the TEXTME study. The proportion of smokers decreased from 42% to 23% in the control group in this study and from 54% to 44% in the TEXTME study. This may contribute to the difference in the effect size between this study and the TEXTME trial.

This study shows that the SMS text messaging program was effective in improving the PA, diet, and medication adherence in patients with CHD. Regarding lifestyle modification, these findings were supported by those of previous trials [6]. Pfaeffli et al [17] reported that in an RCT, SMS text messages and a supporting website effectively promoted healthy behavior of PA, diet, alcohol, and smoking. Khonsari et al [21] showed that a short, automated message enhanced medication adherence in patients with acute coronary syndrome. On the other hand, the differences in LDL-C levels, SBP, and BMI between the intervention and control groups were not significant. Similar findings to those of our study were found, but there are only a few reports on the objective measures to draw a conclusion. Maddison et al [22] reported that SMS text and video messages could increase PA in leisure and walking and self-efficacy but did not enhance the objective measure of the exercise capacity—peak oxygen uptake—upon 24-week follow-up in the patients with CHD. Pfaeffli et al [17] also failed to demonstrate statistical significance regarding the BMI, waist to hip ratio, blood pressure, and LDL-C levels. Similarly, SMS text messaging helped remind the patients to take their medication without skipping but was ineffective in helping

control SBP and LDL-C upon 6-month follow-up [23]. Therefore, further research will be needed to determine the effects of SMS text messaging programs on the physiological measures and clinical outcomes, such as recurrent cardiovascular events and mortality [21].

A method is needed to enhance the self-management of lifestyle and risk factors in patients with CHD. In this study, only a small proportion of patients who underwent PCI achieved the guideline level of all 5 risk factors: 12.7% for the intervention group and 7.7% for the control group. Although the SMS text messages improved PA and diet, many participants who underwent PCI still did not follow the guidelines. As much as 45.2% of the participants were considered physically inactive; only 23.3% of the participants ate fruit/vegetables up to ≥2 times per day in the intervention group. The corresponding proportion was 37.4% for physically inactive participants in the TEXTME trial [10]. Moreover, the proportion of patients who achieved all 5 guideline levels was worse in the TEXTME trial than in this study: 4.7% for the intervention group and 1.8% for the control group. This suggests that risk factor modification still has a long way to go, and various programs will be needed to improve lifestyle among patients with CHD.

SMS text messaging was chosen in this study because it was the most affordable and accessible among the elderly. It required lower cost and effort with and without an automated computer program than alternative methods. On the other hand, this study showed that its effect was not significant on the objective measures. Therefore, it is important to enhance the program by using interactive text messaging, personalized messaging, other smartphone apps, and wearable devices [24]. For example, previous review studies suggested that smartphone apps including goal-setting, self-monitoring of diet and activity, and feedback through SMS text messages could help lower calories, lower fat levels, increase PA, and reduce more weight in the general and individuals with obesity [8,25].

Limitations

This study had some limitations. First, those without literacy or a mobile phone were excluded, which may be an entry bias. Although SMS text messaging is the most affordable and accessible method using mobile Health (mHealth), there could still be a barrier to individuals without literacy or a mobile phone [26]. Second, this study was conducted using a single-blind design, even though a computer program was developed for random allocation. The study personnel who collected the data upon 6-month follow-up were blinded to the allocation. Therefore, the self-reported measures of PA, diet, and medication adherence may be biased because the participants could be subject to the desirability and expectation of the research. Although the evaluator was blinded to the allocation, caution should be taken when interpreting positive subjective outcomes and negative objective outcomes. Similarly, program evaluation may have been subject to information bias because the utility and perceived acceptability were subjective and could be influenced by expectations, even though it was evaluated as a self-report. Therefore, the utility and acceptability of the program may also have been overestimated. Furthermore, the questionnaire on lifestyle was not comprehensive. For example, although the frequency of fruit or vegetable intake was queried, the portion of fruit or vegetable intake was not evaluated.

Conclusions

This SMS text messaging program resulted in an improvement in self-reported lifestyle modifications such as PA, fruit and vegetable intake, and medication adherence among patients who required strict self-management after PCI. In contrast, its impact on physiological measures (LDL-C levels, SBP, and BMI) was not significant. One-way SMS text messaging may be used as an affordable adjuvant method for lifestyle modification to prevent the recurrence of cardiovascular disease. Because many patients still did not achieve guideline levels, future research will need to evaluate other interventions using mHealth tools (interactive messages, personalized messages, and wearable devices to facilitate the self-management of patient behaviors) with the objective measures of risk factor management.

Acknowledgments

WKL received funding for this study from the National Research Foundation of Korea (NRF) (NRF-2017R1C1B5017736). The funders had no role in study design, data collection, analysis, decision to publish, or manuscript preparation. We are grateful to Prof Clara Chow, who inspired us, and the George Institute, which provided the SMS text messages. The authors also wish to thank Sohyun Choi (statistician) and Bo Ram Yang (epidemiologist), who analyzed the data and made these findings sound and robust.

Authors' Contributions

JWB and SSH contributed to conceptualization and study ideas. JWB, SDP, SWK, GSY, SHC, and SIW contributed to data collection. MSK performed data visualization, and JL wrote the first draft of the manuscript. JWB, SDP, SWK, GY, SHC, MSK, and SIW revised the first draft of the manuscript. WKL is the principal investigator, contributed to the study idea and design, and wrote the first draft of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Program evaluation.
[DOCX File , 17 KB - mhealth_v9i9e29928_app1.docx]

Multimedia Appendix 2
CONSORT-eHEALTH checklist (V 1.6.1).
[PDF File (Adobe PDF File), 2905 KB - mhealth_v9i9e29928_app2.pdf]

References

1. Heran BS, Chen JM, Ebrahim S, Moxham T, Oldridge N, Rees K, et al. Exercise-based cardiac rehabilitation for coronary heart disease. Cochrane Database Syst Rev 2011 Jul 06(7):CD001800 [FREE Full text] [doi: 10.1002/14651858.CD001800.pub2] [Medline: 21735386]
2. Critchley J, Capewell S. Smoking cessation for the secondary prevention of coronary heart disease. Cochrane Database Syst Rev 2004(1):CD003041. [doi: 10.1002/14651858.CD003041.pub2] [Medline: 14974003]
3. Cardiovascular diseases (CVDs). World Health Organization. 2021. URL: https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds) [accessed 2021-09-07]
4. Chow CK, Jolly S, Rao-Melacini P, Fox KA, Anand SS, Yusuf S. Association of diet, exercise, and smoking modification with risk of early cardiovascular events after acute coronary syndromes. Circulation 2010 Feb 16;121(6):750-758. [doi: 10.1161/CIRCULATIONAHA.109.891523] [Medline: 20124123]
5. Beratarrechea A, Lee AG, Willner JM, Jahangir E, Ciapponi A, Rubinstein A. The impact of mobile health interventions on chronic disease outcomes in developing countries: a systematic review. Telemed J E Health 2014 Jan;20(1):75-82 [FREE Full text] [doi: 10.1089/tmj.2012.0328] [Medline: 24205809]
6. Pfaeffli Dale L, Dobson R, Whittaker R, Maddison R. The effectiveness of mobile-health behaviour change interventions for cardiovascular disease self-management: A systematic review. Eur J Prev Cardiol 2016 May;23(8):801-817. [doi: 10.1177/2047487315613462] [Medline: 26490093]
7. Vodopivec-Jamsek V, de Jongh T, Gurol-Urganci I, Atun R, Car J. Mobile phone messaging for preventive health care. Cochrane Database Syst Rev 2012 Dec 12;12:CD007457 [FREE Full text] [doi: 10.1002/14651858.CD007457.pub2] [Medline: 23235643]
8. Coughlin S, Whitehead M, Sheats J, Mastromonico J, Hardy D, Smith S. Smartphone applications for promoting healthy diet and nutrition: A literature review. Jacobs J Food Nutr 2015;2(3):021 [FREE Full text] [Medline: 26819969]
9. de Jongh T, Gurol-Urganci I, Vodopivec-Jamsek V, Car J, Atun R. Mobile phone messaging for facilitating self-management of long-term illnesses. Cochrane Database Syst Rev 2012 Dec 12;12:CD007459 [FREE Full text] [doi: 10.1002/14651858.CD007459.pub2] [Medline: 23235644]
10. Chow CK, Redfern J, Hillis GS, Thakkar J, Santo K, Hackett ML, et al. Effect of lifestyle-focused text messaging on risk factor modification in patients with coronary heart disease: A randomized clinical trial. JAMA 2015;314(12):1255-1263. [doi: 10.1001/jama.2015.10945] [Medline: 26393848]
11. Kim A, Yoon S, Kim Y, Kim EJ. The burden of acute myocardial infarction after a regional cardiovascular center project in Korea. Int J Qual Health Care 2015 Oct;27(5):349-355. [doi: 10.1093/intqhc/mzv064] [Medline: 26271544]
12. Redfern J, Thiagalingam A, Jan S, Whittaker R, Hackett ML, Mooney J, et al. Development of a set of mobile phone text messages designed for prevention of recurrent cardiovascular events. Eur J Prev Cardiol 2014 Apr;21(4):492-499. [doi: 10.1177/2047487312449416] [Medline: 22605787]
13. Chungcheongbuk-Do Regional Cardiocerebrovascular Center. URL: https://www.cbnuh.or.kr/ccvc/ [accessed 2021-06-27]
14. Incheon Regional Cardiocerebrovascular Center. URL: https://www.inha.com/site/iccvc/main [accessed 2021-06-27]
15. World Health Organization. Regional Office for the Western Pacific. The Asia-Pacific perspective: redefining obesity and its treatment. Sydney, NSW: Health Communications Australia; 2000.
16. Antypas K, Wangberg SC. An internet- and mobile-based tailored intervention to enhance maintenance of physical activity after cardiac rehabilitation: Short-term results of a randomized controlled trial. J Med Internet Res 2014 Mar 11;16(3):e77 [FREE Full text] [doi: 10.2196/jmir.3132] [Medline: 24618349]
17. Pfaeffli Dale L, Whittaker R, Jiang Y, Stewart R, Rolleston A, Maddison R. Text message and internet support for coronary heart disease self-management: Results from the Text4Heart randomized controlled trial. J Med Internet Res 2015 Oct 21;17(10):e237 [FREE Full text] [doi: 10.2196/jmir.4944] [Medline: 26490012]
18. Lee Y, Kim RB, Lee HJ, Kim K, Shin M, Park H, et al. Relationships among medication adherence, lifestyle modification, and health-related quality of life in patients with acute myocardial infarction: a cross-sectional study. Health Qual Life Outcomes 2018 May 22;16(1):100 [FREE Full text] [doi: 10.1186/s12955-018-0921-z] [Medline: 29788961]
19. Virani SS, Smith SC, Stone NJ, Grundy SM. Secondary prevention for atherosclerotic cardiovascular disease: Comparing recent US and European guidelines on dyslipidemia. Circulation 2020 Apr 07;141(14):1121-1123. [doi: 10.1161/CIRCULATIONAHA.119.044282] [Medline: 32250694]

20. Zou G. A modified poisson regression approach to prospective studies with binary data. Am J Epidemiol 2004 Apr 01;159(7):702-706. [doi: 10.1093/aje/kwh090] [Medline: 15033648]
21. Khonsari S, Subramanian P, Chinna K, Latif LA, Ling LW, Gholami O. Effect of a reminder system using an automated short message service on medication adherence following acute coronary syndrome. Eur J Cardiovasc Nurs 2015 Apr;14(2):170-179. [doi: 10.1177/1474515114521910] [Medline: 24491349]
22. Maddison R, Pfaeffli L, Whittaker R, Stewart R, Kerr A, Jiang Y, et al. A mobile phone intervention increases physical activity in people with cardiovascular disease: Results from the HEART randomized controlled trial. Eur J Prev Cardiol 2015 Jun;22(6):701-709. [doi: 10.1177/2047487314535076] [Medline: 24817694]
23. Wald DS, Bestwick JP, Raiman L, Brendell R, Wald NJ. Randomised trial of text messaging on adherence to cardiovascular preventive treatment (INTERACT trial). PLoS One 2014;9(12):e114268 [FREE Full text] [doi: 10.1371/journal.pone.0114268] [Medline: 25479285]
24. Piette JD, List J, Rana GK, Townsend W, Striplin D, Heisler M. Mobile health devices as tools for worldwide cardiovascular risk reduction and disease management. Circulation 2015 Nov 24;132(21):2012-2027 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.114.008723] [Medline: 26596977]
25. Liu F, Kong X, Cao J, Chen S, Li C, Huang J, et al. Mobile phone intervention and weight loss among overweight and obese adults: a meta-analysis of randomized controlled trials. Am J Epidemiol 2015 Mar 01;181(5):337-348 [FREE Full text] [doi: 10.1093/aje/kwu260] [Medline: 25673817]
26. Khan SU, Lone AN, Alkhouli M. Entry bias in cardiovascular clinical trials. Eur Heart J Qual Care Clin Outcomes 2021 Mar 15;7(2):216-217. [doi: 10.1093/ehjqcco/qcaa083] [Medline: 33118016]

Abbreviations

ACEi: angiotensin-converting enzyme inhibitor
ANCOVA: analysis of covariance
ARB: angiotensin receptor blocker
CHD: coronary heart disease
IPAQ: International Physical Activity Questionnaire
LDL-C: low-density-lipoprotein cholesterol
MET: metabolic equivalent
MMS: Modified Morisky Scale
PCI: percutaneous coronary intervention
RCCVC: regional cardiocerebrovascular center
RR: relative risk
SBP: systolic blood pressure
TEXTME: the Tobacco, Exercise, and Diet Messages

Edited by L Buis; submitted 26.04.21; peer-reviewed by A Akinosun, K Santo; comments to author 11.06.21; revised version received 07.07.21; accepted 27.08.21; published 24.09.21.

Please cite as:
Bae JW, Woo SI, Lee J, Park SD, Kwon SW, Choi SH, Yoon GS, Kim MS, Hwang SS, Lee WK
mHealth Interventions for Lifestyle and Risk Factor Modification in Coronary Heart Disease: Randomized Controlled Trial
JMIR Mhealth Uhealth 2021;9(9):e29928
URL: https://mhealth.jmir.org/2021/9/e29928
doi:10.2196/29928
PMID:34559058

Original Paper

Psychoeducational Social Anxiety Mobile Apps: Systematic Search in App Stores, Content Analysis, and Evaluation

Trent Ernest Hammond[1], BSc (Hons), BSocSc (Psy); Lisa Lampe[2], MBBS, PhD; Andrew Campbell[3], BA, MAppSc, PhD; Steve Perisic[4], DipSocWrk; Vlasios Brakoulias[5], MBBS (Hons), PhD

[1]Nepean Clinical School, Faculty of Medicine and Health, The University of Sydney, Kingswood, New South Wales, Australia

[2]School of Medicine and Public Health, College of Health, Medicine and Wellbeing, University of Newcastle, Callaghan, New South Wales, Australia

[3]Cyberpsychology Research Group, Biomedical Informatics and Digital Health, School of Medical Sciences, Faculty of Medicine and Health, The University of Sydney, Camperdown, New South Wales, Australia

[4]Mental Health Outreach Centre, South Western Sydney Local Health District, Campbelltown, New South Wales, Australia

[5]School of Medicine, Western Sydney University, Blacktown, New South Wales, Australia

Corresponding Author:
Trent Ernest Hammond, BSc (Hons), BSocSc (Psy)
Nepean Clinical School
Faculty of Medicine and Health
The University of Sydney
62 Derby Street
Kingswood, New South Wales, 2747
Australia
Phone: 61 431931448
Email: trent.hammond@sydney.edu.au

Abstract

Background: The wide use of mobile health apps has created new possibilities in social anxiety education and treatment. However, the content and quality of social anxiety apps have been quite unclear, which makes it difficult for people to choose appropriate apps to use on smartphones and tablets.

Objective: This study aims to identify the psychoeducational social anxiety apps in the two most popular Australian app stores, report the descriptive and technical information provided in apps exclusively for social anxiety, evaluate app quality, and identify whether any apps would be appropriate for people with social anxiety or others who know someone with social anxiety.

Methods: This systematic stepwise app search was guided by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) standards and entailed searching for, identifying, and selecting apps in the Australian Apple App and Google Play Stores; downloading, using, and reviewing the identified apps; reporting technical and descriptive information in the app stores, an online app warehouse, and individual apps; evaluating app quality; and deciding whether to recommend the use of the apps.

Results: In the app stores, 1043 apps were identified that contained the keywords *social anxiety*, *social phobia*, or *shyness* in their names or descriptions. Of these, 1.15% (12/1043) were evaluated (3 iOS apps and 9 Android apps). At the time of evaluation, the apps were compatible with smartphones and tablet devices; 9 were free to download from the app stores, whereas 3 were priced between US $2.95 (Aus $3.99) and US $3.69 (Aus $5.00). Among the evaluated apps, 3 were intended for treatment purposes, 3 provided supportive resources, 1 was intended for self-assessment, and the remaining 5 were designed for multiple purposes. At the time of downloading, app store ratings were available for 5 apps. The overall app quality was acceptable according to the Mobile App Rating Scale (MARS). On the basis of the MARS *app quality rating* subscale (sections A-D), the apps functioned well in performance, ease of use, navigation, and gestural design. However, app quality was less favorable when rated using the MARS *app subjective quality* subscale (section E).

Conclusions: The psychoeducational social anxiety apps evaluated in our study may benefit people with social anxiety, health professionals, and other community members. However, given that none of the apps appeared to contain empirical information or were shown to clinically reduce social anxiety (or aid in managing social anxiety), we cannot recommend their use. App accessibility could be improved by developing apps that are free and available for a wider range of operating systems, both

between and within countries and regions. Information communication and technology professionals should collaborate with academics, mental health clinicians, and end users (ie, co-design) to develop current, evidence-based apps.

(JMIR Mhealth Uhealth 2021;9(9):e26603) doi:10.2196/26603

KEYWORDS

anxiety; app; cell phone; mobile app; mobile phone; SAD; smartphone; social anxiety; social phobia; tablet

Introduction

Background

Social anxiety disorder (SAD) is characterized by the avoidance of social interactions that involve perceived scrutiny by others and potential embarrassment [1]. Signs and symptoms of SAD include avoidance of social activities because of anxiety, blushing, problems in making conversation, being unable to think of anything to say, reduced eye contact, nausea, rapid heartbeat, sweating, and dizziness [2]. People with SAD tend to experience problems in their daily activities [3], have poorer educational outcomes [4,5], are less productive at work, and subsequently have decreased employment prospects [1]. Surveys have reported that between 82.2% (n unavailable; total survey N=43,093) [6] and 88% (73/83) [7] of people with SAD had a diagnosis of at least one other mental disorder during the 12-month period before completing the surveys.

National surveys in the United States and Australia (countries with similar sex ratios and age structures) [8] suggest that SAD [9] affects tens of millions of people worldwide [10,11]. The US National Comorbidity Survey Replication [10] and Australian National Survey of Mental Health and Wellbeing (NSMHWB) [11] reported that SAD is common, with 12-month prevalence rates estimated at 6.80% (631/9282) for people aged 18 years and over and 4.70% (752,719/16,015,300) for people aged 16-85 years, respectively. At the time of writing this paper, there were no recent nationally representative statistics in Australia (ie, statistics published after 2008) regarding the prevalence of *subclinical* social anxiety. Despite changes to SAD diagnostic criteria in the Diagnostic and Statistical Manual of Mental Disorders, fifth edition [1] since the National Comorbidity Survey Replication and NSMHWB, the prevalence of SAD is considered unlikely to have changed [12].

Approximately half of people with mental health problems (eg, SAD) do not receive treatment for 15 years [1]. Early access to empirical information, clinical assessment, and efficacious treatment has the potential to decrease the severity and pervasiveness of social anxiety. Ubiquitous commercial mobile apps may be helpful if used adjunctively with psychoeducational interventions that provide educational materials, screening assessments, feedback, or advice regarding treatment [13]. Commercial mobile apps are developed by commercial, *for-profit* organizations but are not necessarily paid apps as many contain advertising. Psychoeducational apps have the potential to empower people, promote positive behaviors, facilitate personal symptom management, and enhance communication between clients and mental health professionals [14-16].

We identified no published empirical research specifying the reasons why people use or do not use social anxiety apps. However, 2 studies on health apps in the United States provide some insight. A web-based survey of 811 people showed the importance people place in considering content, ease of use, cost, encryption, responsive features, customization, privacy policy, and research evidence [17]. A cross-sectional survey identified that iPhone (iOS operating system) and Samsung (Android operating system) smartphone users stop using apps because they find them to be boring or they do not trust how their personal information will be used or managed. People also stop using health apps because it is burdensome to enter data, they become disengaged, or there are hidden costs [18].

At Stockholm University, academics developed an app called *Challenger* [19], which was available for iPhones only in Sweden's Apple App Store. The app includes several customizable features to enhance user engagement during internet-based cognitive behavioral therapy (CBT), such as gamification (eg, challenging board games with self-care rewards), personal skills training, goal setting, social interactions between end users, and notifications. Given the strong evidence base for CBT in treating anxiety disorders [20], Challenger has the potential to substantially enhance the ongoing management of social anxiety.

Of 52 anxiety apps, 2 social anxiety apps, including psychological techniques, were identified in a systematic review of the Romanian Apple App Store and Google Play Store (although app and developer names were not published) [21]. The top features of the anxiety apps included text, audio, worksheets, diaries, and animations. The main psychotherapeutic techniques suggested by these apps include progressive muscle relaxation, breathing, and emotional regulation. Two-thirds of the apps were free to download from the app stores; the remainder were between US $0.99 and $8.71 in price.

In another study, 7 psychoeducational and exclusively social anxiety apps in the New Zealand iTunes Store (now Apple App Store), Google Play Store, and Windows Store (now Microsoft Store) were reviewed [22]. Most of these psychoeducational apps were not universally accessible across mobile platforms and devices (ie, Apple [iOS], Google [Android OS], and Microsoft [Windows OS] smartphones and tablets). None contained expert information, had an evaluation of effectiveness published, or were developed by medical or not-for-profit institutions. Although specific psychoeducational apps were not identified in the published article, the names and platforms of all apps reviewed have been provided. Given that most apps are not accessible to people in Australia, it is difficult to substantiate these research findings. Further, the apps in this study were not physically downloaded and reviewed

individually, meaning limited conclusions can be drawn regarding the content, purpose, and media within the apps.

The objectives of this study were to (1) review the social anxiety apps in the two most popular Australian mobile app stores (ie, Apple App Store and Google Play Store); (2) report the content, technical properties, and descriptive features of the apps, including their platforms, purpose, and media; (3) evaluate psychoeducational app quality using the Mobile App Rating Scale (MARS) [23]; and (4) recommend the use of any quality, evidence-based apps to others who may benefit in the community.

What This Paper Adds

This paper makes a significant contribution to the limited academic literature and informs potential end users, mobile app developers, mental health clinicians, and academics about the content and quality of psychoeducational social anxiety apps. To our knowledge, this is the first peer-reviewed study to report the quality evaluation findings of commercial psychoeducational social anxiety apps. We are also the first to publish findings regarding the purposes, mobile platforms, and media of social anxiety apps available for download from Australian app stores. Our study uses an existing methodological framework previously used by New Zealand [22] and Australian [24] researchers to assess the quality of psychoeducational social anxiety apps available in Australia. Our research findings could inform mobile app developers of ways to improve the design, features, and content of social anxiety apps and avoid potential technological problems before development begins. Empirically informed, high-quality apps could then lead to enhanced user engagement, knowledge about social anxiety, increased help-seeking behavior, earlier treatment, and improved mental health.

Methods

Overview

The systematic app search and evaluation described in this paper focused on readily available *commercial* psychoeducational *social anxiety* mobile apps in Australia for smartphones and tablet devices. Therefore, apps were excluded if they were only accessible to specific organizations or clinical research participants; for example, an app in the United States, which is the only app we identified to be clinically proven to reduce social anxiety [25], was excluded because it was only accessible to study participants in that country during the randomized controlled trial.

Design

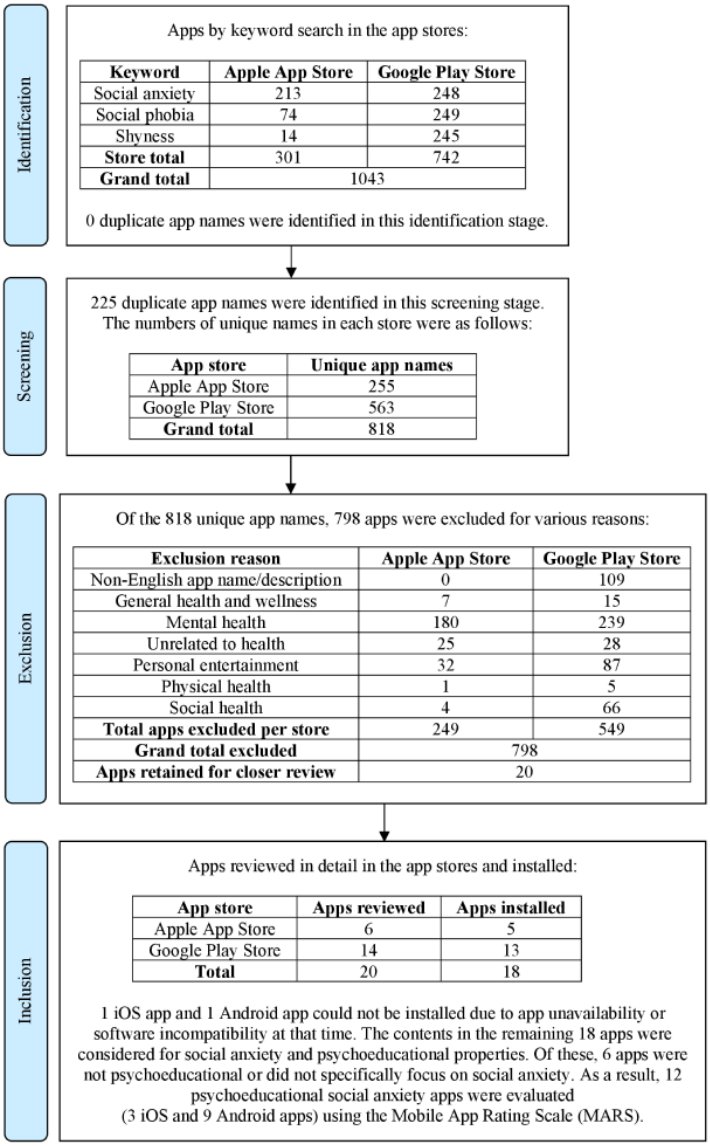
Overview

This systematic stepwise app search and evaluation was informed by the methodology of 2 studies: one investigated commercial social anxiety apps in New Zealand, and the other evaluated the quality of medication adherence apps in Australia [22,24]. The app search and selection strategy in our study was guided by the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [26].

The key steps in our study involved (1) systematically searching for, identifying, and selecting psychoeducational social anxiety apps in the Australian Apple App Store and Google Play Store; (2) downloading, using, and reviewing relevant apps; (3) reporting technical and descriptive information available in the app stores, an online app warehouse, and individual apps; (4) evaluating app quality using the MARS [23]; and (5) deciding whether to recommend apps to end users. Data were managed using Microsoft Excel spreadsheets (Version 16.27 for Mac).

Searches across all categories in the app stores occurred between August 13 and 25, 2019. Depending on the operating system required for app compatibility, apps were downloaded to an Apple iPad (iOS version 12.3.1, sixth Generation MR7F2X/A) or Samsung Galaxy tablet (Android version 4.4.4, SM-T560). At that time, the most recent software packages were installed for these devices before apps were downloaded and installed. The four steps of the app search, review, and selection process, outlined in Figure 1, included app identification, screening, exclusion, and inclusion.

Figure 1. PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) flowchart: app search, review, and selection process.



Step 1: App Identification

TEH conducted 3 separate keyword searches in 2 app stores for app names and descriptions containing the keywords *social anxiety*, *social phobia*, and *shyness*, which yielded 6 keyword groups. The exact app names and version numbers identified (if available) for each of the 6 search groups are listed in Multimedia Appendix 1. The app operating systems, keywords searched, and dates of data extraction have been provided. Owing to character limits of app names in the Apple App Store, the full names of 34 iOS apps were identified by web search using partial app names and other identifying information provided in the stores. English app names were then sorted alphabetically (numbers and special characters [eg, # and %] excluded), and then app names in all languages were compared and counted for each group.

Step 2: App Screening

Data from all groups in Multimedia Appendix 1 were integrated into Multimedia Appendix 2. Columns were added to identify app stores and duplicate app names. English app names were then sorted alphabetically across the 2 app stores. Excel conditional formatting highlighted duplicate app names, which were subsequently reviewed. App names, versions, and developers were considered in determining duplicates. The number of duplicates and unique apps was added up for each app store and overall.

Step 3: App Exclusion

Individual app names, descriptions, languages, purposes, intended audiences, screenshots, and relevant videos in the app stores were reviewed. Exclusion categories were developed based on the World Health Organization's definition of health [27] and the American Psychiatric Association's key SAD diagnostic criteria [1], outlined earlier in this paper. According to the World Health Organization, "health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity" [27]. Therefore, apps were excluded if they did not have an English name or description; were completely unrelated to health; did not have a specific focus on social anxiety but focused more generally on health and wellness, mental health, physical health, or social health;

or if they were intended for personal entertainment. The number of apps excluded for these reasons was counted in each app store and added up overall. The reasons for exclusion for the apps reviewed during step 3 are provided in Multimedia Appendix 2.

Step 4: App Inclusion

The remaining apps were reviewed by viewing app store websites and an app clearing warehouse website, specifically, app images, detailed descriptions, and related commentaries. Psychoeducational apps and apps targeting social anxiety are identified in Multimedia Appendix 2. The apps evaluated using the MARS [23] are listed in Multimedia Appendix 3, including their names, version numbers (if available), platforms, and associated websites.

Collection of App Technical and Descriptive Information

The App Classification section of the MARS [23] was completed for each mobile app evaluated. Although most information was obtained from the app stores and individual apps, some information was collected by searching various app stores and an app warehouse website. Additional information was collected for each app and is available in Multimedia Appendix 2, including the purpose (eg, multipurpose, self-assessment, supportive resources, or therapeutic treatment) and media type (eg, text and audio, text and visual, or text only).

Evaluating Apps Using the MARS

The MARS [23] consists of an *app quality rating* scale (sections A-D), an *app subjective quality* scale (section E), and an *app-specific* scale (section F). The app quality rating scale assesses various dimensions of app quality, including engagement (section A), functionality (section B), esthetics (section C), and information (section D). The 19 items of the scale were rated on a 5-point scale from inadequate to excellent. Sections A to D have an internal consistency of α=.90 and an interrater reliability intraclass correlation coefficient of 0.79. The app subjective quality scale has 4 items with different rating scales that assess whether one would recommend apps to others, how many times apps may be used for a 12-month period, whether one would pay for apps, and overall star ratings. The app-specific scale has 6 items that assess the perceived impact of apps on the user's awareness, knowledge, attitudes, intention to change, help-seeking behaviors, and actual behavior change.

However, given that the interrater reliability has not been determined for the app subjective quality scale and the app-specific scale, the initial quality assessment of apps in our study was predominantly based on the app quality rating scale. Apps were also assessed using the app subjective quality scale to supplement the findings and provide app developers with additional useful information in developing innovative social anxiety apps. The app-specific scale was not used in our study because the items were out of scope; it would be better suited to assess the perceptions of people with social anxiety (ie, how social anxiety apps could impact their knowledge, attitudes, and behaviors).

Two raters (TEH and SP) with knowledge regarding mobile apps, backgrounds in mental health, and training in using the MARS independently evaluated 12 psychoeducational social anxiety apps. This involved using each app for at least 15 minutes, reviewing information about the apps in the app stores, and then completing both the app quality rating scale and the app subjective quality scale. After evaluating all apps, TEH and SP discussed their individual ratings for the individual apps, particularly ratings that were substantially different (ie, 2 or more points different on the rating scales). This consultation stage ensured that key aspects of the apps and different opinions were considered before finalizing app ratings. Individual app scores of TEH and SP for the app quality rating scale and the app subjective quality scale were manually entered into separate Excel spreadsheets. Scores were averaged overall for each dimension of the app quality rating scale and items of the app subjective quality scale.

Results

App Technical and Descriptive Information

A total of 1043 apps were identified in the mobile app stores, of which 12 (1.15%) psychoeducational social anxiety apps were included for evaluation (Apple App Store=301 apps, 3 included for evaluation; Google Play Store=742 apps, 9 included for evaluation). Figure 1 presents the PRISMA flowchart for progress through the selection process and demonstrates why only 12 apps were downloaded and evaluated using the MARS.

The updated version numbers and exact production dates were available for 11 apps. Most were updated between 2016 and 2018. The earliest app was first released in 2011. The most recently updated apps in 2018 and 2019 were downloaded from the Google Play Store. The apps in the Apple App Store were updated during 2016 and 2017.

At the time of download, all apps were affiliated with commercial developers, as opposed to not-for-profit organizations. Only 5 were rated by app users (1 was rated by an Apple app user; 4 were rated by Android users), which had an overall mean rating of 4.2 out of 5 stars (SD 0.7). Most apps (n=9) targeted all age groups; however, 3 targeted adolescents and adults. In total, 9 apps were free to download. An app in the Apple App Store was priced at US $3.32 (Aus $4.49). The prices of the 2 other apps in the Google Play Store were US $2.95 (Aus $3.99) and US $3.69 (Aus $5.00).

All 12 apps were psychoeducational; however, their intended purposes differed. Five apps were for multiple purposes, 1 was for self-assessment, 3 contained supportive resources, and 3 had a therapeutic aim.

Similarly, the media of each app differed markedly. Of the 12 apps evaluated, 25% (3) were text and audio, 42% (5) were text and visual, and 33% (4) were text only. All apps contained advertisements for various products and services. App technical and descriptive information, including content focus, theoretical background, and therapeutic strategies, are listed in Multimedia Appendix 4.

App Quality Evaluation

MARS App Quality Rating Subscale (Sections A-D)

The top 5 ranked apps using the MARS included 3 Apple apps and 2 Android apps. From highest to lowest quality, they included *Beat Social Phobia with Andrew Johnson*, *Social Anxiety Test* (Mood Tools), *Social Anxiety Test* (Eddie Liu), *Social Anxiety Test-Psychological Test* and *How to Overcome Shyness* (Iaks Solutions). Individual and mean app quality ratings for each item in sections A to D are provided in Multimedia Appendix 5. The mean quality ratings of the apps are presented in Table 1.

The measures of central tendency (mean, median, and mode) and dispersion (SD and range) of overall app quality and the four app dimensions are presented in Table 2.

Table 1. Mean app quality ratings (including the 4 dimensions of engagement, functionality, esthetics, and information)[a].

App name	Operating system	Quality	Engagement	Functionality	Esthetics	Information
Beat Social Phobia with Andrew Johnson	iOS	4.10	4.40	4.50	3.67	3.83
Social Anxiety Test (Mood Tools)	Android	4.07	3.20	4.75	4.33	4.00
Social Anxiety Test (Eddie Liu)	iOS	3.94	3.00	4.50	4.00	4.25
Social Anxiety Test - Psychological Test	iOS	3.56	2.70	4.00	4.17	3.38
How To Overcome Shyness (Iaks Solutions)	Android	3.54	3.10	4.38	3.33	3.33
Beat Social Phobia	Android	3.33	3.60	3.88	2.83	3.00
Social Anxiety Disorder (Afradad Media)	Android	3.25	2.00	3.88	3.83	3.30
Social Anxiety Hypnosis	Android	3.10	2.20	4.13	2.33	3.75
How To Overcome Shyness (The Almighty Dollar)	Android	3.09	2.40	3.63	3.33	3.00
How To Overcome Shyness (Dierre09)	Android	2.75	1.90	4.00	2.50	2.60
Social Anxiety Disorder (Bedieman)	Android	2.69	1.70	3.63	2.17	3.25
Recognize Social Anxiety Disorder	Android	2.49	1.60	3.63	2.50	2.25

[a]Developer names are in parentheses for apps that share the same name.

Table 2. Measures of central tendency and dispersion of aggregate app quality ratings (including the four dimensions of engagement, functionality, esthetics, and information).

Variable	Quality	Engagement	Functionality	Esthetics	Information
Value, mean (SD)	3.33 (0.54)	2.65 (0.84)	4.07 (0.38)	3.25 (0.76)	3.33 (0.58)
Value, median (range)	3.29 (4.10-2.49)	2.55 (4.40-1.60)	4.00 (4.75-3.63)	3.33 (4.33-2.17)	3.32 (4.25-2.25)
Value, mode	N/A[a]	N/A	3.63	2.50 and 3.33[b]	3.00

[a]N/A: not applicable.

[b]Bimodal.

Regarding the central tendency measures for the apps, the quality ratings were mean 3.33 (SD 0.54) and median 3.29; there was no mode. These ratings showed that the quality of the apps according to the MARS was acceptable. The measures of dispersion demonstrated the similarity between app quality in Australian app stores. The highest-ranked app had a quality score of 4.10 and the lowest-ranked app had a score of 2.49 (range 1.61).

These results indicated that the strongest determinant of higher quality social anxiety apps was their functionality, as evidenced by a mean rating of 4.07 (SD 0.38). Specifically, most apps scored positively for performance, ease of use, and responsive taps, swipes, and scrolls (gestural design). Most components and functions of the top 5 social anxiety apps functioned correctly. These apps were relatively easy to use because of their clear labels, icons, and instructions.

Poorer app quality ratings were recorded for end user engagement, specifically for entertainment, customization, and interactivity. These results suggest that commercial psychoeducational social anxiety apps are not particularly fun to use. None of the apps evaluated in our review used gamification to entertain end users. Furthermore, most of these apps do not allow users to customize settings and preferences for app features, such as sound, content, and notifications.

MARS App Subjective Quality Subscale (Section E)

The mean subjective quality app ratings in Tables 3 and 4 are similar to the mean app quality ratings in Tables 1 and 2.

Table 3. Mean subjective quality ratings of the apps[a].

App name	Operating system	Subjective quality	Recommendation	12-month usage	Payment	Star rating
Beat Social Phobia with Andrew Johnson	iOS	3.88	4.00	3.50	4.00	4.00
Social Anxiety Test (Eddie Liu)	iOS	3.50	4.00	3.50	3.00	3.50
Social Anxiety Test (Mood Tools)	Android	3.50	4.00	3.50	3.00	3.50
Social Anxiety Hypnosis	Android	3.38	3.50	3.50	3.00	3.50
How To Overcome Shyness (Iaks Solutions)	Android	2.75	3.00	3.00	2.00	3.00
Beat Social Phobia	Android	2.50	3.00	3.00	1.00	3.00
Social Anxiety Disorder (Afradad Media)	Android	2.25	2.50	2.50	1.00	3.00
How To Overcome Shyness (The Almighty Dollar)	Android	2.13	2.50	2.50	1.00	2.50
Social Anxiety Disorder (Bedieman)	Android	2.00	2.50	2.00	1.00	2.50
Recognize Social Anxiety Disorder	Android	1.50	1.50	1.50	1.00	2.00
Social Anxiety Test - Psychological Test	iOS	1.50	2.00	1.50	1.00	1.50
How To Overcome Shyness (Dierre09)	Android	1.38	1.50	1.50	1.00	1.50

[a]Developer names are in parentheses for apps that share the same name.

Table 4. Measures of central tendency and dispersion for app aggregate subjective quality ratings.

Variable	Subjective quality	Recommendation	12-month usage	Payment	Star rating
Value, mean (SD)	2.52 (0.88)	2.83 (0.91)	2.63 (0.83)	1.83 (1.11)	2.79 (0.81)
Value, median (range)	2.38 (3.88-1.38)	2.75 (4.00-1.50)	2.75 (3.50-1.50)	1.00 (4.00-1.00)	3.00 (4.00-1.50)
Value, mode	1.50 and 3.50[a]	2.50 and 4.00[a]	3.50	1.00	3.00 and 3.50[a]

[a]Bimodal.

The app subjective quality ratings for section E of the MARS are available in Multimedia Appendix 6.

Of the top 5 apps, 4 (identified in Table 1) were rated moderately in terms of app subjective quality. From the highest- to the lowest-ranked app, they included *Beat Social Phobia with Andrew Johnson*, *Social Anxiety Test* (Eddie Liu), *Social Anxiety Test* (Mood Tools), *Social Anxiety Hypnosis*, and *How to Overcome Shyness* (Iaks Solutions). The mean app subjective quality of the 12 apps was 2.52 (SD 0.88), the median was 2.38, and there were 2 modes (1.50 and 3.50). These measures of central tendency showed that most of the apps' subjective quality ratings were similar and centered around the mean. The measures of dispersion, including the range of 2.50 (3.88-1.38), showed little variability between the subjective quality scores.

The app subjective quality ratings showed that independent raters (TEH and SP) would consider recommending 50% (6/12) of the psychoeducational apps to others, based on their personal experiences (mean 2.83, SD 0.91). However, none of the apps would definitely be recommended, and given that this measure has not been tested for validity and consistency, we cannot professionally recommend that others use the apps. Raters (TEH and SP) would consider using 50% (6/12) of these apps once in a 12-month period if they were relevant to their needs and wants (mean 2.63, SD 0.83). However, raters typically did not want to pay for the apps (mean 1.83, SD 1.11). Most apps received a star rating of at least 3 out of 5 (mean 2.79, SD 0.81).

Discussion

Overview

In our discussion, we outline the principal research results, considering our research objectives, which were to review the psychoeducational social anxiety mobile apps in the Australian Apple App Store and Google Play Store; describe the apps, their platforms, purpose, and media; evaluate the apps using the MARS [23]; and recommend the use of any quality evidence-based apps to others. We describe the problems encountered when researching psychoeducational social anxiety apps and challenge commercial app developers to enhance existing apps and design new apps that meet users' needs and wants. App development opportunities include enhancing the efficiency of locating social anxiety apps in app stores, improving international access to apps, and decreasing app-device incompatibilities. Further, we discuss descriptive and technical considerations for psychoeducational social anxiety apps, the lack of empirical evidence, and some potential limitations of mobile app reviews.

Principal Findings

There was a large number of apps available in the app stores, and they varied in quality. It is of interest that on the MARS [23], the range of scores between the top 5 and lowest quality apps was greater for the app subjective quality scale (section E) than the app quality rating scale (sections A-D). Considering

all MARS ratings in our study, the 3 highest quality apps were *Beat Social Phobia with Andrew Johnson*, *Social Anxiety Test* (Mood Tools), and *Social Anxiety Test* (Eddie Liu).

The difference in app ratings between the 2 subscales of the MARS could be attributable to scale design and the type and number of questions in each subscale. For example, the app subjective quality scale consists of 4 items, each with different rating scales. Averaging scores across these 4 items makes it challenging to interpret the results. The app quality rating scale has a higher level of face validity because it taps into a wider range of app dimensions, as opposed to only user preferences and potential future actions.

However, there is more to consider when evaluating apps than features and content. For instance, none of the social anxiety apps evaluated in this study appear to have been designed based on empirical evidence, and none have been evaluated to determine clinical effectiveness. App end users should provide constructive feedback to developers by contacting them directly (eg, by locating their contact details in the app stores), providing star ratings, and writing informative and honest reviews in the app stores. Similarly, app developers should work collaboratively with potential end users and clinical populations to identify and best meet their needs and desires.

Challenges for Commercial Social Anxiety App Developers

Our results highlighted the challenges in locating good-quality psychoeducational apps focused specifically on social anxiety in Australian mobile app stores. The initial search of the Google Play Store and Apple App Store revealed 1043 app names, based on keywords in app names and descriptions selected for their apparent relevance to social anxiety, thus representing the type of keywords consumers might use. However, the hit rate (*true positives*) for psychoeducational social anxiety apps was only 1.15% (12/1043). This poor hit rate is consistent with a review of 1154 apps in the New Zealand mobile app stores, which identified 13 psychoeducational social anxiety apps [22] with a hit rate of 1.13%. As the proportion of misses (*false positives*) exceeds 98% in both studies, the keywords in social anxiety app names and descriptions need to be improved. Unlike researchers and app developers, other app users do not typically have the time and patience to run systematic searches for relevant apps and often do not have knowledge regarding how to identify all relevant apps in the app stores.

Access to commercial psychoeducational social anxiety apps is limited because of end users' geographic locations. On the basis of the findings of the New Zealand review [22], only 2 of the 12 apps in our research were available in June 2016 in the New Zealand Google Play Store and iTunes Store (potentially different versions). It is noteworthy that in Australia, we were unable to download apps from the New Zealand Google Play Store and Apple App Store. The social anxiety apps identified by our New Zealand colleagues include *Social Anxiety Hypnosis* and *Beat Social Phobia with Andrew Johnson*. Developers, legislators, and intellectual property regulators should consider opening up web-based markets to increase access to apps. This is particularly important for well-conceptualized psychoeducational social anxiety apps. For example, *Challenger*, which was developed at Stockholm University, is only available in Sweden's Apple App Store [19].

Psychoeducational social anxiety apps are less common for Apple iPhones and iPads than for Android smartphones and tablets. In our study, only 25% (3/12) of the apps could be used on iOS devices, compared with 75% (9/12) of the apps for Android devices. Only 2 apps were available for both Apple and Android devices, namely, *Beat Social Phobia* and *Social Anxiety Test*. This has created an additional barrier for people with Apple devices to access psychoeducational social anxiety apps. Although the Android software for apps was relatively current at the time of review, iOS software for the Apple apps was updated 3 years before the review. Considering potential app-device incompatibility problems, people may have difficulty in downloading and accessing apps on newer, recently updated iPhones and iPads. However, given that 9 of the apps were free and the other 3 were reasonably priced, according to our standards, it is unlikely that there are financial barriers to access the apps.

Descriptive and Technical Considerations

The technical properties and descriptive content of the psychoeducational social anxiety apps, including their platforms, purposes, and media, have been described in the *Results* section of this paper. However, we believe it is important to highlight the need for empirically informed apps and reliable app quality assessment tools.

The apps reviewed in our study were intended to be psychoeducational; however, most contained outdated information about social anxiety and did not appear to be informed by empirical research evidence. Similar to earlier findings [22], there is no evidence that social anxiety apps in Australian app stores were developed by reputable not-for-profit institutions. All apps evaluated in our research were developed by commercial, *for-profit* organizations and contained advertising to promote often unrelated products and services. Although advertising may be included in mobile apps for financial gain, it may also be necessary to cover the costs of developing and maintaining apps, as app development can be expensive, and funding is limited. Further, the social anxiety apps contained insufficient information to confirm whether the content was based on expert knowledge and experience.

Challenger is the only publicly available, empirically-based psychoeducational social anxiety app we identified in our literature review [19]. On the basis of the authors' description of Challenger's features and technical specifications, we could not identify any social anxiety apps in Australian app stores of comparable quality. Raters TEH and SP were unable to evaluate the quality of Challenger using the MARS because the app was not available for download from the Australian app stores. Although the clinical effectiveness of Challenger in managing social anxiety has not been evaluated, it is the first commercial social anxiety app that uses gamification, goal setting, and CBT to engage end users. App developers should consider Challenger's technical and descriptive features, particularly when conceptualizing new evidence-based and empirically evaluated apps.

The completion rate, validity, and reliability of app store reviews in terms of determining app quality are questionable. Only 41.7% (5/12) of the social anxiety apps evaluated in our research were reviewed and rated by Apple App or Google Play Store app users. The moderated reviews in the app stores allow users to provide star ratings and post comments about the apps. However, only 2 of the 5 apps rated to be the highest in quality in our study (out of the 12 apps) received high-quality app store ratings in the app stores. For example, the average MARS app *quality* ratings in our study for *Beat Social Phobia with Andrew Johnson* (mean 4.10) and *Social Anxiety Test* (Mood Tools) (mean 4.07) were similar to the Apple App Store and Google Play Store, with ratings of 5 and 4.5 stars, respectively. As these apps have been available in the Australian app stores for several years and have been downloaded more than 10,000 times, we expected more than a few dozen reviews. Further, most app store reviews consisted only of star ratings rather than ratings and comments; therefore, it was not possible to verify quantitative ratings with qualitative data.

One potential explanation for these varied findings could be the different backgrounds (socioeconomic, cultural, and linguistic), perceptions, and expectations of Australian app users. Although the app raters TEH and SP are highly experienced in using mobile health apps and work in mental health settings, others in the Australian community may not be as objective and could find it difficult to provide detailed feedback when reviewing mobile apps. Furthermore, unlike the mobile app stores which allow app users to provide general written feedback, the items in the MARS allow raters to focus on specific aspects of apps, thereby allowing for more rigorous evaluation. Another potential reason for the varied quality ratings between the app stores and the MARS could be that raters (TEH and SP) reviewed all psychoeducational social anxiety apps in the two most popular Australian app stores. People in the community may be more selective in providing positive feedback for apps if they are more engaged app users (people who are disengaged may simply delete the app). It is also possible that the small number of web-based app reviews could be related to commercial developers wanting to positively market their apps to increase downloads and, therefore, advertising revenue. However, it is important to note that this comment is based on anecdotal feedback from app users and the professional experiences of the app raters (TEH and SP).

Strengths and Limitations

This is the first peer-reviewed systematic app review in which a published search strategy [26] was used to identify commercial social anxiety mobile apps in Australia. To the best of our knowledge, no other study has evaluated the quality of commercial social anxiety apps using a validated assessment inventory. The MARS is a suitable and useful tool for the purpose of quality assessment of social anxiety apps. However, mobile app raters need to have a good understanding of the terminology of the MARS items and substantial experience in using apps. The findings in our study would be very useful for the public in identifying suitable social anxiety apps and for developers in designing high-quality apps that are engaging, functional, esthetically pleasing, and contain appropriate information.

As our study involved reviewing the two most popular app stores in Australia, we believe that most of the psychoeducational social anxiety apps in Australia have been reviewed. However, the Apple App Store and Google Play Store have limited search parameters, making it challenging to identify relevant apps using keywords alone. Although our research findings are relevant to international app developers, they may only be directly transferrable to smartphone and tablet apps available in those app stores. We also acknowledge that other psychoeducational social anxiety apps are available, such as Microsoft, Amazon, Aptoide, F-Droid, and AppBrain apps.

Conclusions and Future Recommendations

The psychoeducational social anxiety apps available for download from the Australian Apple App Store and Google Play Store are of acceptable quality, either free or inexpensive to access, and contain some useful features and content that may assist people experiencing social anxiety or those who know someone with social anxiety. However, these apps appear to contain substantial amounts of outdated and nonempirical information, which is concerning given there is no evidence to suggest that they are practically useful in managing social anxiety. These apps could potentially cause harm to end users by indirectly or directly encouraging people to self-diagnose psychiatric disorders using web-based inventories (which were developed more than 20 years ago) and make personal decisions based on anecdotal evidence and untested treatment options.

Further work should focus on the development of tailored apps to meet the needs and desires of end users through researchers and app developers working collaboratively with mental health professionals, people with social anxiety, and people who know others with social anxiety. Information gleaned from co-design workshops, interviews, focus groups, web-based panels, and evidence-based information from peer-reviewed academic research could lead to the development of app prototypes that are high in quality; contain best practice strategies to manage social anxiety; and designed with suitable data collection, security, and sharing capabilities. Data collection options to better understand social anxiety could include (1) optimizing validated screening tools (eg, questionnaires and inventories) by considering internet and social media use, (2) mental health tracking systems to allow for efficacy testing, and (3) qualitative data collection tools to understand end users' experiences, for example, personal diaries and open-ended questionnaires.

Furthermore, it is time consuming to locate all the potential psychoeducational social anxiety apps in popular app stores, which could potentially increase stress and anxiety for those who are trying to find the right app for their circumstances. App developers could enhance the ease of locating apps by incorporating keywords more specific to social anxiety in app descriptions and names and avoiding broad keywords that encompass mental health problems more generally.

Finally, app compatibility can be enhanced across several mobile app platforms (not just iOS and Android devices). App developers could consider designing hybrid apps to be used across different devices and operating systems or separate versions of the same native app to be downloaded to devices with specific operating systems.

Acknowledgments

Funding for JMIR's Article Processing Fee (APF) US $2,500.00 (AUD $3,453.11) was provided to TEH to publish this paper from a combination of the following three funding sources: (1) JMIR Karma Credits US $171.00 (AUD $236.19), JMIR Publications Inc, 2021; (2) GO Scholarship US $1,809.96 (AUD $2,500), GO Foundation, 2020; and (3) Alan Duncan Memorial Grant US $519.04 (AUD $716.92), New South Wales (NSW) Education Council, 2020. We thank each of those organizations for supporting the publication of our valuable research. The authors also acknowledge The University of Sydney Nepean Clinical School and NSW Health for providing in-kind access to electronic resources, including library databases and computer hardware. The authors thank the collaborators at Western Sydney University and the University of Newcastle, who provided guidance for the app search methodology.

Authors' Contributions

TEH reviewed the published literature and social anxiety apps in mobile app stores, collected and analyzed the data, and wrote the vast majority of this paper. TEH and SP evaluated the quality of commercial psychoeducational social anxiety mobile apps. SP, LL, AC, and VB provided feedback and comments regarding the manuscript before the publication of this paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Apps searched by keyword, device, and app store.
[XLSX File (Microsoft Excel File), 52 KB - mhealth_v9i9e26603_app1.xlsx]

Multimedia Appendix 2
App list per app store and exclusion reasons.
[XLSX File (Microsoft Excel File), 95 KB - mhealth_v9i9e26603_app2.xlsx]

Multimedia Appendix 3
Psychoeducational social anxiety apps evaluated.
[DOCX File , 17 KB - mhealth_v9i9e26603_app3.docx]

Multimedia Appendix 4
App technical and descriptive information.
[DOCX File , 32 KB - mhealth_v9i9e26603_app4.docx]

Multimedia Appendix 5
Quality ratings (sections A-D) aggregate and individual raters.
[XLSX File (Microsoft Excel File), 3142 KB - mhealth_v9i9e26603_app5.xlsx]

Multimedia Appendix 6
Subjective ratings (section E): aggregate and individual raters.
[XLSX File (Microsoft Excel File), 19 KB - mhealth_v9i9e26603_app6.xlsx]

References

1. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition. Arlington, VA: American Psychiatric Association; May 18, 2013.
2. United States National Institute of Mental Health. Social Anxiety Disorder: More Than Just Shyness. Bethesda, MD: National Institute of Mental Health, Office of Science Policy, Planning, and Communications, Science Writing, Press, and Dissemination Branch; 2016. URL: https://www.nimh.nih.gov/health/publications/social-anxiety-disorder-more-than-just-shyness [accessed 2021-07-30]
3. Buist-Bouwman MA, De Graaf R, Vollebergh WAM, Alonso J, Bruffaerts R, Ormel J. Functional disability of mental disorders and comparison with physical disorders: a study among the general population of six European countries. Acta Psychiatr Scand 2006 Jun;113(6):492-500. [doi: 10.1111/j.1600-0447.2005.00684.x] [Medline: 16677226]
4. Van Ameringen M, Mancini C, Farvolden P. The impact of anxiety disorders on educational achievement. J Anxiety Disord 2003;17(5):561-571. [doi: 10.1016/s0887-6185(02)00228-1] [Medline: 12941366]

5. Russell G, Topham P. The impact of social anxiety on student learning and well-being in higher education. J Ment Health 2012 Aug;21(4):375-385. [doi: 10.3109/09638237.2012.694505] [Medline: 22823093]
6. Grant BF, Hasin DS, Blanco C, Stinson FS, Chou SP, Goldstein RB, et al. The epidemiology of social anxiety disorder in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions. J Clin Psychiatry 2005 Nov;66(11):1351-1361. [doi: 10.4088/jcp.v66n1102] [Medline: 16420070]
7. Fehm L, Beesdo K, Jacobi F, Fiedler A. Social anxiety disorder above and below the diagnostic threshold: prevalence, comorbidity and impairment in the general population. Soc Psychiatry Psychiatr Epidemiol 2008 Apr;43(4):257-265. [doi: 10.1007/s00127-007-0299-4] [Medline: 18084686]
8. United Nations Department of Economic and Social Affairs. World Population Prospects 2019. New York, NY: USA: United Nations; 2019. URL: https://population.un.org/wpp/DataQuery/ [accessed 2021-07-30]
9. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR). Washington, DC: USA: American Psychiatric Publishing; 2000.
10. Kessler RC, Chiu WT, Demler O, Merikangas KR, Walters EE. Prevalence, severity, and comorbidity of 12-month DSM-IV disorders in the National Comorbidity Survey Replication. Arch Gen Psychiatry 2005 Jun;62(6):617-627 [FREE Full text] [doi: 10.1001/archpsyc.62.6.617] [Medline: 15939839]
11. Australian Bureau of Statistics. National Survey of Mental Health and Wellbeing 2007: Summary of Results. Belconnen, ACT: Australia: Australian Bureau of Statistics; 2008 Oct 23. URL: https://www.abs.gov.au/statistics/health/mental-health/national-survey-mental-health-and-wellbeing-summary-results/latest-release [accessed 2021-07-30]
12. Crome E, Grove R, Baillie AJ, Sunderland M, Teesson M, Slade T. DSM-IV and DSM-5 social anxiety disorder in the Australian community. Aust N Z J Psychiatry 2015 Mar;49(3):227-235. [doi: 10.1177/0004867414546699] [Medline: 25122449]
13. Donker T, Griffiths KM, Cuijpers P, Christensen H. Psychoeducation for depression, anxiety and psychological distress: a meta-analysis. BMC Med 2009;7:79 [FREE Full text] [doi: 10.1186/1741-7015-7-79] [Medline: 20015347]
14. Coles ME, Coleman SL. Barriers to treatment seeking for anxiety disorders: initial data on the role of mental health literacy. Depress Anxiety 2010;27(1):63-71. [doi: 10.1002/da.20620] [Medline: 19960488]
15. Olfson M, Guardino M, Struening E, Schneier FR, Hellman F, Klein DF. Barriers to the treatment of social anxiety. Am J Psychiatry 2000 Apr;157(4):521-527. [doi: 10.1176/appi.ajp.157.4.521] [Medline: 10739410]
16. Chartier-Otis M, Perreault M, Bélanger C. Determinants of barriers to treatment for anxiety disorders. Psychiatr Q 2010 Jun;81(2):127-138. [doi: 10.1007/s11126-010-9123-5] [Medline: 20131095]
17. Schueller SM, Neary M, O'Loughlin K, Adkins EC. Discovery of and interest in health apps among those with mental health needs: Survey and focus group study. J Med Internet Res 2018 Jun 11;20(6):e10141 [FREE Full text] [doi: 10.2196/10141] [Medline: 29891468]
18. Krebs P, Duncan DT. Health app use among US mobile phone owners: A national survey. JMIR Mhealth Uhealth 2015 Nov 04;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
19. Miloff A, Marklund A, Carlbring P. The challenger app for social anxiety disorder: New advances in mobile psychological treatment. Internet Interventions 2015 Nov;2(4):382-391. [doi: 10.1016/j.invent.2015.08.001]
20. Hofmann SG, Asnaani A, Vonk IJJ, Sawyer AT, Fang A. The efficacy of cognitive behavioral therapy: A review of meta-analyses. Cognit Ther Res 2012 Oct 1;36(5):427-440 [FREE Full text] [doi: 10.1007/s10608-012-9476-1] [Medline: 23459093]
21. Sucala M, Cuijpers P, Muench F, Cardoş R, Soflau R, Dobrean A, et al. Anxiety: There is an app for that. A systematic review of anxiety apps. Depress Anxiety 2017 May 15;34(6):518-525. [doi: 10.1002/da.22654] [Medline: 28504859]
22. Alyami M, Giri B, Alyami H, Sundram F. Social anxiety apps: A systematic review and assessment of app descriptors across mobile store platforms. Evid Based Ment Health 2017 Aug;20(3):65-70. [doi: 10.1136/eb-2017-102664] [Medline: 28666986]
23. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: A new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
24. Santo K, Richtering SS, Chalmers J, Thiagalingam A, Chow CK, Redfern J. Mobile phone apps to improve medication adherence: A systematic stepwise process to identify high-quality apps. JMIR Mhealth Uhealth 2016 Dec 02;4(4):e132 [FREE Full text] [doi: 10.2196/mhealth.6742] [Medline: 27913373]
25. Enock PM, Hofmann SG, McNally RJ. Attention bias modification training via smartphone to reduce social anxiety: A randomized, controlled multi-session experiment. Cogn Ther Res 2014 Mar 4;38:200-216. [doi: 10.1007/s10608-014-9606-z]
26. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. PLoS Med 2009 Jul 21;6(7):e1000097 [FREE Full text] [doi: 10.1371/journal.pmed.1000097] [Medline: 19621072]
27. World Health Organization. Preamble to the constitution of WHO. In: World Health Organization. 1946 Jul 22 Presented at: International Health Conference; 22 July 1946; New York, NY: USA.

Abbreviations

CBT: cognitive behavioral therapy
MARS: Mobile App Rating Scale
NSMHWB: National Survey of Mental Health and Wellbeing
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
SAD: social anxiety disorder

Edited by L Buis; submitted 18.12.20; peer-reviewed by H Alyami, O El-Gayar, M Hamidzadeh; comments to author 24.03.21; revised version received 17.06.21; accepted 25.06.21; published 21.09.21.

Please cite as:
Hammond TE, Lampe L, Campbell A, Perisic S, Brakoulias V
Psychoeducational Social Anxiety Mobile Apps: Systematic Search in App Stores, Content Analysis, and Evaluation
JMIR Mhealth Uhealth 2021;9(9):e26603
URL: https://mhealth.jmir.org/2021/9/e26603
doi:10.2196/26603
PMID:34546179

Original Paper

Toward a Better Understanding of the Intention to Use mHealth Apps: Exploratory Study

Pedro R Palos-Sanchez[1*], PhD; Jose Ramon Saura[2*], PhD; Miguel Ángel Rios Martin[1*], PhD; Mariano Aguayo-Camacho[1*], PhD

[1]Department of Financial Economy and Operations Management, Faculty of Economic and Business Sciences, University of Seville, Seville, Spain
[2]Department of Business Economics, Rey Juan Carlos University, Madrid, Spain
[*]all authors contributed equally

Corresponding Author:
Pedro R Palos-Sanchez, PhD
Department of Financial Economy and Operations Management
Faculty of Economic and Business Sciences
University of Seville
Av. Ramon y Cajal, 1
Seville, 41017
Spain
Phone: 34 954557515
Email: ppalos@us.es

Abstract

Background: An increasing number of mobile health (mHealth) apps are becoming available for download and use on mobile devices. Even with the increase in availability and use of mHealth apps, there has still not been a lot of research into understanding the intention to use this kind of apps.

Objective: The purpose of this study was to investigate a technology acceptance model (TAM) that has been specially designed for primary health care applications.

Methods: The proposed model is an extension of the TAM, and was empirically tested using data obtained from a survey of mHealth app users (n=310). The research analyzed 2 additional external factors: promotion of health and health benefits. Data were analyzed with a PLS–SEM software and confirmed that gender moderates the adoption of mHealth apps in Spain. The explanatory capacity (R^2 for behavioral intention to use) of the proposed model was 76.4%. Likewise, the relationships of the external constructs of the extended TAM were found to be significant.

Results: The results show the importance of healthy habits developed by using mHealth apps. In addition, communication campaigns for these apps should be aimed at transferring the usefulness of eHealth as an agent for transforming attitudes; additionally, as more health benefits are obtained, ease of use becomes greater. Perceived usefulness (PU; β=.415, $t_{0.001;4999}$=3.442, P=.001), attitude toward using (β=.301, $t_{0.01;499}$=2.299, P=.02), and promotion of health (β=.210, $t_{0.05;499}$=2.108, P=.03) were found to have a statistically significant impact on behavior intention to use eHealth apps (R^2=76.4%). Perceived ease of use (PEOU; β=.179, $t_{0.01;499}$=2.623, P=.009) and PU (β=.755, $t_{0.001;499}$=12.888, P<.001) were found to have a statistically significant impact on attitude toward using (R^2>=78.2%). Furthermore, PEOU (β=.203, $t_{0.01;499}$=2.810, P=.005), health benefits (β=.448, $t_{0.001;499}$=4.010, P<.001), and promotion of health (β=.281, $t_{0.01;499}$=2.393, P=.01) exerted a significant impact on PU (R^2=72.7%). Finally, health benefits (β=.640, $t_{0.001;499}$=14.948, P<.001) had a statistically significant impact on PEOU (R^2=40.9%), while promotion of health (β=.865, $t_{0.001;499}$=29.943, P<.001) significantly influenced health benefits (R^2=74.7%).

Conclusions: mHealth apps could be used to predict the behavior of patients in the face of recommendations to prevent pandemics, such as COVID-19 or SARS, and to track users' symptoms while they stay at home. Gender is a determining factor that influences the intention to use mHealth apps, so perhaps different interfaces and utilities could be designed according to gender.

(JMIR Mhealth Uhealth 2021;9(9):e27021) doi:10.2196/27021

KEYWORDS

mHealth apps; mobile apps; eHealth; promotion of health; TAM; PLS–SEM; COVID-19

Introduction

Overview

The use of mobile health (mHealth) apps increased during the first decade of the 21st century [1] and this has led to an increase in the amount of time that users devote to improve their health using mHealth app(s). New ways of monitoring and controlling health indicators and daily activities using new technologies and improvements on the internet have now become available [2].

The increasing use of technology and the internet has forced companies to adapt their marketing strategies to this digital ecosystem. This growth has led to an increase in the use of smartphones around the world [3,4].

For this reason, user behavior and consumption habits with mobile apps have become important fields of research [3,5].

Alharbi et al [6] reported that one type of app which has been increasingly used in recent years is mHealth apps. Support for patients has become more widespread due to the use of these apps. However, users sometimes stop using these apps because they perceive that their usefulness may not cover health quality standards or because the service is not of the same quality as, for example, a visit to the doctor offline [2].

Telemedicine and eHealth have duly become important factors for the analysis, study, improvement, and development of patients' medical and health care. Electronic health or eHealth was defined by Eysenbach [7] as "health services and information provided by the Internet and related technologies."

Many mHealth apps provide direct communication links between patients and health care professionals, health education, health portals, wellness management for measuring calories and following a diet, management of diseases such as diabetes and asthma, self-diagnosis to identify symptoms and early diagnosis, medication reminders, and rehabilitation processes and therapies. Therefore, this kind of app could be used to predict what the behavior of patients would be in the face of recommendations to prevent pandemics, such as COVID-19 or SARS, and to track users' symptoms while they stay at home and follow doctors' recommendations [8].

The term "application" or "app" refers to a self-contained program or piece of software that is designed to fulfill a particular purpose, and is usually optimized to run on mobile devices, such as smartphones, tablet computers, and wearable devices such as smart watches [3].

Therefore, mHealth apps can improve users' health by monitoring risks, symptoms, and health care programs. Consumer interest in mHealth apps has increased at the same rate as new technology use in the health care sector. Taking the characteristics shown by mHealth apps into consideration, the technology acceptance model (TAM) was chosen for this study [9].

TAM is a computational system, presented by Davis [10], which analyses users' decision-making processes when adopting a new technology. The TAM was used in this research paper to investigate the adoption of mHealth apps. External factors that help describe the user adoption of mHealth apps were incorporated into the TAM.

This research therefore fills a gap in the information currently available because it incorporates innovative factors for the adoption of mHealth apps that creators and developers should take into account for successful acceptance and adoption of new mHealth apps. This information duly adds to the existing literature that can be consulted by professionals and researchers.

Therefore, this study addresses the following research question: What factors, including the innovative TAM variables such as promotion of health and health benefits, determine the acceptance of mHealth apps?

This paper is divided into 5 sections. First, the theoretical framework for adoption of mHealth apps is explained. TAM is analyzed and the hypotheses to be studied are formulated. The next section explains the methodology used in the study. The characteristics of the chosen research technique, a survey, are given. This section covers all aspects of questionnaire design and data collection.

Finally, the results of PLS–SEM analysis of the hypotheses and relationships are presented. This section also includes the interpretation, discussion, and implications of the results obtained. The conclusions of the study and the main theoretical and practical implications of the results are also presented.

Theoretical Background

As stated above, in recent years, researchers have become interested in the adoption of mHealth apps. Research by Housman [11] investigated health information on social media by studying how mHealth apps share results on social media platforms. The increase in use of social networks and the factors that affect the relationship and use of mHealth app were also studied by investigating the social acceptance of mHealth apps by internet user communities [3].

Likewise, Li et al [12] studied emotional bonding of patients with mHealth apps. They showed that users accept this type of apps from an emotional perspective, keeping the disease more in mind and, therefore, applying better monitoring protocols.

Handel [13] studied the use of mobile apps for health and wellness and identified the uses of mHealth apps for health, weight loss, consumption of healthy diet and food, monitoring glucose levels and diabetes, calculating calories consumed, disease diagnosis, meditation, yoga, monitoring sleep quality, and tracking sports activities [14,15]. Therefore, these categories of health care have already been accepted as interesting topics for scientific research in the area of mHealth apps.

Atienza and Patrick [16] studied the acceptance of mHealth apps for the care industry. Furthermore, Grundy et al [17] studied the use of high-quality mHealth apps with

innovation-based systems and systematically described the characteristics of recent apps.

Following this line of research, Mueller [18] studied the types of mHealth apps recommended by doctors to their patients, concluding that this type of app is a valid technological support for disease monitoring and treatment.

Likewise, Bloomfield et al [19] studied the influence of SMART goals on the behavior of mHealth app users. Cho [20] investigated the impact of postadoption sentiments on mHealth app use with the postacceptance model and the TAM to find the users' continued intention to use health apps.

Bort-Roig et al [21] investigated how mHealth apps could improve employees' sedentary lifestyles while at work and studied the users' acceptance and continued use of mHealth app. In a similar way, Ashurst and Jones [22] studied the acceptance of mHealth apps among people with diabetes who used one to check and control their condition. It can be seen that the diagnosis and control of medical conditions with technology is an accepted area of scientific research.

Accordingly, Gorkem et al [23] investigated what factors may influence users' behavioral intentions to adopt and use mHealth apps. To this end, the authors extended the TAM with external factors such as price value, trust factors, and perceived risk and evaluated users' technology acceptance. The results of this study showed that the first 2 presented a statistical significance with intention to use.

Deng et al [24] studied which determinants influence the adoption of mHealth services among Chinese patients using the TAM extended with trust, perceived risks, and patients' age and chronic diseases. All external variables were found to be positively correlated with mHealth service adoption.

The study carried out by Mao et al [25] highlighted the importance of studying the recommendations made by patients who have used this type of app to predict what the behavior of patients would be in the face of a change in medical treatment.

In this context, aiming to understand the main advances of mHealth apps, this study takes as a reference the apps regulated by the Food and Drug Administration (FDA). As noted by Humphries et al [2], the FDA is a leading international institution in the regulation of new health products and services and serves as a guide and institutional leader for all other regulatory institutions in the health field around the world, including Spain. Table 1 shows the main mHealth app categories related to this study's objectives.

Table 1. mHealth App categories regulated by the Food and Drug Administration.

Mobile health app (category)	Description	Functions
Slendertone Connect (health and wellness)	Allows users to measure physical exercise intensity by connecting an intelligent device.	Measures patients' resistance, tracks distance traveled, and allows patients to monitor calories.
Kardia (medicine)	Allows patients to measure blood glucose levels to detect possible risks and evaluate the patient's condition.	Tracks patients' heartbeat, measures the glucose and oxygen level in blood, and shares the patient's data with the doctor.
Diasend (health and fitness)	Measures the patient's diabetes constants.	Shares data in real time with other users of the app, allows patients to track exercise and calories, and connects patients' data with other health and medicine apps.
LibreLink (medicine)	Checks blood glucose without extracting blood from the finger using a small external device that connects to the app.	Can add notes about food, insulin, and exercise; gives blood glucose readings; and shares information with family, friends, and doctors in real time.
Qardio heart health (health and fitness)	Allows patients to control blood pressure, heart rate, and weight. A small external device is used to send the data.	Measurement of blood pressure and patient's weight, monitoring of heart rate and prediction of heart attacks, and helps share patient information in real time with family and friends.

Conceptual Framework and Hypothesis Elaboration

The TAM was used to explain the relationship between the acceptance and adoption of technology and the users' intention to use it [26]. Au and Zafar [27] and Chen and Tan [28] used TAM to demonstrate that perceived usefulness (PU) and perceived ease of use (PEOU) are the most critical factors in the process of adoption and use of new technology. In the TAM, PU and PEOU are considered beliefs and evaluations, respectively, given by users, which influence their attitude toward and intention to use the product (in this case, an app) [29], and finally result in behavior change [30,31].

In the study by Davis [10], TAM was used to explain and predict the use of information systems; in other words, TAM was used to understand the influence of the variables *PU* and *PEOU* on the use of technology. The PU is the belief that a certain technology can improve users' performance while using it. The PEOU is defined as the degree to which a person believes that using any particular technological system is simple and stress free.

The TAM consistently explains a large part of the variance, 40% according to many authors such as Venkatesh and Davis [32], in the intention to use different information and communication technologies by users in different environments and countries [27,33,34]. Since its appearance, the TAM has been widely analyzed and expanded in different ways [35].

The most important evolutions of TAM have been the TAM2 model by Venkatesh and Davis [32], the Unified Theory of the Acceptance and the Use of Technology by Venkatesh et al [36],

the model for the acceptance of technology and user satisfaction by Wixom and Todd [37], and the TAM3 model.

The reasons for choosing the TAM are its tremendous popularity and besides many studies have used the model. The TAM is often considered a common and robust model to address consumer acceptance of an innovative technology [38]. Scherer et al [39] confirmed that the TAM successfully predicts user behavior and can thus be of interest to all potential users of a new technology [29,40,41]. TAM is widely used, with its application extending to a multitude of technologies, especially websites and apps [38].

The TAM has found relevant support in the literature: there are more than 14,870 citations regarding this model within the core collection of the Web of Science database, and more than 51,495 citations have been retrieved from Google Scholar for the article by Davis [10] as of June 2020, 30 years after his first theory.

Therefore, TAM has established itself well as a robust, powerful, and parsimonious model for predicting user acceptance. However, it has been modified through different extensions.

The first of the TAM extensions, the so-called TAM2 [42], is based on the expansion of the PU background. Subsequently, with the same intention as in TAM2, but to complete the model by incorporating the background of the original TAM, Venkatesh and Bala [35] developed the TAM3. More specifically, while TAM2 added the history of PU, TAM3 was expanded into the constructions that precede the PEOU and that were already established in [43,44].

Legris et al [45] made an important critical review of the model and concluded that TAM is useful, but it has to be integrated into a broader one that includes variables related to social and human processes of change.

Similarly, Tang and Chen [46] concluded that current studies on TAM and its extended models have made great progress and recommended paying more attention for future research on new variables that come from other theories or topics that must be introduced in the new model to make it easier to interpret.

Thus, in many health care studies where TAM was applied, the authors have added variables to extend the original TAM to better adapt it to the context of health care [47].

We can find research studies that have used the TAM, such as [48], in which the authors evaluated the acceptance of home telemedicine services by elderly patients. Within the health care domain, the TAM has been used to examine the determinants of adoption or the intention to adopt health technologies [49,50] and to know the effects of cognitive and contingent factors on the health adoption of smartphone apps [51].

The use of computing in the health care field is increasing, but adoption remains a challenge. To understand and introduce the health information technology, a series of behavioral models and innovation acceptance models have been studied and specifically applied the TAM to understand the acceptance of technology [52].

In addition, as in our work, TAM was developed with a focus on technology that can be used voluntarily without the assistance of professional health staff [47].

Furthermore, a recent study in the field of mHealth, in which extended TAM was used [53], indicated that the findings in the literature are contradictory regarding the adoption of mHealth self-monitoring tools, thereby suggesting a gap in the literature that must be covered.

Besides, Thies et al [54] justified that the lack of adoption of a mobile app to support patients in self-management of chronic diseases was mainly due to problems related to the usability of the app and that patients are not comfortable with the technology.

Likewise, Paré et al [55] indicated that people who declare themselves ill are less likely to use digital or traditional tools to monitor their well-being/health than people in good health. Therefore, it is especially important to investigate the adoption of these instruments by consumers considering the characteristics of both the technology and the individuals (users), especially those related to their health [53], as well as the reliability of the model. Extended TAM is decisive in using unused constructs to cover this gap identified in the literature.

The hypotheses below were chosen after reviewing research studies on mHealth apps by Cho [20], Kim and Park [56], and Jeon and Park [57].

Cho [20] and Jeon and Park [57] demonstrated the influence of PEOU on the use of mHealth apps and its effect on PU. Veer et al [58] explained how the intention to use mHealth app influences PU in communities of older people. The following hypothesis was therefore proposed:

> *H1: Perceived ease of use has a positive influence on perceived usefulness*

Veer et al [58], Hu and Bentler [59], and Deng [60] explored the influence and effect of PEOU on attitude toward using. Thompson et al [61] studied the effect of attitude toward using a technology on the intention to use it. Based on their study, the following hypothesis was proposed:

> *H2: Perceived ease of use has a positive influence on attitude toward use*

With the emergence of mHealth, some studies [49,62] confirm the influence of the PU of patients' intention to adopt a mHealth management service in other cultural contexts [53].

Chauhan and Jaiswal [63] showed that PU influences attitude toward using an mHealth app. The influence of different variables for using different types of mHealth app was also reported. PU demonstrates how a user feels that a particular technology can have a positive effect on his/her life. This influences the user's attitude toward using the technology [64]. Consequently, the following hypothesis was proposed:

> *H3: Perceived usefulness has a positive influence on attitude toward use*

To investigate PU, Chang et al [65] analyzed the acceptance of a hospital-based eHealth service. The influence of PU on the behavioral intention to use this service by hospital users was

also found. Likewise, Klein [66] concluded that PU has a positive effect on behavioral intention to use in his research on patient psychology and the use of eHealth services. Therefore, the following hypothesis was proposed:

H4: Perceived usefulness has a positive influence on behavioral intention to use

Moores [67] concluded that the attitude toward use variable influences the adoption of technological health care services. In addition, Mun et al [68] concluded that behavioral intention to use has a positive effect on the PU of technology by eHealth professionals. From these investigations, the following hypothesis was proposed:

H5: Attitude toward use has a positive influence on behavioral intention to use

Lin and Yang [69] and Buntin et al [70] examined the health benefits of mHealth apps and reported on the main positive health benefits of mHealth apps by applying the TAM for the PEOU construct. Beldad and Hegner [51] studied health benefits with the "health valuation" construct after users tried a fitness app. The confidence that users have in the app was found by extending the TAM with trust, social influence, and health valuation variables. Consequently, the following hypothesis was proposed:

H6: Health benefits have a positive influence on perceived ease of use

Jeon and Park [57] investigated the factors that affect the acceptance of mHealth apps for obesity and found the influence and effect of health benefits on PU. They suggested that more studies should be carried out with the TAM to find out how mHealth apps can help manage and reduce problems with health and chronic diseases [67]. Based on this, the following hypothesis was proposed:

H7: Health benefits have a positive influence on perceived usefulness

Kim and Park [56] improved the TAM with the promotion of health external variable to apply it for evaluating health information technology. Melzner et al [71] studied the influence of mHealth apps on the promotion of health at the workplace and also the attitude of employees toward using an mHealth app. The effects on productivity and health benefits at work upon using an mHealth app were studied by Kelly et al [72]. Ramtohul [73] performed a comprehensive analysis of the decision to adopt eHealth services from the user's perspective. Therefore, the following hypothesis was proposed:

H8: Promotion of health has a positive effect on the health benefits of mHealth apps

Bert et al [74] studied the influence of mobile phones on promotion of health and concluded that some mHealth apps can help prevent diseases and also influence changes in the users' health behavior. Ramtohul [73] investigated promotion of health with a construct called "Health Needs," which expresses the benefits for mHealth app users. Consequently, the following hypothesis was proposed:

H9: Promotion of health has a positive effect on behavioral intention to use an mHealth app

Ramtohul [73] also analyzed the influence of promotion of health on PU in a study on psychological variables. Cho et al [51] analyzed the influence of PU on health benefits for workers who use smartwatch apps [75]. Moores [67] linked the PU of an mHealth app with the promotion of health of app users [8]. Therefore, the following hypothesis was proposed:

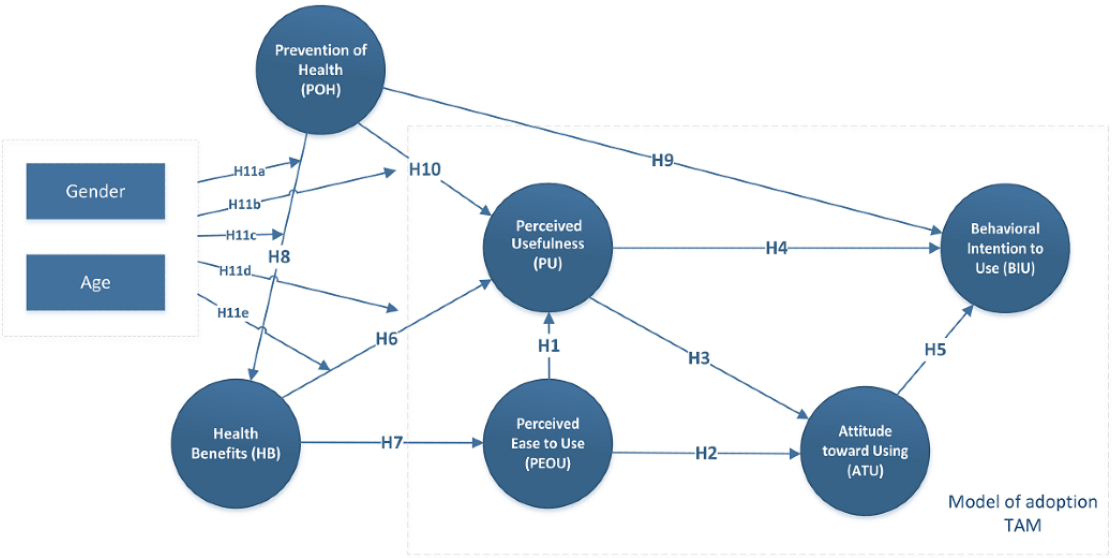
H10: Promotion of health has a positive effect on perceived usefulness of an mHealth app

Venkatesh et al [36] pointed out that men and women have different perceptions of usefulness when deciding on technology acceptance. Shabani [76] studied the importance of gender as a moderating variable for adolescents' emotional health. Bidmon et al [77] indicated that although both men and women use mHealth app, men tend to use it more on mobile devices. Dyck et al [78] studied the moderating effects of age, gender, and education variables and the influence of these on patients' physical activity. Based on the studies by Venkatesh et al [36] and Shabani [76], the following hypothesis was proposed:

H11: Gender and age moderate all the relationships of constructs in the research model

The research model in Figure 1 was formulated to explore the influence of health benefits and promotion of health on the mHealth app adoption model.

Figure 1. Research model to explore the influence of health benefits and promotion of health on the mHealth app adoption model. TAM: technology acceptance model.



Methods

Measurement

A questionnaire was created with 24 questions on attitudes and behavior and 5 questions for group classification. The classification questions were for gender, age, job, residence, and education level. The questionnaire was divided into 3 sections. The first section dealt with questions on the users' behavior, beliefs, and attitudes toward an mHealth app.

Before answering this section and the next one, users could watch a video on different mHealth apps and try them out. A total of 12 different FDA-approved mHealth apps were suggested for trial purposes.

The apps can be found in Google Play or Apple Store by searching their names: my mhealth, Mhealth Medical App, MHealth, Babylon, HealthForYou, Medipal mHealth app, Walking: Pedometer, Medical ID: ICE, Symptom Tracker, ContinuousCare, Medical Record, and ManageMyHealth. All sample members were selected because they indicated that they had previously used mHealth apps and were aware of their functionality and traceability. They were informed about the other apps so that they could take into account additional features of the apps.

The first section of the questionnaire contained 15 questions on PU (n=5), PEOU (n=3), attitude toward using (n=3), and behavioral intention to use (n=4).

The second section consisted of a block of questions on health and disease prevention. These were grouped into health benefits (n=4) and promotion of health (n=4). The last section consisted of 5 questions on the demographic profile of the sample.

Adapted items were used to measure the variables in the TAM [10]. The behavioral items for health were adapted from the studies by Lin and Yang [69] and Jeon and Park [57]. Lin and Yang [69] studied the influence of mHealth app on patients with asthma problems and Jeon and Park [57] studied the influence of mHealth apps on patients with obesity problems. Altogether, there were 24 items in the questionnaire.

All the items, except the demographic profile, were measured using a 5-point Likert scale that ranged from total disagreement (1) to total agreement (5).

A pilot survey was conducted to find the pilot sample's opinions about the content and structure of the questionnaire, so that the questions could be refined if needed. The pilot survey was conducted on a subsample of 31 individuals whose answers were not added to the final sample.

The subsample followed all the instructions and answered all the questions. Participants were asked to provide comments and suggestions to improve both the instructions and the questions in the questionnaire.

The most important comments were made regarding the items with unclear wording, which were not easily understood, which could cause confusion about the question, or with possible ambiguity in the answers. The wording of these erroneous items was later modified or changed.

The psychometric properties of the proposed scale were then evaluated, along with its ability to identify theoretical concepts and constructs from the data extracted from the questionnaire. The criteria, procedures, and validation techniques for scales proposed by Mackenzie et al [79] were used to create the validation process for the scale used. The measurement model gave satisfactory results.

Recruitment

The questionnaires were distributed in Spain, both in Madrid and in towns and cities in nearby regions. The prerequisite for the sample was that the user had 4G or Wi-Fi connectivity to the internet. In total, 442 valid questionnaires were collected from the interviewees between January and February 2020.

The sampling was nonprobabilistic and convenient. Google Forms (Google LLC/Alphabet Inc.) was used to prepare the

questionnaire, which was then distributed on different social networks, especially LinkedIn (Microsoft).

The SPSS 24 statistical software (IBM) was used to calculate the frequency tables and statistics generated by the sample.

Demographic Information

The results from the questionnaires showed that 242/442 members (54.8%) were men, 195/442 (44.1%) were women, and 5/442 were others (1.1%).

Of these, 186 participants live in small populations of less than 5000 inhabitants (42.1%), which makes the sample interesting, as getting to hospitals and health centers may be difficult for them. Furthermore, 336 participants were aged between 18 and 30 years (76.0%) and 291 had studied at a university (65.9%); 64.9% (n=287) of the sample were students.

Statistical Analysis

Data analysis and hypothesis testing were carried out using structural equation modeling (SEM) with variance, which allowed for a statistical examination of the interrelated dependency relationships between the latent variables and the indicator variables of the research model by directly measuring observable variables [80].

SEM was used together with partial least squares (PLS). PLS trajectory modeling can be understood as a complete SEM method to study composite factor models by measuring constructs, estimating structural models, and performing model fitting tests [81].

The PLS–SEM statistical analysis technique, based on the structural equation model, was used, as it is especially recommended for exploratory research. It allows the modeling of latent constructs with both formative and reflective indicators to analyze the collected data [82]. In addition, PLS is appropriate for the prediction and analysis of relatively new phenomena [83]. The SmartPLS 3 software (SmartPLS GmbH) was used in this study [84].

Reinartz et al [85] investigated the conditions under which PLS–SEM should be used in research analysis, and concluded that the technique can be applied for a relatively new object of research with a model that is not fully consolidated. As these were the conditions in this research, we chose to use PLS–SEM. Besides, ours is an exploratory approach [86] for which this type of data analysis is highly recommended [87].

The PLS–SEM technique was also used because one of the aims of this research was to check whether the model was predictive. Chin and Newsted [83], Fornell and Larcker [88], and Hair et al [89] had already shown that PLS–SEM can be used for this purpose.

Fornell and Bookstein [90] state that PLS explicitly defines the latent variables, constructs, or combinations, which can easily be measured. The use of these factors is another point that justifies the use of SEM, as shown in similar studies by Sarstedt et al [80], Henseler [91], and Rigdon et al [92].

Based on the research studies by Sarstedt et al [80], Hair et al [89], and Cepeda-Carrion et al [93], the choice of the best SEM approach depends on the type of latent variables being measured, with the aforesaid studies recommending PLS for reflective or common factor constructs. The information required to analyze these factors was found from other related variables, which is another condition for which PLS–SEM is recommended [80]. Investigation and adoption of mHealth apps is a recent area of research. Because this study is exploratory, PLS–SEM is recommended.

The Harman single-factor test was used as an indicator in the subsequent common method bias test [94,95]. Using this test, no single factor was detected that could explain most of the total variance, which suggests that it is very unlikely that any selection bias exists.

Results

Measurement Model

The measurement model was tested for internal reliability, convergent validity, and discriminant validity. The internal reliability was evaluated using Cronbach α which needs a value of at least .70 for acceptable internal consistency [96]. Causality was analyzed using indicator loadings. Composite reliability was also used to investigate causality [97]. All the constructs had internal consistency, as their Cronbach α values were higher than .7 [86,88,98]. To assess convergent validity, Fornell and Larcker [88] used the average variance extracted (AVE) method and stated that an acceptable value for this factor is 0.50 or more.

The structural model was then analyzed using a bootstrapping technique configured to readjust 5000 subsamples to estimate the statistical significance of the path coefficients [99].

Table 2 shows the element loads, Cronbach α, and AVE which were found for the constructs. Cronbach α values ranged from .899 to .789, which is higher than the recommended level of .70, and therefore indicates strong internal reliability for the constructs. The composite reliability ranged between 0.930 and 0.877 and the AVE between 0.651 and 0.783, which are higher than the recommended levels. The conditions for convergent validity were therefore met. The discriminant validity was calculated with the square root of the AVE and the cross-loading matrix. For satisfactory discriminant validity, the square root of the AVE of a construct should be greater than the correlation with other constructs [88].

These researchers carried out simulation studies to demonstrate that a lack of discriminant validity is better detected by means of another technique, the heterotrait-monotrait ratio, which they had discovered earlier. All the heterotrait-monotrait ratios for each pair of factors was less than 0.90.

Table 2. Reliability, validity of the constructs, Fornell–Larcker criterion, and HTMT.

Construct	Cronbach α alpha	CR[a]	AVE[b]	Fornell-Larcker Criterion						HTMT[c]				
				ATU[d]	HB[e]	BIU[f]	PEOU[g]	POH[h]	PU[i]	ATU	HB	BIU	PEOU	POH
ATU	.861	0.915	0.783	0.898										
HB	.899	0.930	0.768	0.703	0.867					0.828				
BIU	.883	0.920	0.742	0.743	0.711	0.887				0.777	0.788			
PEOU	.789	0.877	0.703	0.596	0.555	0.556	0.854			0.682	0.663	0.641		
POH	.844	0.906	0.762	0.722	0.814	0.719	0.544	0.851		0.788	0.776	0.865	0.632	
PU	.866	0.903	0.651	0.771	0.740	0.771	0.576	0.728	0.811	0.872	0.844	0.822	0.668	0.795

[a]CR: composite reliability.

[b]AVE: average variance extracted.

[c]HTMT: heterotrait-monotrait.

[d]ATU: attitude toward using.

[e]HB: health benefits.

[f]BIU: behavioral intention to use.

[g]PEOU: perceived ease of use.

[h]POH: promotion of health.

[i]PU: perceived usefulness.

Structural Model

In this next stage, the proposed model was analyzed in detail. The structural model was built up from the different relationships between the constructs. The hypotheses for the study were tested by analyzing the relationships between the different constructs in the model to see if they were supported [83,85,100].

The assessment of the significance of structural model is usually preceded by performing an analysis of the indicator reliability and the internal consistency reliability to prove the lack of multicollinearity. The variance inflation factor values obtained were less than 5 and ranged from 1.603 (PEOU3) to 3.496 (behavioral intention to use 3).

The variance is found from the values for the reflective indicators given by the constructs [101,102]. This was found numerically by calculating the R^2 values, which are a measure of the amount of variance for the construct in the model. The bootstrap method was used to test the hypotheses. The detailed results (path coefficient, β, and *t* statistic) are summarized in Table 3 and Figure 2.

PEOU is positively associated with PU (β=.203, $t_{0.01;499}$=2.810, *P*=.005) and attitude toward using (β=.179, $t_{0.01;499}$=2.623, *P*=.009), and therefore, H1 and H2 were compatible with the proposed model with a 99% level of confidence.

Likewise, PU, another relationship established in the TAM, positively influenced the variable attitude toward using. This relationship was therefore confirmed and was compatible with the proposed model (β=.755, $t_{0.001;499}$=12.888, *P*<.001) with a high level of confidence (99.9%).

The TAM constructs that influence behavioral intention to use, such as PU (β=.415, $t_{0.001;4999}$=3.442, *P*=.001), have a significant influence on the intention to use an mHealth app. Therefore, H4 was supported for the proposed model with a confidence level of 99.9%.

The results also indicated that the research model explains 76.4% of the variance of the intention to use an mHealth app (R^2 for behavioral intention to use=76.4%, R^2 values for attitude toward using, health benefits, PEOU, and PU are 78.2%, 74.7%, 40.9%, and 72.7%, respectively). The result of a single linear regression from attitude toward using mHealth apps and behavioral intention to use confirmed that attitude toward using is positively associated with behavioral intention to use an mHealth app (β=.301, $t_{0.01;499}$=2.299, *P*=.02). This means that H5 was supported (99%).

The hypotheses for the external variable health benefits of the original TAM were all supported with the same level of confidence (99.9%). Therefore, the health benefits variable was shown to have a significant influence on PU (β=.448, $t_{0.001;499}$=4.010, *P*<.001) and therefore H6 was supported.

Likewise, health benefits also positively influenced PEOU (β=.640, $t_{0.001;499}$=14.948, *P*<.001), which shows that H7 was supported. The other external variable (ie, promotion of health) was found to significantly influence health benefits (β=.865, $t_{0.001;499}$=29.943, *P*<.001), which means that H8 is supported with the highest values in this research model (99.9%). H7 and H8 had the highest *t* statistic value of all the studied hypotheses (Table 3).

H9 and H10 studied the association of promotion of health with behavioral intention to use (β=.210, $t_{0.05;499}$=2.108, *P*=.03) and PU (β=.281, $t_{0.01;499}$=2.393, *P*=.01) with a 99% level of confidence. H9 had the lower *t* statistic value of all the studied hypotheses (95%).

Table 3. Results of hypothesis: path coefficients and statistical significance (n=5000 subsamples).[a]

Hypothesis	β (coefficient path)	*t* statistic	*P* value	Supported
H1: Perceived ease of use → Perceived usefulness	.203	2.810	.005	Yes[b]
H2: Perceived ease of use → Attitude toward using	.179	2.623	.009	Yes[b]
H3: Perceived usefulness → Attitude toward using	.755	12.888	<.001	Yes[c]
H4: Perceived usefulness → Behavioral intention to use	.415	3.442	.001	Yes[c]
H5: Attitude toward using → Behavioral intention to use	.301	2.299	.02	Yes[b]
H6: Health benefits → Perceived usefulness	.448	4.010	<.001	Yes[c]
H7: Health benefits → Perceived ease of use	.640	14.948	<.001	Yes[c]
H8: Promotion of health → Health benefits	.865	29.943	<.001	Yes[c]
H9: Promotion of health → Behavioral intention to use	.210	2.108	.03	Yes[d]
H10: Promotion Of Health → Perceived usefulness	.281	2.393	.01	Yes[b]

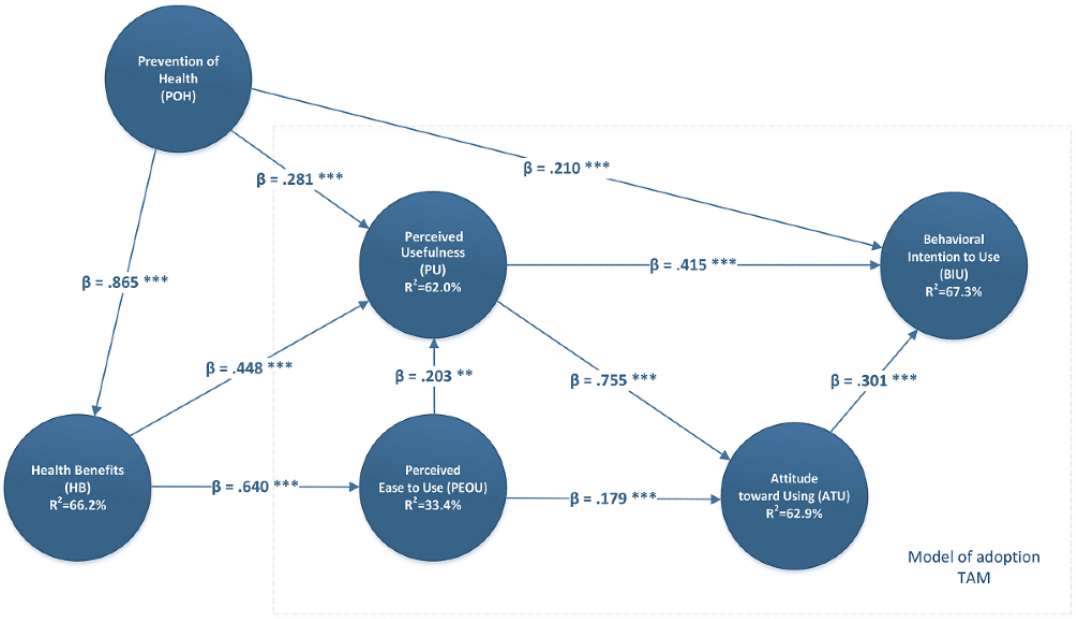
[a]For 5000 subsamples, we used a *t* distribution (4999) of students in single queue.

[b]$P<.01$ ($t_{0.01;499}=2.333843952$).

[c]$P<.001$ ($t_{0.001;499}=3.106644601$).

[d]$P<.05$ ($t_{0.05;499}=1.64791345$).

Figure 2. Analysis results (path coefficient, β, and *t* statistic are presented). TAM: technology acceptance model.



The measurements for approximate adjustments of the model [81,91] are given by the standardized root mean square residual (SRMR) value [103], which measures the difference between the observed correlation matrix and the implied correlation matrix of the model. SRMR shows the average magnitude of these differences.

A low value of SRMR means that the fit is better. In our case SRMR=0.023, which was within the recommendations for a model with a good fit. A good fit is considered to be shown with an SRMR value of less than 0.08 [103].

Regarding the evaluation of the overall fit of the model, Benitez et al [104] recommend evaluating a saturated structural model by investigating discrepancy between empirical and model-implied indicator variance–covariance matrix. Bootstrapping results show that the SRMR sample mean for the saturated model (0.023) is below the 95% mark of its corresponding reference distribution (0.027).

The blindfolding procedure omits part of the data for a given construct during the estimation of parameters. Estimated parameters are then used to try to recreate the omitted data [101]. It is possible to study the predictive relevance of the model in this way using the Stone–Geisser (Q^2) test [105,106]. This test revealed that the model has predictive capability. As can be seen in Table 4, all endogenous constructs fulfill $Q^2 > 0$. Values

of 0.02, 0.15, and 0.35 for Q^2 in the Stone–Geisser test indicate small, medium, and great predictive relevance [107].

As per the R^2 (see Table 4 and Figure 2) values reported by Chin [101], we conclude the following: If R^2=0.67, the result is considered substantial; 0.33, the result is considered moderate, and 0.19, the result is considered weak. The R^2 obtained for the main dependent variable of the model, behavioral intention to use, was 76.4%

This value shows that this model is "substantially" applicable for the adoption of an mHealth app. The variables that are not endogenous do not have a value for R^2.

The blindfolding technique consists in omitting part of the data for a given construct during the estimation of parameters, and then trying to estimate what was omitted from the estimated parameters [83].

In this way the predictive relevance of the model was studied and using the Stone–Geisser (Q^2) test the model was shown to have predictive capacity [105].

Therefore, all constructs, except PEOU, in the studied model have great predictive relevance, as the values of Q^2 are greater than 0.35 (Table 4). The proposed research model thus has good predictive power when explaining behavioral intention to use an mHealth app.

Effect size shows the strength of the relationship between 2 variables in the research model on a numeric scale. The effect size (f^2) shows how much an exogenous latent variable contributes to the R^2 value of an endogenous latent variable. The f^2 values 0.02, 0.15, and 0.35 indicate small, medium, and large effect size [100]. Cohen's tables [107] showed that for 95.2% statistical power and an average effect size of f^2=0.15, a minimum of 107 questionnaires would be needed. In our case the number of samples was 442, showing that this research has adequate statistical power.

Table 4. R^2 and Q^2 results.

Construct	Q^2	R^2 (%)
Attitude toward using	0.478	78.2
Health benefits	0.465	74.7
Behavioral intention to use	0.491	76.4
Perceived ease to use	0.229	40.9
Promotion of health	N/A[a]	N/A
Perceived usefulness	0.381	72.7

[a]N/A: not applicable.

PLS–SEM Results With Moderator (Gender and Age)

In order to check H11 and measure the potential moderating influence of gender and age, we performed a multigroup analysis [108].

First, the sample was divided by gender into men and women. The following process was then repeated, dividing members of the sample into old and young people.

However, before doing this test it is necessary to analyze the measurement invariance of the composite models (MICOM) technique [80]. This test will ensure that the effect of gender is restricted to the trajectory coefficients of the structural model and not to the parameters of the measurement model [109]. As described in Tables 5 and 6, we find the invariance of the measurement in the case of gender, but not in the case of age (Table 6) for the variables attitude toward using, health benefits, behavioral intention to use, perceived ease to use (PEOU), promotion of health, and PU.

Table 5. Results of the measurement invariance of composite models (MICOM) procedure (gender).

Construct	Step 1	Step 2			Step 3a				Step 3b			
	Configural invariance	Compositional invariance			Equal variances				Mean original difference (men–women)	Equal means		
		Original correlation	5%	Partial measurement invariance established	Variance original difference (men–women)	2.5%	97.5%	Equal		2.5%	97.5%	Equal
ATU[a]	Yes	1.000	1.000	Yes	0.209	–0.182	0.182	No	–0.070	–0.308	0.303	Yes
BIU[b]	Yes	1.000	1.000	Yes	–0.011	–0.211	0.176	Yes	0.234	–0.278	0.265	Yes
HB[c]	Yes	1.000	1.000	Yes	0.061	–0.197	0.185	Yes	0.146	–0.277	0.282	Yes
PEOU[d]	Yes	1.000	0.998	Yes	–0.054	–0.202	0.187	Yes	0.095	–0.267	0.270	Yes
POH[e]	Yes	1.000	1.000	Yes	0.092	–0.189	0.167	Yes	0.242	–0.288	0.279	Yes
PU[f]	Yes	.999	0.999	No	0.152	–0.198	0.174	Yes	–0.123	–0.287	0.265	Yes

[a]ATU: attitude toward using.

[b]BIU: behavioral intention to use.

[c]HB: health benefits.

[d]PEOU: perceived ease of use.

[e]POH: promotion of health.

[f]PU: perceived usefulness.

Table 6. Results of the measurement invariance of composite models (MICOM) procedure (age).

Construct	Step 1	Step 2			Step 3a				Step 3b			
	Configural invariance	Compositional invariance			Equal variances				Mean original difference (young people–old people)	Equal means		
		Original correlation	5%	Partial measurement invariance established	Variance original difference (young people–old people)	2.5%	97.5%	Equal		2.5%	97.5%	Equal
ATU[a]	Yes	1.000	0.999	Yes	–0.399	–0.298	0.307	No	–0.665	–0.398	0.537	No
BIU[b]	Yes	1.000	0.999	Yes	–0.520	–0.298	0.298	No	–0.590	–0.423	0.538	No
HB[c]	Yes	1.000	0.998	Yes	–0.461	–0.300	0.304	No	–0.624	–0.410	0.538	No
PEOU[d]	Yes	0.999	0.993	Yes	–0.324	–0.315	0.298	No	–0.432	–0.400	0.512	Yes
POH[e]	Yes	1.000	0.998	Yes	–0.470	–0.299	0.298	No	–0.239	–0.419	0.539	Yes
PU[f]	Yes	1.000	0.997	Yes	–0.606	–0.312	0.293	No	–0.360	–0.401	0.500	Yes

[a]ATU: attitude toward using.

[b]BIU: behavioral intention to use.

[c]HB: health benefits.

[d]PEOU: perceived ease of use.

[e]POH: promotion of health.

[f]PU: perceived usefulness.

Discussion

Principal Findings

The results of this study confirmed that the variable that has the strongest impact on the behavioral intention to use of mHealth apps in Spain is PU. This variable also has a very high predictive capacity as its determination coefficient is high [81,108]. The next most important variable in the model is health benefits.

The results of this research could be applicable to other EU countries with similar levels of internet access. However, it must be taken into account that most of the participants lived in areas with less than 5000 inhabitants (186/442 participants, 42.1%), where acceptance of mHealth apps is also determined by the close social environment. In this type of environment, users of mHealth apps can offer an effective short-term consultation for families and acquaintances before they make a decision to visit hospitals or health clinics.

Comparison With Prior Work

These findings are consistent with previous studies on PU for the acceptance of medical information systems [57,69,101]. These studies also found that PU significantly influences the adoption of medical information systems.

Promotion of health was also found to have a significant effect on health benefits of using mHealth apps in this study, as mHealth apps positively promote and improve the health of mHealth app users in Spain. This relationship was the strongest among all the relationships studied in this research and shows the usefulness of mHealth apps for improving health.

This is important when promoting the idea of preventing diseases and other ailments with mHealth apps, such as controlling continued physical exercise, consumption of certain foods, monitoring the evolution of potential and current patients, and using smartphones or tablet PCs to help prevent health problems. These results are consistent with the findings from a previous study [110].

H8 has been revealed as the relationship with the greatest burden and confirms the extraordinary influence it has on health benefits (β=.865, $t_{0.001;499}$=29.943, P<.001). This means that eHealth apps that take care of nutrition, improve sports activity, or make mealtimes more respectful are perceived by respondents as favoring aspects related to blood pressure, weight loss, blood sugar levels, or mood. In other words, users consider that apps related to healthy habits should be developed. This means that H8 is the most reliable and significant relationship among all.

The second hypothesis with the greatest burden and influence was H7. The relationship between health habits and PEOU of apps indicates that the more beneficial the eHealth app is, the easier it should be to use. Furthermore, the third hypothesis with the greatest intensity is H3, which shows that the perception of usefulness of an eHealth app has an extraordinary influence on the attitude of use. This means that the selling strategy of these apps must be aimed at transferring 2 very important aspects to the user: on the one hand, the usefulness of eHealth as an attitude transforming agent, and on the other, the more health benefits are obtained, the easier it is to use. In addition, these 3 relationships (ie, H3, H7, and H8) were very significant (99.9%).

The TAM is applicable to the use of eHealth apps as was the case with other studies, but with the influence of the "health promotion" and "health benefits" constructs. In addition, health promotion is directly related to the main dependent variable in the behavioral intention to use model. Therefore, health promotion is a construct that should be considered in future research, as it is also directly related to the final construct of the behavioral intention to use as well as indirectly to the PU.

In this study it was demonstrated that mHealth apps were easy to use and that users were familiar with the basic functions and applications of the internet. This is justified by the fact that health benefits had a very significant influence on the perceived usability (PEOU). This is an important point to highlight when explaining mHealth apps, as this can help ensure that mHealth apps are used as often as necessary to achieve effective results. However, the influence of PEOU on PU is the relationship with the lowest load among all (β=.179, $t_{0.01;499}$=2.623, P=.009) and a 99% confidence level. Likewise, the PEOU has a moderate variance (R^2=40.9%), which is why a moderately atypical result was obtained in this research. PEOU has a positive relationship with PU, which suggests that users will not need to learn new skills to use mHealth apps. The sample in this study, however, did not consider it an important factor in this model. In all probability, the advancement of usability of smartphone interfaces reduces the influence of PEOU, so people might need to use smartphones to be able to use these types of apps [111]. These results could be explained by the fact that the Spanish population is already familiar with health promotion and also that current mHealth apps are easy to use and accessible.

The remaining endogenous variables had a very high explanatory capacity (>70%). This gives the model a great capacity to explain the reality of the users' behavior before using eHealth apps, as in the case of behavioral intention to use it was 76.4%.

The results obtained for the relationship between the PEOU and the attitude toward use predict a smooth learning curve. This suggests that the adoption of mHealth app will be permanent and stable in the future. The use of mHealth apps will not present any significant difficulties that may cause users to abandon it.

Our study also confirmed that health promotion has a positive influence on behavioral intention to use and perception of usefulness (PU). In both cases, the level of trust is high, which shows that health promotion is an important factor in this model. Health promotion was also found to have an indirect influence on health benefits. This result supports the previously reported finding that app titles influence behavioral intention to use [112]. Specifically, we found that apps with titles related to symptoms have a significantly lower number of installs as compared with those whose titles are not related to symptoms.

Finally, a moderating capacity was found with a 95% confidence level regarding gender. We found that the 2 relationships with the lowest level of confidence in the model (Table 7), H1 or the relationship of the perception of ease of use with PU (β=–.422, P=.015) and H9 or the relationship of health promotion with

behavioral intention of use (β=–.239, *P*=.04), show significant differences between men and women. Furthermore, gender-moderated behaviors were found in H10, indicating that health promotion also influences the perception of usefulness differently according to gender (β=.178, *P*=.01).

Table 7. PLS[a]–SEM[b] results with moderator (gender).

Hypothesis	β (Coefficient path)	*P* value	Support
H1: Perceived ease of use → Perceived usefulness	–.422	.01	Yes[c]
H2: Perceived ease of use → Attitude toward using	–.100	.24	No
H3: Perceived usefulness → Attitude toward using	.318	.18	No
H4: Perceived usefulness → Behavioral intention to use	.166	.21	No
H5: Attitude toward using → Behavioral intention to use	–.107	.49	No
H6: Health benefits → Perceived usefulness	.266	.11	No
H7: Health benefits → Perceived ease of use	.003	.94	No
H8: Promotion of health → Health benefits	–.318	.22	No
H9: Promotion of health → Behavioral intention to use	–.239	.04	Yes[c]
H10: Promotion of health → Perceived usefulness	.178	.01	Yes[c]

[a]PLS: partial least squares.

[b]SEM: structural equation modeling.

[c]For 500 subsamples, we used a *t* distribution (4999) of students in a single queue: $P<.05$ ($t_{0.05;4999}=1.64791345$).

The other moderating variable (ie, age) was not supported, coinciding with the results of similar studies [113].

Therefore, mHealth app is an effective way to promote good health and habits in the population. Participants in the study believed that mHealth apps could help them improve their health, maintain a meal schedule, take part in more sporting activities, or improve the hours slept at night. Thus, mHealth apps can promote healthy habits and improve the users' quality of life.

Conclusions

Theoretical Implications

As has often been addressed in previous mHealth studies [114,115], health apps on smartphones can serve as very realistic health care alternatives, helping people save on medical expenses and being more effective in managing their personal health. Therefore, we agree with a previous work [20] that the potential advantages of using health apps (mHealth) in terms of improving overall health can be harmed without the use of apps.

The extended TAM adoption model was found to be fully valid for the study of mHealth app use and acceptance in Spain. This result could be extrapolated to other EU countries with similar levels of internet accessibility and sociodemographic characteristics.

This study identified the variables that influence people's intention to use mHealth apps. Using an extended TAM, PU was found to be the most significant variable influencing adoption of mHealth apps in Spain. This means that the most important factor for users are the ways in which mHealth apps can help them. This result is important because users of this type of apps must first understand the utility of the use of these apps, so that they can become cognizant about how they can improve treatment of their diseases and their control.

Practical Implications

Other external variables, such as promotion of health, have a significant effect on the health benefits of mHealth app use. This result showed that users consider maintenance or improvement of health as an additional health benefit provided by these apps.

The predictive capacity of the model and the predictive capacity can be very useful in preventing diseases that need controlled habits. Examples are indulging in regular physical exercise; consumption of certain foods; monitoring the evolution of current and potential patients; and using smartphones, tablets, and other medical devices to prevent health problems. Besides, care centers should have Wi-Fi access so that patients can carry out real-time diagnostic tests.

The results of this research show that gender is neither completely decisive nor moderating in the behavioral intention to use mHealth apps. This means that adoption of mHealth apps for promotion of health was moderated only by gender. Another important factor influencing mHealth app use is PEOU.

Therefore, user-friendliness and health promotion should be gender sensitive when applying utilities to apps. Accordingly, app developers should take into account users' gender and introduce some changes in usage and health promotion levels.

The results obtained using the extended TAM show that promotion of health and health benefits are important variables for mHealth apps users because they indirectly influence the adoption of the technology. This means that mHealth apps could be an alternative way to promote and improve health and could

become a service that minimizes primary care consultations for simple cases.

This is because PU and PEOU are not the only mediators for the final intention to use. Promotion of health is directly related to behavioral intention to use. This was a highly significant relationship and means that users prefer mHealth apps that promote health. This recommendation is important for designers, developers, and start-ups creating new mHealth apps. Therefore, we could start thinking that barriers such as standards, security, and interoperability [116] could be overcome by the activities derived from promotion of health.

The significance of the association between PU and behavioral intention to use explains the importance of mHealth apps for the users. This could explain the evolution of mHealth apps that offer an increasing number of benefits to the user.

An example is that the users' health information can now be transmitted online. This could help health centers have real-time information and minimize visits to health centers for primary care. To increase the adoption and use of mHealth apps, there should be an approved catalog of health service providers and an adoption strategy for citizens.

Based on our study results, the authorities could take the following as indicators for the use of mHealth apps: connectivity of the mHealth app, interaction between the patient and the health professional via the app, the need to prescribe additional quality hardware that allows measurements and analyses, and the personalized and nonautomated accessibility of these apps to the use and analysis of patient data remotely. These tools could be key indicators to measure the quality of this type of apps by health authorities.

In conclusion, gender is a determining factor that influences the intention to use eHealth apps, and therefore, different interfaces and utilities could be designed according to gender.

The findings of this study are beneficial for organizations, governments, and policymakers to provide strategies and policies to improve mHealth app in different hospitals and Spanish primary health care centers.

Limitations

The limitations of the research are those related to the analysis technique used, the country under study, and the size of the sample.

Conflicts of Interest

None declared.

References

1. Baldwin JL, Singh H, Sittig DF, Giardina TD. Patient portals and health apps: Pitfalls, promises, and what one might learn from the other. Healthc (Amst) 2016 Oct 3;5(3):81-85 [FREE Full text] [doi: 10.1016/j.hjdsi.2016.08.004] [Medline: 27720139]
2. Humphries T, Kushalnagar P, Mathur G, Napoli DJ, Rathmann C. Global Regulatory Review Needed for Cochlear Implants: A Call for FDA Leadership. Matern Child Health J 2020 Nov 02;24(11):1345-1359. [doi: 10.1007/s10995-020-03002-5] [Medline: 32876813]
3. Kao C, Liebovitz DM. Consumer Mobile Health Apps: Current State, Barriers, and Future Directions. PM R 2017 May;9(5S):S106-S115. [doi: 10.1016/j.pmrj.2017.02.018] [Medline: 28527495]
4. Saura JR, Reyes-Menendez A, Thomas SB. Gaining a deeper understanding of nutrition using social networks and user-generated content. Internet Interv 2020 Apr;20:100312 [FREE Full text] [doi: 10.1016/j.invent.2020.100312] [Medline: 32300536]
5. Reyes-Menendez A, Palos-Sanchez PR, Saura JR, Martin-Velicia F. Understanding the Influence of Wireless Communications and Wi-Fi Access on Customer Loyalty: A Behavioral Model System. Wireless Communications and Mobile Computing 2018 Dec 02;2018:1-16. [doi: 10.1155/2018/3487398]
6. Alharbi A, Alzuwaed J, Qasem H. Evaluation of e-health (Seha) application: a cross-sectional study in Saudi Arabia. BMC Med Inform Decis Mak 2021 Mar 18;21(1):103 [FREE Full text] [doi: 10.1186/s12911-021-01437-6] [Medline: 33736622]
7. Eysenbach G. Towards ethical guidelines for e-health: JMIR theme issue on eHealth ethics. J Med Internet Res 2000 Mar 31;2(1):E7 [FREE Full text] [doi: 10.2196/jmir.2.1.e7] [Medline: 11720926]
8. Del Campo Tejedor A. Presentación. Análisis y retos antropológicos en tiempos del coronavirus. RAA. 2020. URL: https://theconversation.com/retos-del-teletrabajo-en-tiempos-de-coronavirus-133098 [accessed 2021-08-18]
9. Palos-Sanchez P, Martin-Velicia F, Saura JR. Complexity in the Acceptance of Sustainable Search Engines on the Internet: An Analysis of Unobserved Heterogeneity with FIMIX-PLS. Complexity 2018 Oct 09;2018:1-19. [doi: 10.1155/2018/6561417]
10. Davis FD. Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
11. Housman LT. "I'm Home(screen)!": Social Media in Health Care Has Arrived. Clin Ther 2017 Nov;39(11):2189-2195. [doi: 10.1016/j.clinthera.2017.10.007] [Medline: 29100731]
12. Li J, Zhang C, Li X, Zhang C. Patients' emotional bonding with MHealth apps: An attachment perspective on patients' use of MHealth applications. International Journal of Information Management 2020 Apr;51:1-17. [doi: 10.1016/j.ijinfomgt.2019.102054]

13. Handel MJ. mHealth (mobile health)-using Apps for health and wellness. Explore (NY) 2011 Jul;7(4):256-261. [doi: 10.1016/j.explore.2011.04.011] [Medline: 21724160]
14. Goetzel RZ, Shechter D, Ozminkowski RJ, Marmet PF, Tabrizi MJ, Roemer EC. Promising Practices in Employer Health and Productivity Management Efforts: Findings From a Benchmarking Study. Journal of Occupational and Environmental Medicine 2007;49(2):111-130. [doi: 10.1097/jom.0b013e31802ec6a3]
15. Liu C, Zhu Q, Holroyd KA, Seng EK. Status and trends of mobile-health applications for iOS devices: A developer's perspective. J Syst Software 2011;84(11):2022-2033. [doi: 10.1016/j.jss.2011.06.049]
16. Atienza AA, Patrick K. Mobile health: the killer app for cyberinfrastructure and consumer health. Am J Prev Med 2011 May;40(5 Suppl 2):S151-S153. [doi: 10.1016/j.amepre.2011.01.008] [Medline: 21521588]
17. Grundy QH, Wang Z, Bero LA. Challenges in Assessing Mobile Health App Quality: A Systematic Review of Prevalent and Innovative Methods. Am J Prev Med 2016 Dec;51(6):1051-1059. [doi: 10.1016/j.amepre.2016.07.009] [Medline: 27659122]
18. Mueller RC. Exploring family nurse practitioners' practices in recommending mhealth apps to patients. CIN: Computers, Informatics, Nursing 2020;38(2):71-79. [doi: 10.1097/cin.0000000000000580]
19. Bloomfield RA, Polo-Wood F, Mandel JC, Mandl KD. Opening the Duke electronic health record to apps: Implementing SMART on FHIR. International Journal of Medical Informatics 2017 Mar;99:1-10. [doi: 10.1016/j.ijmedinf.2016.12.005]
20. Cho J. The impact of post-adoption beliefs on the continued use of health apps. Int J Med Inform 2016 Mar;87:75-83. [doi: 10.1016/j.ijmedinf.2015.12.016] [Medline: 26806714]
21. Bort-Roig J, Puig-Ribera A, Contreras RS, Chirveches-Pérez E, Martori JC, Gilson ND, et al. Monitoring sedentary patterns in office employees: validity of an m-health tool (Walk@Work-App) for occupational health. Gaceta Sanitaria 2018 Nov;32(6):563-566. [doi: 10.1016/j.gaceta.2017.05.004]
22. Ashurst EJ, Jones RB. Is the Health App Challenge approach of patient-led application conception, development, and review worthwhile? Health Policy and Technology 2017 Mar;6(1):83-92. [doi: 10.1016/j.hlpt.2016.12.001]
23. Akdur G, Aydin MN, Akdur G. Adoption of Mobile Health Apps in Dietetic Practice: Case Study of Diyetkolik. JMIR Mhealth Uhealth 2020 Oct 02;8(10):e16911-e16912 e16911 [FREE Full text] [doi: 10.2196/16911] [Medline: 33006566]
24. Deng Z, Hong Z, Ren C, Zhang W, Xiang F. What Predicts Patients' Adoption Intention Toward mHealth Services in China: Empirical Study. JMIR Mhealth Uhealth 2018 Aug 29;6(8):e172 [FREE Full text] [doi: 10.2196/mhealth.9316] [Medline: 30158101]
25. Mao X, Zhao X, Liu Y. mHealth App recommendation based on the prediction of suitable behavior change techniques. Decision Support Systems 2020 May;132:113248. [doi: 10.1016/j.dss.2020.113248]
26. Autry CW, Grawe SJ, Daugherty PJ, Richey RG. The effects of technological turbulence and breadth on supply chain technology acceptance and adoption. Journal of Operations Management 2010 Mar 27;28(6):522-536. [doi: 10.1016/j.jom.2010.03.001]
27. Au YA, Zafar H. A multi-country assessment of mobile payment adoption. Working Paper Series 0055IS-296-2008. 2008. URL: https://tinyurl.com/nxdheksp [accessed 2021-06-01]
28. Chen L, Tan J. Technology Adaptation in E-commerce. European Management Journal 2004 Feb;22(1):74-86. [doi: 10.1016/j.emj.2003.11.014]
29. Cheung R, Vogel D. Predicting user acceptance of collaborative technologies: An extension of the technology acceptance model for e-learning. Computers & Education 2013 Apr;63:160-175. [doi: 10.1016/j.compedu.2012.12.003]
30. Joo J, Sang Y. Exploring Koreans' smartphone usage: An integrated model of the technology acceptance model and uses and gratifications theory. Computers in Human Behavior 2013 Nov;29(6):2512-2518. [doi: 10.1016/j.chb.2013.06.002]
31. Son H, Park Y, Kim C, Chou J. Toward an understanding of construction professionals' acceptance of mobile computing devices in South Korea: An extension of the technology acceptance model. Automation in Construction 2012 Dec;28:82-90. [doi: 10.1016/j.autcon.2012.07.002]
32. Venkatesh V, Davis FD. A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Management Science 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
33. Gefen D, Karahanna E, Straub DW. Trust and TAM in Online Shopping: An Integrated Model. MIS Quarterly 2003;27(1):51. [doi: 10.2307/30036519]
34. Hong S, Thong JY, Tam KY. Understanding continued information technology usage behavior: A comparison of three models in the context of mobile internet. Decision Support Systems 2006 Dec;42(3):1819-1834. [doi: 10.1016/j.dss.2006.03.009]
35. Venkatesh V, Bala H. Technology Acceptance Model 3 and a Research Agenda on Interventions. Decision Sciences 2008 May;39(2):273-315. [doi: 10.1111/j.1540-5915.2008.00192.x]
36. Venkatesh V, Morris MG, Davis GB, Davis FD. User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly 2003;27(3):425. [doi: 10.2307/30036540]
37. Wixom BH, Todd PA. A Theoretical Integration of User Satisfaction and Technology Acceptance. Information Systems Research 2005 Mar;16(1):85-102. [doi: 10.1287/isre.1050.0042]
38. Rondan-Cataluña J, Arenas-Gaitán J, Ramírez-Correa P. A comparison of the different versions of popular technology acceptance models. Kybernetes 2015;44(5):788-805. [doi: 10.1108/k-09-2014-0184]

39. Scherer R, Siddiq F, Tondeur J. The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. Computers & Education 2019 Jan;128:13-35. [doi: 10.1016/j.compedu.2018.09.009]
40. Briz-Ponce L, García-Peñalvo FJ. An Empirical Assessment of a Technology Acceptance Model for Apps in Medical Education. J Med Syst 2015 Nov 28;39(11):176-175. [doi: 10.1007/s10916-015-0352-x] [Medline: 26411928]
41. Wang Y, Wang S, Wang J, Wei J, Wang C. An empirical study of consumers' intention to use ride-sharing services: using an extended technology acceptance model. Transportation 2018 May 30;47(1):397-415. [doi: 10.1007/s11116-018-9893-4]
42. Venkatesh V, Davis FD. A Theoretical Extension of the Technology Acceptance Model: Four Longitudinal Field Studies. Management Science 2000 Feb;46(2):186-204. [doi: 10.1287/mnsc.46.2.186.11926]
43. Venkatesh V, Davis FD. A Model of the Antecedents of Perceived Ease of Use: Development and Test. Decision Sciences 1996 Sep;27(3):451-481. [doi: 10.1111/j.1540-5915.1996.tb00860.x]
44. Venkatesh V. Determinants of Perceived Ease of Use: Integrating Control, Intrinsic Motivation, and Emotion into the Technology Acceptance Model. Information Systems Research 2000 Dec;11(4):342-365. [doi: 10.1287/isre.11.4.342.11872]
45. Legris P, Ingham J, Collerette P. Why do people use information technology? A critical review of the technology acceptance model. Information & Management 2003 Jan;40(3):191-204. [doi: 10.1016/S0378-7206(01)00143-4]
46. Tang D, Chen L. A review of the evolution of research on information Technology Acceptance Model. New York, NY: IEEE; 2011 May Presented at: 2011 International Conference on Business Management and Electronic Information; May 13-15, 2011; Guangzhou, China p. 588-591. [doi: 10.1109/icbmei.2011.5917980]
47. Ammenwerth E. Technology Acceptance Models in Health Informatics: TAM and UTAUT. Stud Health Technol Inform 2019 Jul 30;263:64-71. [doi: 10.3233/SHTI190111] [Medline: 31411153]
48. Cimperman M, Makovec Brenčič M, Trkman P. Analyzing older users' home telehealth services acceptance behavior-applying an Extended UTAUT model. Int J Med Inform 2016 Jun;90:22-31. [doi: 10.1016/j.ijmedinf.2016.03.002] [Medline: 27103194]
49. Hung M, Jen W. The adoption of mobile health management services: an empirical study. J Med Syst 2012 Jun 29;36(3):1381-1388. [doi: 10.1007/s10916-010-9600-2] [Medline: 20878452]
50. Chau PY, Hu PJ. Investigating healthcare professionals' decisions to accept telemedicine technology: an empirical test of competing theories. Information & Management 2002 Jan;39(4):297-311. [doi: 10.1016/s0378-7206(01)00098-2]
51. Beldad AD, Hegner SM. Expanding the Technology Acceptance Model with the Inclusion of Trust, Social Influence, and Health Valuation to Determine the Predictors of German Users' Willingness to Continue using a Fitness App: A Structural Equation Modeling Approach. International Journal of Human–Computer Interaction 2017 Nov 30;34(9):882-893. [doi: 10.1080/10447318.2017.1403220]
52. Middlemass JB, Vos J, Siriwardena AN. Perceptions on use of home telemonitoring in patients with long term conditions - concordance with the Health Information Technology Acceptance Model: a qualitative collective case study. BMC Med Inform Decis Mak 2017 Jun 26;17(1):89-13 [FREE Full text] [doi: 10.1186/s12911-017-0486-5] [Medline: 28651588]
53. Ake A, Arcand M. The impact of mobile health monitoring on the evolution of patient-pharmacist relationships. IJPHM 2020 Jan 28;14(1):1-19. [doi: 10.1108/ijphm-04-2019-0030]
54. Thies K, Anderson D, Cramer B. Lack of Adoption of a Mobile App to Support Patient Self-Management of Diabetes and Hypertension in a Federally Qualified Health Center: Interview Analysis of Staff and Patients in a Failed Randomized Trial. JMIR Hum Factors 2017 Oct 03;4(4):e24. [doi: 10.2196/humanfactors.7709] [Medline: 28974481]
55. Paré G, Leaver C, Bourget C. Diffusion of the Digital Health Self-Tracking Movement in Canada: Results of a National Survey. J Med Internet Res 2018 May 02;20(5):e177 [FREE Full text] [doi: 10.2196/jmir.9388] [Medline: 29720359]
56. Kim J, Park H. Development of a health information technology acceptance model using consumers' health behavior intention. J Med Internet Res 2012 Oct 01;14(5):e133 [FREE Full text] [doi: 10.2196/jmir.2143] [Medline: 23026508]
57. Jeon E, Park H. Factors affecting acceptance of smartphone application for management of obesity. Healthc Inform Res 2015 Apr;21(2):74-82 [FREE Full text] [doi: 10.4258/hir.2015.21.2.74] [Medline: 25995959]
58. de Veer AJE, Peeters JM, Brabers AE, Schellevis FG, Rademakers JJJ, Francke AL. Determinants of the intention to use e-Health by community dwelling older people. BMC Health Serv Res 2015 Mar 15;15(1):103-109 [FREE Full text] [doi: 10.1186/s12913-015-0765-8] [Medline: 25889884]
59. Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]
60. Deng Z. Understanding public users' adoption of mobile health service. IJMC 2013;11(4):351-373. [doi: 10.1504/ijmc.2013.055748]
61. Thompson RL, Higgins CA, Howell JM. Personal Computing: Toward a Conceptual Model of Utilization. MIS Quarterly 1991 Mar;15(1):125. [doi: 10.2307/249443]
62. Sun Y, Wang, X, Guo X, Peng Z. Understanding the acceptance of mobile health services: a comparison integration of alternative models? J. Electron. Commer. Res 2013;14(2):183-200 [FREE Full text]
63. Chauhan S, Jaiswal M. A meta-analysis of e-health applications acceptance. JEIM 2017 Mar 06;30(2):295-319. [doi: 10.1108/JEIM-08-2015-0078]

64. Davis FD. User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. International Journal of Man-Machine Studies 1993 Mar;38(3):475-487. [doi: 10.1006/imms.1993.1022]
65. Chang M, Pang C, Michael Tarn J, Liu T, Yen DC. Exploring user acceptance of an e-hospital service: An empirical study in Taiwan. Computer Standards & Interfaces 2015 Feb;38:35-43. [doi: 10.1016/j.csi.2014.08.004]
66. Klein R. An empirical examination of patient-physician portal acceptance. European Journal of Information Systems 2017 Dec 19;16(6):751-760. [doi: 10.1057/palgrave.ejis.3000719]
67. Moores TT. Towards an integrated model of IT acceptance in healthcare. Decision Support Systems 2012 Jun;53(3):507-516. [doi: 10.1016/j.dss.2012.04.014]
68. Yi MY, Jackson JD, Park JS, Probst JC. Understanding information technology acceptance by individual professionals: Toward an integrative view. Information & Management 2006 Apr;43(3):350-363. [doi: 10.1016/j.im.2005.08.006]
69. Lin SP, Yang HY. Exploring key factors in the choice of e-health using an asthma care mobile service model. Telemed J E Health 2009 Nov;15(9):884-890. [doi: 10.1089/tmj.2009.0047] [Medline: 19919195]
70. Buntin MB, Burke MF, Hoaglin MC, Blumenthal D. The benefits of health information technology: a review of the recent literature shows predominantly positive results. Health Aff (Millwood) 2011 Mar;30(3):464-471. [doi: 10.1377/hlthaff.2011.0178] [Medline: 21383365]
71. Melzner J, Heinze J, Fritsch T. Mobile Health Applications in Workplace Health Promotion: An Integrated Conceptual Adoption Framework. Procedia Technology 2014;16:1374-1382. [doi: 10.1016/j.protcy.2014.10.155]
72. Kelly M, Wills J, Sykes S. Do nurses' personal health behaviours impact on their health promotion practice? A systematic review. Int J Nurs Stud 2017 Nov;76:62-77. [doi: 10.1016/j.ijnurstu.2017.08.008] [Medline: 28938104]
73. Ramtohul I. The adoption of e-health services: Comprehensive analysis of the adoption setting from the user s perspective. Health Policy and Technology 2015 Sep;4(3):286-293. [doi: 10.1016/j.hlpt.2015.04.007]
74. Bert F, Giacometti M, Gualano MR, Siliquini R. Smartphones and health promotion: a review of the evidence. J Med Syst 2014 Jan 16;38(1):9995-9911. [doi: 10.1007/s10916-013-9995-7] [Medline: 24346929]
75. Crilly P, Hassanali W, Khanna G, Matharu K, Patel D, Patel D, et al. Community pharmacist perceptions of their role and the use of social media and mobile health applications as tools in public health. Res Social Adm Pharm 2019 Jan;15(1):23-30. [doi: 10.1016/j.sapharm.2018.02.005] [Medline: 29501431]
76. Shabani J, Damavandi AJ. The Importance of Gender as a Moderator for the Relationship between Emotional Intelligence and Mental Health of Adolescents. ASS 2011 Aug 31;7(9):142-148. [doi: 10.5539/ass.v7n9p142]
77. Bidmon S, Terlutter R, Röttl J. What explains usage of mobile physician-rating apps? Results from a web-based questionnaire. J Med Internet Res 2014 Jun 11;16(6):e148 [FREE Full text] [doi: 10.2196/jmir.3122] [Medline: 24918859]
78. Van Dyck D, Cerin E, De Bourdeaudhuij I, Salvo D, Christiansen LB, Macfarlane D, et al. Moderating effects of age, gender and education on the associations of perceived neighborhood environment attributes with accelerometer-based physical activity: The IPEN adult study. Health Place 2015 Nov;36:65-73 [FREE Full text] [doi: 10.1016/j.healthplace.2015.09.007] [Medline: 26454247]
79. MacKenzie, Podsakoff, Podsakoff. Construct Measurement and Validation Procedures in MIS and Behavioral Research: Integrating New and Existing Techniques. MIS Quarterly 2011;35(2):293. [doi: 10.2307/23044045]
80. Sarstedt M, Hair JF, Ringle CM, Thiele KO, Gudergan SP. Estimation issues with PLS and CBSEM: Where the bias lies!. Journal of Business Research 2016 Oct;69(10):3998-4010. [doi: 10.1016/j.jbusres.2016.06.007]
81. Henseler J, Hubona G, Ray PA. PLS path modeling in new technology research: updated guidelines. Ind. Manag. Data Syst 2016;116(1):2-20. [doi: 10.1108/imds-09-2015-0382]
82. Son J, Benbasat I. Organizational Buyers' Adoption and Use of B2B Electronic Marketplaces: Efficiency- and Legitimacy-Oriented Perspectives. Journal of Management Information Systems 2014 Dec 08;24(1):55-99. [doi: 10.2753/mis0742-1222240102]
83. Chin WW, Newsted PR. Structural equation modeling analysis with small samples using partial least squares. In: Hoyle RH, editor. Statistical Strategies for Small Sample Research. Thousand Oaks, CA: SAGE Publications; 1999:307-341.
84. Ringle CM, Wende S, Becker JM. SmartPLS 3. Boenningstedt, Germany: SmartPLS GmbH; 2015. URL: http://www.smartpls.com [accessed 2021-06-01]
85. Reinartz W, Haenlein M, Henseler J. An empirical comparison of the efficacy of covariance-based and variance-based SEM. International Journal of Research in Marketing 2009 Dec;26(4):332-344. [doi: 10.1016/j.ijresmar.2009.08.001]
86. Hair JF, Ringle CM, Sarstedt M. PLS-SEM: Indeed a Silver Bullet. Journal of Marketing Theory and Practice 2014 Dec 08;19(2):139-152. [doi: 10.2753/mtp1069-6679190202]
87. Saura JR, Palos-Sanchez P, Blanco-González A. The importance of information service offerings of collaborative CRMs on decision-making in B2B marketing. JBIM 2019 Sep 23;35(3):470-482. [doi: 10.1108/jbim-12-2018-0412]
88. Fornell C, Larcker DF. Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. Journal of Marketing Research 2018 Nov 28;18(1):39-50. [doi: 10.1177/002224378101800104]
89. Hair JF, Black B, Babin BJ, Anderson RE, Tatham R. Multivariate Data Analysis (5th Edition). Upper Saddle River, NJ: Prentice Hall; 2015.
90. Fornell C, Bookstein FL. Two Structural Equation Models: LISREL and PLS Applied to Consumer Exit-Voice Theory. Journal of Marketing Research 2018 Dec 20;19(4):440-452. [doi: 10.1177/002224378201900406]

91. Henseler J. Bridging Design and Behavioral Research With Variance-Based Structural Equation Modeling. Journal of Advertising 2017 Feb 07;46(1):178-192. [doi: 10.1080/00913367.2017.1281780]
92. Rigdon EE, Sarstedt M, Ringle CM. On Comparing Results from CB-SEM and PLS-SEM: Five Perspectives and Five Recommendations. MAR 2017;39(3):4-16. [doi: 10.15358/0344-1369-2017-3-4]
93. Cepeda-Carrion G, Cegarra-Navarro J, Cillo V. Tips to use partial least squares structural equation modelling (PLS-SEM) in knowledge management. JKM 2019 Jan 14;23(1):67-89. [doi: 10.1108/jkm-05-2018-0322]
94. Guo Z, Xiao L, Van Toorn C, Lai Y, Seo C. Promoting online learners' continuance intention: An integrated flow framework. Information & Management 2016 Mar;53(2):279-295. [doi: 10.1016/j.im.2015.10.010]
95. Kim J, Park EH, Baskerville RL. A model of emotion and computer abuse. Information & Management 2016 Jan;53(1):91-108. [doi: 10.1016/j.im.2015.09.003]
96. Hair JF, Ringle CM, Sarstedt M. Partial Least Squares Structural Equation Modeling: Rigorous Applications, Better Results and Higher Acceptance. Long Range Planning 2013 Feb;46(1-2):1-12. [doi: 10.1016/j.lrp.2013.01.001]
97. Werts C, Linn R, Jöreskog K. Intraclass Reliability Estimates: Testing Structural Assumptions. Educational and Psychological Measurement 2016 Jul 02;34(1):25-33. [doi: 10.1177/001316447403400104]
98. Bagozzi RP, Yi Y. On the evaluation of structural equation models. JAMS 1988 Mar;16(1):74-94. [doi: 10.1007/bf02723327]
99. Tenenhaus M, Vinzi VE, Chatelin Y, Lauro C. PLS path modeling. Computational Statistics & Data Analysis 2005 Jan;48(1):159-205. [doi: 10.1016/j.csda.2004.03.005]
100. Hair JF, Sarstedt M, Ringle CM, Mena JA. An assessment of the use of partial least squares structural equation modeling in marketing research. J Acad Mark Sci 2011 Jun 7;40(3):414-433. [doi: 10.1007/s11747-011-0261-6]
101. Chin WW. The partial least squares approach to structural equation modeling. In: Marcoulides GA, editor. Modern Methods for Business Research. East Sussex, UK: Psychology Press; 1998:295-336.
102. Barclay D, Higgins C, Thompson R. The Partial Least Squares (PLS) Approach to Causal Modelling: Personal Computer Adoption and Use as an Illustration. Technol Stud 1995;2(2):285-309.
103. Hu L, Bentler PM. Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychological Methods 1998;3(4):424-453. [doi: 10.1037/1082-989x.3.4.424]
104. Benitez J, Henseler J, Castillo A, Schuberth F. How to perform and report an impactful analysis using partial least squares: Guidelines for confirmatory and explanatory IS research. Information & Management 2020 Mar;57(2):103168. [doi: 10.1016/j.im.2019.05.003]
105. Stone M. Cross-Validatory Choice and Assessment of Statistical Predictions. Journal of the Royal Statistical Society: Series B (Methodological) 2018 Dec 05;36(2):111-133. [doi: 10.1111/j.2517-6161.1974.tb00994.x]
106. Geisser S. A Predictive Approach to the Random Effects Model. Biometrika 1974;61(1):101-107. [doi: 10.2307/2334290]
107. Cohen J. Quantitative methods in psychology: A power primer. Psychol. Bull 1992;112(1):1155-1159 [FREE Full text]
108. Henseler J. Testing moderating effects in PLS path models: An illustration of available procedures. In: Fassott G, editor. Handbook of Partial Least Squares. Berlin, Heidelberg: Springer; 2010:713-735.
109. Felipe C, Roldán J, Leal-Rodríguez A. Impact of Organizational Culture Values on Organizational Agility. Sustainability 2017 Dec 17;9(12):2354. [doi: 10.3390/su9122354]
110. Zhao Y, Ni Q, Zhou R. What factors influence the mobile health service adoption? A meta-analysis and the moderating role of age. International Journal of Information Management 2018 Dec;43:342-350. [doi: 10.1016/j.ijinfomgt.2017.08.006]
111. Yuan S, Ma W, Kanthawala S, Peng W. Keep Using My Health Apps: Discover Users' Perception of Health and Fitness Apps with the UTAUT2 Model. Telemed J E Health 2015 Sep;21(9):735-741. [doi: 10.1089/tmj.2014.0148] [Medline: 25919238]
112. Huang H, Bashir M. Users' Adoption of Mental Health Apps: Examining the Impact of Information Cues. JMIR Mhealth Uhealth 2017 Jun 28;5(6):e83 [FREE Full text] [doi: 10.2196/mhealth.6827] [Medline: 28659256]
113. Balapour A, Reychav I, Sabherwal R, Azuri J. Mobile technology identity and self-efficacy: Implications for the adoption of clinically supported mobile health apps. International Journal of Information Management 2019 Dec;49:58-68. [doi: 10.1016/j.ijinfomgt.2019.03.005]
114. Luxton DD, McCann RA, Bush NE, Mishkind MC, Reger GM. mHealth for mental health: Integrating smartphone technology in behavioral healthcare. Professional Psychology: Research and Practice 2011 Dec;42(6):505-512. [doi: 10.1037/a0024485]
115. Lewis TL, Wyatt JC. mHealth and mobile medical Apps: a framework to assess risk and promote safer use. J Med Internet Res 2014 Sep 15;16(9):e210 [FREE Full text] [doi: 10.2196/jmir.3133] [Medline: 25223398]
116. Gleason AW. mHealth — Opportunities for Transforming Global Health Care and Barriers to Adoption. Journal of Electronic Resources in Medical Libraries 2015 Jun 19;12(2):114-125. [doi: 10.1080/15424065.2015.1035565]

Abbreviations

PLS: partial least squares
SEM: structural equation modeling
TAM: technology acceptance model

Edited by R Kukafka; submitted 07.01.21; peer-reviewed by A Infante Moro, S Rostam Niakan Kalhori; comments to author 10.02.21; revised version received 18.03.21; accepted 30.05.21; published 09.09.21.

Please cite as:
Palos-Sanchez PR, Saura JR, Rios Martin MÁ, Aguayo-Camacho M
Toward a Better Understanding of the Intention to Use mHealth Apps: Exploratory Study
JMIR Mhealth Uhealth 2021;9(9):e27021
URL: https://mhealth.jmir.org/2021/9/e27021
doi:10.2196/27021
PMID:34499044

Original Paper

Nursing Interns' Attitudes Toward, Preferences for, and Use of Diabetes Virtual Simulation Teaching Applications in China: National Web-Based Survey

Fang Liu[1,2], RN, MD; Huiting Weng[1,2], RN, MNurs; Rong Xu[2], RN, MNurs; Xia Li[1], MD; Zhe Zhang[1,2], RN; Kuaile Zhao[1,2], RN; Zhiguang Zhou[1], MD; Qin Wang[2,3], RN, MNurs

[1]National Clinical Research Center for Metabolic Diseases, Department of Metabolism and Endocrinology, The Second Xiangya Hospital, Central South University, Changsha, China

[2]Clinical Nursing Teaching and Research Section, The Second Xiangya Hospital, Central South University, Changsha, China

[3]Department of Ophthalmology, The Second Xiangya Hospital, Central South University, Changsha, China

Corresponding Author:
Qin Wang, RN, MNurs
Clinical Nursing Teaching and Research Section
The Second Xiangya Hospital
Central South University
139 Middle Renmin Road
Changsha, 410011
China
Phone: 86 18774806226
Email: wangqin3421@csu.edu.cn

Abstract

Background: Diabetes has placed heavy social and economic burdens on society and families worldwide. Insufficient knowledge and training of frontline medical staff, such as nurses, interns, and residents, may lead to an increase in acute and chronic complications among patients with diabetes. However, interns have insufficient knowledge about diabetes management. The factors that affect interns' current level of diabetes-related knowledge are still unclear. Therefore, understanding the behavioral intentions of interns is essential to supporting the development and promotion of the use of virtual simulation teaching applications.

Objective: This study aimed to identify the determinants of nursing interns' intentions to use simulation-based education applications.

Methods: From December 1, 2020, to February 28, 2021, the web-based survey tool Sojump (Changsha Xingxin Information Technology Co) was used to survey nursing interns in hospitals across China. Two survey links were sent to 37 partner schools in 23 major cities in China, and they were disseminated through participants' WeChat networks. Multiple regression analysis was used to determine the association between demographic information and basic disease information and the use of the application for treating adult patients.

Results: Overall, 883 nursing interns from 23 provinces in China responded to the survey. Among them, the virtual simulation utilization rate was 35.6% (314/883) and the awareness rate was 10.2% (90/883). In addition, among the interns, only 10.2% (90/883) correctly understood the concept of virtual simulation, and most of them (793/883, 89.8%) believed that scenario-simulation training or the use of models for teaching are all the same. Multiple regression analysis showed that the educational level, independent learning ability, and professional identity of the interns were related to use of the application (P<.05). Skills and knowledge that the interns most wanted to acquire included the treatment of hypoglycemia (626/883, 70.9%), functional test simulation (610/883, 69.1%), and blood glucose monitoring technology (485/883, 54.9%). A total of 60.5% (534/883) of the interns wanted to acquire clinical thinking skills, while 16.0% (141/883) wanted to acquire operational skills. Nursing trainees believed that the greatest obstacles to virtual simulation included limited time (280/883, 31.7%), the degree of simulation (129/883, 14.6%), the demand for satisfaction (108/883, 12.2%), and test scores (66/883, 7.5%).

Conclusions: The understanding and usage rate of diabetes virtual simulation teaching applications by Chinese nursing interns is very low. However, they have high requirements regarding this teaching method. Conducting high-quality randomized controlled trials and designing applications that are suitable for the needs of different nurse trainees will increase students' interest in learning and help improve diabetes knowledge among nursing interns.

(JMIR Mhealth Uhealth 2021;9(9):e29498) doi:10.2196/29498

KEYWORDS

nursing interns; virtual simulation; China; nursing education; diabetes

Introduction

Background

With the rapid development of the social economy, continuous changes in modern people's behaviors and lifestyles, and the aging of the population, the incidence of diabetes is also increasing rapidly in all parts of the world. According to forecasts, from 1995 to 2030, the number of patients with diabetes worldwide will increase from 135 million to 472 million, among which more than 75% are in developing countries [1]. According to the World Health Organization [2], as of 2015, the prevalence of diabetes in China had reached 10%, and the estimated prevalence of diabetes in adults over 18 years old in China was 11.6%. The number of adult patients with diabetes in China has reached 92.4 million [3]; China is now the country with the largest number of patients with diabetes.

Effective diabetes education for patients is indispensable; it is a necessary means to ensure that patients receive effective therapies. Nurses are the most important providers of diabetes education in China. Due to China's national conditions, there are no specialized diabetes educators that provide health guidance and dietary education to patients with diabetes. Such work is often undertaken by clinical staff (doctors, nurses, interns, etc). Nurses and interns have the most contact with patients with diabetes and are most likely to provide patients with diabetes-related knowledge [4]. In nursing programs, Chinese students often need to go to designated hospitals for 8- to 12-month internships in their senior year. During these internships, they learn basic knowledge about diseases, professional skills, and communication skills at the hospital. At the end of the internship, they need to pass a unified examination jointly organized by the hospital and the school before they can graduate. The clinical internship is a critical period for nursing students to transition from student to nurse. At this stage, interns often cannot perform nursing activities alone, and all of these require the demonstration and guidance of clinical teachers. In most clinical teaching sessions, the teacher and the interns are apprentices. In addition to regular and unified theoretical training and operational training in the nursing department, teachers often use a one-to-one teaching mode. The importance of nursing interns in disease prevention is enormous, as they are the members of health care teams who spend the most time with the patients [5]. They also serve as resources for patients with diabetes seeking information about the early detection of diabetes complications [6]. The knowledge and practice they acquire during their studies and internships play important roles in providing accurate and up-to-date information to improve the health behaviors and outcomes of patients with diabetes. Nursing interns must possess the necessary knowledge to enable them to care for patients with diabetes, helping them to achieve a high quality of life devoid of complications [7]. In addition, an intern can detect hypoglycemia for the first time when measuring a patient's blood sugar. If nursing interns know how to deal with hypoglycemic events, they can instruct patients to eat the correct glucose-increasing food immediately, thereby reducing the time interval from discovery to treatment of hypoglycemia and reducing the occurrence of adverse events [8].

The cultivation of self-learning ability by interns is inseparable from the application of self-learning methods and tools [9]. With the goal of improving the self-learning ability of nursing students, many scholars at home and abroad have carried out various studies in this field. Some studies have shown that project-based learning methods, preceptorship programs, and reflective diaries have improved students' abilities for critical thinking, clinical decision making, and humanistic care [10-13]. However, these methods focus on cultivating students' information-seeking and cooperation abilities in order to enhance the autonomous learning ability of interns, and the effect is not lasting. Without the supervision of teachers, the internal motivation of students to learn independently is still insufficient [14].

Interns' apprenticeships have always constituted a challenge faced by the government, health educators, health managers, and the students themselves to ensure the quality and safety of learning and clinical practice [15]. Students in the 21st century are using information and communications technology (ICT) every day [16,17]. The use of ICT has led to different learning processes and information structure processes. The development of digital and virtual technology has simplified the ability to reconstruct reality using virtual patients depicted on a computer touch screen (ie, virtual simulation) [18].

A virtual simulation is a real-life reproduction depicted on a computer screen, and it involves a real person operating the simulation system. This type of simulation puts people at the center of a situation by exercising decision making, motor control, and communication skills [19]. Virtual simulation uses virtual patients in dynamic and immersive clinical environments, ranging from prehospital to community environments [20]. The latest technological advances in virtual simulation have improved their authenticity and dynamic interaction, and it is possible to display thousands of clinical situations on a touch screen or on the web [21,22]. However, little is known about their effect on students' learning satisfaction, self-efficacy, knowledge retention, and clinical reasoning, especially when using the latest developments in virtual simulation [21].

This study aimed to assess the knowledge needs of nursing students for managing diabetes mellitus. By evaluating the self-learning ability of nursing students and the degree of demand for diabetes-related knowledge, the demand for virtual simulation teaching applications for nursing students was explored. Therefore, the purpose of this study was to evaluate the level of understanding of diabetes specialist knowledge and

the demand for virtual simulation teaching among nursing students in China.

Objectives

We aimed to investigate the use of virtual simulation teaching applications by nursing interns as well as their perspectives, attitudes, and associated factors regarding these teaching applications. We also aimed to investigate interns' needs for these applications in order to provide information for the design of virtual simulation teaching applications and to learn how best to promote their use, which will help teachers to further improve their teaching methods and strengthen the willingness of nursing students to learn independently.

Methods

Questionnaire Design

An expert group consisting of five nursing educators and five clinical nursing staff members searched for applications on the national, virtual simulation, education platform; they then designed a questionnaire based on the current diabetes guidelines and the problems encountered in clinical practice. These questions were presented in a selective format. If the respondent disagreed with the listed options, they could select "other options" and write their answer in the "remarks" column. The questionnaire collected information about respondents' demographics and their views, attitudes, and needs for virtual simulation education applications.

To determine the validity of the questionnaire content, a total of 15 experts, consisting of 12 nursing education experts and three diabetes education nurses with at least 5 years of experience, rated the relevance and clarity of the items on a 4-point scale ranging from 1 (irrelevant) to 4 (highly relevant), with a content validity index of 0.91. Before administering the questionnaire survey, we conducted a pilot test on 18 interns at Xiangya Second Hospital in China. The Cronbach α value of the questionnaire was .83.

Survey Platform and Methods

WeChat has become one of the largest mobile traffic platforms in China. It provides many services, including messaging, free phone calls, browsing and publishing for instant sharing of information, and mobile payments [23]. It has been installed on more than 90% of mobile phones and has become part of the daily lives of most people [24]. As of 2019, the number of monthly active accounts on WeChat reached 1.15 billion, and the number of daily active accounts of mini programs exceeded 300 million [25]. As the most commonly used social media tool in China, WeChat has an expansive network of contacts. The network makes it possible for administrators to manage questionnaires through WeChat.

From December 1, 2020, to February 28, 2021, we used Sojump (Changsha Xingxin Information Technology Co), a web-based survey tool, to conduct snowball sampling through the WeChat contact network and to conduct convenience sampling through WeChat public accounts to recruit interns. The survey link was initially sent to 35 universities in 23 representative major cities in China. We asked the teachers at these universities to post the survey link on their WeChat account to reach their network contacts.

Survey respondents were all nursing trainees in China. Other nursing students who did not take part in internships at hospitals were excluded from our survey. Before administering the survey, we introduced the purpose of the survey, and the questionnaire was filled out by respondents voluntarily without any compensation.

Ethical Approval

This study was approved by the ethics committee of the Second Xiangya Hospital, Central South University, China (ID: 2020-S790).

Statistical Data

The data were analyzed using SPSS, version 23.0 (IBM Corp). Quantile-quantile (Q-Q) charts were used to check the normality of all continuous variables and express them as the mean (SD) or median (IQR) where appropriate. Categorical variables were expressed as frequencies and percentages. The chi-square test was used to assess the differences between groups. The generalized logic model was used to obtain the odds ratio (OR) and its 95% CI at the same time. First, we conducted a univariate analysis to analyze the OR of the potential correlation between demographic factors and autonomous learning ability. Then, we inputted all important factors into the multivariate analysis to obtain the multivariate adjusted OR. Questionnaires with missing values were excluded from the multivariate analysis. Statistical significance was defined as $P<.05$.

Results

Sample Characteristics

A total of 883 interns distributed among 26 provinces in China (Figure 1) responded to the patient survey. The respondents' characteristics are shown in Table 1. Among the respondents, 10.1% (89/883) were male, and respondents had a mean age of 20.64 (SD 2.1) years. Overall, 56.9% (502/883) had a bachelor's degree. A total of 83.0% (733/883) of the respondents had been an intern for more than 8 months, and 46.5% (411/883) did not know their reason for choosing to study nursing (Table 1).

Figure 1. Distribution of the nursing intern sample in China by province. The numbers represent how many questionnaires were collected in each corresponding province.

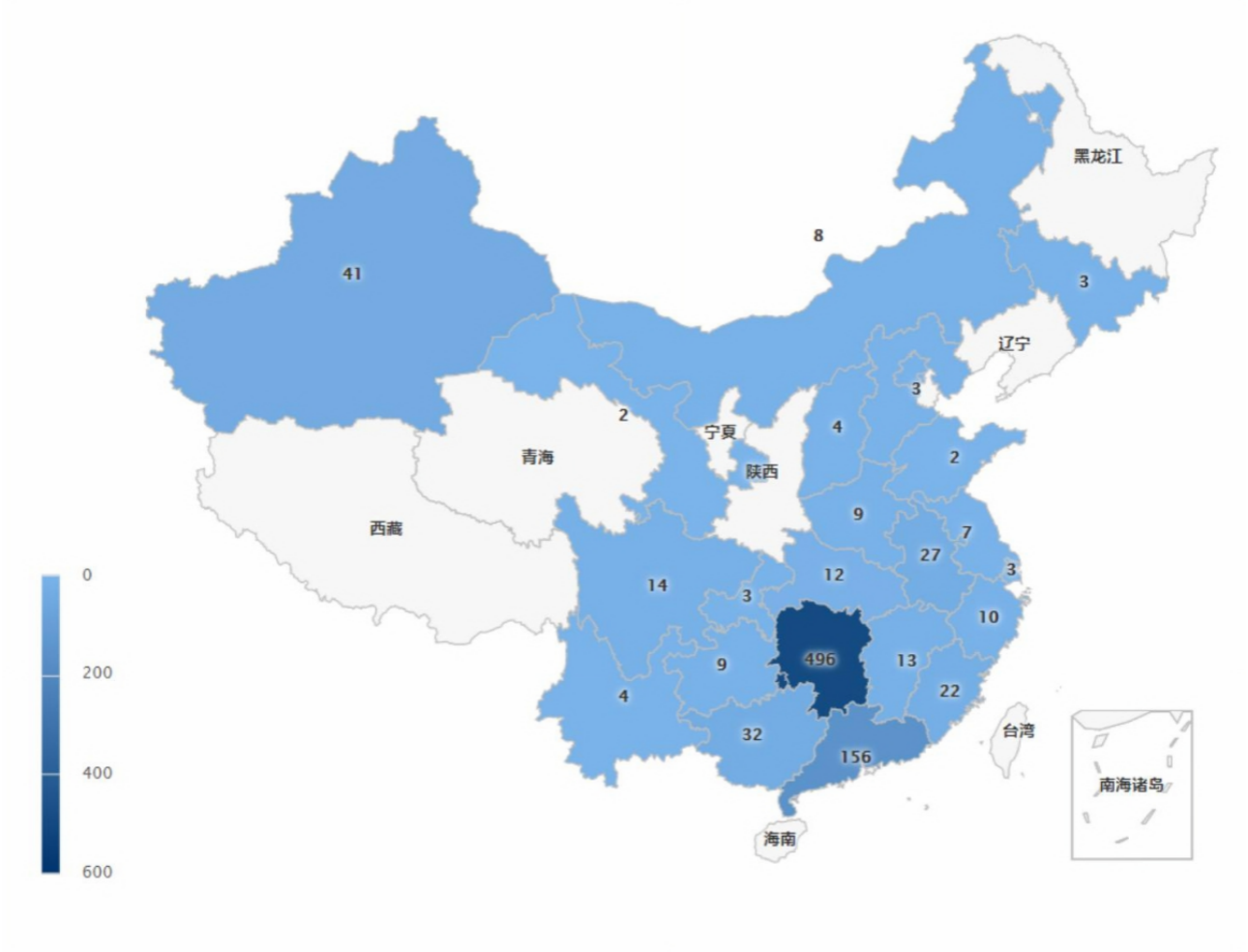


Table 1. Characteristics of nursing interns.

Characteristic	Respondents (N=883)
Gender, n (%)	
Male	89 (10.1)
Female	794 (89.9)
Age (years), mean (SD)	20.64 (2.1)
Educational level, n (%)	
Middle school	26 (2.9)
High school	102 (11.6)
Technical college	502 (56.9)
Bachelor's degree	242 (27.4)
Master's degree or higher	11 (1.2)
Internship time (months), mean (SD)	6.02 (1.6)
Reason for choosing nursing, n (%)	
I like nursing	287 (32.5)
Parents' suggestion	247 (28.0)
Acquaintances' recommendation	55 (6.2)
The school transferred me	99 (11.2)
Good employment	195 (22.1)
Feelings about nursing, n (%)	
I love the nursing career	410 (46.4)
Not sure	197 (22.3)
I can accept as a job, but not as a career	260 (29.4)
I don't like nursing	16 (1.8)
Employment intention, n (%)	
Nurse	747 (84.6)
Nursing-related industries	115 (13.0)
Others	21 (2.4)

Assessment of the Self-Learning Ability of Interns

All of the interns (N=883) were able to fill out the self-learning ability scale. The Q-Q normality was the sum of the total scores of the self-learning ability of interns (Figures 2 and 3). The total score is represented by the diagonal line, so it is considered that the total score of the autonomous learning ability of nursing students conforms to the normal distribution. The data were analyzed using the Pearson correlation coefficient. Age, gender, educational level, and length of internship were not related to the self-learning ability of interns (P>.05). The correlation coefficient between the "reason for choosing nursing" and the "autonomous learning ability scale score" was 0.993; the correlation between them was statistically significant (P<.001). This correlation was also reflected with "feelings about nursing" (P=.02), which showed that interns who love nursing had stronger self-learning ability. In addition, the correlation between "feelings about nursing" and the "score of the learning strategy scale" was statistically significant (P=.001). This indicates that the more positive feelings the nursing student interns had toward the nursing profession, the higher their scores were on the learning strategy scale (Table 2).

Figure 2. The normal quantile-quantile (Q-Q) chart for the score of autonomous learning ability.

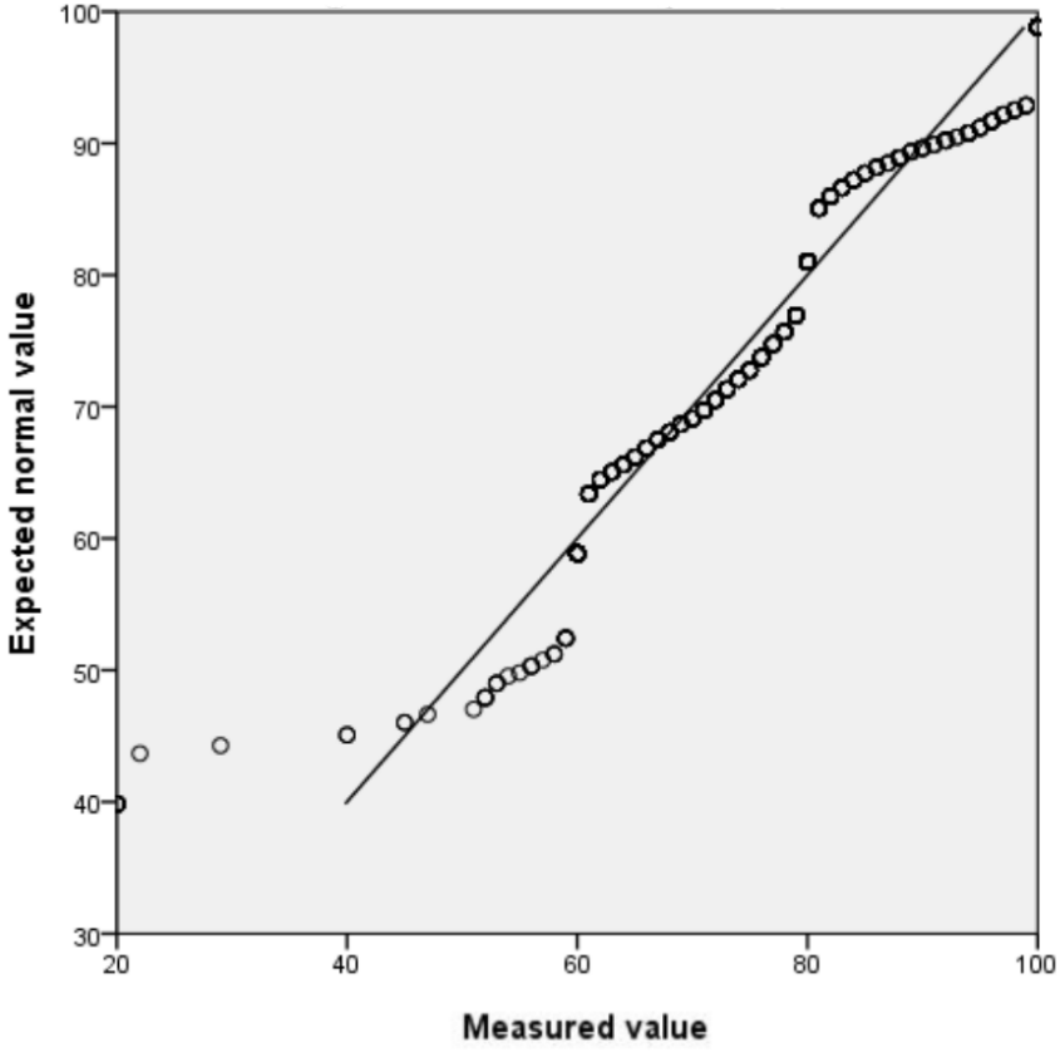


Figure 3. The normal quantile-quantile (Q-Q) chart for the score of the learning strategy scale.

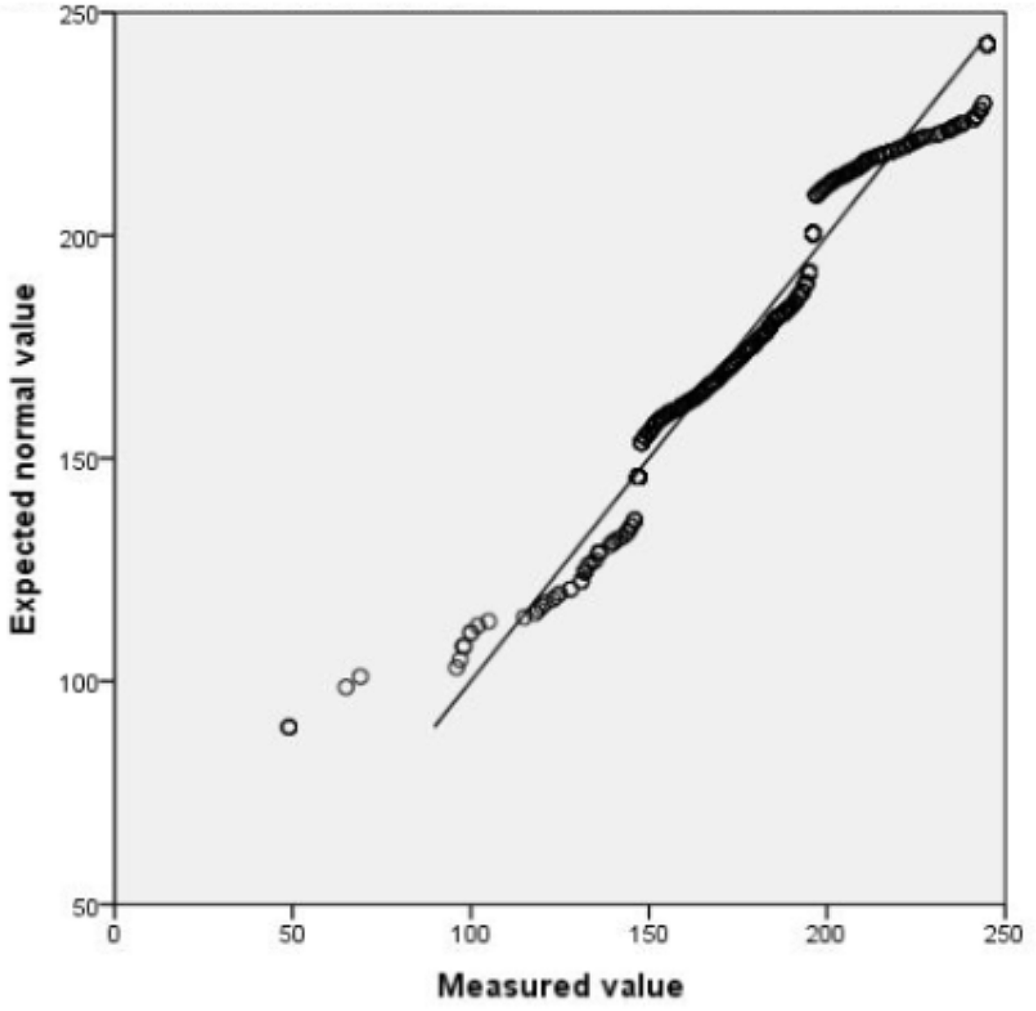


Table 2. Correlation analysis of self-learning ability of interns (N=883).

Characteristic	Autonomous learning ability		Score of learning strategy scale	
	r	*P* value	*r*	*P* value
Age	–0.020	.56	–0.014	.67
Gender	–0.101	.10	0.018	.59
Educational level	0.008	.81	0.008	.81
Internship time	0.410	.22	–0.044	.19
Reason for choosing nursing	0.993	<.001	0.174	<.001
Feelings about nursing	0.595	.02	0.298	<.001
Employment intention	0.011	.75	–0.175	<.001

Interns' Needs and Expectations of Diabetes Virtual Simulation Applications

Nursing trainees believed that important functions of a diabetes virtual simulation application are to help them treat patients with hypoglycemia and the simulation of functional tests. Almost all respondents believed the listed functions were important or very important. However, most interns believed that oral administration, venofusion, and intramuscular injection were important (Figure 4). When comparing teaching methods with the expectations of nurse interns, PowerPoint presentations (222/883, 25.1%) and face-to-face teaching (219/883, 24.8%) were the most-used teaching methods, while students expected to use more virtual simulations (204/883, 23.1%) and to reduce the use of PowerPoint presentations (148/883, 16.8%) (Figure 5).

Figure 4. Importance of different simulation scenes on a diabetes virtual simulation application as reported by interns.

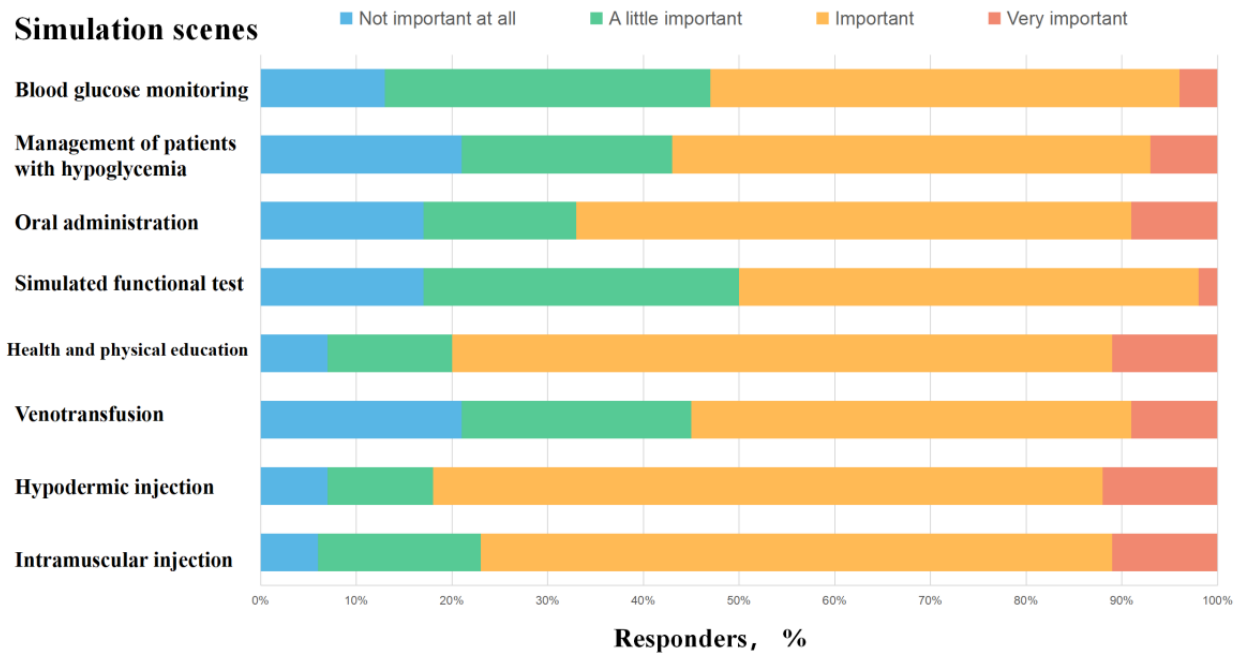
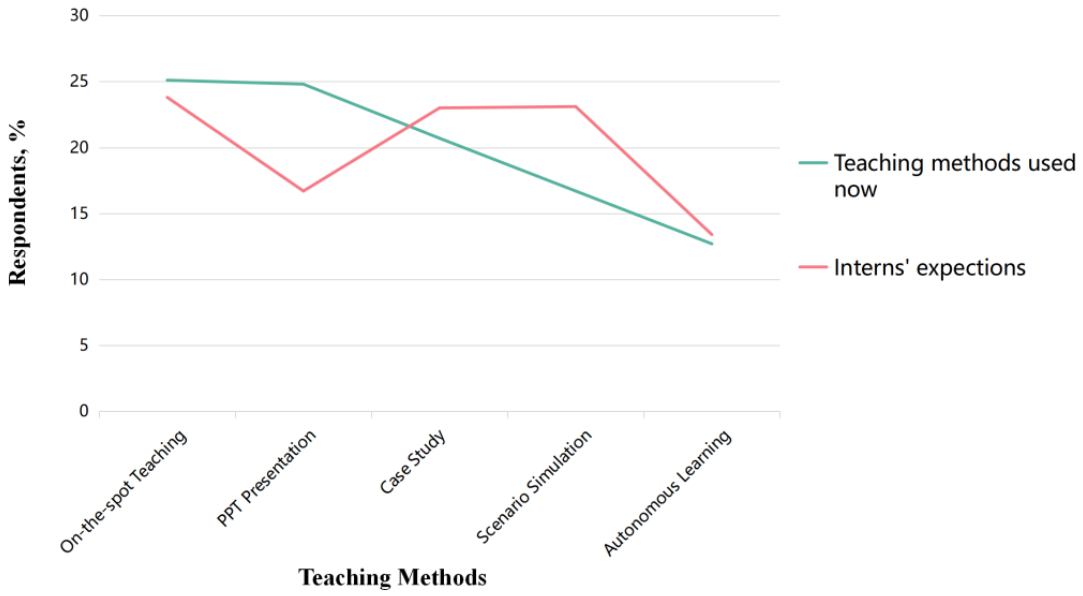


Figure 5. Comparison of teaching methods with interns' expectations. PPT: PowerPoint.



In this study, out of 883 interns, 569 (64.4%) had never participated in virtual simulation teaching and 793 (89.8%) had not heard of the concept of virtual simulation before this survey. Table 3 shows that through the virtual simulation application, what interns most want to improve is their clinical thinking ability (534/883, 60.5%), followed by their comprehension ability (156/883, 17.7%).

Table 3. Nurse interns' usage and preferences of a diabetes virtual simulation application.

Question	Respondents (N=883), n (%)
Have you participated in virtual simulation teaching?	
No	569 (64.4)
Yes	231 (26.2)
Have you participated in similar activities? (yes)	83 (9.4)
Do you know about virtual simulation teaching?	
No	793 (89.8)
Yes	90 (10.2)
What do you think of virtual simulation teaching?	
Very good	138 (15.6)
Good	251 (28.4)
Neutral	76 (8.6)
Bad	7 (0.8)
Very bad	2 (0.2)
Do not know	409 (46.3)
What is your acceptance level of virtual simulation teaching?	
Very good	354 (40.1)
Good	393 (44.5)
Neutral	126 (14.3)
Bad	7 (0.8)
Very bad	3 (0.3)
Which ability do you most want to improve in virtual simulation teaching?	
Comprehension skills	156 (17.7)
Analytical skills	141 (16.0)
Judgment skills	48 (5.4)
Clinical thinking ability	534 (60.5)
Others	4 (0.5)
What do you think is appropriate for the average duration of each session? (minutes)	
0-10	209 (23.7)
11-30	483 (54.7)
31-60	169 (19.1)
61-90	22 (2.5)

Discussion

Principal Findings

The Use of a Virtual Simulation Teaching Application and its Influencing Factors Among Interns

Among the interns, 26.2% (231/883) had participated in virtual simulation education, and 9.4% (83/883) had participated in similar activities. These rates are comparable to results from surveys conducted in New York [17] and Florida [16], and higher than the rate (7%) found in a 2011 survey in Canada [18]. In China, more nursing interns in Southern China (87.3%) participated in virtual simulation teaching than in Northern China (12.7%). One possible reason is that China's economic development is uneven, and medical resources are unevenly distributed. These resources are more highly concentrated in economically developed areas. In these developed areas, the economy is developing well, and the government and families attach great importance to education [26]. In addition, nursing students who participated in virtual simulation teaching preferred it (271/883, 30.7% vs 43/883, 4.9%), which is consistent with previous studies [27,28]. This could be the case because virtual simulation teaching caters to the thinking skills of young people more than traditional teaching. Through the use of virtual simulations, nursing trainees have the opportunity to practice skills and deal with difficult situations. Virtual simulation teaching allows greater access rights, and it allows interns to

appear “virtually” only as participants. In addition, the virtual environment provides a safe environment for practicing nontechnical skills such as teamwork.

Suggestions to Promote the Use of Virtual Simulation Teaching Applications

The utilization rate of virtual simulation teaching applications in China is low because of the low awareness of this teaching method among interns. Only 10.2% (90/883) of interns had heard about virtual simulation teaching. In 2008, Tsinghua University launched a medical-related virtual simulation project for the first time to help doctors complete neurosurgery operations [29]. In 2018, China established a national virtual simulation experiment teaching platform, and virtual simulation teaching began to develop [30].

Through virtual simulation, clinical thinking was the ability that interns wanted to acquire the most (534/883, 60.5%); the second most desired ability was analytical skills (141/883, 16.0%). This result is consistent with a study in Canada [31]. This indicates that virtual simulation is a supplementary teaching strategy that provides opportunities to improve students’ clinical reasoning ability through exposure to a large number of clinical situations. The use of clinical virtual simulation as a teaching strategy should be integrated and coordinated with other teaching strategies in the classroom and other resources (eg, the high-, medium-, and low-tech simulators used in our simulation laboratory) to maximize the development of students’ cognitive, emotional, and psychomotor skills [32,33].

Nursing trainees believed that the scenarios that should be included in the virtual simulation of diabetes care are the treatment of patients with hypoglycemia (626/883, 70.9%), functional test simulation (610/883, 69.1%), and blood glucose monitoring technology (485/883, 54.9%).

Several studies have also shown that nursing interns lack the knowledge to properly handle patients with hypoglycemia, especially elderly patients with diabetes, which could increase the risk of acute complications in these patients [34,35]. This reminds us that virtual simulation is an interactive learning strategy that can increase students’ intrinsic motivation and satisfaction. It focuses on the application of basic knowledge to clinical learning challenges that reproduce the clinical scenarios that students will face in the future. It allows for competency-based education and assessment to enable deeper learning and the development of clinical expertise. Virtual simulation can help reduce clinical errors and improve the safety and quality of health care. When designing diabetes virtual simulations, we should focus on the design of scenarios for patients with hypoglycemia.

Comparison With Previous Work

To the best of our knowledge, no large-scale survey on the use and demand of virtual simulation has been previously conducted among Chinese nursing interns. An Indian survey showed that the need for diabetes knowledge by interns is urgent, consistent with our research, but that study did not identify what kinds of teaching tools the interns wanted. The survey only investigated the needs of first-year nursing students in one city in regard to virtual simulation [36], while our research collected information about the understanding of virtual simulation among interns in various provinces of China. Our research found that students who received virtual simulation teaching tended to be younger, more educated, and have a stronger autonomous learning ability, which is consistent with a survey conducted in Canada [37].

Strengths and Limitations

A strength of our research is that the initial survey links for patients and diabetes experts were sent to 37 partner schools in 23 representative major cities in China, and these were disseminated through their WeChat contact networks. In addition to this snowball-sampling method, the survey was also carried out through three convenience-sampling methods on WeChat Moments.

Our research also has some limitations. First, the sample of 883 nurse interns could not fully represent all interns in China. Our sample came from 23 provinces in China; thus, not all provinces were represented. Second, our sampling was not stratified by geographic area, urban or rural area, school level, or hospital level where internships were based. Certain selection biases were inevitable. Finally, our sampling was based on the WeChat network. Although WeChat has 1.04 billion monthly active users [38], some people rarely use WeChat or surf the internet. Our research methods included a cross-sectional survey. Although the views and attitudes of interns are very important in developing teaching methods for them, people’s attitudes toward the usefulness of simulations and their possible effects depend to a large extent on their current technological development and implementation methods. Therefore, with the development of technology and changes in people’s perceptions, these findings must be updated over time. In addition, many factors affect the use of teaching methods. Although we adjusted for some factors in the multivariate analysis, other potential confounding factors still exist.

Conclusions

Chinese nursing interns’ awareness and usage of diabetes virtual simulation teaching methods are low. However, interns desire the knowledge they would gain by using these methods. Designing virtual simulations of diabetes that are suitable for the needs of different nurse trainees will increase students’ interest in learning and help improve diabetes knowledge among nursing interns. High-quality randomized controlled trials can be conducted to improve the effectiveness of virtual simulation teaching of diabetes, provide evidence for teachers to choose suitable teaching tools, and help with the promotion of the correct management of diabetes. China should improve people’s understanding of virtual simulation teaching in universities, and relevant policies and regulations should be published to support teachers in using virtual simulation teaching tools in schools or hospitals. Virtual simulation is a potentially effective supplement for teaching. It can be used anywhere and at any time to improve the self-learning methods of Chinese nursing interns.

Acknowledgments

This work was supported by the Education and Teaching Reform Research Project of Central South University (2020jy166-11) and the Clinical Nursing Research Fund Project of the Second Xiangya Hospital, Central South University (2019-HLKY-25), the People's Republic of China.

Conflicts of Interest

None declared.

References

1. American Diabetes Association. Position Statement 15. Diabetes care in the hospital: Standards of medical care in diabetes—2019. Diabetes Care 2019 Jan;42(Supplement 1):S173-S181 [FREE Full text] [doi: 10.2337/dc19-s015]
2. American Diabetes Association. Position Statement 15. Diabetes care in the hospital: Standards of medical care in diabetes—2020. Diabetes Care 2020 Jan;43(Supplement 1):S193-S202. [doi: 10.2337/dc20-s015]
3. Cho N, Shaw J, Karuranga S, Huang Y, da Rocha Fernandes J, Ohlrogge A, et al. IDF Diabetes Atlas: Global estimates of diabetes prevalence for 2017 and projections for 2045. Diabetes Res Clin Pract 2018 Apr;138:271-281. [doi: 10.1016/j.diabres.2018.02.023] [Medline: 29496507]
4. Choi TS, Davidson ZE, Walker KZ, Lee JH, Palermo C. Diabetes education for Chinese adults with type 2 diabetes: A systematic review and meta-analysis of the effect on glycemic control. Diabetes Res Clin Pract 2016;116:218-229. [doi: 10.1016/j.diabres.2016.04.001] [Medline: 27321339]
5. Biagetti B, Ciudin A, Portela M, Dalama B, Mesa J. Interns' viewpoints and knowledge about management of hyperglycemia in the hospital setting. Endocrinol Nutr (Engl Ed) 2012 Aug;59(7):423-428. [doi: 10.1016/j.endoen.2012.08.010]
6. Brenol CV, Monticielo OA, Xavier RM, Brenol JCT. Artrite reumatóide e aterosclerose. Rev Assoc Med Bras 2007 Oct;53(5):465-470. [doi: 10.1590/S0104-42302007000500026]
7. Henry B, Duellman M, Smith T. Nutrition-based standardized patient sessions increased counseling awareness and confidence among dietetic interns. Top Clin Nutr 2009;24(1):25-34. [doi: 10.1097/tin.0b013e3181978050]
8. Hemanthkumar B, Keertish N, Sathyanarayana M. Is there any difference between attitude of interns toward psychiatric illness and other chronic medical conditions? A comparative study. Indian J Psychiatry 2018;60(2):195. [doi: 10.4103/psychiatry.indianjpsychiatry_107_17]
9. Bektaş İ, Yardımcı F. The effect of web-based education on the self-confidence and anxiety levels of paediatric nursing interns in the clinical decision-making process. J Comput Assist Learn 2018 Aug 17;34(6):899-906. [doi: 10.1111/jcal.12298]
10. Raparla N, Davis D, Shumaker D, Kumar A, Hafiz S, Sava J, et al. A pilot program to improve nursing and surgical intern collaboration: Lessons learned from a mixed-methods study. Am J Surg 2017 Feb;213(2):292-298. [doi: 10.1016/j.amjsurg.2016.10.014] [Medline: 28017298]
11. Aboshaiqah A, Qasim A. Nursing interns' perception of clinical competence upon completion of preceptorship experience in Saudi Arabia. Nurse Educ Today 2018 Sep;68:53-60. [doi: 10.1016/j.nedt.2018.05.021] [Medline: 29886285]
12. Min Zhi L, Dan WF. Innovation and research of teaching way of clinical nursing interns in Internal Neurology Department. Iran J Public Health 2016 Jun;45(6):824-825 [FREE Full text] [Medline: 27648430]
13. Zanchetta M, Schwind J, Aksenchuk K, Gorospe FF, Santiago L. An international internship on social development led by Canadian nursing students: Empowering learning. Nurse Educ Today 2013 Jul;33(7):757-764. [doi: 10.1016/j.nedt.2013.04.019] [Medline: 23684525]
14. Slater P, Hasson F, Gillen P, Gallen A, Parlour R. Virtual simulation training: Imaged experience of dementia. Int J Older People Nurs 2019 Sep;14(3):e12243. [doi: 10.1111/opn.12243] [Medline: 31070851]
15. Paul P, Olson J, Jackman D, Gauthier S, Gibson B, Kabotoff W, et al. Perceptions of extrinsic factors that contribute to a nursing internship experience. Nurse Educ Today 2011 Nov;31(8):763-767. [doi: 10.1016/j.nedt.2010.11.016] [Medline: 21131105]
16. Foronda C, Budhathoki C, Salani D. Use of multiuser, high-fidelity virtual simulation to teach leadership styles to nursing students. Nurse Educ 2014;39(5):209-211. [doi: 10.1097/NNE.0000000000000073] [Medline: 25137445]
17. Mabry J, Lee E, Roberts T, Garrett R. Virtual simulation to increase self-efficacy through deliberate practice. Nurse Educ 2019 Dec 5;45(4):202-205. [doi: 10.1097/nne.0000000000000758]
18. Jetté S, Charette S. The next step in virtual simulation. Can Nurse 2011 Jun;107(6):8. [Medline: 21761800]
19. Karvonen T, Muranishi Y, Yamamoto G, Kuroda T, Sato T. Evaluation of a novel multi-articulated endoscope: Proof of concept through a virtual simulation. Int J Comput Assist Radiol Surg 2017 Jul;12(7):1123-1130. [doi: 10.1007/s11548-017-1599-0] [Medline: 28534312]
20. Polivka BJ, Anderson S, Lavender SA, Sommerich CM, Stredney DL, Wills CE, et al. Efficacy and usability of a virtual simulation training system for health and safety hazards encountered by healthcare workers. Games Health J 2019 Apr;8(2):121-128 [FREE Full text] [doi: 10.1089/g4h.2018.0068] [Medline: 30234397]
21. Palancia Esposito C, Sullivan K. Maintaining clinical continuity through virtual simulation during the COVID-19 pandemic. J Nurs Educ 2020 Sep 01;59(9):522-525. [doi: 10.3928/01484834-20200817-09] [Medline: 32865587]

22. Coyne E, Calleja P, Forster E, Lin F. A review of virtual-simulation for assessing healthcare students' clinical competency. Nurse Educ Today 2021 Jan;96:104623. [doi: 10.1016/j.nedt.2020.104623] [Medline: 33125979]
23. Zhang X, Wen D, Liang J, Lei J. How the public uses social media WeChat to obtain health information in China: A survey study. BMC Med Inform Decis Mak 2017 Jul 05;17(Suppl 2):66 [FREE Full text] [doi: 10.1186/s12911-017-0470-0] [Medline: 28699549]
24. Maymone MBC, Du T, Dellavalle RP. Healthcare and dermatology on WeChat. Dermatol Online J 2019 Apr 15;25(4):1-3 [FREE Full text] [doi: 10.5070/D3254043546] [Medline: 31046905]
25. Zhang L, Jung EH. How does WeChat's active engagement with health information contribute to psychological well-being through social capital? Univers Access Inf Soc 2021 Feb 18:1-17 [FREE Full text] [doi: 10.1007/s10209-021-00795-2] [Medline: 33619431]
26. Yahong W, Khan S. A Cross-Sectional Analysis of Employment Returns to Education and Health Status in China: Moderating Role of Gender. Front Psychol 2021;12:638599 [FREE Full text] [doi: 10.3389/fpsyg.2021.638599] [Medline: 33716907]
27. Foronda CL, Hudson KW, Budhathoki C. Use of virtual simulation to impact nursing students' cognitive and affective knowledge of evidence-based practice. Worldviews Evid Based Nurs 2017 Apr;14(2):168-170. [doi: 10.1111/wvn.12207] [Medline: 28207994]
28. Kalisch BJ, Aebersold M, McLaughlin M, Tschannen D, Lane S. An intervention to improve nursing teamwork using virtual simulation. West J Nurs Res 2015 Feb;37(2):164-179. [doi: 10.1177/0193945914531458] [Medline: 24763705]
29. Yu H, Zhang CY, Zhang SH, Cheng H, Chen J. Virtual simulation teaching centre in dental education: A report from Fujian Medical University, China. Chin J Dent Res 2017;20(3):173-177 [FREE Full text] [doi: 10.3290/j.cjdr.a38773] [Medline: 28808702]
30. Xia S, Fu B, Wang B, Wu J, Cui Y, Wang X. Computed tomography imaging-based preoperative virtual simulation for calcaneal fractures reduction. J Foot Ankle Surg 2019 Mar;58(2):248-252. [doi: 10.1053/j.jfas.2018.08.054] [Medline: 30850096]
31. Chircop A, Cobbett S. Gett'n on the bus: Evaluation of Sentinel City®3.0 virtual simulation in community/population health clinical placement. Int J Nurs Educ Scholarsh 2020 Jun 08;17(1):1. [doi: 10.1515/ijnes-2019-0109] [Medline: 32543461]
32. Liu W. The effects of virtual simulation on undergraduate nursing students' mental health literacy: A prospective cohort study. Issues Ment Health Nurs 2021 Mar;42(3):239-248. [doi: 10.1080/01612840.2020.1793248] [Medline: 32783682]
33. Breen H, Jones M. Experiential learning: Using virtual simulation in an online RN-to-BSN program. J Contin Educ Nurs 2015 Jan;46(1):27-33. [doi: 10.3928/00220124-20141120-02] [Medline: 25401340]
34. Foronda C, Lippincott C, Gattamorta K. Evaluation of virtual simulation in a master's-level nurse education certificate program. Comput Inform Nurs 2014 Nov;32(11):516-522. [doi: 10.1097/CIN.0000000000000102] [Medline: 25251860]
35. Tilton KJ, Tiffany J, Hoglund BA. Non-acute-care virtual simulation: Preparing students to provide chronic illness care. Nurs Educ Perspect 2015;36(6):394-395. [doi: 10.5480/14-1532] [Medline: 26753300]
36. Saeed N, Jaffery T, Ansari FA, Hamid B, Khan BA. Knowledge and perceptions about pre-diabetes amongst doctors, medical students, and patients in a tertiary care hospital of Islamabad. J Pak Med Assoc 2019 Apr;69(4):527-532 [FREE Full text] [Medline: 31000858]
37. Tyerman J, Luctkar-Flude M, Baker C. Rapid development of a COVID-19 assessment and PPE virtual simulation game. Clin Simul Nurs 2021 Jul;56:125-132 [FREE Full text] [doi: 10.1016/j.ecns.2021.03.002] [Medline: 33727989]
38. Xu H, Huang S, Qiu C, Liu S, Deng J, Jiao B, et al. Monitoring and management of home-quarantined patients with COVID-19 using a WeChat-based telemedicine system: Retrospective cohort study. J Med Internet Res 2020 Jul 02;22(7):e19514 [FREE Full text] [doi: 10.2196/19514] [Medline: 32568727]

Abbreviations

ICT: information and communications technology
OR: odds ratio
Q-Q: quantile-quantile

Edited by G Eysenbach; submitted 09.04.21; peer-reviewed by T Hebda; comments to author 30.04.21; revised version received 03.07.21; accepted 27.07.21; published 09.09.21.

Please cite as:
Liu F, Weng H, Xu R, Li X, Zhang Z, Zhao K, Zhou Z, Wang Q
Nursing Interns' Attitudes Toward, Preferences for, and Use of Diabetes Virtual Simulation Teaching Applications in China: National Web-Based Survey
JMIR Mhealth Uhealth 2021;9(9):e29498
URL: https://mhealth.jmir.org/2021/9/e29498
doi:10.2196/29498
PMID:34499047

Original Paper

Upper-Limb Motion Recognition Based on Hybrid Feature Selection: Algorithm Development and Validation

Qiaoqin Li[1], PhD; Yongguo Liu[1], PhD; Jiajing Zhu[1], PhD; Zhi Chen[1], PhD; Lang Liu[1], MD; Shangming Yang[1], PhD; Guanyi Zhu[2], BA; Bin Zhu[3], BSc; Juan Li[4], PhD; Rongjiang Jin[4], PhD; Jing Tao[5], PhD; Lidian Chen[5], PhD

[1]Knowledge and Data Engineering Laboratory of Chinese Medicine, School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu, China

[2]College of Electrical and Information Engineering, Hunan University, Changsha, China

[3]Chengdu Chronic Diseases Hospital, Chengdu, China

[4]College of Health Preservation and Rehabilitation, Chengdu University of Traditional Chinese Medicine, Chengdu, China

[5]College of Rehabilitation Medicine, Fujian University of Traditional Chinese Medicine, Fuzhou, China

Corresponding Author:
Yongguo Liu, PhD
Knowledge and Data Engineering Laboratory of Chinese Medicine
School of Information and Software Engineering
University of Electronic Science and Technology of China
No 4, Section 2
Jianshe North Road
Chengdu, 610054
China
Phone: 86 13980786625
Email: liuyg@uestc.edu.cn

Abstract

Background: For rehabilitation training systems, it is essential to automatically record and recognize exercises, especially when more than one type of exercise is performed without a predefined sequence. Most motion recognition methods are based on feature engineering and machine learning algorithms. Time-domain and frequency-domain features are extracted from original time series data collected by sensor nodes. For high-dimensional data, feature selection plays an important role in improving the performance of motion recognition. Existing feature selection methods can be categorized into filter and wrapper methods. Wrapper methods usually achieve better performance than filter methods; however, in most cases, they are computationally intensive, and the feature subset obtained is usually optimized only for the specific learning algorithm.

Objective: This study aimed to provide a feature selection method for motion recognition of upper-limb exercises and improve the recognition performance.

Methods: Motion data from 5 types of upper-limb exercises performed by 21 participants were collected by a customized inertial measurement unit (IMU) node. A total of 60 time-domain and frequency-domain features were extracted from the original sensor data. A hybrid feature selection method by combining filter and wrapper methods (FESCOM) was proposed to eliminate irrelevant features for motion recognition of upper-limb exercises. In the filter stage, candidate features were first selected from the original feature set according to the significance for motion recognition. In the wrapper stage, k-nearest neighbors (kNN), Naïve Bayes (NB), and random forest (RF) were evaluated as the wrapping components to further refine the features from the candidate feature set. The performance of the proposed FESCOM method was verified using experiments on motion recognition of upper-limb exercises and compared with the traditional wrapper method.

Results: Using kNN, NB, and RF as the wrapping components, the classification error rates of the proposed FESCOM method were 1.7%, 8.9%, and 7.4%, respectively, and the feature selection time in each iteration was 13 seconds, 71 seconds, and 541 seconds, respectively.

Conclusions: The experimental results demonstrated that, in the case of 5 motion types performed by 21 healthy participants, the proposed FESCOM method using kNN and NB as the wrapping components achieved better recognition performance than the traditional wrapper method. The FESCOM method dramatically reduces the search time in the feature selection process. The results also demonstrated that the optimal number of features depends on the classifier. This approach serves to improve feature

selection and classification algorithm selection for upper-limb motion recognition based on wearable sensor data, which can be extended to motion recognition of more motion types and participants.

(JMIR Mhealth Uhealth 2021;9(9):e24402) doi:10.2196/24402

KEYWORDS

feature selection; inertial measurement unit; motion recognition; rehabilitation exercises; machine learning

Introduction

Background

The combination of wearable devices and wireless network technologies enables modern health care service providers to ubiquitously monitor patients out of hospital who require long-term exercise [1-3]. Motion recognition plays an important role in maintaining the intensity and quality of autonomous training with no or reduced supervision [4]. O'Brien et al [5] investigated the performance of action recognition based on signals collected by accelerometer, gyroscope, and barometer sensors in a mobile phone in a home setting for stroke patients. Zhang et al [6] proposed a fuzzy kernel motion classifier to address the overlapping motion class issue caused by irregular motion samples performed by patients with different functional impairments. Cui et al [7] developed an automatic gait analysis system for stroke patients based on multimodal fusion architecture. Cai et al [8] investigated the feasibility of a support vector machine (SVM) classifier for motion recognition of the upper-limb exercises via surface electromyogram (sEMG) signals [8]. Huang et al [9] proposed a knowledge-driven multimodal activity recognition framework that exploits external knowledge to fuse multimodal data.

Feature Selection in Motion Recognition

Most motion recognition methods are based on feature extraction and machine learning algorithms [10]. Time-domain and frequency-domain features are extracted from original time series data [11,12]. Castiblanco et al [13] exploited myoelectric signals (EMG) to identify finger and hand motions through pattern recognition techniques. Several methods for feature extraction, ranking, and classification from EMG signals were implemented, and the performance of motion identification was compared. Shawen et al [14] developed 4 classifiers that use accelerometer and gyroscope data collected by mobile phone from able-bodied individuals to detect falls in individuals with a lower limb amputation. A set of 40 features was computed from the original sensor data, and classifiers were trained to detect falls. Lin et al [15] used 2 sensors on the arm and wrist to collect acceleration and angular velocity of 6 types of upper-limb exercises performed by 13 volunteers. Motor features were used to train a back-propagation neural network (BPNN) algorithm for motion recognition. Wu et al [16] developed a method to identify upper-limb motion for community rehabilitation. The feature vector space was established by variance, mean absolute value, the fourth-order autoregressive, zero crossings, and root mean square. Various feature sets were extracted for classification.

Feature selection is an essential step to eliminate redundant or irrelevant features for specific classification task so as to deal with high-dimensional data [17,18]. Its task is to find the most representative feature subset from the original feature set. Ramezani et al [19] analyzed physical activity sensor features and activities with regard to indoor localization. Random forest (RF) was used to build a predictive model based on the most significant features. The study demonstrated that a subset of features can better distinguish between at-risk patients that can gain independence versus patients that will be rehospitalized. Wang et al [20] proposed 2 feature selection methods to improve activity recognition. Experimental results showed that the proposed methods reduce the dimensionality of the original feature space and contribute to the enhancement of overall recognition accuracy. Fang et al [21] compared feature selection methods based on interclass distance for human activity recognition in smart home environments. The experimental results showed that activity recognition accuracy is related to the feature set selected and an unsuitable feature set increases computational complexity and degrades activity recognition accuracy. Zhou et al [22] proposed a feature selection method for human motion recognition based on open human motion data. The experimental results showed that their feature selection method yields better recognition accuracy than nonfeature selection models.

Feature selection methods can be categorized into filter and wrapper methods [23]. For filter methods, the selection of the feature subset is independent of the classification algorithm. The feature fitness is evaluated via the statistical characteristics of the dataset, and the features with top ranking fitness are selected [24,25]. Banos et al [26] proposed a feature selection method for physical activity recognition using a feature quality group ranking via statistical criteria based on discrimination and robustness. Satisfactory results were achieved in both laboratory and seminaturalistic activity living datasets for real problems using several classification models. Hong et al [27] proposed a motion gesture recognition system via accelerometer (MGRA) implemented on mobile devices. The best feature vector including 27 items was selected using the minimal-redundancy-maximal-relevance criterion taking both static and mobile scenarios into consideration. The experimental results confirmed that MGRA can accommodate a broad set of gesture variations within each class and achieve higher accuracy than previous methods [28]. As for wrapper methods, the feature subset is selected simultaneously with the estimation of its goodness in a specific classification task [29,30]. Camargo and Young [31] implemented motion classification from sEMG signals for prosthetic control by exploiting Chow-Liu trees for selecting features and evaluating 6 different classification algorithms as the wrapping component [32]. The results demonstrated that feature selection is critical for improving classification accuracy. Xue et al [33] presented a novel wrapper feature selection algorithm that utilizes a generic algorithm to

wrap an extreme learning machine to search for the optimum feature subset. Experiments were conducted on benchmark datasets and compared with 4 filter methods and 2 hybrid wrapper methods. The results revealed that the presented wrapper method is useful for feature selection problems and outperforms other algorithms in comparison. Chen and Chen [34] introduced a wrapper method to eliminate irrelevant features during classifier construction by introducing the cosine distance into SVM. The feature selection method has been applied to fault diagnosis of rolling element bearings and diagnosis of mild cognitive impairment. The results showed that the proposed method has great capacity for feature selection and pattern recognition.

The Hybrid Feature Selection Method

Filter methods are often time-efficient, but the results are not always satisfactory. On the other hand, wrapper methods usually achieve better performance, but could be computationally intensive and the obtained feature subset optimized only for the specific learning algorithm [35]. As a result, hybrid feature selection methods take advantages of both filter and wrapper methods. Manbari et al [36] presented a hybrid feature selection algorithm based on the combination of clustering and the modified binary ant system to overcome the search space and high-dimensional data processing challenges. A damped mutation strategy was introduced to avoid local optima, and a new redundancy reduction policy was adopted to estimate the correlation between the selected features so as to further improve the algorithm.

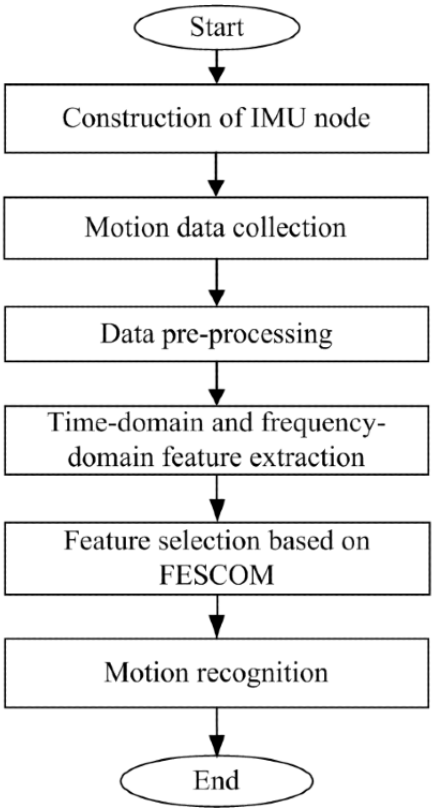
Existing feature selection methods usually select feature sets that are relevant for specific classification tasks. To select the most representative features for motion recognition of upper-limb exercises, we propose a hybrid feature selection method combining the filter and wrapper methods called FESCOM in this paper. In the filter stage, candidate features are selected by ranking the feature significance index, which reflects the importance of each feature for motion recognition. In the wrapper stage, a classifier-specific feature selection algorithm is applied to further refine the candidate features. Classifiers including kNN, NB, and RF are constructed as the wrapping components. To the best of our knowledge, FESCOM is the first method that exploits hybrid feature selection for motion recognition of upper-limb exercises.

Methods

Workflow

The general workflow of this work is illustrated in Figure 1. An inertial measurement unit (IMU) node was customized for motion data collection. Motion data including acceleration and angular velocity from 5 types of upper-limb exercises were collected. Original data were preprocessed by applying a median filter to remove outliers. Time-domain and frequency-domain features were extracted from the preprocessed acceleration and angular velocity data on each axis. The feature selection method was built to select the most representative features for motion recognition. Then, motion recognition was implemented using the optimal feature set and corresponding classifier.

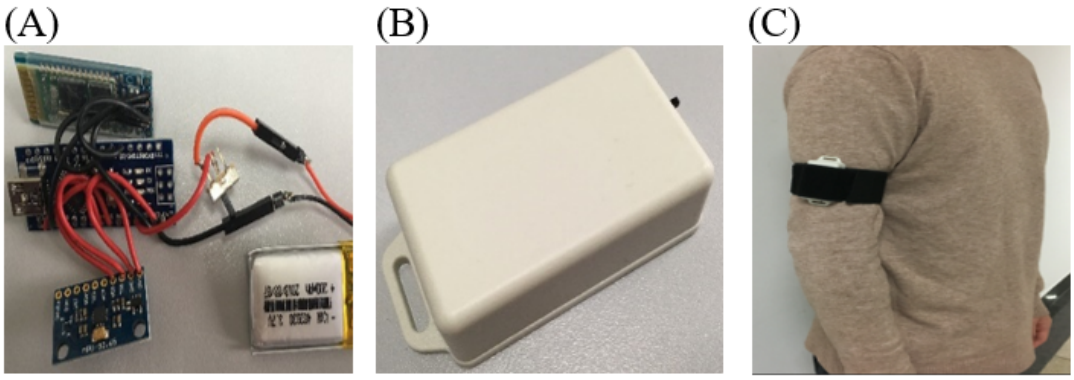
Figure 1. General workflow. FESCOM: hybrid feature selection method by combining filter and wrapper methods; IMU: inertial measurement unit.



Construction of the IMU

The IMU module consists of 1 inertial sensor (MPU9250), 1 8-bit low power consumption micro-controller (ATmega328P on board nano V3.0), 1 Bluetooth wireless transmitter (HC-06), and 1 battery, shown in Figure 2. The inertial sensor MPU9250 is comprised of a 3-axis accelerometer, gyroscope, and magnetometer. The built-in digital motion processing engine in the MPU9250 can reduce the complex computation and load of the microcontroller. The measurement ranges of the accelerometer and gyroscope in the MPU9250 are ±16 *g* and ±2000 °/s, respectively, where *g* represents gravitational acceleration. These specifications meet the needs of upper-limb exercises. Sampled motion data can be transmitted to a PC station by Bluetooth in real time. The battery is 3.7 V and 200 mAh. No recharge module is used. It is convenient for wearable devices. The scale ranges of the accelerometer and the gyroscope can be adjusted using the programming interface of the inertial sensor and were set at ±2 *g* and ±250 °/s, respectively, in this study. The sampling frequency was set at 20 Hz, which is suitable for upper-limb exercises by patients with motion functionality impairment. The baud rate of Bluetooth was set at 19200 bps. Angular velocity was computed based on the gyroscope data. Magnetometer data were not used in this study. These components were connected and embedded into a 58 mm x 32 mm x 19 mm box. The IMU node was attached to the outside of the right upper limb of the participant with a stretchable 350 mm x 38 mm rubber belt, shown in Figure 2. The positive direction of the y-axis points to the wrist.

Figure 2. (A) inertial measurement unit (IMU) components, (B) box, and (C) belt.



Experimental Protocol

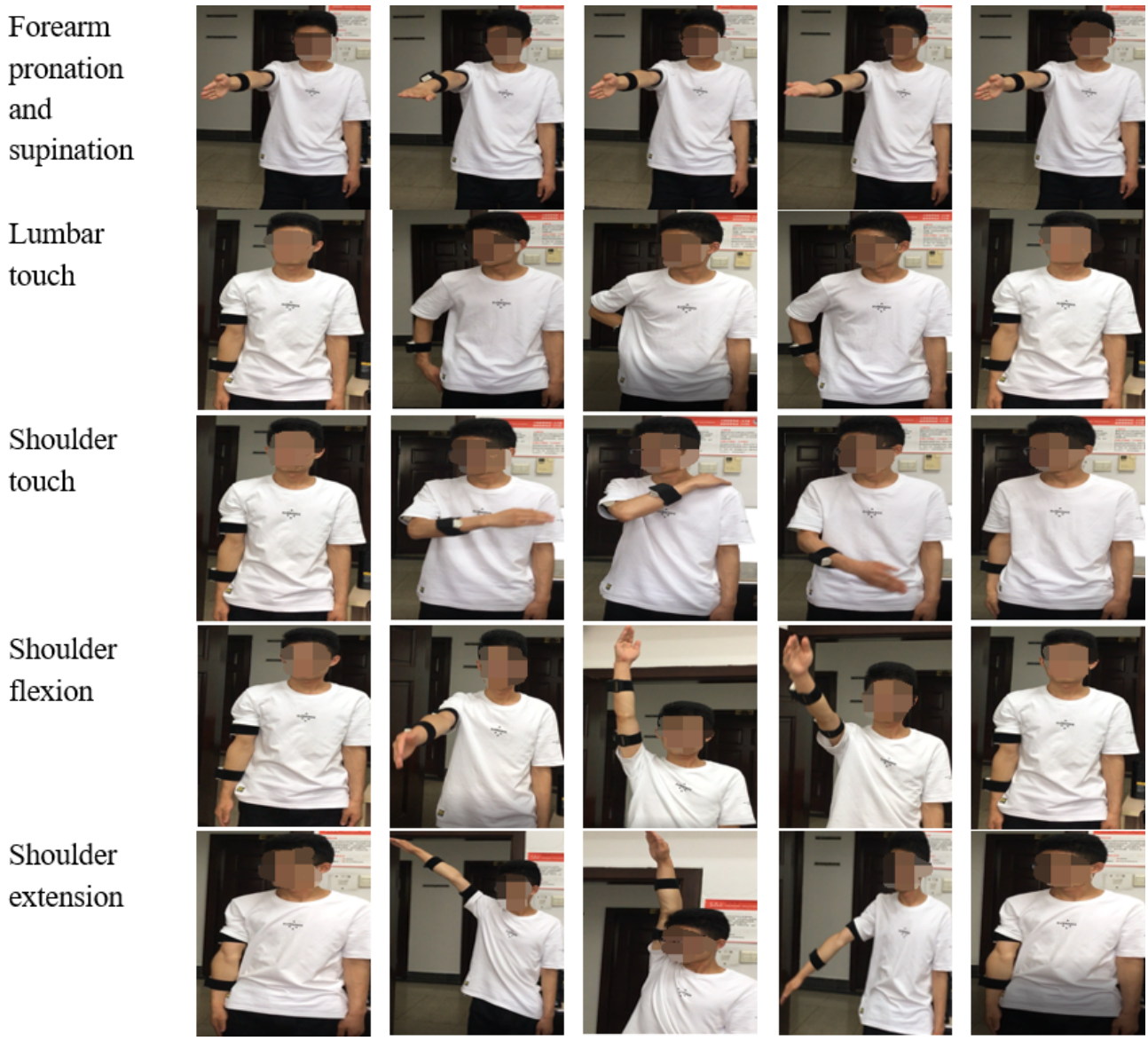
In this study, upper-limb exercises for post-stroke rehabilitation training were considered. From the clinical point of view, a subset of the training items can represent the 33 upper limb–related training items in the Fygl-Meyer Assessment (FMA) scale [37]. In this experiment, 5 representative upper-limb exercises based on the FMA scale were selected:

1. Forearm pronation and supination: Raise the right arm to the horizontal position in the sagittal plane. Then, carry out forearm pronation and supination.
2. Lumbar touch: The right arm hangs naturally. Move the right arm back to touch the back of the waist with the hand. Then, slowly move back to the initial position.
3. Shoulder touch: The right arm hangs naturally. Raise the right arm to the horizontal position in the sagittal plane. Then, carry out an elbow adduction motion and rotate the wrist to touch the opposite shoulder with the hand. Finally, put the arm down to the initial position.
4. Shoulder flexion: The right arm hangs naturally. Raise the right arm in the sagittal plane as high as possible. Then, hold for 3 seconds and move back to the initial position.
5. Shoulder extension: The right arm hangs naturally. Raise the right arm in the coronal plane as high as possible. Then, hold for 3 seconds and move back to the initial position.

Figure 3 includes 5 photos of each exercise taken during the execution process.

Motion data were collected from 21 healthy participants (15 men, 6 women; age, mean 33.2, SD 12.7 years; height, mean 172.5, SD 7.1 cm; weight, mean 62.8, SD 17.5 kg) instead of actual patients who are post-stroke. The study was approved by the institutional review board of the Eighth People's Hospital of Chengdu. Written informed consent was obtained from all participants. In the sampling experiment, participants were first asked to rest for a while. Before the sampling began, they were invited to perform each exercise several times with the guidance of a guiding video until they performed the motions fluidly. Then, they were required to complete 3 valid repetitions of each exercise independently. Each repetition followed an interval of about 3 seconds. A valid repetition was a coherent movement, and each repetition was completed in 1-4 seconds.

Figure 3. Sequence of each exercise performed by the participants.



FESCOM Method

In this study, 10 types of time-domain and frequency-domain features were extracted from the motion data from the upper-limb exercises. The time-domain features included the mean, standard deviation, maximum, and minimum values of the signal as well as the kurtosis, skewness, and interquartile range of the signal, which may reflect the exercise frequency, regularity, and symmetry, respectively. The frequency-domain features included average power, average frequency, and median frequency of the signal. As each sample included acceleration and angular velocity data in 3 axes, the dimension of the original feature vector was 60.

The original feature set contains not only the features that are relevant for classification but also some redundancy features, which decrease the computational efficiency and classification accuracy. We propose a hybrid feature selection method, called FESCOM, to remove redundant features so as to improve the computational efficiency and classification accuracy. Figure 4 shows the procedure of the FESCOM method. In the filter stage, the statistical *t* test method was adopted to compute the statistical significance value (*P* value) of each feature, reflecting the capability of motion recognition [38,39]. For samples *x* and *y*, a two-sample *t* test was considered for analysis, which is defined as:

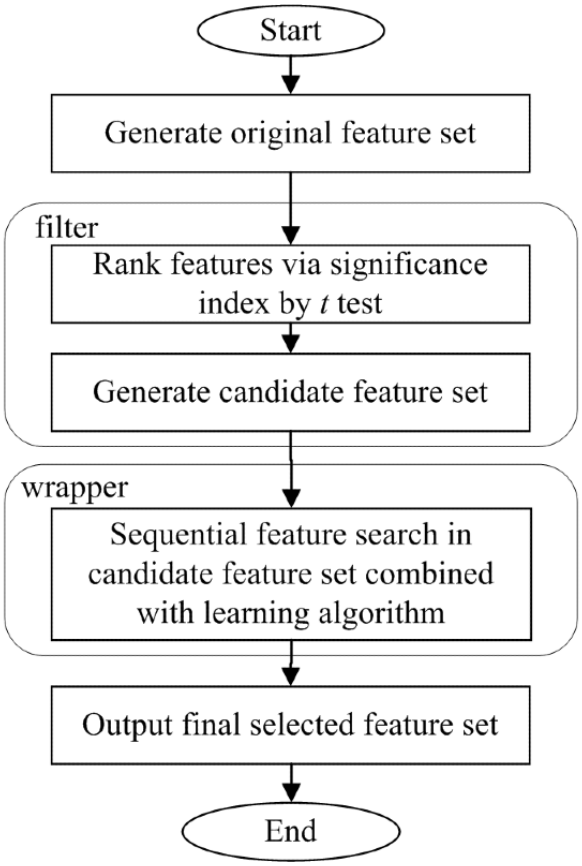
where  and  are the sample means, s_x and s_y are the sample standard deviations, and *n* and *m* are the sample size. As there are 5 types of motion, the *t* test method was applied to each class pair. Let $P_k(i,j)$ represent the *P* value of feature *k* on class *i* and *j*, the average *P* value of feature *k* on all class pairs is computed as:

where $i=1,\ldots,C$, $j=i+1,\ldots,C$, and *C* is the number of motion classes. The standard deviation of the *P* value of feature *k* on all class pairs is:

Then, the significance index of feature *k* is computed as:

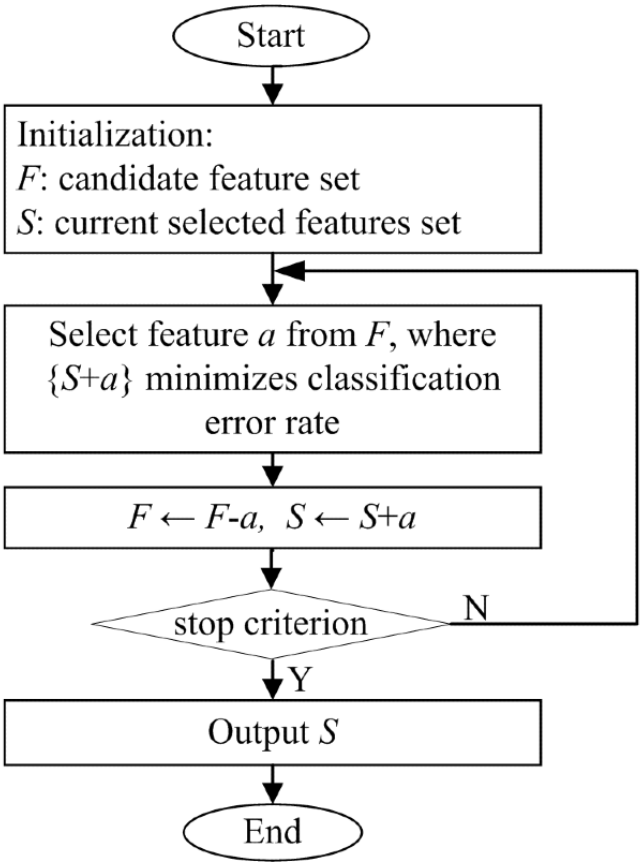
A smaller *s* value means stronger classification capacity. The features with *s* smaller than a threshold were selected for the candidate feature set, ranked in ascending order. The threshold value was a compromise between time efficiency and the classification accuracy of FESCOM.

Figure 4. The hybrid feature selection method by combining filter and wrapper methods (FESCOM).



In the wrapper stage, a sequential feature selection (SFS) method was used to refine the features from the candidate feature set obtained in the filter stage. SFS includes a search algorithm and an objective function, also called criterion [40]. In this study, the search algorithm was sequential forward selection, and the criterion was the classification error rate. Starting from an empty feature set, SFS selects a subset of features from the candidate feature set by sequentially selecting features using the abovementioned criterion until there is no improvement in classification performance, evaluated by each classification algorithm (ie, the wrapping component). The procedure of SFS is illustrated as Figure 5.

Figure 5. Sequential feature selection process.



Classification Algorithms

In the experiment, kNN, NB, and RF were adopted as classification algorithms in the wrapper stage of FESCOM.

Classifier kNN is a simple classification algorithm based on the calculation of the distance (usually the Euclidean distance) between the new sample to be classified and the closest samples in the training set. The training samples are sorted in descending order according to their distance from the new object. Then, the new sample is assigned to the class that most of its k-nearest neighbors belong to [41].

NB is a probabilistic classifier based on the assumption that all features are independent of each other, given the category variable [42]. For discrete features, multinomial or Bernoulli distributions are popular. Despite apparently over-simplifier assumptions, NB classifier works quite well in many complex real-world applications, such as medical diagnosis, key phrase extraction, and text classification. The NB classifier is particularly useful for handling incomplete data and could yield good predictions even with a small data size.

RF is a type of ensemble learning method that is formed through the combination of multiple decision trees trained on the training dataset. When applied to the test dataset, the predictions of individual tree models within the RF are combined into an overall classification decision through means of a majority vote or the application of weights. The RF model can avoid overfitting and provide robust classification performances [43]. The number of decision trees in this experiment was set at 20.

Results

Overview

In this study, MATLAB 2016a was used to develop the proposed FESCOM method for motion recognition of upper-limb exercises. The original dataset was randomly partitioned into a training set and testing set. The training set was applied to train each classifier by using five-fold cross validation. For each iteration, one of the partitions was held back as the validation set, whereas the other partitions were used to train the classification model. The model was then validated by the validation set. This process was repeated 5 times, so that each subset was used as a validation set once. The results were averaged over all rounds. Finally, the performance of each classifier was evaluated on the testing set.

The performance of the proposed algorithm was evaluated using the metrics of classification error rate, computed as the ratio of number of instances classified incorrectly to the total number of instances.

Experimental Data

Acceleration and angular velocity in 3 axes of 5 exercises performed by 1 female participant are illustrated in Figure 6 and Figure 7, respectively. Exe1, Exe2, Exe3, Exe4, and Exe5 in Figures 6 and 7 represent the 5 types of motion defined in the experimental protocol section. The time-domain waveforms of the 5 exercises showed different characteristics. For example, the acceleration and angular velocity of forearm pronation and supination, lumbar touch, and shoulder touch showed totally different trends. Although the time-domain features of the

acceleration in the y-axis of shoulder flexion were similar with those of shoulder extension, the acceleration in the x-axis of shoulder flexion exhibited a higher peak compared with that of shoulder extension. Moreover, angular velocity in both the x-axis and y-axis of shoulder flexion exhibited a smaller peak than that of shoulder extension. These differences between exercises can be used for motion recognition.

Figure 6. Acceleration in each axis of the 5 exercises.

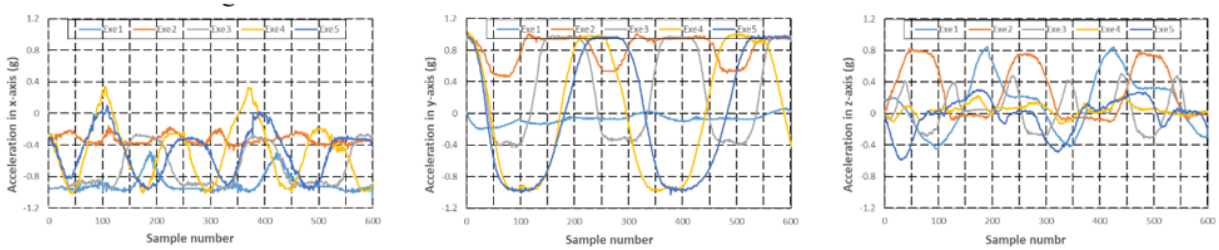
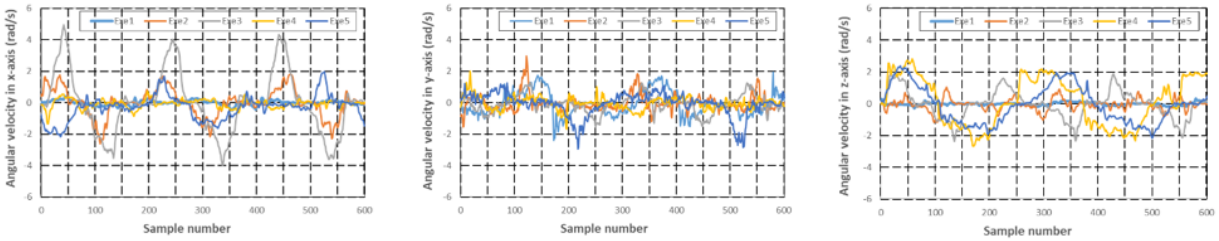


Figure 7. Angular velocity in each axis of the 5 exercises.



Feature Significance Index

Table 1 shows the top-10 feature significance index value and rank order computed by the statistical *t* test method. An extended version of Table 1, including the significance index of all 60 features, is presented in Multimedia Appendix 1. The significance index value was computed on 2 types of signals (ie, acceleration and angular velocity), represented in parentheses following the feature name, with a suffix representing on which axis it is. The significance index value of the minimum angular_velocity_y ranks the highest, whereas the average acceleration_z ranks the lowest of all 60 features. A smaller significance index means stronger classification capacity.

Table 1. Top 10 feature significance index.

Feature name	Significance index	Rank order
minimum (angular_velocity_y)	0.00030	1
average power (angular_velocity_x)	0.00035	2
average power (acceleration_x)	0.00049	3
standard deviation (angular_velocity_z)	0.00058	4
skewness (acceleration_z)	0.00130	5
average power (acceleration_y)	0.00132	6
median frequency (acceleration_y)	0.00155	7
median frequency (angular_velocity_y)	0.00174	8
maximum (angular_velocity_x)	0.00240	9
standard deviation (angular_velocity_y)	0.00283	10

Experimental Results

To analyze the impact of the feature number on the performance of motion recognition of upper-limb exercises, experiments were conducted on the training set including a different number of features. The results of the classification error rate are shown in Figure 8. The general trends for the 3 classifiers are similar. With an increase in feature number, the classification error rate decreases rapidly. With a further increase in feature number, the classification error rate shows an increasing trend. The trend for kNN is not as stable as that of NB and RF. For kNN, there are several local optima with an increase in feature number.

Figure 8. Classification error rate vs feature number. kNN: k-nearest neighbor; NB: Naïve Bayes; RF: random forest.

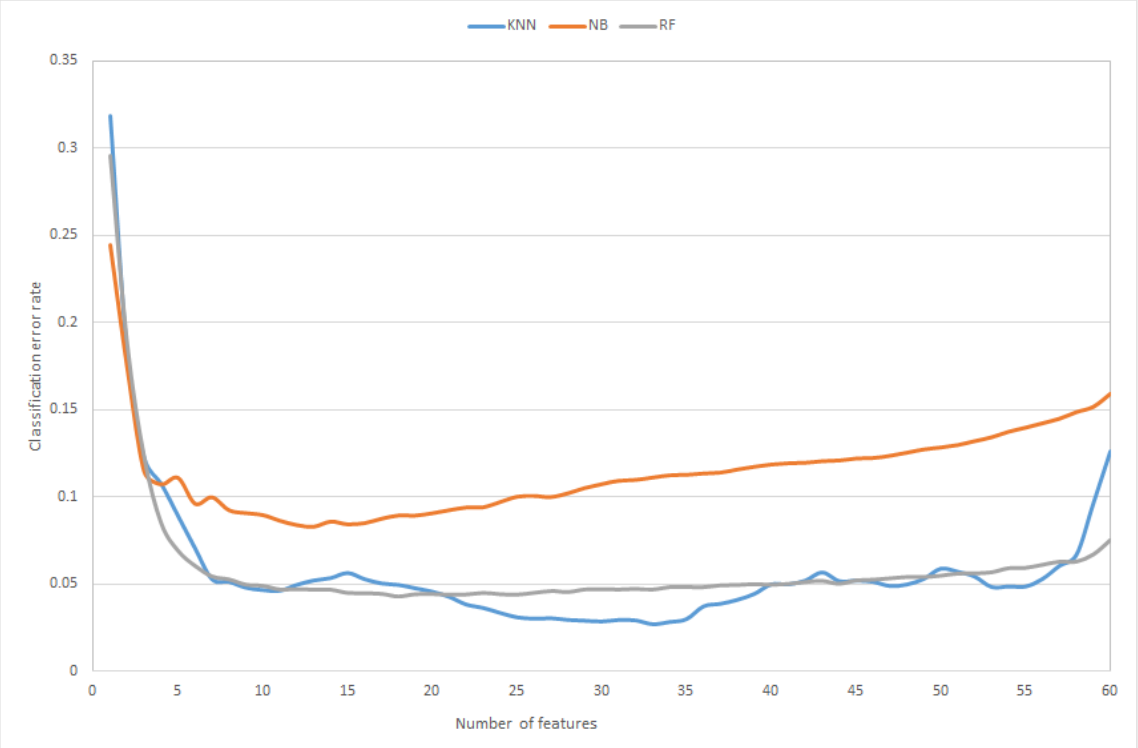


Table 2 shows the optimal number of features for the different classifiers. There is an obvious distinction between the optimal number of features for the different classifiers. Classifier kNN needs more features to achieve the minimum classification error rate than classifiers NB and RF.

Table 2. Optimal number of features.

Classifier	Optimal number of features
kNN[a]	33
NB[b]	13
RF[c]	18

[a]kNN: k-nearest neighbor.

[b]NB: Naïve Bayes.

[c]RF: random forest.

As wrapper methods usually achieve better classification performance than filter methods, and until now, there has been no hybrid feature selection method for motion recognition of rehabilitation exercises, we compared the motion recognition performance of the proposed FESCOM method with the traditional wrapper method. Experiments were conducted on the testing set by selecting the optimal feature set. For the traditional wrapper method, SFS was used to search for the optimal feature set from all 60 original features. For FESCOM, SFS was used to refine the features from the candidate feature set, composed of features with a significance index value smaller than 0.05 and ranked in ascending order. The criterion to set the threshold of the significance index was an assumption that the number of candidate features increased 30% (10 out of 33 features) from the highest optimal number of features in Table 2. The classification error rate is shown in Table 3. For both feature selection methods, the classification performance of kNN was better than that of NB and RF. The classification error rate of FESCOM using kNN and NB as the wrapping component was lower than the corresponding wrapper methods.

Table 3. Classification error rate.

Feature selection method	Classification error rate (%)		
	kNN[a]	NB[b]	RF[c]
Wrapper	2.2	13.4	6.7
FESCOM[d]	1.7	8.9	7.4

[a]kNN: k-nearest neighbor.

[b]NB: Naïve Bayes.

[c]RF: random forest.

[d]FESCOM: hybrid feature selection method by combining filter and wrapper methods.

The time consumed on feature selection in each iteration for both feature selection methods is listed in Table 4. As the candidate feature set of FESCOM is smaller than that of the wrapper method, the search time for FESCOM was much less than that of the wrapper method for all classifiers. For the same feature selection method, kNN needs much less search time than NB and RF.

Table 4. Search time for each iteration.

Feature selection method	Search time (seconds)		
	kNN[a]	NB[b]	RF[c]
Wrapper	23	159	876
FESCOM[d]	13	71	541

[a]kNN: k-nearest neighbor.

[b]NB: Naïve Bayes.

[c]RF: random forest.

[d]FESCOM: hybrid feature selection method by combining filter and wrapper methods.

Discussion

Principal Findings

This paper presents a hybrid feature selection method for motion recognition of upper-limb exercises. For motion recognition based on feature extraction and feature selection, the feature set used for the classification algorithm had a direct impact on the performance of classification. The experimental results in this study verified that recognition performance depends on the feature set. For all 3 classifiers in this study, the same trends existed: The classification error rate decreased to an optimum value when the number of features increased and increased with a further increase in feature number due to overfitting. The optimal number of features depended on the classifier. The optimal numbers of features for classifiers kNN, NB, and RF were 33, 13, and 18, respectively.

Each feature contributes differently to the motion classification task. Take the proposed FESCOM method as an example: The frequency-domain features contribute more than other features to recognition performance. When using the classifier kNN as the wrapping component, the top 3 significant features for motion recognition were average power of angular_velocity_x, average acceleration_x, and mean frequency of angular_velocity_x. When using the classifier NB as the wrapping component, the top 3 significant features for motion recognition were average power of acceleration_x, standard deviation of acceleration_y, and kurtosis of angular_velocity_x. When using the classifier RF as the wrapping component, the top 3 significant features for motion recognition were average power of angular_velocity_x, average power of acceleration_y, and mean frequency of angular_velocity_y.

Motion recognition performance also depends on the classifier. For both feature selection methods, the classification performance of kNN was the best, while NB was the worst classification performance. The proposed FESCOM method reduces the feature space and improves the time efficiency by filtering irrelevant features for motion classification.

Comparison With Previous Works

The common methods for motion recognition combine wearable sensing techniques and machine learning algorithms. Acceleration, angular velocity, or sEMG signals collected by wearable sensors are used to represent the motion characteristics. Cai et al [8] exploited sEMG signals and the SVM classifier for motion recognition of upper-limb exercises; 5 healthy participants participated in the experiments. The average recognition accuracy of 5 motions was 93.34%. Motion recognition of upper-limb exercises in [15] was based on acceleration, angular velocity data, and BPNN algorithm; 13 volunteers participated in the experiments. Five upper-limb exercises involving simple swinging and stretching movements were recognized with an accuracy of 85%-95%, while exercises consisting of spiral rotations were recognized with an accuracy of 60%. The knowledge-driven activity recognition method in [9] focused on egocentric video and accelerometer/gyroscope

data. Experiments were conducted on 3 public datasets, with a best recognition accuracy of 76.1%.

Using kNN, NB, and RF as the wrapping components, the recognition performance of FESCOM in this study achieved 98.3%, 91.1%, and 92.6%, respectively. Compared with previous studies on upper-limb motion recognition, the recognition performance of FESCOM is at the same level or even better than that in previous works. Time efficiency is one of the main concerns especially in real-time applications, such as motion recognition in autonomous rehabilitation systems. However, previous works seldom considered time efficiency. The FESCOM method in this study reduced the feature space and improved the time efficiency by filtering irrelevant features for motion classification. Compared with the search time of the traditional wrapper method, the search time of FESCOM using kNN, NB, and RF classifiers as the wrapping component reduced the search time by 43% (from 23 seconds to 13 seconds), 55% (from 159 seconds to 71 seconds), and 38% (from 876 seconds to 541 seconds), respectively. Hence, this study contributes by evaluating the number and types of features for different classification algorithms that achieve acceptable performance for motion recognition of upper-limb exercises.

Limitations

The FESCOM method proposed in this study has some limitations. It was only evaluated based on data from 21 healthy participants, and only 5 types of upper-limb exercises were considered in the experiments. However, the behavior of patients with a central nervous system lesion, such as that caused by stroke, may be very different from that of healthy participants. The experimental results may be different in such cases. The number of samples for training and testing is not high enough for machine learning algorithms, which may also affect the reliability. The customized IMU module in this work is just a prototype. The components in the sensor node are connected with cables. This may lead to unreliable connections, especially when used in movement conditions. Another drawback is that the validation of the system did not use real-time exercise examples.

In our future work, to further confirm the feasibility of FESCOM, we plan to extend our experiment considering the following aspects. First, we will evaluate and compare the performance of different methods in the filter and wrapper stage of FESCOM. Second, we will evaluate the performance of FESCOM considering more classifiers as the wrapping component in the wrapper stage, such as SVM and latent Dirichlet allocation. Third, we will evaluate the performance of FESCOM on more datasets, such as public datasets including more motion types and datasets including not only healthy participants but also real patients with different functional impairments in the recovery stage in a clinical situation. Fourth, we plan to improve the IMU node as an embedded system on a circuit board for real-time data collection and validate the whole system by real-time prediction of upper-limb exercises.

Conclusions

In this study, a hybrid feature selection method, FESCOM, was proposed for motion recognition of upper-limb exercises and evaluated using 5 types of upper-limb exercises performed by 21 healthy participants. The experimental results demonstrate that FESCOM is feasible for motion recognition of upper-limb exercises performed by healthy participants. FESCOM improves the recognition accuracy when using kNN and NB as the wrapping component and improves the time efficiency in the wrapper stage. The results also demonstrate that, for different classifiers, different feature sets are selected to achieve optimal performance. This work can be extended to provide motion recognition of more motion types and participants including healthy people and actual patients with minor motor damage.

Acknowledgments

The authors would like to thank the participants who made data collection possible. This study was supported in part by the National Key R&D Program of China (grant 2019YFC1710300) and the Sichuan Science and Technology Program (grants 2019YFS0019, 2020YFS0283, 2021YJ0184, and 2021YFS0152).

Conflicts of Interest

None declared.

Multimedia Appendix 1
Extended version of Table 1 with significance index of all 60 features.
[DOCX File , 16 KB - mhealth_v9i9e24402_app1.docx]

References

1. Chen Y, Abel KT, Janecek JT, Chen Y, Zheng K, Cramer SC. Home-based technologies for stroke rehabilitation: A systematic review. Int J Med Inform 2019 Mar;123:11-22 [FREE Full text] [doi: 10.1016/j.ijmedinf.2018.12.001] [Medline: 30654899]
2. Chae SH, Kim Y, Lee K, Park H. Development and Clinical Evaluation of a Web-Based Upper Limb Home Rehabilitation System Using a Smartwatch and Machine Learning Model for Chronic Stroke Survivors: Prospective Comparative Study. JMIR Mhealth Uhealth 2020 Jul 09;8(7):e17216 [FREE Full text] [doi: 10.2196/17216] [Medline: 32480361]

3. Parker J, Powell L, Mawson S. Effectiveness of Upper Limb Wearable Technology for Improving Activity and Participation in Adult Stroke Survivors: Systematic Review. J Med Internet Res 2020 Jan 08;22(1):e15981 [FREE Full text] [doi: 10.2196/15981] [Medline: 31913131]
4. Agarwal V, Smuck M, Tomkins-Lane C, Shah NH. Inferring Physical Function From Wearable Activity Monitors: Analysis of Free-Living Activity Data From Patients With Knee Osteoarthritis. JMIR Mhealth Uhealth 2018 Dec 18;6(12):e11315 [FREE Full text] [doi: 10.2196/11315] [Medline: 30394876]
5. O'Brien MK, Shawen N, Mummidisetty CK, Kaur S, Bo X, Poellabauer C, et al. Activity Recognition for Persons With Stroke Using Mobile Phone Technology: Toward Improved Performance in a Home Setting. J Med Internet Res 2017 May 25;19(5):e184 [FREE Full text] [doi: 10.2196/jmir.7385] [Medline: 28546137]
6. Zhang Z, Liparulo L, Panella M, Gu X, Fang Q. A Fuzzy Kernel Motion Classifier for Autonomous Stroke Rehabilitation. IEEE J Biomed Health Inform 2016 May;20(3):893-901. [doi: 10.1109/JBHI.2015.2430524] [Medline: 25956000]
7. Cui C, Bian G, Hou Z, Zhao J, Su G, Zhou H, et al. Simultaneous Recognition and Assessment of Post-Stroke Hemiparetic Gait by Fusing Kinematic, Kinetic, and Electrophysiological Data. IEEE Trans Neural Syst Rehabil Eng 2018 Apr;26(4):856-864. [doi: 10.1109/TNSRE.2018.2811415] [Medline: 29641390]
8. Cai S, Chen Y, Huang S, Wu Y, Zheng H, Li X, et al. SVM-Based Classification of sEMG Signals for Upper-Limb Self-Rehabilitation Training. Front Neurorobot 2019;13:31 [FREE Full text] [doi: 10.3389/fnbot.2019.00031] [Medline: 31214010]
9. Huang Y, Yang X, Gao J, Sang J, Xu C. Knowledge-driven Egocentric Multimodal Activity Recognition. ACM Trans. Multimedia Comput. Commun. Appl 2021 Jan 28;16(4):133. [doi: 10.1145/3409332]
10. Davoudi A, Wanigatunga AA, Kheirkhahan M, Corbett DB, Mendoza T, Battula M, et al. Accuracy of Samsung Gear S Smartwatch for Activity Recognition: Validation Study. JMIR Mhealth Uhealth 2019 Feb 06;7(2):e11270 [FREE Full text] [doi: 10.2196/11270] [Medline: 30724739]
11. Li K, Habre R, Deng H, Urman R, Morrison J, Gilliland FD, et al. Applying Multivariate Segmentation Methods to Human Activity Recognition From Wearable Sensors' Data. JMIR Mhealth Uhealth 2019 Feb 07;7(2):e11201 [FREE Full text] [doi: 10.2196/11201] [Medline: 30730297]
12. Kopke JV, Hargrove LJ, Ellis MD. Applying LDA-based pattern recognition to predict isometric shoulder and elbow torque generation in individuals with chronic stroke with moderate to severe motor impairment. J Neuroeng Rehabil 2019 Mar 05;16(1):35 [FREE Full text] [doi: 10.1186/s12984-019-0504-1] [Medline: 30836971]
13. Castiblanco JC, Ortmann S, Mondragon IF, Alvarado-Rojas C, Jöbges M, Colorado JD. Myoelectric pattern recognition of hand motions for stroke rehabilitation. Biomedical Signal Processing and Control 2020 Mar;57:101737. [doi: 10.1016/j.bspc.2019.101737]
14. Shawen N, Lonini L, Mummidisetty CK, Shparii I, Albert MV, Kording K, et al. Fall Detection in Individuals With Lower Limb Amputations Using Mobile Phones: Machine Learning Enhances Robustness for Real-World Applications. JMIR Mhealth Uhealth 2017 Oct 11;5(10):e151 [FREE Full text] [doi: 10.2196/mhealth.8201] [Medline: 29021127]
15. Lin H, Chiang S, Lee K, Kan Y. An activity recognition model using inertial sensor nodes in a wireless sensor network for frozen shoulder rehabilitation exercises. Sensors (Basel) 2015 Jan 19;15(1):2181-2204 [FREE Full text] [doi: 10.3390/s150102181] [Medline: 25608218]
16. Wu Q, Shao J, Wu X, Zhou Y, Liu F, Xiao F. Upper Limb Motion Recognition Based on LLE-ELM Method of sEMG. Int. J. Patt. Recogn. Artif. Intell 2017 Mar 30;31(06):1750018. [doi: 10.1142/s0218001417500185]
17. Li J, Cheng K, Wang S, Morstatter F, Trevino RP, Tang J, et al. Feature Selection: A data perspective. ACM Comput. Surv 2018 Jan 12;50(6):94. [doi: 10.1145/3136625]
18. Chaudhry MU, Lee J. Feature Selection for High Dimensional Data Using Monte Carlo Tree Search. IEEE Access 2018;6:76036-76048. [doi: 10.1109/access.2018.2883537]
19. Ramezani R, Zhang W, Xie Z, Shen J, Elashoff D, Roberts P, et al. A Combination of Indoor Localization and Wearable Sensor-Based Physical Activity Recognition to Assess Older Patients Undergoing Subacute Rehabilitation: Baseline Study Results. JMIR Mhealth Uhealth 2019 Jul 10;7(7):e14090 [FREE Full text] [doi: 10.2196/14090] [Medline: 31293244]
20. Wang A, Chen G, Wu X, Liu L, An N, Chang C. Towards Human Activity Recognition: A Hierarchical Feature Selection Framework. Sensors (Basel) 2018 Oct 25;18(11):3629 [FREE Full text] [doi: 10.3390/s18113629] [Medline: 30366461]
21. Fang H, He L, Si H, Liu P, Xie X. Human activity recognition based on feature selection in smart home using back-propagation algorithm. ISA Trans 2014 Sep;53(5):1629-1638. [doi: 10.1016/j.isatra.2014.06.008] [Medline: 25016308]
22. Zhou H, You M, Liu L, Zhuang C. Sequential data feature selection for human motion recognition via Markov blanket. Pattern Recognition Letters 2017 Jan;86:18-25. [doi: 10.1016/j.patrec.2016.12.008]
23. Hsu H, Hsieh C, Lu M. Hybrid feature selection by combining filters and wrappers. Expert Systems with Applications 2011 Jul;38(7):8144-8150. [doi: 10.1016/j.eswa.2010.12.156]
24. Yao C, Liu YF, Jiang B, Han J, Han J. LLE Score: A New Filter-Based Unsupervised Feature Selection Method Based on Nonlinear Manifold Embedding and Its Application to Image Recognition. IEEE Trans Image Process 2017 Nov;26(11):5257-5269. [doi: 10.1109/TIP.2017.2733200] [Medline: 28767370]

25. Lazar C, Taminau J, Meganck S, Steenhoff D, Coletta A, Molter C, et al. A survey on filter techniques for feature selection in gene expression microarray analysis. IEEE/ACM Trans Comput Biol Bioinform 2012;9(4):1106-1119. [doi: 10.1109/TCBB.2012.33] [Medline: 22350210]
26. Banos O, Damas M, Pomares H, Prieto A, Rojas I. Daily living activity recognition based on statistical feature quality group selection. Expert Systems with Applications 2012 Jul;39(9):8013-8021. [doi: 10.1016/j.eswa.2012.01.164]
27. Hong F, You S, Wei M, Zhang Y, Guo Z. MGRA: Motion Gesture Recognition via Accelerometer. Sensors (Basel) 2016 Apr 13;16(4):530 [FREE Full text] [doi: 10.3390/s16040530] [Medline: 27089336]
28. Peng H, Long F, Ding C. Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans Pattern Anal Mach Intell 2005 Aug;27(8):1226-1238. [doi: 10.1109/TPAMI.2005.159] [Medline: 16119262]
29. Mansbridge N, Mitsch J, Bollard N, Ellis K, Miguel-Pacheco GG, Dottorini T, et al. Feature Selection and Comparison of Machine Learning Algorithms in Classification of Grazing and Rumination Behaviour in Sheep. Sensors (Basel) 2018 Oct 19;18(10):3532 [FREE Full text] [doi: 10.3390/s18103532] [Medline: 30347653]
30. Wuniri Q, Wei H, Liu Y, Lin X, Liu L, Yu Z. A Generic-Driven Wrapper Embedded With Feature-Type-Aware Hybrid Bayesian Classifier for Breast Cancer Classification. IEEE Access 2019;7:119931-119942. [doi: 10.1109/access.2019.2932505]
31. Camargo J, Young A. Feature Selection and Non-Linear Classifiers: Effects on Simultaneous Motion Recognition in Upper Limb. IEEE Trans Neural Syst Rehabil Eng 2019 Apr;27(4):743-750. [doi: 10.1109/TNSRE.2019.2903986] [Medline: 30869626]
32. Schaffernicht E, Stephan V, Gross HM. An Efficient Search Strategy for Feature Selection Using Chow-Liu Trees. In: de Sá JM, Alexandre LA, Duch W, Mandic D, editors. Artificial Neural Networks. ICANN 2007. Lecture Notes in Computer Science, vol 4669. Berlin Heidelberg: Springer; Sep 09, 2007:190-199.
33. Xue X, Yao M, Wu Z. A novel ensemble-based wrapper method for feature selection using extreme learning machine and genetic algorithm. Knowl Inf Syst 2017 Nov 17;57(2):389-412. [doi: 10.1007/s10115-017-1131-4]
34. Chen G, Chen J. A novel wrapper method for feature selection and its applications. Neurocomputing 2015 Jul;159:219-226. [doi: 10.1016/j.neucom.2015.01.070]
35. Zhang R, Nie F, Li X, Wei X. Feature selection with multi-view data: A survey. Information Fusion 2019 Oct;50:158-167. [doi: 10.1016/j.inffus.2018.11.019]
36. Manbari Z, AkhlaghianTab F, Salavati C. Hybrid fast unsupervised feature selection for high-dimensional data. Expert Systems with Applications 2019 Jun;124:97-118. [doi: 10.1016/j.eswa.2019.01.016]
37. Yu L, Xiong D, Guo L, Wang J. A remote quantitative Fugl-Meyer assessment framework for stroke patients based on wearable sensor networks. Comput Methods Programs Biomed 2016 May;128:100-110. [doi: 10.1016/j.cmpb.2016.02.012] [Medline: 27040835]
38. Raya MA, Gailey RS, Gaunaurd IA, Jayne DM, Campbell SM, Gagne E, et al. Comparison of three agility tests with male servicemembers: Edgren Side Step Test, T-Test, and Illinois Agility Test. J Rehabil Res Dev 2013;50(7):951-960 [FREE Full text] [doi: 10.1682/JRRD.2012.05.0096] [Medline: 24301432]
39. Shri TKP, Sriraam N. Comparison of t -test ranking with PCA and SEPCOR feature selection for wake and stage 1 sleep pattern recognition in multichannel electroencephalograms. Biomedical Signal Processing and Control 2017 Jan;31:499-512. [doi: 10.1016/j.bspc.2016.09.016]
40. Pudil P, Novovičová J, Kittler J. Floating search methods in feature selection. Pattern Recognition Letters 1994 Nov;15(11):1119-1125. [doi: 10.1016/0167-8655(94)90127-9]
41. Altman NS. An Introduction to Kernel and Nearest-Neighbor Nonparametric Regression. The American Statistician 1992 Aug;46(3):175-185. [doi: 10.2307/2685209]
42. Kuncheva LI. On the optimality of Naïve Bayes with dependent binary features. Pattern Recognition Letters 2006 May;27(7):830-837. [doi: 10.1016/j.patrec.2005.12.001]
43. Cutler DR, Edwards TC, Beard KH, Cutler A, Hess KT, Gibson J, et al. Random forests for classification in ecology. Ecology 2007 Nov;88(11):2783-2792. [doi: 10.1890/07-0539.1] [Medline: 18051647]

Abbreviations

BPNN: back-propagation neural network
FESCOM: hybrid feature selection method by combining filter and wrapper methods
FMA: Fygl-Meyer Assessment
IMU: inertial measurement unit
kNN: k-nearest neighbor
MGRA: motion gesture recognition system via accelerometer
NB: Naïve Bayes
RF: random forest
sEMG: surface electromyogram

SFS: sequential feature selection
SVM: support vector machine

Edited by L Buis; submitted 17.09.20; peer-reviewed by EJ Yang, J Ruokolainen, E Vanegas; comments to author 29.01.21; revised version received 30.04.21; accepted 15.07.21; published 02.09.21.

Please cite as:
Li Q, Liu Y, Zhu J, Chen Z, Liu L, Yang S, Zhu G, Zhu B, Li J, Jin R, Tao J, Chen L
Upper-Limb Motion Recognition Based on Hybrid Feature Selection: Algorithm Development and Validation
JMIR Mhealth Uhealth 2021;9(9):e24402
URL: https://mhealth.jmir.org/2021/9/e24402
doi:10.2196/24402
PMID:34473067

Original Paper

Psychological Effects of Heart Rate and Physical Vibration on the Operation of Construction Machines: Experimental Study

Nobuki Hashiguchi[1], PhD; Jianfei Cao[1], MSc; Yeongjoo Lim[2], PhD; Shinichi Kuroishi[3], MSc; Yasuhiro Miyazaki[4], MSc; Shigeo Kitahara[4], BSc; Shintaro Sengoku[5], PhD; Katsushi Matsubayashi[6], PhD; Kota Kodama[1], PhD

[1]Graduate School of Technology Management, Ritsumeikan University, Ibaraki, Japan

[2]Faculty of Business Administration, Ritsumeikan University, Ibaraki, Japan

[3]Metropolitan Area Branch Civil Engineering Department, Kumagai Gumi Co, Ltd, Shinjuku-ku, Japan

[4]Civil Engineering Business Headquarters, Kumagai Gumi Co, Ltd, Shinjuku-ku, Japan

[5]School of Environment and Society, Tokyo Institute of Technology, Minato-ku, Japan

[6]National Institute of Technology, Tokyo College, Tokyo, Japan

Corresponding Author:
Kota Kodama, PhD
Graduate School of Technology Management
Ritsumeikan University
2-150
Iwakuracho
Ibaraki, 5678570
Japan
Phone: 81 0726652448
Email: kkodama@fc.ritsumei.ac.jp

Abstract

Background: A construction method has emerged in which a camera is installed around a construction machine, and the operator remotely controls the machine while synchronizing the vibration of the machine with the images seen from the operator's seat using virtual reality (VR) technology. Indices related to changes in heart rate (HR) and physical vibration, such as heart rate variability (HRV) and multiscale entropy (MSE), can then be measured among the operators. As these indices are quantitative measures of autonomic regulation in the cardiovascular system, they can provide a useful means of assessing operational stress.

Objective: In this study, we aimed to evaluate changes in HR and body vibration of machine operators and investigate appropriate methods of machine operation while considering the psychological load.

Methods: We enrolled 9 remote operators (18-50 years old) in the experiment, which involved 42 measurements. A construction machine was driven on a test course simulating a construction site, and three patterns of operation—riding operation, remote operation using monitor images, and VR operation combining monitor images and machine vibration—were compared. The heartbeat, body vibration, and driving time of the participants were measured using sensing wear made of a woven film-like conductive material and a three-axis acceleration measurement device (WHS-2). We used HRV analysis in the time and frequency domains, MSE analysis as a measure of the complexity of heart rate changes, and the ISO (International Standards Organization) 2631 vibration index. Multiple regression analysis was conducted to model the relationship among the low frequency (LF)/high frequency (HF) HRV, MSE, vibration index, and driving time of construction equipment. Efficiency in driving time was investigated with a focus on stress reduction.

Results: Multiple comparisons conducted via the Bonferroni test and Kruskal-Wallis test showed statistically significant differences (P=.05) in HRV-LF/HF, the vibration index, weighted acceleration, motion sickness dose value ($MSDV_z$), and the driving time among the three operation patterns. The riding operation was found to reduce the driving time of the machine, but the operation stress was the highest in this case; operation based on the monitor image was found to have the lowest operation stress but the longest operation time. Multiple regression analysis showed that the explanatory variables (LH/HF), RR interval, and vibration index ($MSDV_z$ by vertical oscillation at 0.5-5 Hz) had a negative effect on the driving time (adjusted coefficient of determination R^2=0.449).

Conclusions: A new method was developed to calculate the appropriate operating time by considering operational stress and suppressing the physical vibration within an acceptable range. By focusing on the relationship between psychological load and

physical vibration, which has not been explored in previous studies, the relationship of these variables with the driving time of construction machines was clarified.

(JMIR Mhealth Uhealth 2021;9(9):e31637) doi:10.2196/31637

KEYWORDS

heart rate variability; complexity; vital signs; vibration at work; stress; wearable technology; remote operation; monitoring

Introduction

Background

Construction work in Japan, which is often affected by natural disasters such as large-scale earthquakes, windstorms, floods, and volcanic disasters, has been attracting attention for initiating the remote operation of construction machines from a safe location in cases where the actual location faces a risk of secondary disasters such as mudslides [1,2]. Compared to actual machine operation, remote operation requires care and consideration because it is difficult to ascertain the situation of the construction machine and the working environment; this tends to place a higher psychological load on the operator. In this construction method, the operator recognizes the actual situation of the construction machine (inclination and shaking) through a monitor image and the vibration of the construction machine. The key challenge is to operate the machine reliably and efficiently without increasing the psychological load of the operation. Very few reports on typical construction focus on the psychological load because of the high priority given to avoiding physical hazards for workers [3]. Therefore, this study focused on the psychological load and stress experienced by technicians operating construction equipment to determine efficient and appropriate operations.

Stress resulting from physical and psychological loads in any job can reduce job efficiency by decreasing the sense of satisfaction and well-being. When analyzing stress, observing variations in the heartbeat interval provides a quantitative measure of the autonomic regulation of the cardiovascular system in response to stressors [4,5]. Regarding psychological workload, workers need to be consistently aware of many things, and an imbalance in the resources contributing to psychological workload could pose a safety risk. The assessment of psychological load has gained precedence in many job tasks [6]. Therefore, the relationship of heart rate variability (HRV) and multiscale entropy (MSE) with psychological workload has been investigated in several studies [6-9].

Recent advances in wearable technology have provided an opportunity to easily monitor the biometric information and physical condition of the subjects. The use of wearable devices to monitor autonomic nervous system activity through heart rate (HR) observations is economical, with easy access to data [10,11]. Evaluation of daily changes in HR provides useful information for understanding heart health status with respect to workloads [12], mental states [13,14], and physical conditions [15]. Previous studies have reported that a 5-minute HRV measurement provides a highly accurate analysis [5].

Life Events and Stress

Stress can be attributed to multiple factors, such as physical, chemical, and biological stressors. Holmes and Rahe [16] have pointed out the psychological and social stressors in social life. Factors that contribute to these stressors include relationships, family problems, and occupational problems. Lazarus and Cohen [17] also argued that the daily hustle of life, comprising minor daily irritations, contributes more to the negative effects on our physical and mental health than the less frequent serious life events. Psychological and social stressors are complex and diversify annually with the changes in our environment and social conditions.

When the human body encounters an unpleasant or harmful event, a defensive reaction of the body and mind occurs. The level of arousal increases to alert us to the outside world, and anxious feelings emerge. In the body, the autonomic nervous system, called the sympathetic nervous system, and the endocrine system, which secretes adrenal cortical and other hormones, becomes more active [5,18]. Through experiments on animals, Hance Selye [19] revealed that in contrast to the usual level of resistance to stress, the warning response phase includes a shock phase in which resistance decreases immediately after encountering the stressor, and then shifts to an antishock phase in which resistance increases. The defensive response of the body and mind after encountering a stressor changes significantly over time. The physical activity and resistance of individuals drop significantly below their usual levels during the stress phase. Then, in the antishock phase, adrenaline is secreted, and the sympathetic nervous system becomes more active, resulting in higher levels of arousal and activity [20]. The liver produces glucose to supply the whole body with energy for activity, and the bronchi tend to become thicker, and respiration becomes faster to take in more oxygen. Fluctuations such as an increased HR occur to pump large amounts of nutrients and oxygenated blood throughout the body [21].

Research Objectives

In this study, we aimed to investigate the effects of stress and body vibrations on HR and consider machine operation that accounts for the load caused by work vibration during riding and remote operation based on the characteristics of HR information.

The results of this analysis are expected to lead to a new computational model for evaluating operation stress and driving time according to the widely adopted sensing wear and vital signs collected using HR sensors and three-axis accelerometers.

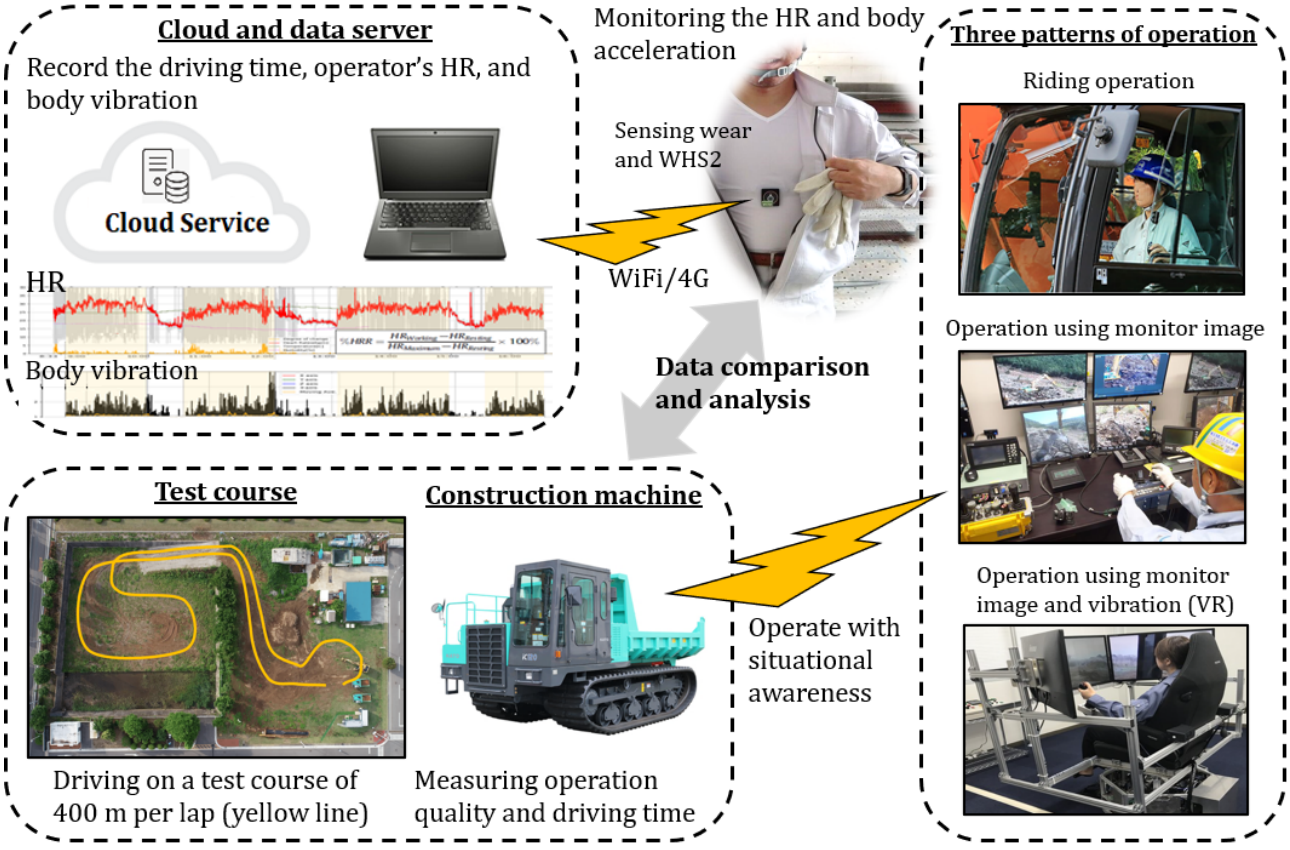
Methods

Measurement Tools

Multimedia Appendix 1 lists the measurement devices and infrastructure considered in this study. We measured the HR and physical activity of machine operators on the basis of the electrocardiogram (ECG) signals captured using the sensing wear worn by the operators. Sensing wear is an underwear-type shirt fitted with a biometric information sensor (for detecting HR). As sensing wear clothing is made of stretchable fabric, stretchable ECG electrodes were integrated with the hardware for measuring the HR. The HR was detected using the RR intervals (RRIs) in the ECG signals. The RRI and body acceleration extracted from the ECGs were measured to evaluate the load of the operator in the work environment. The devices used for physiological measurements were WHS-2 for HR measurements and three-axis accelerometers for measuring vital signs (Union Tool Co Ltd), COCOMI (Toyobo Co Ltd) as the sensing wear, and a CC2650 data acquisition device (Texas Instruments).

Using a Bluetooth low-energy device, the HR and three-axis acceleration data were sent to the data acquisition device used by the operators. Subsequently, the data from the acquisition device were transmitted to and stored on the cloud server installed on the network using the established wireless access point (using transfer devices based on WiFi and 4G) in the work area. The measurement device and system configuration used in this study are shown in Figure 1.

Figure 1. Schematic representation of the monitoring platform. HR: heart rate; VR: virtual reality.



Participants

The data were collected at the Tsukuba Technical Research Institute of Kumagai Corporation (Kumagai-gumi, Inc, Chiyoda-ku, Tokyo), a construction company, on July 29 and 30, 2020, and February 18, 2021. A construction company employee (who is also a member of our research team) recruited 20 operators within the company, and engineers who responded to the call participated in this experiment. Participants were included if they were healthy adults aged 18-50 years and excluded if they had any neurological or cardiovascular diseases. As the operators participating in this experiment had highly specialized knowledge and skills and were busy with their daily work, it was difficult for many applicants to participate in the experiment. Hence, 9 construction technicians trained in remote control were selected from construction companies (age: 35.6, SD 12.8 years; height: 168.7, SD 4.1 cm; weight: 71.1, SD 13.2 kg; BMI: 24.9, SD 4.4). Participants drove a crawler carrier (IC120-2, KATO WORKS Co, Ltd) while wearing a device on a test course that mimicked a construction site with a 400-meter lap. All participants were familiar with the experimental procedures.

Protocol

Our research group investigated the potential risks and discomfort of the participants as well as the privacy issues relating to data collection prior to commencing data collection. The sensing garment was confirmed to be a noninvasive device that does not interfere with machine operation. In accordance with the Declaration of Helsinki, the human genome, and the Universal Declaration of Human Rights, the protocol for data collection was approved by the Ritsumeikan University's

Research Ethics Review Board (number BKC-2019-038). In addition, an explanation of the participants' rights was included in the informed consent form distributed to all participants before data collection, ensuring the confidentiality of the participants' data. Employee names were not used in the experiments and data analysis to minimize the risk of disclosure of personal information. Instead, a personal identification code (identifier) was assigned to each participant.

Data Collection and Analysis

The SPSS Version 26 for Windows (IBM Corp) and Excel add-in software Bell Curve (Social Survey Research Information Co, Ltd) for Excel version 3.21(Microsoft Corporation) were used as tools for conducting statistical analyses.

The measurement time was from 9 AM to 4 PM, and the HR information and body vibration data of the operators were collected at any time. We collected 42 data sets from 9 participants measured in approximately 5 minutes, excluding preparation time and breaks. The measurement data of the operators in the working environment are presented in Multimedia Appendix 2. All participants were male and were asked to provide information on their age, height, and weight. As the cardiopulmonary function was intended to exclude unhealthy participants from the measurement, participants were also asked about their history of cardiovascular disease and their current health status.

We checked for the Hawthorne effect when participants were examined [22]. In this experiment, our research team did not monitor the participants' activities. It stayed away from the remote-control seat area and recorded the work using two cameras installed in the control area. Our study focused on the mental load that occurs during daily construction machine operation tasks. Therefore, before starting the measurements, we explained to the participants that this study was intended to measure the load during operation but not their operating skills, and we instructed them not to deviate from their daily operating mindset.

HRV Metrics

HRV is associated with other aspects of health that are directly affected by autonomic function, such as self-regulation, and psychological and physiological stress [5,9,10]. A low HRV indicates inappropriate coordination between the sympathetic and parasympathetic nervous systems and is a reliable predictor of future cardiovascular disease [4,5]. Therefore, HRV measurements provide important information for assessing physical functioning and help identify the risk of physical fatigue and debilitation [23,24].

HR and HRV metrics have recently shown promise in multiple applications for health care providers [25-27]. Although studies performing HRV analysis are being reported since a long time, further improvements in technology and the interest of many researchers and physicians have brought more attention to this field [28]. Despite concerns about the validity of certain metrics of HRV data for measuring sympathetic balance [29,30] (eg, low frequency [LF] power of 0.04-0.15 Hz and the ratio of LF to high frequency [HF] power of 0.15-0.4 Hz in HRV), a number of previous studies support the notion that HRV analysis could reveal the balance of sympathetic and parasympathetic tones in the body [31]. The autonomic response to psychological stress has been studied using HRV [7]. Notable studies on HRV have emphasized the value of objective measures of stress in health care workers, and Joseph et al [27] found that self-reported stress was associated with proportionally elevated physiological levels of stress. These results provide compelling evidence for physicians, especially those who routinely perform surgical medical duties under tight time constraints, to assess their own stress. The widespread use of objective and ecologically valid measures of stress might provide important clues for understanding and reducing the psychological burden of stressful situations [32].

There are several widely accepted HRV metrics [32,33]. HRV measurements are classified into two categories: time-domain measures and frequency-domain measures. HRV metrics include the root mean square of the continuous difference (RMSSD; time domain), SD of the RRI (SDRR; time domain) and the LF/HF ratio (frequency domain). In a previous study involving corporate employees, RMSSD values were found to be related to perceived mental stress [34], with lower values indicating higher stress. The RMSSD metric is less sensitive to the number of missing data points. Therefore, the RMSSD can be seen as a more robust metric for evaluating patients with low data quality. The SDRR is calculated from the SD of normal RR intervals, and the lower the SDRR, the lower the HRV [33].

It is important to note that HRV measurements are derived from RR data and affected by the duration of the time series (number of data points), time of day, body orientation, and activity being performed. Where possible, these factors are derived using 5-minute RRIs and provide values for each activity, although Troubat et al [35] found that even brief periods of mental stress are associated with lower mean HRV values.

MSE Metrics

MSE is an analytical algorithm that has gained popularity in the last 20 years to evaluate the complexity of time series at various time scales [36]. The physiological systems involved in maintaining stable health and well-being are complex and are affected by multiple interactions within and between system components. The complexity of the time series data being analyzed is reflected in the temporal structure of the variability of the output signal [37,38]. Entropy has been recognized as an excellent indicator of system complexity by applying and calculating the dynamics related to the HR, brain waves, and body sway [39]. Low entropy is associated with frailty, fatigue, aging, and functional impairment, whereas high entropy is associated with a greater ability to adapt to a changing environment [40,41]. Entropy has been reported as a reliable marker of neurophysiological complexity and adaptability in autonomic and somatic nervous systems [38]. In this study, the numerical value of entropy confirmed that adaptive capacity reduced because of task fatigue. The entropy value is obtained by plotting the entropy value of each coarse-grained time series as a function of the scale. The cardiac entropy index shows the area under the corresponding MSE curve (area calculated using the trapezoidal formula), and this area is treated as the entropy value [37,38,41].

Since its conception, the MSE algorithm has been applied to several analyses with significant success [42,43]. However, concerns have been raised about the statistical unreliability of the sample entropy of the coarse-grained series as the time scale factor of the MSE increases [44]. In recent years, a number of improved algorithms have been presented to address this concern, and these can be applied with satisfactory accuracy in the analysis of relatively small time series data sets having 750 points or less [41,43]. To calculate the complexity index, the time scale in our study was chosen from 1 to 14 [38]. In the analysis of the HR and MSE data of the participants, m=2 (vector length of time series) and r=0.15% (the similarity criterion used to compare vectors) of the SD of the original time series were used to calculate the sample entropy [38,43,44]; moreover, the refined composite multiscale entropy (RCMSE) [41] was used as the calculation algorithm of the MSE.

Physical Vibration

Vibration occurs when the body is exposed to internal or external forces. Physical factors such as noise, heat, vibration, and radiation are environmental stressors with many stimuli that are detrimental to health and can alter bodily functions [45]. Vibrations can be harmful to human health, depending on the intensity and duration of exposure. The ISO (International Standards Organization) 2631-1 developed in 1997 [46] provides guidance on the use of methods to assess human exposure to vibration. For this purpose, frequency weighting and magnification of each evaluation axis are applied, as human response to vibration and its effects depend on the frequency of the vibration, its direction, and the studied effect (health, comfort, task) [45-47].

The transmission of vibrations from external systems to the human body has a significant impact on comfort, performance, and health. As the actual operation of the construction machinery and the remote-control seat including virtual reality (VR) are dynamic systems, the related transmission depends on the frequency and direction of the input motion. The transmission rate of vibration also depends on the characteristics of the seat exposed to the vibrations. On-road and off-road vehicles are exposed to vibrations caused by uneven road and soil profiles, and by moving elements in the machine. This is also the case for technical vehicles and wheelchair systems. Vibrations in the frequency range below 10-12 Hz affect the entire human body, whereas vibrations above 12 Hz have only localized effects [48]. LF (4-6 Hz) cyclical movements, such as vehicle tires rolling over an uneven road, may cause the body to resonate. Exposure to vibration in a seated position can cause muscle fatigue, weaken soft tissues, and increase the strain on the operator's back and whole body [49]. Continued external forced body vibration might lead to unpleasant symptoms such as lassitude, discomfort, and in severe cases, vomiting [48,50]. The vibration indices used for measuring physical vibration from the output of the three-axis accelerometer attached to the sensing wear while the machine operator sat on the remote-control seat are listed in Multimedia Appendix 3.

The physical vibration of the operator at the observation time (exposure time) T can be expressed by the root mean square weighted acceleration (Aw). The total vibration acceleration at each sample time, Aw(t), is the instantaneous value of the frequency-corrected acceleration (m/s^2), and it is the composite of the accelerations along each axis occurring in the vertical, horizontal, and lateral directions. Further, acc is the vibration acceleration along each axis (m/s^2) and a function of time. Aw is the basis for evaluating the effect of vibration on the human body according to ISO 2631. Health hazards and discomfort caused by vibration acceleration are affected not only by steady vibration but also by occasional shocks [45,46]. However, because Aw is an effective value and is averaged over the observation time, the impact of shocks can be possibly underestimated.

The vibration doses value (VDV) index defines the amount of vibration exposure. Instead of determining the change in acceleration over time by squaring the acceleration, it is determined by quadrature; the VDV determined using the root of the fourth power is more sensitive to peak values than the root of the second power; it is not averaged over the observation time and represents the entire vibration exposure during the observation time T [45,46].

The motion sickness dose value ($MSDV_{z)}$ is calculated by correcting the vibration acceleration of the vertical axis of the operator using the frequency correction factor Wf [46,50]. $MSDV_z$ is an index of vertical vibration of approximately 0.5-5 Hz, and experiments have shown that it is affected by the discomfort and stress of the ride in the passenger seat of the vehicle [48].

The vibration index of the operators' body vibration in this study is expressed using Aw, VDV, and $MSDV_z$.

Workload (Percentage Heart Rate Reserve [%HRR])

Hwang et al [51] suggested that caution should be exercised while sustaining a 30-40% HRR among construction workers, and Norton et al [52] suggested that a 40-60% HRR lasting 30-60 minutes is equivalent to a moderate physical load for adequate health care of sedentary persons. Compared to construction workers (eg, scaffolders and steel handlers), who are often exposed to physical loads that exceed workload limits, construction equipment operators are exposed to higher psychological loads and stresses. Although psychological load has a negligible effect on the HR when measured over a long period, it may affect the %HRR for a short period of time [52].

HRR is a measure of the workload or pressure intensity at work, associated with muscle activity [53]. Equation 1 depicts how it is estimated:

$$HRR = (HR_{working} - HR_{resting}) / (HR_{maximum} - HR_{resting}) \times 100\ (\%) \quad \textbf{(1)}$$

where $HR_{working}$ is the mean working heart rate, $HR_{resting}$ it the resting heart rate, and $HR_{maximum}$ is the maximum heart rate based on age [51,53].

Removal of Artifacts

Two types of outliers are commonly found in heartbeat interval time series because of error beats and artifacts. These outliers have no physiological significance. However, artifacts can

significantly distort measurements in the time and frequency domains, increasing the power in all frequency bands [37]. For HRV data, a value can be considered valid if the clean segment is long enough in the time series to calculate the power in the frequency band. For example, it has been pointed out that at least 2.5 minutes of clean data is needed to estimate LF power [54]. Furthermore, for MSE, if the RRIs of the heartbeats differ by several orders of magnitude from the mean of the time series, it may have a significant impact on the entropy calculation [37]. The data set collected in this research was filtered to exclude artifacts, ventricular extrasystoles, and undetected heartbeats [37,44,55]. Briefly, at the center point of a moving window of length l, anything outside the interval  was excluded.  represents the mean of the data points within that moving window, calculated excluding the center point, and a is a positive number less than or equal to 1. In this study, we used l=41 and a=0.2 [37,44,55].

Hypotheses Development

Based on the research objectives and literature review, the following hypotheses were developed:

H1: In the driving time of the construction machine, differences in the operating environments during the riding operation, remote operation using the monitor image, and remote operation by VR appear in the parameters of the HRV and MSE.

H2: The magnitude of vibration in the operating environment has a negative relationship with the parameters of the HRV and MSE.

H3: The driving time of the construction machine has a negative relationship with the magnitude of vibration of the operating environment and the parameters of the HRV and MSE.

Results

Normality Tests for Data

The normality of the collected data was evaluated. When the sample size is greater than 50, only the Kolmogorov-Smirnov test is suitable to determine normality [56]. However, in this research, the sample size was not sufficiently large, and the data for which the normality could be confirmed using the Shapiro-Wilk and Kolmogorov-Smirnov tests were considered to be normally distributed (see Multimedia Appendix 4). In both the tests, the null hypothesis assumes that the dataset is normally distributed, with an alpha null hypothesis going further to assume that the data set is normally distributed with P=.05 [57]. The data on LF, MSE, %HRR, and VDV satisfied the conditions of normal distribution.

Descriptive Statistics and Intergroup Comparisons

Descriptive statistics, means, and SDs were used to determine if there were any significant differences among the data collected for each group in the three operating environments: actual machine operation, remote operation using only monitor images, and VR operating environment. For the analysis of normally distributed data, multiple comparisons using the Bonferroni test were employed in the primary allocation analysis of variance. For the analysis of non-normally distributed data, multiple comparisons using the Steel-Dwass method in the Kruskal-Wallis test were conducted. The results of the analysis are presented in Table 1.

The Kruskal-Wallis test does not require a normal distributed dataset [58]. Its null hypothesis is that there is “no difference between the three groups” at a significance level of .05. If P<.05, the null hypothesis is rejected, indicating that there is a statistically significant difference in the means of the different groups. The Bonferroni test has the same hypothesis as the Kruskal-Wallis test, but it relies on the assumptions of normality and homogeneity of the population [59].

The HRV time-domain parameter of HR RRI, and HRV frequency-domain parameters of LF, LF/HF, HR variability, complexity of MSE, workload %HRR, body vibration Aw, VDV, $MSDV_z$, and driving time of the construction equipment were statistically significant between the two operating environments. Statistically significant differences were also found between the three operating environments for the LF/HF HRV parameters in the frequency domain, Aw of physical vibration, $MSDV_z$, and driving time of the construction equipment.

The riding operation of the construction machine resulted in the highest stress indices, LF/HF HRV, physical vibration Aw, and $MSDV_z$, and the shortest driving time. The riding operation shortened the driving time, but it increased the operator's stress. In contrast, remote control using the monitor image showed the smallest LF/HF, Aw, and $MSDV_z$, and the longest driving time. Thus, in remote operation using monitor images, the stress of the operator was lower, but the driving time was longer.

Table 1. Mean, variance, and *P* value of each parameter for heart rate variability, physical workload, work vibration, and machine operation.

Parameter	Riding operation	Remote operation	VR[a] operation	Norm test	*P* value between each operation		
					1-2	2-3	1-3
HRV[b] time domain, mean (SD)							
RRI[c] (ms[d])	664.8 (92.6)	828.9 (140)	803.0 (152)	NP[e]	.008	.05	.85
SDRR[f] (ms)	66.5 (22.6)	70.7 (36.3)	60.0 (22.0)	NP	.96	.66	.72
RMSSD[g] (ms)	24.2 (7.96)	25.3 (10.1)	26.4 (8.02)	NP	.90	.57	.78
HRV frequency domain, mean (SD)							
LF^{h}_{nu}	77.2 (9.44)	63.6 (9.70)	73.0 (13.0)	P[i]	<.001	.005	.38
LF_{power}	31.1 (7.13)	29.9 (8.69)	32.7 (5.36)	NP	.87	.63	.98
LF/HF[j]	4.72 (1.52)	1.85 (0.78)	3.27 (0.99)	NP	<.001	.004	.04
MSE[k]	7.13 (1.26)	11.3 (2.78)	6.88 (2.10)	P	<.001	<.001	.99
Physical workload, mean (SD)							
%HRR[l]	13.0 (6.30)	3.70 (4.01)	5.17 (5.45)	P	<.001	.001	.99
Work vibration, mean (SD)							
Aw[m]	60.7 (18.1)	11.6 (12.8)	56.5 (21.4)	NP	<.001	<.001	.004
VDV[n]	5678 (4067)	1.80 (2.25)	828.2 (2377)	P	<.001	.99	<.001
$MSDV_z$[o]	133.9 (2.59)	126.5 (2.26)	131.3 (2.56)	NP	<.001	<.001	.04
Machine operation, mean (SD)							
Running time	306.3 (48.7)	436.2 (81.7)	369.8 (77.6)	NP	<.001	.04	.05

[a]VR: virtual reality.
[b]HRV: heart rate variability.
[c]RRI: RR interval.
[d]ms: milliseconds.
[e]NP: nonparametric.
[f]SDRR: SD of RRI.
[g]RMSSD: root mean square of the continuous difference.
[h]LF: low frequency.
[i]P: parametric.
[j]HF: high frequency.
[k]MSE: multiscale entropy.
[l]HRR: heart rate reserve.
[m]Aw: vibration index.
[n]VDV: vibration doses value.
[o]$MSDV_z$: motion sickness dose value.

Relationships Between Psychological and Working Loads, and Physical Vibration

Data collected in the three operating environments were combined to analyze their effects on the psychological load and workload. Multiple regression analysis was conducted to evaluate the significant relationships between the psychological load, workload, and physical vibration. The results are presented in Table 2. In the multiple regression analysis, we checked for multicollinearity in the independent variables (indicators of physical vibration). All three indices of physical vibration, (Aw, VDV, and $MSDV_z$) had a variance inflation factor (VIF) less than 10. Subsequently, a significant relationship was found between the operational load and body vibration.

As shown in Table 3, the large oscillation of the VDV determined by the quadrature oscillation dose method has a positive effect on the workload %HRR. Additionally, the time average of the root mean square weighted acceleration, Aw, had a positive effect on the stress index, LF/HF HRV, as observed in Table 4.

Aw had a negative effect on the MSE, a measure of adaptability inferred from the complexity of the heartbeats, as indicated in Table 5. The adjusted R^2 value for this regression equation was 0.189.

These results show that the %HRR has a significant relationship with the VDV. However, the other vibration indices, Aw and $MSDV_z$, did not show any significant relationship with the %HRR. LF/HF and MSE showed significant relationships with Aw. The effects of some vibration indices on the %HRR, which indicates the workload in the operating environment, LF/HF, which indicate the psychological load, and MSE, are shown.

Table 2. Relationship between the independent variable (work vibration) and dependent variables (workload, low frequency/high frequency, and multiscale entropy).

Independent variables	Dependent variables								
	%HRR[a]			LF[b]/HF[c]			MSE[d]		
	β[e]	SE	*P* value	β	SE	*P* value	β	SE	*P* value
Work vibration									
Aw[f]	–0.002	0.039	.96	0.0238	0.0093	.01	–0.031	0.0175	.37
VDV[g]	0.0006	0.0003	.06	0.000	0.0001	.82	–0.0001	0.0001	.24
$MSDV_z$[h]	0.515	0.294	.09	0.0983	0.0693	.16	–0.156	0.130	.08

[a]HRR: heart rate reserve.
[b]LF: low frequency.
[c]HF: high frequency.
[d]MSE: multiscale entropy.
[e]β: beta coefficient.
[f]Aw: vibration indices.
[g]VDV: vibration doses value.
[h]$MSDV_z$: motion sickness dose value.

Table 3. Relationships between independent variable (work vibration) and dependent variable (workload).

Model 1-1: independent variable	Dependent variable: %HRR[a]			
	Estimated	SE	*t* value[b]	*P* value
VDV[c]	0.0008	0.0003	3.29	.002
Intercept	5.469	1.07	5.11	<.001
Multiple R^{2d}	0.213	—[e]	—	—
Adjusted R^2	0.193	—	—	—
F static value[f]	10.8	—	—	.002

[a]HRR: heart rate reserve.
[b]*t* value: result of the student *t* test.
[c]VDV: vibration doses value.
[d]R^2: coefficient of determination.
[e]Not available
[f]*F* static value: variance ratio.

Table 4. Relationship between independent variables (work vibration) and dependent variables (low frequency/high frequency).

Model 1-2: independent variable	Dependent variable: LF[a]/HF[b]			
	Estimated	SE	*t* value[c]	*P* value
Aw[d]	0.0329	0.0079	4.18	<.001
(Intercept)	1.817	0.405	4.49	<.001
Multiple R^{2e}	0.305	—[f]	—	—
Adjusted R^2	0.287	—	—	—
F static value[g]	17.5	—	—	<.001

[a]LF: low frequency.

[b]HF: high frequency.

[c]*t* value: result of the student's *t* test.

[d]Aw: vibration indices.

[e]R^2: coefficient of determination.

[f]Not available

[g]*F* static value: variance ratio.

Table 5. Relationship between independent variables (work vibration) and dependent variables (multiscale entropy).

Model 1-3: independent variable	Dependent variable: MSE[a]			
	Estimated	SE	*t* value[b]	*P* value
Aw[c]	–0.0485	0.00149	–3.24	.02
Intercept	10.55	0.768	13.7	<.001
Multiple R^{2d}	0.208	—[e]	—	—
Adjusted R^2	0.189	—	—	—
F static value[f]	10.5	—	—	.002

[a]MSE: multiscale entropy.

[b]*t* value: result of the student *t* test.

[c]Aw: vibration indices.

[d]R^2: coefficient of determination.

[e]Not available.

[f]*F* static value: variance ratio.

Significant Statistic of Each Parameter for Driving Time

Data collected in the three operating environments were combined, and multiple regression analysis was performed to evaluate the relationships among several HRV indices and the parameters of MSE, HR RRI, and physical vibration, which indicated the complexity of HR changes, with the driving time of construction equipment. Two multiple regression equations were used to confirm a statistically significant relationship. First, in the multiple regression analysis, we found no multicollinearity among the independent variables. As a result, it was confirmed that the parameters among the two sets of dependent variables used, namely, LF/HF, RRI, $MSDV_z$, and MSE, and RRI and $MSDV_z$, had VIFs between 1 and 2, and there was no possibility of multicollinearity. In the subsequent analysis of the physical and psychological loads and physical vibrations during the operation, two significant relationships were found.

The first was the effect of the explanatory variables LF/HF, HR RRI, and oscillation $MSDV_z$ on the driving time; the adjusted R^2 of this regression equation was 0.449. Second, the driving time was affected by the explanatory variables, namely MSE (complexity of HR change), HR RRI, and vibration $MSDV_z$; the adjusted R^2 of this regression equation was 0.400. The results of the analysis are presented below in Tables 6 and 7.

Multiple regression analysis of the data for each construction machine operation suggested that the driving time affects LF/HF, which indicates operational stress, and MSE, which indicates adaptability; it also affects the RRI and vibration index $MSDV_z$. Equations 2 and 3 are the multiple regression equations obtained for the driving time HRV and driving time MSE. In both these regression equations, as the construction machine runs faster

and the driving time becomes shorter, the operator's stress increases, the adaptability to the task decreases, and the RRI and $MSDV_z$ also increase.

$$\text{Driving time}_{HRV} = -24.5 \times LF/HF - 0.350 \times RRI - 10.7 \times MSDV_z + 2115 \quad \textbf{(2)}$$

$$\text{Driving time}_{MSE} = 10.5 \times MSE - 0.259 \times RRI - 10.3 \times MSDV_z + 1823 \quad \textbf{(3)}$$

Table 6. Relationship among independent variables (low frequency/high frequency, RR interval, and work vibration) and dependent variable (driving time).

Independent variables	Dependent variable: driving time				
	Estimated	SE	*t* value[a]	*P* value	VIF[b]
LF[c]/HF[d]	–24.5	7.32	–3.34	.002	1.40
RRI[e]	–0.350	0.0814	–4.30	<.001	1.40
$MSDV_z$[f]	–10.7	3.07	–3.49	.001	1.44
(Intercept)	2115.6	420.0	5.04	<.001	__[g]
Multiple R^2[h]	0.490	—	—	—	—
Adjusted R^2	0.449	—	—	—	—
F static value[i]	12.1	—	—	<.001	—

[a]*t* value: result of the student *t* test.
[b]VIF: variance inflation factor.
[c]LF: low frequency.
[d]HF: high frequency.
[e]RRI: RR interval.
[f]$MSDV_z$: motion sickness dose value.
[g]Not available.
[h]R^2: coefficient of determination.
[i]*F* static value: variance ratio.

Table 7. Relationship among independent variables (multiscale entropy, RR interval, and work vibration) and dependent variable (driving time).

Independent variables	Dependent variable: driving time				
	Estimated	SE	*t* value[a]	*P* value	VIF[b]
MSE[c]	10.5	4.01	2.63	.01	1.27
RRI[d]	–0.259	0.0807	–3.21	.003	1.24
$MSDV_z$[e]	–10.3	3.31	–3.11	.004	1.53
Intercept	1823.3	476.6	3.83	<.001	__[f]
Multiple R^2[g]	0.444	—	—	—	—
Adjusted R^2	0.400	—	—	—	—
F static value[h]	10.1	—	—	<.001	—

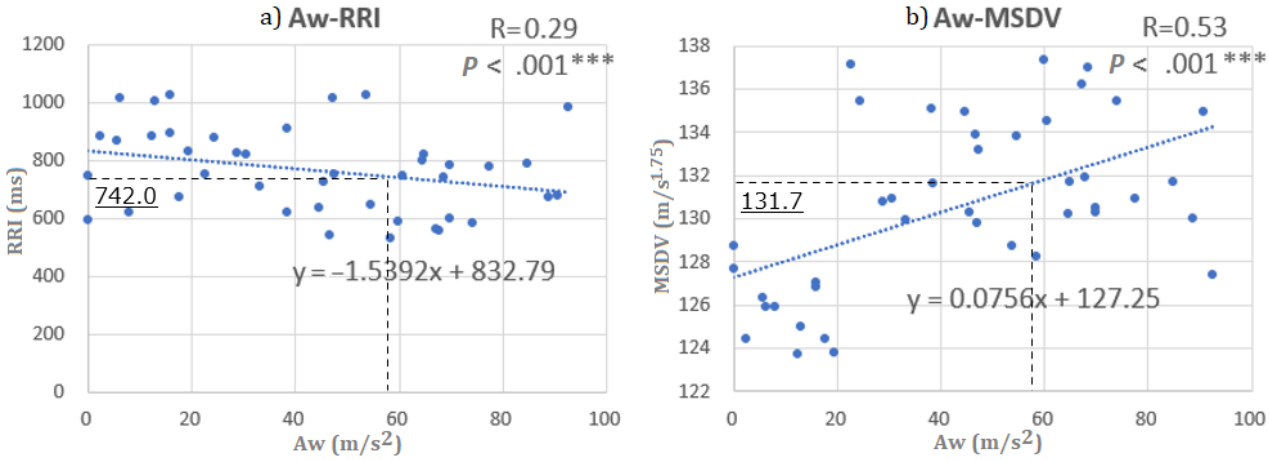
[a]*t* value: result of the student *t* test.
[b]VIF: variance inflation factor.
[c]MSE: multiscale entropy.
[d]RRI: RR interval.
[e]$MSDV_z$: motion sickness dose value.
[f]Not available.
[g]R^2: coefficient of determination.
[h]*F* static value: variance ratio.

Driving Construction Machines With Acceptable Operational Stress

Using equation 2 regarding the driving time of construction machinery obtained in this study, we suggest reducing the driving time of the construction machine while suppressing the operation stress. Equation 2 shows the relationship between the physical vibration and psychological load during machine operation, and it is expected that reducing the physical vibration during operation will reduce the operational stress. By reducing the stress caused by Aw, the time average of the squared vibration, to an acceptable level, it is possible to reduce the driving time and operator stress.

There are no previous reports where LF/HF for stress levels has been quantitatively determined. In this study, an LF/HF of value 2 was considered an acceptable stress level based on reports investigating the stress of participants in a sitting posture [60-62]. Using the relationship presented in Table 5, the average squared vibration acceleration during traveling $Aw_{LF/HF=2}$ is approximately 59 m/s^2 according to Aw = 2/0.0329 – 1.82. From the relationship between equation 2 and Figure 2, the traveling time of the construction machine = –24.5 × 2 – 0.35 × 742.0 –10.7 × 131.7 + 2115 = 397.1 seconds (approximately), which is the driving time for one lap that is acceptable for the psychological load of the operator. It was also estimated that the running speed at an acceptable psychological load = 400 m /397.1 s ≈ 1.01 m/s ≈ 3.63 km/h.

Figure 2. A. Correlation between RR interval and vibration index Aw and B. Correlation between motion sickness doses value and vibration index Aw. The dashed lines show the estimated values of the RR interval and motion sickness dose value for the allowed Aw. In the figure, the linear relationship equation, correlation coefficient R, and *P* values related to the two axes are shown. MSDV: motion sickness dose value; RRI: RR interval.



Discussion

Principal Findings

In this research, we investigated the effects of stress on psychological health during the operation of construction equipment, and the relationship between the appropriate stress and the driving time of construction machines. We found a significant relationship between LF/HF HRV [7,31] and MSE [6,40-42], which indicates the complexity of the HR, and body vibration [48-50]; thus, our hypothesis was supported. In addition, it was found that by keeping the operational stress caused by Aw, which is the time average of the squared vibration acceleration, at an acceptable level, an appropriate driving time that takes the operator into account can be obtained.

Stress and Remote Operation

Indices related to stress characterize the activity of the sympathetic part of the autonomous nervous system and can be appropriately applied to estimate psychological load as well as the intensity of physical workload [63]. Therefore, a similar relationship can be inferred between several HRVs related to stress and MSE. This is evident in the relationships shown in equations 2 and 3. The study results show that the uncertainty of what constitutes an acceptable limit of psychological load can be resolved by analyzing the stress index and some vibration indices in the working environment. The operating technicians were exposed to different stress levels owing to the physical vibrations of the construction equipment. In addition, the workload by %HRR was found to be related to physical vibration, but not to the driving time. Riding operators with the highest vibration exhibited multiple lower HRV indices and MSEs than operators exposed to stress when remotely controlling the construction machine with only monitor images, which had the lowest vibration. Therefore, the hypothesis that the physical vibration experienced by construction workers has a negative effect on the HRV and MSE was supported. In addition, the hypothesis that the driving time of the construction machine has a negative relationship with the magnitude of vibration in the operating environment, and the parameters of the HRV and MSE, was also confirmed.

In an environment with large vibrations, the operator has lower multiple HRV indices because of a higher sympathetic nervous system tone (LF power) and sympathetic balance (LF/HF), and a lower parasympathetic nervous system tone (ie, RRI, SDRR, RMSSD, and HF power) [33]. This result is consistent with previous studies estimating that increased work stress is associated with lower parasympathetic activation as sympathetic activity increases [64,65]. LF/HF provides insight into the stress classification of participants. Operators in this study were exposed to a low physical load and high mental stress. This is in comparison to physical workers who engage in production through physical activity, which may result in higher

psychological load due to the nature of their work, as it involves paying full attention to the safety of their surroundings and ensuring work quality through machine operation. The results of this study reflect the findings of Boschman et al [66]. Operators experience high psychological job demands and a high need for recovery. Hence, job-specific psychosocial work factors need to be assessed.

The measurement system and sensing wear used in this study are reliable [67] and provide valid HRV data. However, care should be taken when using them for implementing frequency-domain analysis to interpret cardiac autonomic modulation [68]. For accurate measurement of indices related to LFs and HFs in the frequency domain, continuous recording with a stable HR measurement period of at least 3 minutes is recommended [5]. This application can be socially implemented as a useful tool for monitoring the cardiac autonomic health status of operation workers. It is useful in managing the stress levels of operation technicians during machine operations by efficiently using short-term HRV and MSE beat information and body vibration recordings.

Theoretical and Practical Contributions

This study contributes theoretically by demonstrating the influence of psychological load as measured by the HRV and MSE on the operation of construction machines, and the effect of the psychological load of skilled workers on the HR interval, vibration in the working environment, and driving time of the construction machine. In addition, the study presents a new relational model using biometric information on HR and vibration indices in the work environment for the driving time of construction equipment.

Regarding the practical contributions, we quantified the vibration in the work environment of the driving operation and clarified the psychological workload of the operation. The evidence connecting the physical vibration in the work environment and the psychological and physical fatigue of workers could cause construction companies to improve their working environment and workforce management [69]. Furthermore, a new concept considering the psychology of the operator and the efficiency of the operation from the perspective of health psychology was introduced by comparing riding operation and remote operations, assuming a construction site where construction machines could not be operated.

Limitations

This study had several limitations. First, the number of construction machine operators employed was disproportionate; hence, the age and gender of the operators were not considered. There are reports that stress varies with age and gender [70,71], but this study was conducted on healthy males aged 18-50 years; hence, the study results may not be generalizable to all technicians in the construction industry. The number of male workers in the Japanese construction industry is very high, and further research might be beneficial in countries where there are a promising number of women in the construction and operations engineering professions. Second, this study was a cross-sectional analysis, and data were collected from the operators over 3 days. The collection of data over a longer period may provide more definitive results. Third, it would be desirable to analyze the operability and productivity of construction machines in relation to stress, as this research was limited to evaluation considering the driving time. The present study was conducted on construction machines that are typically used in construction work. A study of the psychological load during operation using machines with more fine-grained operational needs and controls could provide a comparison of the effects on the operator. Finally, as frequency-based metrics have been reported to represent the balance between sympathetic and parasympathetic activities more accurately [72], it is critical to improve the quality of HR interval recordings in wearable devices. Among the data collected in this research, there were some missing heartbeat intervals, which affected the selection of HRV metrics and necessitated the removal of a sample of participants from statistical analysis. To conduct a large sample study over a long period, future research aimed at furthering sensing wear and wearable technologies such as the WHS-2 to improve recording quality (eg, further minimizing motion artifacts) is essential. This will enhance the usefulness of the devices used.

Conclusions

A new method was developed in this study to calculate the appropriate operating time considering operational stress and maintaining the physical vibration within an acceptable range. The participants had to be alert while operating the machine in an environment that could expose them to high stress from vibration. Although this research is based on a limited number of participants in a special environment, by focusing on the relationship between psychological load and physical vibration, which remains unexplored in previous studies, the relationship of these variables with the operation time of construction machines was clarified.

Acknowledgments

This research was financially supported by the Ministry of Land, Infrastructure, Transport and Tourism (FY2019-FY2021 research and development for construction technology subsidy program policy issue solving type “Analytical evaluation system for improving productivity using lifelog information in unmanned construction”).

Conflicts of Interest

None declared.

Multimedia Appendix 1
List of devices and infrastructure used for the measurements of heart rate and acceleration.
[PDF File (Adobe PDF File), 134 KB - mhealth_v9i9e31637_app1.pdf]

Multimedia Appendix 2
Measurement data of the operators collected in the work environment.
[PDF File (Adobe PDF File), 229 KB - mhealth_v9i9e31637_app2.pdf]

Multimedia Appendix 3
Measurement indices for physical vibrations (root mean square weighted acceleration, vibration doses value , and motion sickness dose value) in operating work.
[PDF File (Adobe PDF File), 444 KB - mhealth_v9i9e31637_app3.pdf]

Multimedia Appendix 4
Results of normality tests for heart rate variability and vibration specifications (root mean square weighted acceleration, vibration dose value, and motion sickness dose value).
[PDF File (Adobe PDF File), 554 KB - mhealth_v9i9e31637_app4.pdf]

References

1. Chikushi S, Moriyama Y, Fujii H, Tamura Y, Yamakawa H, Nagatani K, et al. Automated image presentation for backhoe embankment construction in unmanned construction site. In: Proceedings of the 2020 IEEE/SICE International Symposium on System Integration. 2020 Jan Presented at: 2020 IEEE/SICE International Symposium on System Integration; Jan 12-15, 2020; Hawai p. 22-27. [doi: 10.1109/sii46433.2020.9026298]
2. Tripicchio P, Ruffaldi E, Gasparello P, Eguchi S, Kusuno J, Kitano K, et al. A Stereo-Panoramic Telepresence System for Construction Machines. Procedia Manufacturing 2017 Jan;11(1):1552-1559. [doi: 10.1016/j.promfg.2017.07.292]
3. Brandt M, Madeleine P, Samani A, Ajslev JZ, Jakobsen MD, Sundstrup E, et al. Effects of a participatory ergonomics intervention with wearable technical measurements of physical workload in the construction industry: cluster randomized controlled trial. J Med Internet Res 2018 Dec;20(12):e10272 [FREE Full text] [doi: 10.2196/10272] [Medline: 30567694]
4. Ernst G. Heart-rate variability-more than heart beats? Front Public Health 2017 Sep;5:240 [FREE Full text] [doi: 10.3389/fpubh.2017.00240] [Medline: 28955705]
5. Shaffer F, Ginsberg JP. An overview of heart rate variability metrics and norms. Front Public Health 2017 Sep;5:258 [FREE Full text] [doi: 10.3389/fpubh.2017.00258] [Medline: 29034226]
6. Tiwari A, Albuquerque I, Parent M, Gagnon J, Lafond D, Tremblay S, et al. Multi-scale heart beat entropy measures for mental workload assessment of ambulant users. Entropy (Basel) 2019 Aug;21(8):783 [FREE Full text] [doi: 10.3390/e21080783] [Medline: 33267496]
7. Castaldo R, Melillo P, Bracale U, Caserta M, Triassi M, Pecchia L. Acute mental stress assessment via short term HRV analysis in healthy adults: a systematic review with meta-analysis. Biomed Signal Process Control 2015 Apr;18:370-377. [doi: 10.1016/j.bspc.2015.02.012]
8. Wen W, Liu G, Mao Z, Huang W, Zhang X, Hu H, et al. Toward constructing a real-time social anxiety evaluation system: exploring effective heart rate features. IEEE Trans Affective Comput 2020 Jan;11(1):100-110. [doi: 10.1109/taffc.2018.2792000]
9. Patel M, Lal S, Kavanagh D, Rossiter P. Applying neural network analysis on heart rate variability data to assess driver fatigue. Expert Syst Appl 2011 Jun;38(6):7235-7242. [doi: 10.1016/j.eswa.2010.12.028]
10. Salahuddin L, Cho J, Jeong M, Kim D. Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings. In: Proceedings of the 9th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2007 Aug Presented at: 9th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; August 22-26, 2007; Lyon, France p. 4656-4659. [doi: 10.1109/iembs.2007.4353378]
11. Landreani F, Martin-Yebra A, Casellato C, Frigo C, Pavan E, Migeotte P, et al. Beat-to-beat heart rate detection by smartphone's accelerometers: validation with ECG. In: Proceedings of the 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2016 Aug Presented at: 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); August 16-20, 2016; Orlando, FL p. 525-528. [doi: 10.1109/EMBC.2016.7590755]
12. Hashiguchi N, Kodama K, Lim Y, Che C, Kuroishi S, Miyazaki Y, et al. Practical judgment of workload based on physical activity, work conditions, and worker's age in construction site. Sensors (Basel) 2020 Jul;20(13):3786 [FREE Full text] [doi: 10.3390/s20133786] [Medline: 32640611]
13. Castaldo R, Montesinos L, Melillo P, James C, Pecchia L. Ultra-short term HRV features as surrogates of short term HRV: a case study on mental stress detection in real life. BMC Med Inform Decis Mak 2019 Jan;19(1):12 [FREE Full text] [doi: 10.1186/s12911-019-0742-y] [Medline: 30654799]

14. Melo HM, Martins TC, Nascimento LM, Hoeller AA, Walz R, Takase E. Ultra - short heart rate variability recording reliability: the effect of controlled paced breathing. Ann Noninvasive Electrocardiol 2018 Jun;23(5):e12565. [doi: 10.1111/anec.12565]
15. Matsuura H, Mukaino M, Otaka Y, Kagaya H, Aoshima Y, Suzuki T, et al. Validity of simplified, calibration-less exercise intensity measurement using resting heart rate during sleep: a method-comparison study with respiratory gas analysis. BMC Sports Sci Med Rehabil 2019 Nov;11(1):27 [FREE Full text] [doi: 10.1186/s13102-019-0140-x] [Medline: 31700643]
16. Holmes TH, Rahe RH. The social readjustment rating scale. J Psychosom Res 1967 Aug;11(2):213-218. [doi: 10.1016/0022-3999(67)90010-4]
17. Lazarus R, Cohen J. Environmental stress. In: Human Behavior and Environment. Boston, MA: Springer; 1997.
18. Chelidoni O, Plans D, Ponzo S, Morelli D, Cropley M. Exploring the effects of a brief biofeedback breathing session delivered through the biobase app in facilitating employee stress recovery: randomized experimental study. JMIR Mhealth Uhealth 2020 Oct;8(10):e19412 [FREE Full text] [doi: 10.2196/19412] [Medline: 33055072]
19. Selye H. A syndrome produced by diverse nocuous agents. Nature 1936 Jul;138(3479):32-32. [doi: 10.1038/138032a0]
20. Selye H. Stress of Life 2nd edition. United States: McGraw-Hill Inc; 1978.
21. Ministry of Education, Culture, Sports, Science and Technology—Japan. URL: https://www.mext.go.jp/a_menu/shotou/clarinet/002/003/010/003.htm [accessed 2021-08-31]
22. Adair JG. The Hawthorne effect: a reconsideration of the methodological artifact. J Appl Psychol 1984;69(2):334-345. [doi: 10.1037/0021-9010.69.2.334]
23. Mccraty R, Shaffer F. Heart rate variability: new perspectives on physiological mechanisms, assessment of self-regulatory capacity, and health risk. Glob Adv Health Med 2015 Jan;4(1):46-61. [doi: 10.7453/gahmj.2014.073]
24. Graham SA, Jeste DV, Lee EE, Wu T, Tu X, Kim H, et al. Associations between heart rate variability measured with a wrist-worn sensor and older adults' physical function: observational study. JMIR Mhealth Uhealth 2019 Oct;7(10):e13757 [FREE Full text] [doi: 10.2196/13757] [Medline: 31647469]
25. Amirian I, Toftegård Andersen L, Rosenberg J, Gögenur I. Decreased heart rate variability in surgeons during night shifts. Can J Surg 2014 Oct;57(5):300-304 [FREE Full text] [doi: 10.1503/cjs.028813] [Medline: 25265102]
26. Rieger A, Stoll R, Kreuzfeld S, Behrens K, Weippert M. Heart rate and heart rate variability as indirect markers of surgeons' intraoperative stress. Int Arch Occup Environ Health 2014 Feb;87(2):165-174. [doi: 10.1007/s00420-013-0847-z] [Medline: 23370764]
27. Joseph B, Parvaneh S, Swartz T, Haider AA, Hassan A, Kulvatunyou N, et al. Stress among surgical attending physicians and trainees: a quantitative assessment during trauma activation and emergency surgeries. J Trauma Acute Care Surg 2016 Dec;81(4):723-728. [doi: 10.1097/TA.0000000000001162] [Medline: 27389128]
28. Billman GE. Heart rate variability - a historical perspective. Front Physiol 2011;2:86 [FREE Full text] [doi: 10.3389/fphys.2011.00086] [Medline: 22144961]
29. Reyes del Paso GA, Langewitz W, Mulder LJM, van Roon A, Duschek S. The utility of low frequency heart rate variability as an index of sympathetic cardiac tone: a review with emphasis on a reanalysis of previous studies. Psychophysiology 2013 May;50(5):477-487. [doi: 10.1111/psyp.12027] [Medline: 23445494]
30. Billman GE. The LF/HF ratio does not accurately measure cardiac sympatho-vagal balance. Front Physiol 2013;4:26 [FREE Full text] [doi: 10.3389/fphys.2013.00026] [Medline: 23431279]
31. Malliani A, Pagani M, Lombardi F, Cerutti S. Cardiovascular neural regulation explored in the frequency domain. Circulation 1991 Aug;84(2):482-492. [doi: 10.1161/01.CIR.84.2.482]
32. Peters GA, Wong ML, Joseph JW, Sanchez LD. Pulse rate variability in emergency physicians during shifts: pilot cross-sectional study. JMIR Mhealth Uhealth 2019 Oct;7(10):e13909 [FREE Full text] [doi: 10.2196/13909] [Medline: 31579017]
33. Nwaogu JM, Chan AP. Work-related stress, psychophysiological strain, and recovery among on-site construction personnel. Autom Constr 2021 May;125:103629. [doi: 10.1016/j.autcon.2021.103629]
34. Orsila R, Virtanen M, Luukkaala T, Tarvainen M, Karjalainen P, Viik J, et al. Perceived mental stress and reactions in heart rate variability--a pilot study among employees of an electronics company. Int J Occup Saf Ergon 2008;14(3):275-283. [doi: 10.1080/10803548.2008.11076767] [Medline: 18954537]
35. Troubat N, Fargeas-Gluck M, Tulppo M, Dugué B. The stress of chess players as a model to study the effects of psychological stimuli on physiological responses: an example of substrate oxidation and heart rate variability in man. Eur J Appl Physiol 2009 Feb;105(3):343-349. [doi: 10.1007/s00421-008-0908-2] [Medline: 18987876]
36. Costa M, Goldberger AL, Peng C. Multiscale entropy analysis of complex physiologic time series. Phys Rev Lett 2002 Jul;89(6):068102. [doi: 10.1103/physrevlett.89.068102]
37. Costa M, Goldberger AL, Peng C. Multiscale entropy analysis of biological signals. Phys Rev E Stat Nonlin Soft Matter Phys 2005 Feb;71(2 Pt 1):021906. [doi: 10.1103/PhysRevE.71.021906] [Medline: 15783351]
38. Blons E, Arsac L, Gilfriche P, Deschodt-Arsac V. Multiscale entropy of cardiac and postural control reflects a flexible adaptation to a cognitive task. Entropy 2019 Oct;21(10):1024. [doi: 10.3390/e21101024]
39. Busa MA, van Emmerik RE. Multiscale entropy: a tool for understanding the complexity of postural control. J Sport Health Sci 2016 Mar;5(1):44-51 [FREE Full text] [doi: 10.1016/j.jshs.2016.01.018] [Medline: 30356502]

40. Manor B, Costa MD, Hu K, Newton E, Starobinets O, Kang HG, et al. Physiological complexity and system adaptability: evidence from postural control dynamics of older adults. J Appl Physiol (1985) 2010 Dec;109(6):1786-1791 [FREE Full text] [doi: 10.1152/japplphysiol.00390.2010] [Medline: 20947715]
41. Wu S, Wu C, Lin S, Lee K, Peng C. Analysis of complex time series using refined composite multiscale entropy. Phys Lett A 2014 Apr;378(20):1369-1374. [doi: 10.1016/j.physleta.2014.03.034]
42. Pan J, Hu H, Liu X, Hu Y. Multiscale entropy analysis on human operating behavior. Entropy 2015 Dec;18(1):3. [doi: 10.3390/e18010003]
43. Wu S, Wu C, Lin S, Wang C, Lee K. Time series analysis using composite multiscale entropy. Entropy 2013 Mar;15(3):1069-1084. [doi: 10.3390/e15031069]
44. Costa M, Goldberger A. Generalized multiscale entropy analysis: application to quantifying the complex volatility of human heartbeat time series. Entropy 2015 Mar;17(3):1197-1203. [doi: 10.3390/e17031197]
45. Duarte MLM, de Araújo PA, Horta FC, Vecchio SD, de Carvalho LAP. Correlation between weighted acceleration, vibration dose value and exposure time on whole body vibration comfort levels evaluation. Safety Science 2018 Mar;103:218-224. [doi: 10.1016/j.ssci.2017.11.008]
46. ISO/WD 2631-1: Mechanical vibration and shock — Evaluation of human exposure to whole-body vibration — Part 1: General requirements. International Organization for Standardization. 1997. URL: https://www.iso.org/standard/76369.html [accessed 2020-12-01]
47. Subashi G, Nawayseh N, Matsumoto Y, Griffin M. Nonlinear subjective and dynamic responses of seated subjects exposed to horizontal whole-body vibration. J Sound Vib 2009 Mar;321(1-2):416-434. [doi: 10.1016/j.jsv.2008.09.041]
48. Hostens I, Papaioannou Y, Spaepen A, Ramon H. A study of vibration characteristics on a luxury wheelchair and a new prototype wheelchair. J Sound Vib 2003 Sep;266(3):443-452. [doi: 10.1016/s0022-460x(03)00577-7]
49. Pope M, Magnusson M, Wilder D. Lowback pain and whole-body vibration. Clin Orthop Relat Res 1998 Sep;354:241-248.
50. Els P. The applicability of ride comfort standards to off-road vehicles. J Terramech 2005 Jan;42(1):47-64. [doi: 10.1016/j.jterra.2004.08.001]
51. Hwang S, Lee S. Wristband-type wearable health devices to measure construction workers' physical demands. Autom Constr 2017 Nov;83:330-340. [doi: 10.1016/j.autcon.2017.06.003]
52. Norton K, Norton L, Sadgrove D. Position statement on physical activity and exercise intensity terminology. J Sci Med Sport 2010 Sep;13(5):496-502. [doi: 10.1016/j.jsams.2009.09.008] [Medline: 20005170]
53. Ismaila SO, Oriolowo KT, Akanbi OG. Cardiovascular strain of sawmill workers in South-Western Nigeria. Int J Occup Saf Ergon 2013 Jan;19(4):607-611. [doi: 10.1080/10803548.2013.11077016] [Medline: 24321639]
54. Kuusela T. Methodological aspects of heart rate variability analysis. In: Kamath MV, Watanabe MA, Upton ARM, editors. Heart Rate Variability (HRV) Signal Analysis. Boca Raton, FL: CRC Press; 2013:9.
55. Behar JA, Rosenberg AA, Weiser-Bitoun I, Shemla O, Alexandrovich A, Konyukhov E, et al. PhysioZoo: a novel open access platform for heart rate variability analysis of mammalian electrocardiographic data. Front Physiol 2018 Oct;9:1390 [FREE Full text] [doi: 10.3389/fphys.2018.01390] [Medline: 30337883]
56. Mishra P, Pandey C, Singh U, Gupta A, Sahu C, Keshri A. Descriptive statistics and normality tests for statistical data. Ann Card Anaesth 2019;22(1):67-72. [doi: 10.4103/aca.aca_157_18]
57. Darko A, Chan APC. Strategies to promote green building technologies adoption in developing countries: the case of Ghana. Build Environ 2018 Feb;130:74-84. [doi: 10.1016/j.buildenv.2017.12.022]
58. McKight P, Najab J. Kruskal-Wallis test. In: The Corsini Encyclopedia of Psychology. New Jersey: John Wiley & Sons; 2010.
59. Abdi H. The Bonferonni and Šidák corrections for multiple comparisons. Encyclopedia of Measurement and Statistics 2007;3:103-107. [doi: 10.4135/9781412952644.n297]
60. Skibniewski FW, Dziuda Ł, Baran PM, Krej MK, Guzowski S, Piotrowski MA, et al. Preliminary results of the LF/HF ratio as an indicator for estimating difficulty level of flight tasks. Aerosp Med Hum Perform 2015 Jun;86(6):518-523. [doi: 10.3357/amhp.4087.2015] [Medline: 26099122]
61. Seong H, Lee J, Shin T, Kim W, Yoon Y, Yoon Y. The analysis of mental stress using time-frequency distribution of heart rate variability signal. In: Proceedings of the 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2004 Presented at: The 26th Annual International Conference of the IEEE Engineering in Medicine and Biology Society; September 1-5, 2004; San Francisco, CA p. 283-285. [doi: 10.1109/iembs.2004.1403147]
62. Subahni A, Xia L, Malik A. Association of mental stress with video games. In: Proceedings of the 4th International Conference on Intelligent and Advanced Systems. 2012 Presented at: 4th International Conference on Intelligent and Advanced Systems; June 12-14, 2012; Kuala Lumpur p. 82-85. [doi: 10.1109/icias.2012.6306164]
63. Quendler E, Trieb K, Nimmerichter A. Validation of automated detection of physical and mental stress during work in a Hühnermobil 225. Ann Agric Environ Med 2017;24(2):329-331. [doi: 10.5604/12321966.1235183]
64. Garza JL, Cavallari JM, Eijckelhof BHW, Huysmans MA, Thamsuwan O, Johnson PW, et al. Office workers with high effort–reward imbalance and overcommitment have greater decreases in heart rate variability over a 2-h working period. Int Arch Occup Environ Health 2014 Sep;88(5):565-575. [doi: 10.1007/s00420-014-0983-0]

65. Järvelin-Pasanen S, Sinikallio S, Tarvainen MP. Heart rate variability and occupational stress-systematic review. Ind Health 2018 Nov;56(6):500-511 [FREE Full text] [doi: 10.2486/indhealth.2017-0190] [Medline: 29910218]
66. Boschman J, van der Molen H, Sluiter J, Frings-Dresen M. Psychosocial work environment and mental health among construction workers. Appl Ergon 2013 Sep;44(5):748-755. [doi: 10.1016/j.apergo.2013.01.004] [Medline: 23380530]
67. Discover TOYOBO's New Materials. URL: https://www.toyobo-global.com/discover/materials/cocomi/index.html [accessed 2021-08-31]
68. Chen Y, Lu W, Pagaduan JC, Kuo C. A novel smartphone app for the measurement of ultra-short-term and short-term heart rate variability: validity and reliability study. JMIR Mhealth Uhealth 2020 Jul;8(7):e18761 [FREE Full text] [doi: 10.2196/18761] [Medline: 32735219]
69. Hashiguchi N, Cao J, Lim Y, Kubota Y, Kitahara S, Ishida S, et al. The effects of psychological factors on perceptions of productivity in construction sites in Japan by worker age. Int J Environ Res Public Health 2020 May;17(10):3517 [FREE Full text] [doi: 10.3390/ijerph17103517] [Medline: 32443446]
70. Rudolph KD, Hammen C. Age and gender as determinants of stress exposure, generation, and reactions in youngsters: a transactional perspective. Child Dev 1999 May;70(3):660-677. [doi: 10.1111/1467-8624.00048] [Medline: 10368914]
71. Folkman S, Lazarus RS, Pimley S, Novacek J. Age differences in stress and coping processes. Psychology and Aging 1987;2(2):171-184. [doi: 10.1037/0882-7974.2.2.171]
72. Singh N, Moneghetti KJ, Christle JW, Hadley D, Plews D, Froelicher V. Heart rate variability: an old metric with new meaning in the era of using mhealth technologies for health and exercise training guidance. part one: physiology and methods. Arrhythm Electrophysiol Rev 2018 Aug;7(3):193-198 [FREE Full text] [doi: 10.15420/aer.2018.27.2] [Medline: 30416733]

Abbreviations

ECG: electrocardiogram
HF: high frequency
HR: heart rate
HRR: heart rate reserve
HRV: heart rate variability
ISO: International Standards Organization
LF: low frequency
MSDV: motion sickness dose value
MSE: multiscale entropy
RMSSD: root mean square of the continuous difference
RRI: RR interval
SDRR: standard deviation of the RR intervals
VDV: vibration doses value
VIF: variance inflation factor
VR: virtual reality

Edited by G Eysenbach; submitted 29.06.21; peer-reviewed by T Lefèvre, R Ciorap; comments to author 21.07.21; revised version received 29.07.21; accepted 31.07.21; published 15.09.21.

Please cite as:
Hashiguchi N, Cao J, Lim Y, Kuroishi S, Miyazaki Y, Kitahara S, Sengoku S, Matsubayashi K, Kodama K
Psychological Effects of Heart Rate and Physical Vibration on the Operation of Construction Machines: Experimental Study
JMIR Mhealth Uhealth 2021;9(9):e31637
URL: https://mhealth.jmir.org/2021/9/e31637
doi:10.2196/31637
PMID:34524105

Original Paper

Reactivity to UV Radiation Exposure Monitoring Using Personal Exposure Devices for Skin Cancer Prevention: Longitudinal Observational Study

Bridget G Parsons[1], MSPH; Elizabeth S Nagelhout[2], MPH, PhD; Ali P Wankier[1], BS; Nan Hu[3], PhD; Riley Lensink[1], BA; Angela Zhu[1], BS; Katy Nottingham[1], BS; Douglas Grossman[1], MD, PhD; Jakob D Jensen[4], PhD; Yelena P Wu[1,5], PhD

[1]Huntsman Cancer Institute, University of Utah, Salt Lake City, UT, United States

[2]Department of Public Health, Westminster College, Salt Lake City, UT, United States

[3]Department of Biostatistics, Robert Stempel College of Public Health & Social Work, Florida International University, Miami, FL, United States

[4]Department of Communication, University of Utah, Salt Lake City, UT, United States

[5]Department of Dermatology, University of Utah, Salt Lake City, UT, United States

Corresponding Author:
Yelena P Wu, PhD
Huntsman Cancer Institute
University of Utah
2000 Circle of Hope Drive
Salt Lake City, UT, 84112
United States
Phone: 1 801 213 5653
Email: yelena.wu@utah.edu

Abstract

Background: Emerging UV radiation (UVR) monitoring devices may present an opportunity to integrate such technology into skin cancer prevention interventions. However, little is known about the effects of using a wearable UVR monitor on adults' and children's sun protection–related behaviors and attitudes (eg, cancer worry and perceived risk). Understanding the potential role of reactivity and seasonal effects will help inform the use of objective monitors in the context of skin cancer prevention research, including intervention studies.

Objective: The aim of this study is to examine the potential reactivity associated with a wearable personal UVR monitor, specifically the effects associated with reported sun-protective behaviors and skin cancer–related attitudes, which are often the targets of skin cancer preventive interventions.

Methods: Child-parent dyads (n=97 dyads) were asked to wear a UVR monitoring device during waking hours for 2 weeks. Participants were asked to sync the device daily with a smartphone app that stored the UVR exposure data. Participants were blinded to their UVR exposure data during the 2-week period; thus, the smartphone app provided no feedback to the participants on their UVR exposure. Participants completed self-report questionnaires assessing sun-protective behaviors, sunburn, tanning, skin self-examination, skin cancer–related knowledge, perceived risk, cancer worry, response efficacy, and intentions to change behaviors over the 2-week period. Linear regressions were conducted to investigate changes in the outcomes over time and to account for the role of the season of study participation.

Results: Regression results revealed that there was a significant decrease over time for several sun protection outcomes in children, including time spent outdoors on weekends (P=.02) and weekdays (P=.008), sunscreen use (P=.03), reapplication (P<.001), and unintentional tanning (P<.001). There were no significant changes over time in children's and parents' UVR exposure, sunburn occurrence, or sun protection attitudes. Season of participation was associated with several outcomes, including lower sunscreen use (P<.001), reapplication (P<.001), sunburns (P=.01), intentions to change sun-protective behaviors (P=.02), and intentional (P=.008) and unintentional tanning (P=.01) for participants who participated in the fall versus the summer.

Conclusions: The findings from this study suggest that daily use of a UVR monitoring device over a 2-week period may result in changes in certain sun-protective behaviors. These results highlight the importance of identifying and addressing potential reactivity to UVR monitoring devices, especially in the context of skin cancer preventive intervention research. Ultimately,

objectively assessed UVR exposure could be integrated into the outcome assessment for future testing of skin cancer prevention interventions.

(JMIR Mhealth Uhealth 2021;9(9):e29694) doi:10.2196/29694

KEYWORDS

ultraviolet radiation exposure; wearable device; melanoma; melanoma prevention; mHealth; digital health; eHealth; UVR monitoring; mobile phone

Introduction

Background

Mobile or personal health monitors are frequently used by the general public and in the context of health behavior research [1]. For example, FitBit, one of the most popular wearable activity monitors, reported 2 million new users in 2019, bringing its total active users to 29.6 million [2]. In the past decade, >300 studies using FitBit as an outcome measure were registered with the National Institutes of Health [3]. Although physical activity monitors are frequently used in research settings, monitors related to environmental exposures and their associated health behaviors are in the early phases of development, testing, and dissemination [4-6]. In the context of skin cancer prevention, assessment of an individual's UV radiation (UVR) exposure is essential and is becoming more feasible given the availability of personal UVR monitors.

Melanoma is the deadliest type of skin cancer and the fifth most common form of cancer [7]. UVR exposure from the sun or artificial sources is the primary modifiable risk factor for melanoma [8]. Efforts to prevent melanoma focus on decreasing individuals' UVR exposure through sun-protective behaviors, such as wearing sunscreen and clothing that covers the skin, and avoiding peak UVR hours, which are typically assessed through self-report measures [9]. There has been a growing number of objective UVR assessment methods that can quantify personal UVR exposure and be useful for documenting the efficacy of melanoma preventive interventions [4,10]. However, a potential concern with using objective UVR exposure assessments is reactivity or the possibility that simply monitoring a behavior may in and of itself result in behavior change [11,12]. For example, it is possible that individuals who are aware that their UVR exposure is being monitored could choose to use sun-protection methods more consistently. For example, in the physical activity assessment literature, it has been found that after individuals use objective assessments of their exercise behaviors, changes can be seen in both their behaviors and exercise-related attitudes (eg, covering more distance and reporting great perceived exertion) [13]. In addition, it is not known to what extent personal UVR exposure based on objective assessments could differ depending on environmental factors such as the season. For instance, individuals may make different sun-protection choices based on seasonal differences in weather and temperature [14]. Studies have found that individuals who rely on ambient temperature as a method for determining the need for sun protection, particularly in the winter [15,16], are more likely to receive sunburns in cool weather [17] and tend to use clothing such as long sleeves and pants for warmth rather than sun protection [14].

Although prior studies of objective UVR exposure monitors have not yet examined potential reactivity effects, this phenomenon has been more fully explored in physical activity research [18]. In the literature, there are mixed findings regarding whether monitoring behavior alone is associated with behavior change. Most studies reporting on behavioral changes among adults and children related to monitoring used devices such as pedometers that provided feedback to the participants on physical activity outcomes (eg, number of steps taken daily) [19-22]. In the studies of monitors that did not provide feedback, most found that monitoring alone was not associated with behavior change [23-26].

Objective

In relation to melanoma prevention, studies have examined the effects of providing periodic feedback to adults, adolescents, and children on UVR exposure; information on the risk of sunburn; and advice about sun protection methods [27-29]. For example, the provision of UVR exposure feedback is associated with less UVR exposure on weekends and an increase in the use of some forms of sun protection [27-29]. Other studies have examined variability in UVR exposure and sun-protective behaviors among different populations, such as rural and urban children, and periods, such as during vacation [30-34]. However, it remains unknown to what extent UVR measurement itself, in the absence of providing feedback on UVR exposure, is associated with changes in sun-protective behaviors and attitudes (eg, cancer worry and perceived risk) in adults and children. Some sun-protection behaviors (eg, sunscreen use, hat use, and wearing pants) are not directly measured by the device. Our hypothesis, based on the literature on physical activity and limited UVR exposure reactivity, is that wearing a UVR monitoring device could lead individuals to become more aware of their UVR exposure and, as a result, change their behaviors and attitudes related to sun exposure. For example, in the physical activity literature, attitudes related to exercise such as perceived efficiency, intensity of effort, and fatigue have been found to increase with accelerometer use [13]. We hypothesize that sun-protection behaviors increase over the monitoring period and that UVR measured by the device would decrease as a result of increased UVR exposure awareness. Understanding the potential role of reactivity and seasonal effects will help inform the use of objective monitors in the context of skin cancer prevention research, including intervention studies. This pilot study was designed to validate the use of a UVR monitoring device in larger skin cancer prevention interventions designed for parents and children [5]. The goal of this study is to assess the potential reactivity associated with UVR exposure

monitoring among adults and children as it relates to sun-protective behaviors and skin cancer prevention–related attitudes, which are often the targets of skin cancer preventive interventions [35-37].

Methods

Participants

Potential participants were recruited through health and community events (eg, health fairs and farmers' markets) and invitation letters that were mailed to residents with a child aged between of 8 and 17 years. A web-based marketing resource was used to obtain the addresses of potential participants living in Utah [38]. To be eligible to participate in this study, adults who were residents of Utah, aged ≥18 years, had at least one child aged between 8 and 17 years, had a smartphone with Bluetooth and Wi-Fi capabilities, were willing to use a smartphone app that shared their UVR exposure information with the research team, did not have a pacemaker (because of potential interference from the UVR monitor), and were able to read and write in English were included. For children to be eligible to participate, they had to be aged between 8 and 17 years, live with a primary caregiver in Utah, have no previous melanoma diagnosis, and not have a pacemaker. In total, 224 adults expressed interest in participating in the study. Of the 224 adults, 150 (67.0%) were screened for eligibility. Of the 150 adults screened, 116 (77.3%) were screened as eligible and 34 (22.7%) as ineligible. Reasons for ineligibility included not having a child in the desired age range (28/34, 82%), not having a smartphone with the necessary specifications (5/34, 15%), and not being able to read or write in English (1/34, 3%). A total of 97 parents (77% biological mothers) and 97 children (mean age 12.7 years, SD 2.7) were enrolled between June 2018 and October 2018. These months were selected because of the high ambient UVR levels in the region during those months [39]. All study procedures were approved by the relevant institutional review board.

Procedures

Participants were asked to wear a UVR monitoring device (Shade wearable UVR sensor, model V1.00, YouV Labs Inc) [4,40] on their clothing for a 2-week period. They completed the baseline assessment before being given the UVR monitoring devices to wear. A monitoring period of 2 weeks was selected to allow the capture of data from both weekdays and weekends for >1 week (to minimize missing data for either weekdays or weekend days). Participants were asked to synchronize the Shade device daily with a smartphone app that stored the UVR exposure data in a cloud-based server accessed by the research team. Participants were blinded to their UVR exposure data during the 2-week period; thus, the smartphone app provided no feedback to the participants on their UVR exposure. Changes in UVR exposure were analyzed by comparing total UVR exposure in the first week of study participation with total UVR exposure in the second week of study participation. Participants were asked to complete electronically delivered questionnaires via REDCap (Research Electronic Data Capture) [41] at the beginning and end of the 2-week period about sun-protective behaviors—such as wearing sunscreen and protective clothing (10 items) [42], number of sunburns received (one item) [42], tanning (two items) [43], skin self-examination (one item) [44], skin cancer–related knowledge (10 items) [45]—and relevant attitudes such as cancer worry (four items) [46], perceived risk (two items) [47], efficacy of skin cancer preventive behaviors (four items) [48], and intentions to change preventive behaviors (five items) [49-51]. Parents reported on these constructs for both themselves and their child, and children provided self-reports. All items except knowledge (true or false items) were asked on a 5-point Likert scale. The multi-item scales, including skin cancer–related knowledge, cancer worry (α=.87, child report; α=.90, parent report; parents not asked to report on child), perceived risk (α=.75, child report; α=.82, parent report; α=.84, parent report on child), efficacy of preventive behaviors (α=.89, child report; α=.94, parent report; α=.95, parent report on child), and intention to change preventive behaviors (α=.39, child report; α=.55, parent report; α=.66, parent report on child), were summed. After completing their study participation, each participant was provided with a gift card and a report summarizing their UVR exposure over the 2-week period.

Analyses

Descriptive statistics were used to summarize the participants' demographic information and summary statistics for the outcomes of interest. Linear regressions were used to examine potential changes over the 2-week period in UVR exposure measured from the device, sun-protective behaviors, sunburn, tanning, skin self-examination, and in summed scales for skin cancer–related knowledge, perceived risk, cancer worry, response efficacy, and intentions to change behaviors. Additional linear regressions were conducted with seasonal (summer [June-August] vs fall [September-October]) assessment time points (baseline vs exit) and their interactions as predictors and the same outcomes as dependent variables [52]. Season was included in these models based on the existing literature [53], and it was significantly related to multiple outcomes in this study.

Results

Sample Characteristics

In total, 97 parent-child dyads participated in the study. Of the 97 parents, 73 (77%) were biological mothers, 83 (87%) were non-Hispanic White, and 5 (5%) were Hispanic. Of the 97 children (mean age 12.7 years, SD 2.7), 81 (85%) were non-Hispanic White, 8 (8%) were Hispanic, and 56 (59%) were female (Table 1).

Table 1. Demographic characteristics of participating parents and children (N=97)[a].

Characteristics	Parents	Children
Age (years), mean (SD)	41.6 (6.3)	12.7 (2.7)
Gender (female), n (%)	73 (77)	56 (59)
Race, n (%)		
Non-Hispanic White	83 (87)	81 (85)
Hispanic	5 (5)	8 (8)
Asian or Asian American	5 (5)	4 (4)
Other	2 (2)	2 (2)
Fitzpatrick skin type (I-VI)		
Type I	9 (9)	2 (2)
Type II	15 (15)	17 (17)
Type III	42 (43)	41 (42)
Type IV	24 (25)	32 (33)
Type V	5 (5)	3 (3)
Marital status, n (%)		
Married or marriage-like relationship	84 (88)	N/A[b]
Divorced/separated	9 (10)	N/A
Widowed	2 (2)	N/A
Level of education, n (%)		
High school graduate or GED[c]	8 (8)	N/A
Some college, including 2-year degree	41 (42)	N/A
Bachelor's degree	25 (26)	N/A
Master's/doctoral degree	21 (22)	N/A
Household income (US $), n (%)		
<50,000	23 (24)	N/A
>50,000	64 (67)	N/A
Would prefer not to report	8 (8)	N/A

[a]Two families did not complete the baseline questionnaire, thus each of the categories have a total of 95.

[b]N/A: not applicable (children were not asked this question).

[c]GED: General Education Development.

Descriptive Statistics

Descriptive statistics (means and SDs) for UVR exposure, sun-protective and risk behaviors, and skin cancer–related attitudes are presented in Table 2.

Table 2. Descriptive statistics for UV radiation exposure and sun-protective behaviors, knowledge, and attitudes.

	Child self-report		Parent report on child		Parent self-report	
	Baseline, mean (SD)	Exit, mean (SD)	Baseline, mean (SD)	Exit, mean (SD)	Baseline, mean (SD)	Exit, mean (SD)
UVR[a] exposure[b]	6.00 (6.09)	4.76 (5.27)	N/A[c]	N/A	6.89 (6.38)	6.11 (6.62)
UVR exposure (adjusted)[d]	6.47 (5.94)	5.50 (5.07)	N/A	N/A	7.56 (6.58)	6.87 (6.81)
Hours outdoors: weekday	2.69 (1.75)	2.24 (1.65)	2.72 (1.47)	2.09 (1.43)	2.79 (1.80)	2.01 (1.68)
Hours outdoors: weekend	3.22 (2.09)	2.49 (1.57)	3.55 (1.97)	2.76 (1.68)	3.88 (1.91)	2.72 (1.66)
Sunscreen[e]	2.35 (1.45)	1.83 (1.13)	2.52 (1.39)	2.02 (1.21)	2.53 (1.45)	2.11 (1.33)
Reapplication[e]	1.95 (1.15)	1.50 (0.84)	1.98 (1.19)	1.60 (0.86)	1.87 (1.19)	1.60 (0.96)
Long sleeves[e]	1.76 (0.93)	2.12 (1.03)	1.86 (1.00)	2.16 (1.04)	1.87 (1.06)	2.20 (1.15)
Pants[e]	2.76 (1.21)	3.03 (1.36)	2.52 (1.19)	2.82 (1.32)	3.02 (1.34)	3.19 (1.37)
Hat[e]	1.31 (0.82)	1.37 (0.91)	1.37 (0.73)	1.37 (0.90)	1.97 (1.33)	1.84 (1.07)
Shade[e]	2.56 (1.04)	2.51 (1.06)	2.45 (0.96)	2.43 (1.01)	2.67 (1.00)	2.55 (1.02)
Avoid peak hours[e]	2.08 (0.94)	2.14 (1.02)	2.33 (1.19)	2.34 (1.04)	2.55 (1.14)	2.35 (1.00)
Sunglasses[e]	1.92 (1.18)	1.59 (1.01)	1.89 (1.18)	1.60 (0.99)	3.58 (1.40)	3.19 (1.34)
Sunburn	0.23 (0.57)	0.16 (0.43)	0.15 (0.36)	0.16 (0.42)	0.10 (0.30)	0.18 (0.44)
Intentional tanning[e]	1.24 (0.63)	1.15 (0.47)	1.16 (0.45)	1.09 (0.33)	1.13 (0.48)	1.08 (0.31)
Unintentional tanning[e]	2.88 (1.17)	1.80 (1.07)	2.95 (1.33)	1.80 (1.02)	2.93 (1.16)	1.90 (1.05)
Indoor tanning[e]	1.00 (0)	1.00 (0)	1.00 (0)	1.00 (0)	1.00 (0)	1.00 (0)
Skin self-examination[e]	1.88 (0.32)	1.88 (0.32)	0.09 (0.29)	0.17 (0.38)	0.16 (0.37)	0.12 (0.33)
Skin cancer knowledge[f]	2.98 (1.42)	3.23 (1.49)	N/A	N/A	4.08 (1.01)	4.24 (0.95)
UVR knowledge[f]	2.97 (1.27)	3.31 (1.13)	N/A	N/A	3.84 (0.67)	3.88 (0.65)
Total knowledge[f]	5.94 (2.35)	6.55 (2.25)	N/A	N/A	7.92 (1.29)	8.11 (1.22)
Perceived risk[g]	4.52 (1.68)	4.48 (1.79)	5.67 (1.75)	5.83 (1.60)	6.33 (1.88)	6.19 (1.83)
Cancer worry[h]	7.36 (3.34)	7.10 (3.20)	N/A	N/A	9.52 (3.73)	9.12 (3.52)
Response efficacy[i]	15.29 (3.44)	15.40 (3.81)	16.53 (3.84)	17.3 (2.74)	17.00 (3.57)	17.37 (2.89)
Intentions to change[j]	10.66 (3.93)	10.92 (4.53)	11.35 (4.73)	11.74 (4.89)	13.66 (5.04)	13.38 (4.96)

[a]UVR: UV radiation.

[b]First 7 days versus last 7 days of the 2-week wearing period.

[c]N/A: not applicable (parents not asked to report this on children).

[d]First 7 days versus last 7 days of the 2-week wearing period adjusted to include participants that had at least 4 days of UVR data.

[e]Response options included never=1 to always=5.

[f]0=false, 1=true. Possible range of 0-5, or 0-10 for total knowledge.

[g]1=very unlikely, 5=very likely. Possible range of 1-10.

[h]1=not at all, 5=very much. Possible range of 1-20.

[i]1=strongly disagree, 5=strongly agree. Possible range of 1-20.

[j]1=no and I do not intend to start doing so in the next 6 months, 5=yes, I have been for more than 6 months. Possible range of 1-25.

Regression Analyses

Regression results revealed that there was a significant change over time for several outcomes (Tables 3-5). In addition, separate analyses examining the relationship between season alone and these outcomes are provided in the table footnotes.

Table 3. Child self-report regression models examining association between time and season of participation and sun-protective behaviors[a].

Outcome	R^2	*F* test (*df*)	Time					Season				
			ß	SE (ß)	B	*t* test	*P* value	ß	SE (ß)	B	*t* test	*P* value
UVR[b] exposure[c]	0.00	0.17 (181)	0	1.17	0.00	0	.99	–.05	1.29	–0.64	–0.49	.62
UVR exposure adjusted[d]	0.01	0.38 (137)	–.06	1.26	–0.70	–0.56	.58	–.04	1.30	–0.42	–0.33	.75
Hours outdoors: weekday	0.02	1.36 (181)	–.17	0.33	–0.59	–1.77	.08	–.11	0.36	–0.39	–1.06	.29
Hours outdoors: weekend	0.04	2.59 (181)	–.16	0.37	–0.61	–1.67	.09	.01	0.39	0.02	0.04	.96
Sunscreen	0.29	23.62 (181)	–.22	0.22	–0.59	–2.69	.008	–.56	0.24	–1.52	–6.39	<.001
Reapplication	0.18	13.10 (181)	–.28	0.18	–0.57	–3.12	.002	–.43	0.20	–0.92	–4.60	<.001
Long sleeves	0.06	3.83 (181)	.07	0.19	0.14	0.76	.45	–.01	0.20	–0.01	–0.06	.95
Pants	0.05	3.29 (181)	.08	0.24	0.19	0.81	.42	.18	0.26	0.45	1.74	.08
Hat	0.02	1.48 (181)	.01	0.17	0.02	0.11	.92	–.15	0.18	-0.27	-1.46	.15
Shade	0.05	3.22 (181)	.08	0.20	0.17	0.88	.34	–.11	0.22	-0.23	-1.09	.28
Avoid peak hours	0.07	4.63 (181)	.17	0.18	0.34	1.89	.06	–.02	0.20	–0.03	–0.16	.87
Sunglasses	0.06	3.64 (181)	–.14	0.21	–0.32	–1.52	.13	–.14	0.23	–0.32	–1.41	.16
Sunburn	0.09	5.64 (181)	–.12	0.09	–0.12	–1.29	.20	–.36	0.10	–0.36	-3.65	<.001
Intentional tanning	0.04	2.43 (181)	–.09	0.11	–0.10	–0.96	.34	–.18	0.11	–0.20	–1.75	.08
Unintentional tanning	0.18	13.42 (181)	–.43	0.22	–1.06	–4.88	<.001	–.12	0.24	–0.27	–1.15	.25
Indoor tanning	__[e]	—	—	—	—	—	—	—	—	—	—	—
Skin self-exam	0.01	0.70 (181)	.05	0.06	0.03	0.50	.62	.13	0.07	0.09	1.26	.21
Skin cancer knowledge	0.01	0.68 (181)	.03	0.28	0.09	0.33	.75	.01	0.31	0.03	0.09	.93
UVR knowledge	0.02	1.21 (181)	.12	0.24	0.30	1.25	.39	.09	0.26	0.22	0.85	.39
Total knowledge	0.02	1.13 (181)	.08	0.45	0.39	0.87	.39	.05	0.49	0.25	0.51	.61
Perceived risk	0.01	0.07 (181)	–.03	0.34	–0.03	–0.11	.92	.01	0.37	0.01	0.01	.99
Cancer worry	0.01	0.79 (181)	–.08	0.63	–0.52	–0.82	.41	–.11	0.69	–0.76	–1.09	.27
Response efficacy	0.01	0.35 (181)	.01	0.69	0.01	0.01	.99	–.07	0.75	–0.54	–0.72	.48
Intentions to change	0.09	5.50 (181)	.06	0.78	0.46	0.59	.56	–.22	0.85	–1.89	–2.23	.03

[a]Models examining the relationship between season alone and the outcomes of interest indicated that shade seeking, avoidance of peak hours, wearing sunglasses, and intentional tanning were lower in the fall versus summer, whereas pants and hat wearing were higher in the fall.

[b]UVR: UV radiation.

[c]First 7 days versus last 7 days of the 2-week wearing period.

[d]First 7 days versus last 7 days of the 2-week wearing period adjusted to include participants who had at least 4 days of UVR data.

[e]No indoor tanning was reported.

Table 4. Parent report on child regression models examining the association between time and season of participation and sun-protective behaviors[a].

Outcome	R^2	*F* test (df)	Time					Season				
			ß	SE (ß)	B	*t* test	*P* value	ß	SE (ß)	B	*t* test	*P* value
Hours outdoors: weekday	0.08	5.16 (183)	–.22	0.28	–0.68	–2.42	.02	–.16	0.31	–0.48	1.56	.12
Hours outdoors: weekend	0.07	4.30 (183)	–.25	0.35	–0.94	–2.68	.008	–.14	0.38	–0.55	–1.44	.15
Sunscreen	0.23	18.26 (183)	–.18	0.23	–0.51	2.17	.03	–.41	0.26	–1.16	–4.53	<.001
Reapplication	0.20	14.86 (183)	–.29	0.21	–0.69	–3.31	<.001	–.47	0.23	–1.15	–5.02	<.001
Long sleeves	0.04	2.29 (183)	–.02	0.21	0.03	0.16	.87	–.10	0.23	–0.24	–1.02	.31
Pants	0.04	2.47 (183)	–.01	0.24	–0.03	–0.11	.91	–.05	0.28	–0.13	–0.49	.62
Hat	0.01	0.23 (183)	–.02	0.17	0.04	–0.24	.81	.03	0.19	–0.06	–0.33	.74
Shade	0.08	5.19 (183)	.04	0.20	0.09	0.45	.66	–.19	0.22	–0.41	–1.86	.07
Avoid peak hours	0.03	1.97 (183)	.01	0.24	0.03	0.11	.91	–.12	0.26	–0.31	1.17	.25
Sunglasses	0.02	1.33 (183)	–.14	0.24	–0.33	–1.42	.16	–.07	0.26	–0.18	–0.68	.50
Sunburn	0.06	3.57 (183)	–.01	–0.07	–0.01	–0.02	.98	–.26	0.08	–0.20	–2.53	.01
Intentional tanning	0.05	3.35 (183)	–.17	0.09	–0.18	–1.86	.07	–.27	0.11	–0.29	–2.67	.008
Unintentional tanning	0.24	18.58 (182)	–.48	0.23	–1.30	–5.72	<.001	–.24	0.25	0.66	–2.62	.01
Indoor tanning	__[b]	—	—	—	—	—	—	—	—	—	—	—
Skin self-exam	0.01	0.75 (183)	–.12	0.07	–0.08	–1.27	.21	–.03	0.07	–0.02	–0.25	.81
Perceived risk	0.03	2.04 (183)	.10	0.32	0.33	1.04	.30	.23	0.35	0.76	2.20	.03
Response efficacy	0.02	0.94 (183)	.16	0.64	1.05	1.65	.10	.07	0.70	0.48	0.69	.49
Intentions to change	0.07	4.42 (183)	.05	0.89	0.43	0.48	.63	–.24	0.97	–2.35	–2.42	.02

[a]Models examining the relationship between season alone and the outcomes of interest indicated that shade seeking, avoidance of peak hours, and hours spent outside on weekdays were lower in the fall versus summer, whereas pants wearing was higher in the fall.

[b]No indoor tanning was reported.

Table 5. Parent self-report regression models examining the association between time and season of participation and sun-protective behaviors[a].

Outcome	R^2	*F* test (*df*)	Time					Season				
			ß	SE (ß)	B	*t* test	*P* value	ß	SE (ß)	B	*t* test	*P* value
UVR[b] exposure[c]	0.02	0.92 (179)	–.08	1.24	–1.05	–0.85	.40	–.14	1.38	–1.85	–1.34	.18
UVR exposure adjusted[d]	0.01	0.64 (162)	–.08	1.30	–1.05	–0.81	.42	–.14	1.45	–1.87	–1.29	.20
Hours outdoors: weekday	0.01	0.32 (177)	.01	0.33	0.01	0.01	.99	–.07	0.36	–0.25	–0.70	.49
Hours outdoors: weekend	0.02	0.98 (183)	.04	0.33	0.14	0.42	.67	–.06	0.36	–0.19	–0.54	.59
Sunscreen	0.16	11.44 (183)	–.16	0.25	–0.44	–1.79	.08	–.40	0.27	–1.13	–4.21	<.001
Reapplication	0.14	9.91 (183)	–.14	0.19	–0.31	–1.55	.12	–.36	0.22	–0.82	–3.78	<.001
Long sleeves	0.04	2.28 (183)	.04	0.21	0.09	0.43	.67	–.09	0.23	–0.20	–0.87	.39
Pants	0.06	3.32 (183)	–.02	0.25	–0.05	–0.18	.86	.10	0.28	0.26	0.95	.35
Hat	0.01	0.41 (183)	–.03	0.23	–0.08	–0.33	.74	–.05	0.25	–0.13	–0.52	.60
Shade	0.08	5.43 (183)	.01	0.18	0.02	0.13	.89	–.16	0.20	–0.33	–1.61	.11
Avoid peak hours	0.05	2.99 (183)	.04	0.20	0.08	0.39	.70	–.01	0.22	–0.03	–0.12	.91
Sunglasses	0.02	1.37 (183)	–.16	0.26	–0.44	–1.68	.09	–.03	0.29	–0.08	–0.27	.79
Sunburn	0.03	2.03 (183)	.16	0.07	0.12	1.71	.09	–.06	0.08	–0.04	–0.56	.58
Intentional tanning	0.01	0.58 (183)	–.08	0.08	–0.06	–0.83	.41	–.09	0.08	–0.08	–0.94	.35
Unintentional tanning	0.22	16.52 (183)	–.37	0.20	–0.88	–4.33	<.001	–.14	0.22	–0.35	–1.57	.12
Indoor tanning	__[e]	—	—	—	—	—	—	—	—	—	—	—
Skin self-exam	0.01	0.88 (183)	.04	0.07	0.03	0.45	.65	.03	0.07	0.02	0.24	.81
Skin cancer knowledge	0.01	0.26 (183)	.04	0.19	0.09	0.46	.65	–.03	0.21	–0.06	–0.29	.77
UVR knowledge	0.01	0.09 (183)	.01	0.13	0.01	0.08	.94	–.05	0.14	–0.06	–0.44	.66
Total knowledge	0.01	0.27 (183)	.04	0.24	0.09	0.40	.69	–.05	0.26	–0.12	–0.47	.64
Perceived risk	0.01	0.84 (183)	–.03	0.35	–0.12	–0.34	.73	.12	0.39	0.45	1.15	.25
Cancer worry	0.02	0.91 (183)	–.05	0.70	–0.36	–0.52	.60	.12	0.77	0.87	1.13	.26
Response efficacy	0.01	0.33 (183)	.06	0.62	0.41	0.66	.51	–.03	0.68	–0.20	–0.30	.77
Intentions to change	0.02	2.01 (183)	.03	0.95	0.26	0.27	.79	–.07	1.05	–0.68	–0.65	.52

[a]Models examining the relationship between season alone and the outcomes of interest indicated that shade seeking, unintentional, and intentional tanning were lower in the fall versus summer.

[b]UVR: UV radiation.

[c]First 7 days versus last 7 days of the two-week wearing period.

[d]First 7 days versus last 7 days of the two-week wearing period adjusted to include participants that had at least 4 days of UVR data.

[e]No indoor tanning was reported.

There was a significant decrease in children's sunscreen use based on child ($F_{3,178}$=23.62; *P*<.001; R^2=0.29) and parent report ($F_{3,180}$=18.21; *P*<.001; R^2=0.23). When season was held constant, sunscreen use decreased in children over the 2-week study period based on child (ß=–.22; t_3=–2.69; *P*=.008) and parent report (ß=–.18; t_3=–2.17; *P*=.03). There were also decreases in reapplication of sunscreen in children based on child (ß=–.28; t_3=–3.12; *P*=.002) and parent report (ß=–.69; t_3=–3.31; *P*<.001). There was a significant decrease in reported unintentional tanning for children based on child report (ß=–.43; t_3=–4.88; *P*<.001) and parent report (ß=–.48; t_3=–5.72; *P*<.001) and parents (ß=–.37; t_3=–4.33; *P*<.001). In addition, there were decreases in the hours children spent outside on weekdays (ß=–.22; t_3=–2.42; *P*=.02) and weekends (ß=–.25; t_3=–2.68; *P*=.008) based on parent report. There were no significant changes over time in UVR exposure, sunburn, or attitudes.

Season of participation was associated with several outcomes. Reported sunscreen use was lower for children based on child (ß=–.56; t_3=–6.39; *P*<.001) and parent report (ß=–.41; t_3=–4.53; *P*<.001) and parents (ß=–.40; t_3=–4.21; *P*<.001) who participated in fall versus summer. Reapplication of sunscreen was also lower for children based on child (ß=–.43; t_3=–4.60; *P*<.001) and parent report (ß=–.47; t_3=–5.02; *P*<.001) and parents (ß=–.36; t_3=–3.78; *P*<.001) who participated in fall. Reported child sunburns per child (ß=–.36; t_3=–3.65; *P*<.001) and parent report (ß=–.26; t_3=–2.53; *P*=.01) and intentions to

change were lower for children who participated in the fall (ß=–.22; t_3=–2.23; *P*=.03) than in the summer (ß=–.24; t_3=–2.42; *P*=.02). Intentional tanning (ß=–.27; t_3=–2.67; *P*=.008) and unintentional tanning (ß=–.24; t_3=–2.62; *P*=.01) were lower for children who participated in the fall based on parent report. Perceived risk for cancer was higher for children who participated in the fall compared with those who participated in summer based on parent report (ß=.23; t_3=2.20; *P*=.03).

Discussion

Principal Findings

The current findings indicate that it may be important to identify and address the reactivity to UVR monitoring devices among children and parents. Our results provide initial evidence that the use of a UVR monitoring device may be associated with changes in sun-protective behaviors. In the context of intervention studies seeking to improve the use of sun-protection behaviors, the reactivity effects observed, particularly the decreased use of sun-protection behaviors, could potentially dampen the detection of desired intervention effects.

In contrast to most previous studies that found that monitoring alone did not lead to behavior change, some significant changes in skin cancer preventive behaviors over time were detected, including in children's time spent outside on weekdays and weekends, use and reapplication of sunscreen, and unintentional tanning. Notably, almost all of the changes in outcomes observed were decreases in reported sun-protective behaviors over time. This could indicate that wearing the UVR monitoring device initially led participants to increase their use of sun-protective behaviors, and then over time, the effect wore off as participants acclimatized to wearing the device. This has been observed in the physical activity literature, for example, that the first measurement day was the most active day for participants, and this tapered off in subsequent days [20]. Future studies could further explore the role of reactivity associated with UVR monitoring devices by assessing participants over a longer wearing period and incorporating an initial familiarization phase to allow for any reactivity effects to subside. In addition, future studies could consider systematically excluding the first few days of UVR monitoring data during analysis.

We observed seasonal differences in sun-protection behaviors based on both child reports and parent reports, which is consistent with the findings of prior studies that individuals may make different sun-protection choices based on weather or time of year [14-17]. Our findings confirm the importance of controlling for seasonal effects, either through statistical methods or study design. For example, studies could take place during the course of a single season (eg, summer).

This study has several strengths and limitations. Strengths include that our study is one of the first to examine the reactivity to UVR exposure monitoring using a wearable UVR monitoring device. In addition, parents and children were included in this study, and they could both benefit from skin cancer interventions that incorporate wearable UVR monitoring devices. Limitations include a study sample from a single location, which may limit generalizability to populations in other geographic areas. Future studies could include a control group that does not wear a sensor to disentangle the potential contributions of the self-report method of assessment of sun-protection behaviors and other outcomes. Future studies could further assess potential reactivity associated with the use of wearable UVR monitoring devices that do provide feedback on exposure and account for other factors that may impact potential reactivity, such as age, amount of time spent outdoors, and geographic location.

Conclusions

This study is among the first to assess the potential reactivity associated with UVR exposure monitoring. Reactivity effects should be further examined in both intervention and observational contexts to better understand the impact of UVR monitoring on sun-protective behaviors and other relevant clinical outcomes. Ultimately, objectively assessed UVR exposure is an important measure to be integrated into outcome assessment for future testing of skin cancer prevention interventions. In the context of intervention testing, researchers who deploy objective UVR measures may want to compare intervention outcomes between individuals who used a UVR monitor and those who did not. In addition, when examining intervention effects on objectively assessed UVR exposure, researchers may want to analyze UVR data with and without the first few days of UVR data collected to minimize potential reactivity effects.

Acknowledgments

This work was supported by Dave Overholt, the National Cancer Institute of the National Institutes of Health (K07CA196985 to YPW; 3DP2EB022360-01S1 to JDJ), a Pilot Project Award from the American Cancer Society Huntsman Cancer Institute Institutional Research Grant (129785-IRG-16-190-01-IRG), and the Office of Communications, Genetic Counseling Shared Resource, and Cancer Biostatistics Shared Resource supported by grant P30CA042014 to Huntsman Cancer Institute. Data for this project were collected using REDCap (Research Electronic Data Capture), which is supported by the National Institutes of Health (8UL1TR000105, formerly UL1RR025764). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. The authors greatly appreciate Peter Kaplan and Emmanuel Dumont for their guidance on and assistance with using the Shade device and Jared Luther for assistance with data management. The authors are grateful for the study recruitment efforts of Mark Hyde, Garrett Harding, Jennyffer Morales, and Jane Ostler. The authors also thank James Carrington for his assistance in organizing the results.

Conflicts of Interest

None declared.

References

1. Feehan LM, Geldman J, Sayre EC, Park C, Ezzat AM, Yoo JY, et al. Accuracy of Fitbit Devices: Systematic Review and Narrative Syntheses of Quantitative Data. JMIR Mhealth Uhealth 2018 Aug 09;6(8):e10527 [FREE Full text] [doi: 10.2196/10527] [Medline: 30093371]
2. Fitbit reports 2019 fourth quarter and full year results. FitBit, Inc. 2020. URL: https://investor.fitbit.com/press-releases/press-release-details/2020/Fitbit-Reports-2019-Fourth-Quarter-and-Full-Year-Results/default.aspx [accessed 2021-03-22]
3. US National Library of Medicine. ClinicalTrials.gov. 2020. URL: https://clinicaltrials.gov/ [accessed 2021-02-12]
4. Banerjee S, Hoch E, Kaplan P, Dumont E. A comparative study of wearable ultraviolet radiometers. In: Proceedings of the IEEE Life Sciences Conference (LSC). 2017 Presented at: IEEE Life Sciences Conference (LSC); Dec. 13-15, 2017; Sydney, NSW, Australia. [doi: 10.1109/lsc.2017.8268131]
5. Nagelhout ES, Lensink R, Zhu A, Parsons BG, Jensen JD, Wu YP. The feasibility and acceptability of using a wearable uv radiation exposure monitoring device in adults and children: cross-sectional questionnaire study. JMIR Dermatol 2020 Apr 29;3(1):e15711. [doi: 10.2196/15711] [Medline: PMC7546527]
6. Cai S, Xu X, Yang W, Chen J, Fang X. Materials and designs for wearable photodetectors. Adv Mater 2019 Feb 20;31(18):1808138. [doi: 10.1002/adma.201808138] [Medline: 30785644]
7. Cancer stat facts: Melanoma of the skin. NIH Surveillance Epidemiology and End Results Program. 2020. URL: https://seer.cancer.gov/statfacts/html/melan.html [accessed 2020-08-27]
8. Armstrong BK, Kricker A. How much melanoma is caused by sun exposure? Melanoma Res 1993;3(6):395-401. [doi: 10.1097/00008390-199311000-00002] [Medline: 8161879]
9. Hillhouse J, Turrisi R, Jaccard J, Robinson J. Accuracy of self-reported sun exposure and sun protection behavior. Prev Sci 2012 Oct 27;13(5):519-531 [FREE Full text] [doi: 10.1007/s11121-012-0278-1] [Medline: 22855253]
10. Glanz K, Mayer JA. Reducing ultraviolet radiation exposure to prevent skin cancer methodology and measurement. Am J Prev Med 2005 Aug;29(2):131-142. [doi: 10.1016/j.amepre.2005.04.007] [Medline: 16005810]
11. Baumann S, Groß S, Voigt L, Ullrich A, Weymar F, Schwaneberg T, et al. Pitfalls in accelerometer-based measurement of physical activity: The presence of reactivity in an adult population. Scand J Med Sci Sports 2017 Oct 12;28(3):1056-1063. [doi: 10.1111/sms.12977] [Medline: 28921747]
12. Motl RW, McAuley E, Dlugonski D. Reactivity in baseline accelerometer data from a physical activity behavioral intervention. Health Psychol 2012;31(2):172-175. [doi: 10.1037/a0025965] [Medline: 22023436]
13. Vandoni M, Correale L, Del Bianco M, Marin L, Codrons E. Does reactivity to accelerometers occur in a single trial? Brief report in a sample of young adults. J Health Psychol 2016 Feb 15;22(11):1458-1462. [doi: 10.1177/1359105316628758] [Medline: 26880758]
14. Andersen PA, Buller DB, Walkosz BJ, Scott MD, Beck L, Liu X, et al. Environmental variables associated with vacationers' sun protection at warm weather resorts in North America. Environ Res 2016 Apr;146:200-206. [doi: 10.1016/j.envres.2015.12.034] [Medline: 26775001]
15. van der Leun JC, Piacentini RD, de Gruijl FR. Climate change and human skin cancer. Photochem Photobiol Sci 2008;7(6):730. [doi: 10.1039/b719302e] [Medline: 18528559]
16. Andersen PA, Buller DB, Walkosz BJ, Scott MD, Maloy JA, Cutter GR, et al. Environmental cues to UV radiation and personal sun protection in outdoor winter recreation. Arch Dermatol 2010 Nov 15;146(11):200-206. [doi: 10.1001/archdermatol.2010.327] [Medline: 21079060]
17. Hill D, White V, Marks R, Theobald T, Borland R, Roy C. Melanoma prevention: Behavioral and nonbehavioral factors in sunburn among an Australian urban population. Preven Med 1992 Sep;21(5):654-669. [doi: 10.1016/0091-7435(92)90072-p] [Medline: 1438112]
18. Silfee VJ, Haughton CF, Jake-Schoffman DE, Lopez-Cepero A, May CN, Sreedhara M, et al. Objective measurement of physical activity outcomes in lifestyle interventions among adults: A systematic review. Preventive Medicine Reports 2018 Sep;11:74-80. [doi: 10.1016/j.pmedr.2018.05.003] [Medline: 29984142]
19. Clemes SA, Matchett N, Wane SL. Reactivity: an issue for short-term pedometer studies? Br J Sports Med 2007 Jun 05;42(1):68-70. [doi: 10.1136/bjsm.2007.038521] [Medline: 18178685]
20. Dössegger A, Ruch N, Jimmy G, Braun-Fahrländer C, Mäder U, Hänggi J. Reactivity to accelerometer measurement of children and adolescents. Med Sci Sports Exerc 2014;46(6):1140. [doi: 10.1249/mss.0000000000000215] [Medline: 24219978]
21. Bravata DM, Smith-Spangler C, Sundaram V, Gienger AL, Lin N, Lewis R, et al. Using pedometers to increase physical activity and improve health. J Am Med Assoc 2007 Nov 21;298(19):2296. [doi: 10.1001/jama.298.19.2296] [Medline: 18029834]
22. Foley JT, Beets MW, Cardinal BJ. Monitoring children's physical activity with pedometers: reactivity revisited. J Exer Sci Fit 2011 Dec;9(2):82-86. [doi: 10.1016/s1728-869x(12)60002-1]

23. Lubans DR, Morgan PJ, Tudor-Locke C. A systematic review of studies using pedometers to promote physical activity among youth. Prev Med 2009 Apr;48(4):307-315. [doi: 10.1016/j.ypmed.2009.02.014] [Medline: 19249328]
24. Vincent S, Pangrazi R. Does reactivity exist in children when measuring activity levels with pedometers? Pediatr Exerc Sci 2002;14(1):56-63. [doi: 10.1123/pes.14.1.56]
25. Matevey C, Rogers LQ, Dawson E, Tudor-Locke C. Lack of reactivity during pedometer self-monitoring in adults. Meas Phys Educ Exerc Sci 2006 Mar;10(1):1-11. [doi: 10.1207/s15327841mpee1001_1]
26. Behrens TK, Dinger MK. Motion sensor reactivity in physically active young adults. Res Q Exerc Sport 2007 Mar;78(2):1-8. [doi: 10.1080/02701367.2007.10599397] [Medline: 7479568]
27. Hacker E, Horsham C, Vagenas D, Jones L, Lowe J, Janda M. A mobile technology intervention with ultraviolet radiation dosimeters and smartphone apps for skin cancer prevention in young adults: randomized controlled trial. JMIR Mhealth Uhealth 2018 Nov 28;6(11):e199. [doi: 10.2196/mhealth.9854] [Medline: 30487115]
28. Buller DB, Berwick M, Lantz K, Buller MK, Shane J, Kane I, et al. Evaluation of immediate and 12-week effects of a smartphone sun-safety mobile application: a randomized clinical trial. JAMA Dermatol 2015 May;151(5):505-512 [FREE Full text] [doi: 10.1001/jamadermatol.2014.3894] [Medline: 25629819]
29. Horsham C, Antrobus J, Olsen CM, Ford H, Abernethy D, Hacker E. Testing wearable UV sensors to improve sun protection in young adults at an outdoor festival: field study. JMIR Mhealth Uhealth 2020 Sep 16;8(9):e21243 [FREE Full text] [doi: 10.2196/21243] [Medline: 32936083]
30. Bodekær M, Petersen B, Philipsen PA, Heydenreich J, Thieden E, Wulf HC. Sun exposure patterns of urban, suburban, and rural children: a dosimetry and diary study of 150 children. Photochem Photobiol Sci 2015;14(7):1282-1289. [doi: 10.1039/c5pp00052a] [Medline: 26066309]
31. Thieden E, Philipsen PA, Heydenreich J, Wulf HC. UV radiation exposure related to age, sex, occupation, and sun behavior based on time-stamped personal dosimeter readings. Arch Dermatol 2004 Feb 01;140(2):197-203. [doi: 10.1001/archderm.140.2.197] [Medline: 14967793]
32. Thieden E, Ågren M, Wulf H. Solar UVR exposures of indoor workers in a Working and a Holiday Period assessed by personal dosimeters and sun exposure diaries. Photodermatol Photoimmunol Photomed 2001;17(6):249-255. [doi: 10.1111/j.1600-0781.2001.170601.x] [Medline: 11722749]
33. Thieden E, Philipsen P, Wulf H. Ultraviolet radiation exposure pattern in winter compared with summer based on time-stamped personal dosimeter readings. Br J Dermatol 2006;154(1):133-138. [doi: 10.1111/j.1365-2133.2005.06961.x] [Medline: 16403106]
34. Baczynska KA, Khazova M, O'Hagan JB. Sun exposure of indoor workers in the UK – survey on the time spent outdoors. Photochem Photobiol Sci 2019 Jan 15;18(1):120-128. [doi: 10.1039/c8pp00425k] [Medline: 30357230]
35. Wu YP, Aspinwall LG, Conn BM, Stump T, Grahmann B, Leachman SA. A systematic review of interventions to improve adherence to melanoma preventive behaviors for individuals at elevated risk. Prev Med 2016 Dec;88:153-167 [FREE Full text] [doi: 10.1016/j.ypmed.2016.04.010] [Medline: 27090434]
36. Nahar VK, Ford MA, Brodell RT, Boyas JF, Jacks SK, Biviji-Sharma R, et al. Skin cancer prevention practices among malignant melanoma survivors: a systematic review. J Cancer Res Clin Oncol 2016 Jun;142(6):1273-1283. [doi: 10.1007/s00432-015-2086-z] [Medline: 26642962]
37. Finch L, Janda M, Loescher LJ, Hacker E. Can skin cancer prevention be improved through mobile technology interventions? A systematic review. Prev Med 2016 Sep;90:121-132. [doi: 10.1016/j.ypmed.2016.06.037] [Medline: 27374946]
38. Prospect lists - Consumer mailing list builder. Experian. 2018. URL: https://www.experian.com/small-business/mailing-lists [accessed 2019-02-16]
39. Shade UV sensor - Funded by the National Cancer Institute. Shade. URL: https://www.wearshade.com/ [accessed 2021-03-04]
40. UV index: Annual time series. National Weather Service - Climate Prediction Center. URL: https://www.cpc.ncep.noaa.gov/products/stratosphere/uv_index/uv_annual.shtml [accessed 2021-02-04]
41. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]
42. Glanz K, Yaroch AL, Dancel M, Saraiya M, Crane LA, Buller DB, et al. Measures of sun exposure and sun protection practices for behavioral and epidemiologic research. Arch Dermatol 2008 Feb;144(2):217-222. [doi: 10.1001/archdermatol.2007.46] [Medline: 18283179]
43. Brewer J, Mcquinn D, Lohse C, Hassani J. Tanning bed perception survey: A questionnaire-based study. J Clin Aesthet Dermatol 2015 Mar;8(3):23-27 [FREE Full text] [Medline: 25852811]
44. Aspinwall LG, Leaf SL, Dola ER, Kohlmann W, Leachman SA. CDKN2A/p16 Genetic test reporting improves early detection intentions and practices in high-risk melanoma families. Cancer Epidemiol Biomark Prev 2008 Jun 16;17(6):1510-1519. [doi: 10.1158/1055-9965.epi-08-0010] [Medline: 18559569]
45. Køster B, Søndergaard J, Nielsen JB, Allen M, Olsen A, Bentzen J. The validated sun exposure questionnaire: association of objective and subjective measures of sun exposure in a Danish population-based sample. Br J Dermatol 2017 Feb;176(2):446-456. [doi: 10.1111/bjd.14861] [Medline: 27412948]

46. Aspinwall L, Stump T, Taber J, Drummond D, Kohlmann W, Champine M. Genetic test reporting of CDKN2A provides informational and motivational benefits for managing melanoma risk. Transl Behav Med 2017;8(1):29-43. [doi: 10.1093/tbm/ibx011] [Medline: 29385581]
47. Wu YP, Parsons BG, Aspinwall LG, Hay JL, Boucher KM, Caputo H, et al. Parent and child perspectives on perceived barriers to child sun protection and their association with sun protection strategies among children of melanoma survivors. Pediatr Dermatol 2019 Mar 20;36(3):317-323. [doi: 10.1111/pde.13796] [Medline: 30895676]
48. Lescano CM, Rodrigue JR. Skin cancer prevention behaviors among parents of young children. Child Health Care 1997 Apr;26(2):107-114. [doi: 10.1207/s15326888chc2602_4]
49. Nigg CR. Explaining adolescent exercise behavior change: A longitudinal application of the transtheoretical model. Ann Behav Med 2001 Feb;23(1):11-20. [doi: 10.1207/s15324796abm2301_3] [Medline: 11302350]
50. Prochaska JO, Velicer WF, Rossi JS, Goldstein MG, Marcus BH, Rakowski W, et al. Stages of change and decisional balance for 12 problem behaviors. Health Psychol 1994;13(1):39-46. [doi: 10.1037/0278-6133.13.1.39] [Medline: 8168470]
51. Rossi JS, Blais LM, Redding CA, Weinstock MA. Preventing skin cancer through behavior change. Implications for interventions. Dermatol Clin 1995 Jul;13(3):613-622. [doi: 10.1016/S0733-8635(18)30067-6] [Medline: 7554509]
52. Tribby CP, Julian AK, Oh AY, Perna FM, Berrigan D. Associations between ultraviolet radiation, tree cover and adolescent sunburns. Int J Health Geogr 2020 Dec 14;19(1):59. [doi: 10.1186/s12942-020-00253-x] [Medline: 33317555]
53. Holmbeck G. Post-hoc probing of significant moderational and mediational effects in studies of pediatric populations. J Pediatr Psychol 2002;27(1):87-96. [doi: 10.1093/jpepsy/27.1.87] [Medline: 11726683]

Abbreviations

REDCap: Research Electronic Data Capture
UVR: UV radiation

Edited by L Buis; submitted 22.04.21; peer-reviewed by T Stump, J Robinson; comments to author 25.05.21; revised version received 23.06.21; accepted 03.08.21; published 28.09.21.

Please cite as:
Parsons BG, Nagelhout ES, Wankier AP, Hu N, Lensink R, Zhu A, Nottingham K, Grossman D, Jensen JD, Wu YP
Reactivity to UV Radiation Exposure Monitoring Using Personal Exposure Devices for Skin Cancer Prevention: Longitudinal Observational Study
JMIR Mhealth Uhealth 2021;9(9):e29694
URL: https://mhealth.jmir.org/2021/9/e29694
doi:10.2196/29694
PMID:34581683

Corrigenda and Addenda

Correction: Effect of Physician-Pharmacist Participation in the Management of Ambulatory Cancer Pain Through a Digital Health Platform: Randomized Controlled Trial

Lu Zhang[1,2*], MD; Howard L McLeod[3,4,5,6*], Prof Dr; Ke-Ke Liu[1,2], MD; Wen-Hui Liu[7], Prof Dr; Hang-Xing Huang[1,2], MD; Ya-Min Huang[1,2], MD; Shu-Sen Sun[1,2,8], Prof Dr; Xiao-Ping Chen[2,5,6], Prof Dr; Yao Chen[5], Prof Dr; Fang-Zhou Liu[9], Prof Dr; Jian Xiao[1,2], Prof Dr

[1]Department of Pharmacy, Xiangya Hospital, Central South University, Changsha, China

[2]Institute for Rational and Safe Medication Practices, National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, China

[3]Geriatric Oncology Consortium, Tampa, FL, United States

[4]Taneja College of Pharmacy, University of South Florida, Tampa, FL, United States

[5]Department of Clinical Pharmacology, Xiangya Hospital, Central South University, Changsha, China

[6]Hunan Key Laboratory of Pharmacogenetics, Institute of Clinical Pharmacology, Central South University, Changsha, China

[7]Department of Pharmacy, The Second Xiangya Hospital, Central South University, Changsha, China

[8]College of Pharmacy and Health Sciences, Western New England University, Boston, MA, United States

[9]College of Information Science and Engineering, Hunan Normal University, Changsha, China

[*]these authors contributed equally

Corresponding Author:
Jian Xiao, Prof Dr
Department of Pharmacy
Xiangya Hospital
Central South University
Room 404
No. 87 Xiangya Road, Kaifu District
Changsha, 410008
China
Phone: 86 13973168895
Fax: 86 073184327460
Email: admanoas@163.com

Related Article:

Correction of: https://mhealth.jmir.org/2021/8/e24555/

(JMIR Mhealth Uhealth 2021;9(9):e33223) doi:10.2196/33223

In "Effect of Physician-Pharmacist Participation in the Management of Ambulatory Cancer Pain Through a Digital Health Platform: Randomized Controlled Trial" (JMIR Mhealth Uhealth 2021;9(8):e24555), one error was noted.

In the originally published article, the name of Corresponding Author "Jian Xiao" was formatted incorrectly as "Xiao Jian." The original order of authors was listed as follows:

> *Lu Zhang, Howard L McLeod, Ke-Ke Liu, Wen-Hui Liu, Hang-Xing Huang, Ya-Min Huang, Shu-Sen Sun, Xiao-Ping Chen, Yao Chen, Fang-Zhou Liu, Xiao Jian*

This has been corrected to:

> *Lu Zhang, Howard L McLeod, Ke-Ke Liu, Wen-Hui Liu, Hang-Xing Huang, Ya-Min Huang, Shu-Sen Sun, Xiao-Ping Chen, Yao Chen, Fang-Zhou Liu, Jian Xiao*

The correction will appear in the online version of the paper on the JMIR Publications website on September 13, 2021, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 28.08.21; this is a non–peer-reviewed article;accepted 02.09.21; published 13.09.21.

Please cite as:
Zhang L, McLeod HL, Liu KK, Liu WH, Huang HX, Huang YM, Sun SS, Chen XP, Chen Y, Liu FZ, Xiao J
Correction: Effect of Physician-Pharmacist Participation in the Management of Ambulatory Cancer Pain Through a Digital Health Platform: Randomized Controlled Trial
JMIR Mhealth Uhealth 2021;9(9):e33223
URL: https://mhealth.jmir.org/2021/9/e33223
doi:10.2196/33223
PMID:34516388

Corrigenda and Addenda

Correction: Strategies for the Successful Implementation of a Novel iPhone Loaner System (iShare) in mHealth Interventions: Prospective Study

William E Yang[1], MD; Erin M Spaulding[2], RN, BSN; David Lumelsky[3], BS; George Hung[1], MD; Pauline Phuong Huynh[1], BA; Kellen Knowles[1], MD; Francoise A Marvel[1,4], MD; Valerie Vilarino[3], BA; Jane Wang[1], MD; Lochan M Shah[1], BA; Helen Xun[1], BS; Rongzi Shan[4], BS; Shannon Wongvibulsin[1,5], PhD; Seth S Martin[1,4], MD, MHS

[1]School of Medicine, Johns Hopkins University, Baltimore, MD, United States

[2]School of Nursing, Johns Hopkins University, Baltimore, MD, United States

[3]Krieger School of Arts and Sciences, Johns Hopkins University, Baltimore, MD, United States

[4]Ciccarone Center for the Prevention of Cardiovascular Disease, Division of Cardiology, School of Medicine, Johns Hopkins University, Baltimore, MD, United States

[5]Department of Biomedical Engineering, School of Medicine, Johns Hopkins University, Baltimore, MD, United States

Corresponding Author:
William E Yang, MD
School of Medicine
Johns Hopkins University
4940 Eastern Ave
Baltimore, MD, 21224
United States
Phone: 1 410 550 0100
Email: wyang@jhmi.edu

Related Article:

Correction of: https://mhealth.jmir.org/2019/12/e16391/

(JMIR Mhealth Uhealth 2021;9(9):e31472) doi:10.2196/31472

In "Strategies for the Successful Implementation of a Novel iPhone Loaner System (iShare) in mHealth Interventions: Prospective Study" (JMIR Mhealth Uhealth 2019;7(12):e16391) the authors noted two errors.

In the *Results* section of the originally published paper, the first paragraph of the subsection "Evaluation of Process Outcomes" included incorrect numerators and denominators in the statistical values introduced during the copyediting process. The final two sentences of this paragraph were originally published as follows:

> *Compared to iPhone owners, iShare participants were slightly younger (age range 30-81 years, mean 57.4 [SD 11] vs age range 32-89 years, mean 60.8 [SD 11]). They were also more likely to be women (72/200, 36.0% vs 45/200, 23.0%), of black race (50/200, 25.0% vs 28/200, 14.0%), and insured by Medicaid (40/200, 20.0% vs 8/200, 4.0%).*

This has been corrected to:

> *Compared to iPhone owners, iShare participants were slightly younger (age range 30-81 years, mean 57.4 [SD 11] vs age range 32-89 years, mean 60.8 [SD 11]). They were also more likely to be women (33/92, 35.9% vs 25/108, 23.2%), of black race (23/92, 25.0% vs 15/108, 13.9%), and insured by Medicaid (18/92, 19.6% vs 4/108, 3.7%).*

As well, the phone number of the Corresponding Author has been updated from the originally published number to "1 410 550 0100."

The correction will appear in the online version of the paper on the JMIR Publications website on September 20, 2021, together with the publication of this correction notice. Because this was made after submission to PubMed, PubMed Central, and other full-text repositories, the corrected article has also been resubmitted to those repositories.

Submitted 22.06.21; this is a non–peer-reviewed article;accepted 24.06.21; published 20.09.21.

Please cite as:
Yang WE, Spaulding EM, Lumelsky D, Hung G, Huynh PP, Knowles K, Marvel FA, Vilarino V, Wang J, Shah LM, Xun H, Shan R, Wongvibulsin S, Martin SS
Correction: Strategies for the Successful Implementation of a Novel iPhone Loaner System (iShare) in mHealth Interventions: Prospective Study
JMIR Mhealth Uhealth 2021;9(9):e31472
URL: https://mhealth.jmir.org/2021/9/e31472
doi:10.2196/31472
PMID:34543222

Original Paper

User Experiences of the NZ COVID Tracer App in New Zealand: Thematic Analysis of Interviews

Alexei Tretiakov[1], PhD; Inga Hunter[1], MB, MPhil

School of Management, Massey University, Palmerston North, New Zealand

Corresponding Author:
Alexei Tretiakov, PhD
School of Management
Massey University
Tennent Drive
Palmerston North, 4474
New Zealand
Phone: 64 6 951 7905
Email: A.Tretiakov@massey.ac.nz

Abstract

Background: For mobile app–based COVID-19 contact tracing to be fully effective, a large majority of the population needs to be using the app on an ongoing basis. However, there is a paucity of studies of users, as opposed to potential adopters, of mobile contact tracing apps and of their experiences. New Zealand, a high-income country with western political culture, was successful in managing the COVID-19 pandemic, and its experience is valuable for informing policy responses in similar contexts.

Objective: This study asks the following research questions: (1) How do users experience the app in their everyday contexts? and (2) What drives the use of the app?

Methods: Residents of New Zealand's Auckland region, which encompasses the country's largest city, were approached via Facebook, and 34 NZ COVID Tracer app users were interviewed. Interview transcripts were analyzed using thematic analysis.

Results: Interviews ranged in duration from 15 to 50 minutes. Participants ranged in age from those in their late teens to those in their early sixties. Even though about half of the participants identified as White New Zealanders of European origin, different ethnicities were represented, including New Zealanders of South Pacific, Indian, Middle Eastern, South American, and Southeast Asian descent. Out of 34 participants, 2 (6%) identified as Māori (Indigenous New Zealanders). A broad range of careers were represented, from top-middle management to health support work and charity work. Likewise, educational backgrounds ranged broadly, from high school completion to master's degrees. Out of 34 participants, 2 (6%) were unemployed, having recently lost their jobs because of the pandemic. The thematic analysis resulted in five major themes: perceived benefits, patterns of use, privacy, social influence, and need for collective action. Benefits of using the app to society in general were more salient to the participants than immediate health benefits to the individual. Use, however, depended on the alert level and tended to decline for many participants at low alert levels. Privacy considerations played a small role in shaping adoption and use, even though the participants were highly aware of privacy discourse around the app. Participants were aware of the need for high levels of adoption and use of the app to control the pandemic. Attempts to encourage others to use the app were common, although not always successful.

Conclusions: Appeals to civic responsibility are likely to drive the use of a mobile contact tracing app under the conditions of high threat. Under the likely scenario of COVID-19 remaining endemic and requiring ongoing vigilance over the long term, other mechanisms promoting the use of mobile contact tracing apps may be needed, such as offering incentives. As privacy is not an important concern for many users, flexible privacy settings in mobile contact tracing apps allowing users to set their optimal levels of privacy may be appropriate.

(JMIR Mhealth Uhealth 2021;9(9):e26318) doi:10.2196/26318

KEYWORDS

COVID-19; contact tracing; app; New Zealand; adoption; use; civic responsibility; privacy

Introduction

Background

Contact tracing is a nonpharmaceutical intervention commonly used for curbing the spread of COVID-19 [1,2]. Manual contact tracing, conducted by interviewing patients diagnosed with the disease to identify their close contacts, is rather slow, and digital contact tracing involving digitally recording information about individuals' movements has been suggested to be potentially considerably more effective based on simulation evidence [3]. Mobile apps that perform digital contact tracing have been implemented in many countries, such as Australia [4] and Singapore [5]. Based on cross-country comparison, Urbaczewski and Lee [6] asserted that mobile app–based contact tracing is effective in helping countries to keep COVID-19 under control.

For mobile app–based contact tracing to be fully effective, a large majority of the population needs to be using the app on an ongoing basis [3,7]. The importance of mobile contact tracing app adoption prompted several empirical studies. Trang et al [8] conducted an experiment in Germany where they suggested alternative designs and made different appeals about the benefits offered, asking the respondents to rate their intent to install the app. They found that citizens can be divided into three categories: critics, undecided, and advocates. Critics were more likely to accept an app based on an appeal that by using it they would protect society in general, and they were more likely to accept app designs with strong privacy features. Undecided citizens, similar to critics, responded to societal-benefit appeals, but valued convenience in app use more than they valued privacy. Neither critics nor undecided citizens cared about the app offering health benefits to them as individuals. Finally, advocates responded to both societal-benefit and self-benefit appeals and did not care about privacy or convenience.

In a similar experiment conducted in the United Kingdom, Wiertz et al [9] offered citizens four configurations of an app differing by self-benefits offered, privacy, and the entity overseeing the app. Citizens—treated as a single group—preferred an app offering self-benefits and that was overseen by an independent entity, rather than by the government. Wiertz et al [9] found no evidence suggesting that citizens valued the privacy features of an app. Jonker et al [10] conducted a discrete choice experiment in the Netherlands, allowing citizens to rate a range of possible features of a mobile contact tracing app. Citizens preferred an app that would store data locally and give them control over whether to share it with the authorities. Further, they preferred an app that would offer a small financial reward. Thus, the results regarding the effects of privacy features and of self-benefits were not consistent across studies.

Walrave et al [11] conducted a survey in Belgium to determine factors affecting citizen intention to adopt a mobile contact tracing app. The study used the unified theory of acceptance and use of technology (UTAUT) framework [12]; performance expectancy (ie, benefits offered by the app, conceptualized by Walrave at al [11] as societal benefits), effort expectancy, social influence, and facilitating conditions (ie, having the knowledge and resources necessary to use the app) were considered as potential factors. Further, innovativeness, privacy concerns, and COVID-19–related stress were added to the basic UTAUT model. The most important factor was performance expectancy, followed by facilitating conditions and social influence. Innovativeness and privacy concerns had weaker effects on intention to adopt. The results by Walrave et al [11] are complemented by a survey by Altmann et al [13] that was conducted in France, Germany, Italy, the United Kingdom, and the United States. Regarding reasons to install a mobile contact tracing app, the respondents rated benefits to family and friends higher than benefits to the broader community. In addition, they rated concerns about government surveillance and security highly regarding the main reasons against installation, with concerns about government surveillance rated the highest. Further, greater trust in the government was associated with higher app installation intent.

In all the studies introduced above, the participants had no exposure to a real mobile contact tracing app and answered questions with a hypothetical app in mind.

Research Questions

The results of the studies published so far are not entirely consistent regarding factors driving the adoption of a mobile contact trading app. Further, studies of users, as opposed to potential adopters, of mobile contact tracing apps and their experiences are not available.

Better understanding of user experiences with the NZ (New Zealand) COVID Tracer app is of practical interest for countries using mobile contact tracing apps to protect their populations from COVID-19, particularly if COVID-19 becomes endemic [14], possibly necessitating the continued use of contact tracing over the long term. Further, understanding user experiences with a mobile contact tracing app is of broader theoretical interest for epidemiology. Therefore, this study asks the following research questions:

1. How do users experience the app in their everyday contexts?
2. What drives the use of the app?

Methods

Overall Approach and Study Setting

Qualitative design was used, as it is particularly suitable for an exploratory study of user experiences [15-17]. Data were collected via semistructured interviews with users of the NZ COVID Tracer app, New Zealand's official mobile contact tracing app overseen by the Ministry of Health [18].

The study was conducted in the Auckland region; Auckland is the biggest city in New Zealand. New Zealand's COVID-19 outbreak response, as assessed in October 2020, has been recognized as successful [19]. Thus, the New Zealand experience may be of interest. The Auckland region was chosen because it has experienced more COVID-19–related disruption than the rest of the country, as detailed in the following section.

Interviews were conducted 5 months after the app became available, allowing us to explore user experiences in the context of how the pandemic situation in the Auckland region and the

functionality of the app evolved over time. This context is described in the following section.

NZ COVID Tracer App and COVID-19 Pandemic in New Zealand

The NZ COVID Tracer app was released by New Zealand's Ministry of Health on May 20, 2020 [20], simultaneously for iOS and for Android platforms. The app was presented as a "digital diary," allowing the recording of places the users of the app visited by scanning QR (Quick Response) codes. Users could also register their contact details with the app to make it easier for COVID-19 contact tracers to reach them.

Privacy was emphasized in the app design and in the Ministry of Health's communications about the app: information about places visited by the user was stored locally on the phone and was not shared with contact tracing services automatically; in the initial release of the app, the user had to open the app and read out the information to contact tracers. Further, the information was automatically deleted after 31 days. Moreover, for security and privacy reasons, to use the app users had to log on using a strong (ie, sufficiently long and complex) password. The password had to be re-entered every 30 days, resulting in confusion for some of the users, who would not remember the password and, thus, were locked out of the app, as documented in comments on the Apple App Store [21] and on Google Play [22].

On June 15, 2020, the app was updated to allow users to be notified if they visited a venue around the same time as a known COVID-19 case [23]. This feature was implemented without sending users' location data to the Ministry of Health. Further, users could now send their location data to contact tracers if they chose to do so. If the initial version of the app solely supported the contact tracing process, thus offering benefits for the community or for the country as a whole, the updated app offered immediate benefits to the users, who, in case of exposure, could be diagnosed earlier and could receive early treatment, thus improving their prognosis [24]. Moreover, users receiving an alert could self-isolate, thus protecting their family, friends, and colleagues.

Benefits to the community, the user's family, and the user as an individual have been repeatedly highlighted in subsequent communications by the Ministry of Health: "Taking a few seconds to scan in with the app means we can quickly inform you when you may have been exposed to the virus, so you can take steps to protect yourself and your whānau [extended family]," "It also means if you test positive for the virus, you can instantly provide your digital diary to contact tracers to give them a massive head-start," and "The faster we can contact trace, the quicker we can get ahead of the virus and prevent spread in the community" [25].

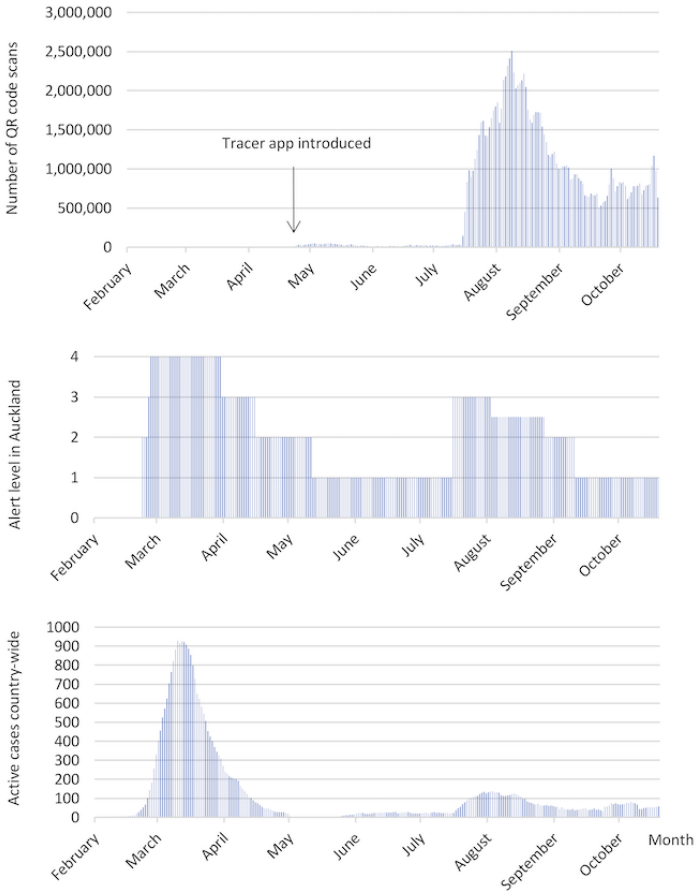
Another major update of the app was on July 30, 2020, when the ability to add manual entries to record visits to locations with no QR codes, such as visits to friends and family, was added [26], allowing one "to maintain a complete – and private – record." Initially, organizations were encouraged but not required to display QR codes compatible with the Ministry of Health's NZ COVID Tracer app [20]. However, starting from August 19, 2020, displaying QR codes became compulsory for most business premises and for many transport services [27].

Even though location data were held locally on users' phones, usage data, including the number of app registrations, the number of active devices, the number of QR code scans, and the number of manual entries, were available to the Ministry of Health, and some aggregate data were routinely shared via media releases (eg, Ministry of Health [28,29]). Further, historical data were available for download, and some of them are presented in Figure 1, where they are combined with historical data on the number of active COVID-19 cases in New Zealand on the COVID-19 data portal from Stats NZ, New Zealand's official data agency [30]. A more detailed graph of the number of active COVID-19 cases in New Zealand, distinguishing import-related and locally acquired cases, can be viewed at the Ministry of Health website [31]. The history of COVID-19 alert levels in the Auckland region—Auckland is the biggest city in New Zealand; the population of the Auckland region centered on Auckland is 1.6 million, about one-third of the total population of New Zealand [32]—is also shown in Figure 1, based on a document released by the New Zealand government [33]; Alert Level 4 corresponds to a lockdown with substantial restrictions on movement, while Alert Level 1 suggests heightened vigilance, but very few restrictions.

As seen in Figure 1, the NZ COVID Tracer app was introduced at the end of the first lockdown, which covered the whole of New Zealand, including Auckland [33], and received very little acceptance over June and July, while the country remained at Alert Level 1. Nonetheless, on August 12, 2020, a COVID-19 case with unknown source was discovered in Auckland, resulting in the alert level being raised to Alert Level 3 in the Auckland region and to Alert Level 2 in the rest of the country. This prompted a steep increase in the use of the NZ COVID Tracer app, with the daily number of QR scans growing by two levels of magnitude. However, the level of use decreased considerably once the country returned to Alert Level 1, although it remained considerably higher than before the second lockdown. Relatively high levels of active cases in October and November were almost exclusively imported cases, reflecting the growth of the pandemic overseas [34], and were not associated with higher use of the app.

New Zealand's COVID-19 outbreak response, as assessed in October 2020, has been recognized as successful [19]. However, in spite of the growth in adoption over the second lockdown in Auckland, the potential of the NZ COVID Tracer app in contributing to this response was not fully realized. As of November 13, 2020, even though 2.3 million users—almost half of the population of the country—were registered with the app, fewer than 1 in 6 of them were using it daily [25]. In an incident in Auckland involving a COVID-19 case visiting business premises on November 7, 2020, the number of potential contacts who could be traced via the app was very low, prompting the Ministry of Health to issue an appeal to citizens to use the app more [35]. In November 2020, improving user engagement with the NZ COVID Tracer app remained a problem for New Zealand.

Figure 1. COVID-19 pandemic in New Zealand and NZ COVID Tracer app use in 2020. QR: Quick Response.



Semistructured Interviews

The semistructured interview guide (Multimedia Appendix 1) was based on the UTAUT framework [11,12] and focused on effort expectancy (ie, effort associated with using the app), social influence (ie, the extent to which important others are perceived as encouraging the use of the app), facilitating conditions (ie, help available), and habit. Following Walrave et al [11], privacy concerns also received focus. Further, the interview guide emphasized perceived severity of COVID-19 (ie, the perceived consequences of being infected) and perceived susceptibility to COVID-19 (ie, the perceived likelihood of getting infected), concepts borrowed from the protection motivation theory (PMT) [36,37]. The benefits of using a mobile contact tracer app were explored at several levels, following Altmann et al [13], distinguishing benefits to the individual, the family, and society in general. Further, the self-reported patterns of use and the associated experiences were explored in detail, focusing both on current use and on how the approach to using the app by the respondent has changed over time. Moreover, the respondents were asked to project how they are anticipating using the app in the future. Respondents were allowed to deviate from the framework suggested by the interview guide, for as long as the interview remained overall relevant to the research questions of the study.

Participants were recruited using an advertising campaign on Facebook targeting Auckland region residents aged 18 to 64 years (55.19% of the New Zealand population are Facebook users [38]). The campaign invited users of the NZ COVID Tracer app to contribute to the fight against COVID-19 by granting an interview. Further, the participants were entered into a draw to win a token prize. All individuals meeting the criteria who expressed interest in being interviewed were interviewed until the desired sample size was reached; thus, a nonprobability consecutive sampling strategy was used. Following Braun and Clarke [39], the sample size was based on the sample sizes found to be sufficient to answer research questions in similar studies, such as Wessels et al [15] and Byambasuren et al [40], and on pragmatic considerations, such as the ability of the researchers to analyze the resulting volume of data within a reasonable time. The interviews were conducted by the first author over Zoom in late October and early November 2020. The interviews were transcribed in full for analysis.

Analysis

Thematic analysis of interview transcripts was conducted following Braun and Clarke [39]. Both deductive and inductive approaches were used, with deductive coding drawing from the UTAUT and the PMT. Following Braun and Clarke, concepts drawn from the UTAUT and the PMT, as introduced in the previous section, were used as a sensitizing device that was used to attract analysts' attention to potentially relevant aspects in the data; the aim was to understand user experiences and drivers of app use, rather than to test the UTAUT or the PMT. NVivo 12 (QSR International) was used for coding.

Both coauthors analyzed the data. Both researchers have higher degrees in information technology–related disciplines, with the first coauthor having a stronger technical background and the second coauthor having a background in medicine. Because of the difference in backgrounds, the researchers provided complementary perspectives. The researchers analyzed the data independently, periodically integrating the findings, and resolved differences via discussion.

Ethics

Following the university's ethics procedures, a low-risk notification was filed. Participants were informed in writing of their rights, such as the right to withdraw from the study at any point and the right to ask questions about the study. After receiving this information, the participants gave consent in writing.

Results

Participants

Interviews were conducted with 34 residents of the Auckland region, with interview durations ranging from 15 to 50 minutes (mean 23, SD 8.9; median 21.4). Participants (Table 1) ranged in age from those in their late teens to those in their early sixties. Even though about half of the participants identified as White New Zealanders of European origin, different ethnicities were represented, including New Zealanders of South Pacific, Indian, Middle Eastern, South American, and Southeast Asian descent. Out of 34 participants, 2 (6%) identified as Māori (Indigenous New Zealanders) and 1 (3%) was a temporary visitor from Europe stranded in New Zealand because of the COVID-19 pandemic (Participant #4). A broad range of careers was represented, from top-middle management to health support work and charity work. Out of 34 participants, 1 (3%) was a female homemaker and 1 (3%) was retired. Likewise, educational backgrounds ranged broadly, from high school completion to master's degrees. Out of 34 participants, 2 (6%) were unemployed (Participants #1 and #2), having recently lost their jobs because of the pandemic.

Table 1. Characteristics of the participants.

Participant No.	Gender	Age range (years)	Ethnicity	Education	Occupation
1	Male	18-29	NZ European[a]	Undergraduate diploma	Sales
2	Female	40-49	NZ European	Bachelor's degree	Pilot
3	Female	40-49	Macedonian	Master of Business Administration	Data scientist
4	Male	18-29	Caucasian	Master's degree	Biologist
5	Male	50-59	NZ European	Master's degree	Training provider
6	Female	50-59	Latin American	Master's degree	Interpreter
7	Male	30-39	Māori	Bachelor's degree	Support worker
8	Female	40-49	NZ European	Nursing degree	Nurse
9	Male	40-49	Middle Eastern	Master's degree	Account manager
10	Female	50-59	NZ European	Postgraduate degree	Teacher
11	Female	40-49	NZ European	Postgraduate degree	Head of compliance
12	Female	50-59	NZ European	Master's degree	Homemaker
13	Female	60-64	NZ European	Bachelor of Arts	Teacher
14	Female	40-49	NZ European	Postgraduate diploma	Teacher
15	Female	40-49	Scottish	Postgraduate degree	Health manager
16	Female	50-59	NZ European	Bachelor's degree	Charity worker
17	Female	50-59	Australian	Undergraduate diploma	Customer service
18	Female	50-59	NZ European	Postgraduate	Administration
19	Male	30-39	NZ European	High school diploma	Manager
20	Female	60-64	NZ European	Undergraduate degree	Retired
21	Female	50-59	European	Postgraduate	Health consultant
22	Female	50-59	Indian	Bachelor's degree	Manager
23	Female	50-59	NZ European	Master's degree	Teacher
24	Female	60-64	NZ European	Bachelor's degree	Dental receptionist
25	Male	40-49	Samoan	Trade certificate	Human resources
26	Male	18-29	Korean	High school diploma	University student
27	Female	30-39	European	Tertiary certificate	Manager
28	Male	50-59	Scottish	High school diploma	Business owner
29	Female	40-49	European	University diploma	Sales
30	Male	18-29	NZ European	High school diploma	Customer service
31	Male	40-49	NZ European	Master's degree	Local government
32	Female	40-49	NZ European	Postgraduate diploma	Health manager
33	Male	30-39	Māori	High school diploma	System engineer
34	Female	50-59	Indian	Postgraduate degree	Travel agent

[a]NZ European: White New Zealanders of European origin.

Themes

The thematic analysis resulted in five major themes: perceived benefits, patterns of use, privacy, social influence, and need for collective action. These themes are depicted along with the underlying subthemes and codes in Table 2. The content of the themes is presented in detail in the following sections.

Table 2. Major themes and the underlying subthemes and codes.

Themes and subthemes	Codes
Perceived benefits	
Society	Contact tracing
	Location data available to the government
Family	Self-isolate
Individual	Does not reduce risk
	Peace of mind from protecting others
	Reduction in uncertainty
	Better experience in contact tracing
	Personal diary
Patterns of use	
Use by yourself	Scanning codes
	Manual entries during a visit
	Manual entries after a visit
Involving others	Manual entries to record information about others
	Complaining to location manager if code not available
Persistence and changes	Persistence in the face of technical difficulties
	Maintaining use irrespective of the level of threat
	Use driven by the level of threat
Privacy	
I don't care	Nothing to hide
	The information is already out there
	Others care
I care, but I comply	Benefits outweigh privacy concerns
Improvements that would reduce privacy	Recording information automatically
	Using GPS location data
Social influence	
Encouragement	Encouraged by others
	Encouraging others
Help	Help from others
	Helping others
	Seeking help from the internet
Need for collective action	
Citizenship	Trust in New Zealand government
	Civic responsibility
Frustration	Ignoring conspiracy theorists
	Frustration by others' lack of use

Perceived Benefits

For most of the participants, the most prominent benefit of the app was supporting contact tracing in the context of controlling the pandemic in a broad sense:

> *Limits the damage and the spread of the virus drastically.* [Participant #4]

> *Feels like a very easy habit to maintain and a very small price to pay, because I can absolutely see the value of having a very quick easy method of tracking people.* [Participant #11]

> *If the app is doing what it says it's doing, then you know, one click, and everybody knows, and you've captured the problem, and you know that much faster, and we don't have to go into the stress of this lockdown business again.* [Participant #22]

Benefits to the family resulting from the individual being able to self-isolate early if at risk of infection were also mentioned:

> *It definitely gives me peace of mind because I have young children; I obviously never want to put them in harm's way.* [Participant #27]

> *I better remain cautious because of my wider family. My twin sister...if she caught it she would probably die.* [Participant #24]

Specific immediate health benefits to the individual using the app as well as greater likelihood to be diagnosed early and, thus, to receive early treatment, resulting in better prospects for the individual, were often not clear to the participants:

> *I wouldn't say I'm protecting myself. Because probably it doesn't reduce my risk in any way.* [Participant #20]

> *It can't prevent me from catching COVID, I would say it probably more protects the people around me.* [Participant #30]

Some of the participants identified reduction in uncertainty as a benefit to the individual using the app (ie, if you are infected, you are likely to know about it faster if you use the app):

> *I can see the benefit if it happens that I ended up being in contact with somebody that's got it. I would rather know quicker.* [Participant #27]

> *I feel like it's protecting me by keeping me in the know.* [Participant #21]

Another individual benefit suggested by the participants was the presumed better experience for the individual that was contract traced in the event contact tracing becomes necessary:

> *If I get sick, I can concentrate on getting well and I can leave contacting people to the government trackers who are paid to do their job.* [Participant #2]

> *I can instantly track back where I've gone, and I can provide that information rather than trying to think back, "Oh where was I, did I do that...?" It's all there.* [Participant #31]

Further, some of the participants found benefits that are not associated with COVID-19 virus control. For them, the app acted as a diary making it easy to recollect where they have been:

> *And it's quite good for me too, I find, because sometimes I forget where I've been. And I look at my app.* [Participant #10]

> *...helps me remember where I have been.* [Participant #26]

Overall, while benefits for contact tracing in the context of protecting society in general arose very naturally in the interviews, benefits to the individual using the app were often mentioned only after specific prompting by the interviewer, and different participants had different views on what they are.

Finally, some of the participants perceived the ability of the government to have access to location data and to conduct research using these data as a benefit, although it was not really a benefit because of the privacy features of the app:

> *Obviously, the government would know where everyone's going. So that's like, you know, we're helping...It is a good thing the government knows where you've been.* [Participant #26]

> *It's giving them data and it's giving them a platform to start developing what could be needed if this pandemic continues, or if there's a future one.* [Participant #24]

Patterns of Use

Simply scanning Ministry of Health QR codes displayed by businesses was the most common use of the app:

> *Just scan it sometimes if I'm not in a hurry. I don't scan it if I can't be bothered. I try to do it every time. But I'm not religious about it, you know...I suppose it's just a habit now.* [Participant #1]

> *I walk into a building, grab my phone out, and swipe the tracer. It doesn't really change my life, it's not that difficult.* [Participant #2]

> *I plan out when I'm getting out of the car. I have my phone in my hand. Anyway, and I'll just open the tracer app. And I just, I just walk by, like, you just hardly even have to stop now.* [Participant #21]

Manual entries were used when a QR code was not available or could not be scanned easily (eg, it was in an inconvenient location or was laminated, so that reflected light inhibited scanning):

> *But if they don't, for some reason, have the QR code available, I always put a manual entry in.* [Participant #23]

> *I just look for the QR codes and just record visits. Every now and then, it hasn't worked and I've recorded manual visit, but that doesn't seem to be happening so much.* [Participant #14]

Another common use of manual entries was to record visits to locations after the fact, when forgetting to scan the QR code:

> *I try to not forget, wherever I enter any place. But I've also caught myself several times, adding it manually after.* [Participant #3]

> *I have forgotten and I've moved away. And I remember like half an hour later, I put a manual entry in.* [Participant #31]

When a QR code could not be found, some of the participants complained to the location manager; others just did nothing:

> *I've asked people if you got a poster and gone and found it.* [Participant #5]

> *I just go and tell the management...If you don't display, I'm not comfortable in coming here...That's what I do if it's not available, or not prominently displayed.* [Participant #34]

> *When I'm entering the store or going to a place and I see a poster, that kind of reminds me to use it. I must admit, when I haven't seen a poster, I haven't used it. It's not automatic...* [Participant #10]

> *If there's nothing I don't bother asking.* [Participant #1]

The app allowed users to enter extra information in addition to scanning a QR code. This was occasionally used to record the

presence of others who did not scan for themselves, such as children. Sometimes the presence of others was recorded, possibly without their explicit approval:

> *If I've been with other people, like, particularly if my family have been with me, but they haven't had the phone with them, like my teenage sons. I'll put down that they were with me.* [Participant #18]

> *...she's trying to download and it doesn't work on her phone. So when she and I are together, that's okay, because I know we have a record that we're out together. And I'm doing it.* [Participant #11]

> *Twice I was with somebody who doesn't have their phone with them...So I just added their name to mine.* [Participant #13]

Patterns of use were impacted by updates rolled out by the app developers. For many of the participants, the app was entirely unusable at the beginning (eg, did not scan QR codes well enough or did not scan them at all). One of the participants reported installing and uninstalling the app multiple times, until a version that worked on her phone had been released:

> *The first couple of versions of it was such crap that each time I would end...I would uninstall it and I scream and shout and say, "I am never putting this back on my phone again." ...Then finally...I try to use it every day, I mean, if I go out. I try to always remember to have my phone.* [Participant #17]

> *Few places I tried to use it, it didn't work. So I put it away for a bit. I think when we went into the next community transmission and they said, or, you know...They seem to have done some work on it and they were raising again that it was good to use. And then I used it a couple of times and it worked. So I thought, "Okay, if it works, then why not?"* [Participant #22]

Some of the participants reported considerable resilience in continuing to use the app even while experiencing difficulties scanning the codes or being forcibly logged out of the app and having to recover the password:

> *Sometimes it's a little bit slow. And not just starting up. It's a little bit slow and sometimes forgets my password, and I have to log in again.* [Participant #1]

> *The only thing that has been a bit annoying would be that sometimes I was logged out...other than that it was a seamless transition. I mean, being logged out means that I need to figure out which was my password and I'm terrible at that. But other than that, I figured it's quite good.* [Participant #3]

Participants could be divided according to how their use of the app related to alert levels. While some of them reported continuing to use the app irrespective of the level of the alert—this behavior tended to be associated with the perception of being highly vulnerable to the virus—others reported less consistent use before and after the second lockdown; this was consistent with the pattern suggested in Figure 1:

> *I am an asthmatic. I would be considered a high-risk group...As a general rule, every time I see a QR code I scan the app.* [Participant #31]

> *I've had pneumonia before and I know it's worse than that...The codes at places didn't work properly...Halfway through the first lockdown. I actually deleted it. And then once they said that they done a few of the bug fixes and I've downloaded it again and I can definitely say that I've used it quite a lot since then.* [Participant #27]

> *I didn't use it for the first couple of weeks. Because I think when it first came out, we were at Level 1 already...I think I forgot the timing, but it might have been after the second lockdown. I've been using it absolutely consistently ever since and continue to do so.* [Participant #11]

> *During the lockdowns I always used the app. I mean, because you know it's a lockdown...But at the moment, with Level 1, I don't really use the app. And it's probably because no one else seems to be, whenever I walk into a place.* [Participant #25]

Very strong belief in the benefits and the necessity of the app and active steps to encourage others to use the app did not rule out reducing use once the alert level went down:

> *I was a true supporter to start off with. I've scanned wherever I could. I suggested to businesses to go and get it [the QR code for the app]. Made sure that our business got it right away, being a real supersupporter. And then we got over the first wave, we got back to work, there were no cases, so that it sort of died down, I've seen businesses removing the scan codes...there were no cases for quite a long time, and my use of the app changed, actually using it a lot less.* [Participant #29]

Privacy

Most of the participants did not worry about the app reducing their privacy. Privacy, however, was very prominent in the interviews, with the participants often discussing it with no prompt from the interviewer. Reasons mentioned for not worrying about privacy included the following: (1) the participant has nothing to hide; (2) the participants already perceive themselves as having no privacy as they are tracked via social media, by mobile phone service providers, via transaction records, or by other means; and (3) the participant relies on the app's privacy features:

> *I don't care that they know where I'm at. I don't think they'd care that much.* [Participant #1]

> *I don't have anything to hide. I'm not cheating anybody.* [Participant #2]

> *At the end of the day, it's like these cameras at your workplace. If you got nothing to hide, you don't have to worry about it.* [Participant #19]

> *...knowing what Google and the likes of Google can do...If someone wants to get something on you, everything is available.* [Participant #3]

> *All of us like to tap into free Wi-Fi everywhere we go. So really, there's a lot of data out there about us. But so, no. No, I don't care.* [Participant #22]

> *I'm not worried about the data that it gathers because that data is mine until it is required.* [Participant #31]

On the other hand, the existence of others who do care about privacy was often acknowledged, with their preferences mostly accepted as legitimate:

> *It [the data being recorded] doesn't bother me. I have a son who's a lawyer who refuses to...use the app. But not me. It does not bother me.* [Participant #16]

> *You don't have to put in your personal details, because there were a few [employees at work] that had privacy concerns with the COVID app, they are worried about people watching them, and we pointed out to them that you do not need to put any personal details into the app. You don't have to put in your first and your last name, you could call yourself Mister 123 if you really wanted to.* [Participant #19]

At the same time, some of the participants expressed negative attitudes toward mainstream and social media discourses overemphasizing privacy issues around the app. One of the participants, when asked about others discouraging her from using the app, pointed at one of the major New Zealand newspapers:

> *Newspapers, like [name of a major New Zealand newspaper], constantly have articles about how it's taking away our privacy and stuff like that.* [Participant #2]

Of the two participants who expressed concerns about privacy, one reported weighing privacy concerns against the benefits of faster contact tracing and deciding that benefits overweigh the risks. The other participant reiterated privacy concerns throughout the interview, but the concerns were not focused on the app and, rather, were about the overall environment, including social media and mobile phone service providers. At the same time, when asked how the app could be improved, the participant suggested an improvement that would reduce, rather than increase, privacy:

> *I am slightly a conspiracy theorist, but I thought weighing it all up, I felt it was more wise for me to embrace it.* [Participant #24]

> *This is this app, this is that app, there is a lot, you are controlling my life...It is annoying that you need to be booking everywhere you go, like keeping a diary of everything. It's a form of controlling, Facebook is a form of controlling, Google knows where you, whatever, is controlling, you have no privacy...Yes, I want to help the government and things like that, but at the same time this is a bit like Animal Farm [a novel by George Orwell].* [Participant #6]

> *You go to places, you need to park your car. Maybe integrate with your car parking...* [Participant #6]

Many participants suggested improvements that would reduce privacy (eg, recording visits automatically using wireless technology, using GPS to track app user location, or using wireless technology to automatically detect and record the proximity of others). None of the participants suggested changes to the app that would increase privacy.

Social Influence and the Need for Collective Action

Overwhelmingly, the participants expressed high levels of trust in the New Zealand government. Often, the fact that information comes from the government was a criterion of its trustworthiness. Using the app was seen as their civic duty and a way to be a good citizen of New Zealand. Moreover, some of the participants framed patterns of behavior in terms of "good" and "bad." A phrase introduced by the Prime Minister and repeatedly used in communications about the pandemic by her and other officials was commonly mentioned: "the team of the five million" [41,42]:

> *I go straight to the New Zealand government COVID-19 website. I try to stay away from the internet.* [Participant #28]

> *I feel like I should do it [use the app]...just to be a good citizen of New Zealand united team of 5 million and all of that.* [Participant #1]

> *I feel like you kind of need to set the example...ultimately, if I get sick I would feel that it is my civic responsibility to make sure that anyone who was near me, came in contact with me, would get the health care that they might need.* [Participant #33]

> *In the beginning there was a lot of talk about it in the news, and people saying, "Oh no, that's stealing all your personal data." But I don't see it this way at all. I think this is good data to be used in a good way.* [Participant #25]

However, trust in the government was not a prerequisite for using the app. One of the participants, a manager in an industry that was highly critical in maintaining the functioning of the city during lockdowns, expressed very low levels of trust in the government, even suggesting that the government purposefully distorted some of the information related to the pandemic; at the same time, he reported not only using the app but also ensuring that it was installed on mobile devices used by employees, as well as establishing procedures to ensure that visitors to company premises used the app:

> *I think there's a lot covered to avoid panic. So, yeah, yeah, it's hard to trust.* [Participant #19]

The flow of social influence was rather complex, involving multiple actors. Participants reported being encouraged by others to use the app. Further, for some, using the app was a requirement at their workplace. Participants also reported encouraging others, in face-to-face settings and online:

> *My parents actually...said, "You should probably get it [the app]." And I said, "Yeah, yeah."* [Participant #26]

> *I installed it I when I still had a job...It was a requirement for me as part of my job to use it. But then when I lost my job, I became more flexible with it. Like, it's not a requirement. Now for me, it's just*

> *something that I grew in, it became a habit.* [Participant #1]
>
> *My husband, he does use it a lot. Whenever we go out, and if sometimes I rush going somewhere, he just stops me and says, "Just scan it." He always reminds me and encourages me to be more vigilant...My friends, I just told them to install it, maybe they have, I do not know...When the app was first introduced, I sent messages to a lot of my contacts.* [Participant #34]

Businesses displaying the codes occasionally encouraged app use; in their turn, some of the participants actively engaged with businesses when a code could not be found or was unusable, as already highlighted in the Patterns of Use section:

> *Now [after the second lockdown ended] I'm not really using it, no. I only use it if the shopkeeper asks me to use it and then I will say, "Of course I'll use it."* [Participant #26]
>
> *We do actively, when we have customers coming to pick items, my receptionist will say..."Can you scan in please?" We literally say it to everyone who walks in...To be honest, people don't scan when they are coming in, the instance you say, "Would you mind scanning," nobody's ever thrown anything back at us, they just say, "No worries," they do it.* [Participant #19]

Some of the participants relied on internet-based resources, such as the Ministry of Health website, when having problems using the app. Help was not always readily available:

> *...just go to the [Ministry of Health] website for help.* [Participant #25]
>
> *I look at the Ministry website...* [Participant #8]
>
> *...and I'm going to COVID-19 website and I cannot even find where the bloody test centers are.* [Participant #9]
>
> *I even tried to contact someone and say, "You know, it's not working." And then I figured out, I guess, you know, it's not working. So they're getting too much communication...but once it has started working, like once I started using it the second time round, haven't had any need to contact anyone for help.* [Participant #22]

Family members and colleagues may have been a readier source of help:

> *I'll ask my 21-year-old son. He is quite tech savvy. So I utilize expertise inside my family.* [Participant #28]
>
> *I would ask my husband.* [Participant #27]
>
> *I was coming back by bus...and it was not scanning...somebody was sitting next to me, my colleague...I passed my phone to her, and she scanned it for me, because she was a little bit closer.* [Participant #34]

Further, strangers helped each other:

> *I was in the supermarket and then was having difficulty. And the guy said, "Oh, you know, you need to turn on the camera."* [Participant #6]
>
> *I do see a lot of young people helping older people use it.* [Participant #8]
>
> *I've helped a couple of people to download it.* [Participant #21]

The existence of "conspiracy theorists" raising, from the perspective of the participants, unreasonable or untrue privacy concerns, was occasionally acknowledged. Although one of the participants described himself as a "conspiracy theorist," he still used the app, judging that the benefits were greater than the risks, as introduced in the Privacy section:

> *...some think that the COVID-19 scanning app was taking the wrong information on each individual to hand more information to the government power. I don't believe that myself. Those conspiracy theorists inside my own family unit, I took it with a grain of salt. So meaning...I believe in the app.* [Participant #28]
>
> *My mother-in-law, she's saying, you know, "Don't use that, don't install it, because they're tracking all your data and everything."...I'm ignoring her for a lot of things.* [Participant #18]

Many of the participants were concerned about the behavior of others in using the app and in reducing the risk of COVID-19 spread in other ways. The realization that protecting the country from the pandemic depended on collective action was rather strong. Often, rather than expecting the authorities to improve technological capabilities or ease of use of the app, the participants highlighted the need to encourage its broader use:

> *I find it quite frustrating going to places and seeing people around me who, you know, are walking on without even bothering to scan.* [Participant #23]
>
> *I think New Zealand's getting very complacent.* [Participant #2]
>
> *My friends or family don't use it. Full stop.* [Participant #19]
>
> *My concern is that we have a lot of people just disregarding the impact of COVID.* [Participant #8]
>
> *I think, some shops are deliberately making it [the QR code] hard to find.* [Participant #5]

Discussion

Principal Findings

The main contribution of this study of adoption and use of a mobile contact tracing app is that it is based on data reflecting real user experiences, rather than on perceptions of individuals who are yet to use such an app. Prior studies predicting mobile contact tracing app adoption and use relied on data obtained from nonusers.

The results of this study are consistent with the finding by Trang et al [8] that perceptions of benefits for society as a whole are likely to drive the use of a mobile contact tracing app. However,

the results also suggest that such benefits are mainly relevant when the level of threat to society is high. For many individuals, but not all, the logic of taking individual action to protect the society on which the individuals depend is powerful enough to drive sustained use only when the threat to the society is salient enough.

The results of our study are consistent with Altmann et al [13] in suggesting that trust in the government helps to promote mobile tracing app use. Nonetheless, the finding by Altmann et al [13] that concerns about government surveillance are very important were not confirmed by our study. This may be, in part, because our study covered only the users of the app, who were likely to fall into the "advocates" category following Trang et al [8]. Assuming the participants of this study were "advocates," the finding that privacy did not matter for them is consistent with the results by Trang et al [8]. The results of our study are consistent with Wiertz et al [9], who also found little evidence that privacy is highly relevant, as well as with the body of literature on the privacy paradox [43], which suggests that in actual use, users are prepared to trade their privacy even for rather small benefits.

The NZ COVID Tracer app was designed in such a way that its use, or nonuse, was highly visible. Thus, social influence, found to have an effect by Walrave et al [11], could be highly influential. Nonetheless, for most of the participants, social influence by peers appeared to play a secondary role in driving their app use. Indeed, some of them continued to use the app while surrounded by nonusers. For them, social influence was coming from the government, not from the peers. At the same time, the results indicate that organizations may be effective in promoting the use of the mobile tracing app by their employees: employees who are not users are likely to comply to become users, rather than resist.

Our study found no indications that an app overseen by an independent entity, rather than by the government, would be better accepted or used more, and in this respect, our results did not confirm the results by Wiertz et al [9]. Indeed, the discourse by the app users around good citizenship and civic duty as reasons for using the app suggested that oversight by the government was a good choice in the New Zealand context. Nonetheless, this conclusion has to be confirmed by a study of nonusers of the app.

The study by Walrave et al [11] did not find effort expectancy to be an important factor. Our results were consistent with this finding. Determined users of the app were prepared to persist in the face of technical difficulties. This is not to suggest that effort expectancy is irrelevant; however, there was little evidence to suggest that, after the initial bugs were fixed, making the app even more effortless to use would result in significantly higher adoption and use.

The implications for practice are that appeals to civic responsibility are likely to drive the use of a mobile tracing app under the conditions of high threat, as citizens "rally around the flag." Under the likely scenario of COVID-19 remaining endemic and requiring ongoing vigilance over the long term, other mechanisms promoting the use of mobile tracing apps may be needed, such as "nudging" [44] (eg, offering incentives). Further, the results suggest that privacy is not an important concern for many users. Having access to more detailed information faster would benefit contact tracing, enabling faster isolation of probable cases and, thus, better control of the pandemic. Therefore, compared with a mobile tracing app with uniformly restrictive privacy features, an app with flexible privacy settings allowing users to set their optimal levels of privacy—thus allowing users who are less concerned about privacy to opt in to provide more detailed information faster—may be more appropriate.

The value of comparing responses between countries in informing decision making in the COVID-19 pandemic has been highlighted by Pearce et al [45], who characterized health policy responses in different jurisdictions as "numerous natural experiments in progress" (page 1059 in Pearce et al [45]). The case of New Zealand is particularly valuable in this respect because it presents an example of a successful response [46,47] achieved in a country with western political culture [48]. As such, the New Zealand experience in managing the pandemic has received a lot of attention in international literature [49-56]. Our study contributes to this body of literature by focusing on the experiences of the users of the NZ COVID Tracer app. However, the results of this study, as well as of other studies of the New Zealand experience in managing the COVID-19 pandemic, cannot be mechanically applied to other contexts. Rather, as for most qualitative studies, the process of case-to-case transfer [57] should apply: the readers and the consumers of the research should compare their context of interest to the New Zealand context and judge the extent to which the findings apply to their situation (page 1453 in Polit and Beck [57]). A broad description of the New Zealand context as it applies to the management of the COVID-19 pandemic is given by Jefferies et al [54], who assert that the New Zealand response to COVID-19 "has international relevance, particularly for other island nations, high-income and western settings" (page e613 in Jefferies et al [54]). Further, aspects of the context immediately relevant to the research questions of our study, such as the app design and the way it was introduced, relying on persuasion rather than on mandates, are described in the initial sections of this paper.

Conclusions

Appeals to civic responsibility are likely to drive the use of a mobile contact tracing app under the conditions of high threat. Under the likely scenario of COVID-19 remaining endemic and requiring ongoing vigilance over the long term, other mechanisms promoting the use of mobile contact tracing apps may be needed, such as offering incentives. As privacy is not an important concern for many users, flexible privacy settings in mobile contact tracing apps allowing users to set their optimal levels of privacy may be appropriate.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Interview guide.
[DOCX File , 21 KB - mhealth_v9i9e26318_app1.docx]

References

1. Park YJ, Choe YJ, Park O, Park SY, Kim Y, Kim J, COVID-19 National Emergency Response Center, Epidemiology and Case Management Team. Contact tracing during coronavirus disease outbreak, South Korea, 2020. Emerg Infect Dis 2020 Oct;26(10):2465-2468 [FREE Full text] [doi: 10.3201/eid2610.201315] [Medline: 32673193]
2. Cheng H, Jian S, Liu D, Ng T, Huang W, Lin H, Taiwan COVID-19 Outbreak Investigation Team. Contact tracing assessment of COVID-19 transmission dynamics in Taiwan and risk at different exposure periods before and after symptom onset. JAMA Intern Med 2020 Sep 01;180(9):1156-1163 [FREE Full text] [doi: 10.1001/jamainternmed.2020.2020] [Medline: 32356867]
3. Ferretti L, Wymant C, Kendall M, Zhao L, Nurtay A, Abeler-Dörner L, et al. Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. Science 2020 May 08;368(6491):1-7 [FREE Full text] [doi: 10.1126/science.abb6936] [Medline: 32234805]
4. COVIDSafe app. Australian Government, Department of Health. URL: https://www.health.gov.au/resources/apps-and-tools/covidsafe-app [accessed 2020-11-18]
5. TraceTogether, safer together. Government of Singapore. URL: https://www.tracetogether.gov.sg/ [accessed 2020-11-18]
6. Urbaczewski A, Lee Y. Information technology and the pandemic: A preliminary multinational analysis of the impact of mobile tracking technology on the COVID-19 contagion control. Eur J Inf Syst 2020 Aug 10;29(4):405-414 [FREE Full text] [doi: 10.1080/0960085x.2020.1802358]
7. Hinch R, Probert W, Nurtay A, Kendall M, Wymant C, Hall M, et al. Effective Configurations of a Digital Contact Tracing App: A Report to NHSX. Oxford, UK: Fraser Group; 2020 Apr 16. URL: https://045.medsci.ox.ac.uk/files/files/report-effective-app-configurations.pdf [accessed 2020-11-18]
8. Trang S, Trenz M, Weiger W, Tarafdar M, Cheung C. One app to trace them all? Examining app specifications for mass acceptance of contact-tracing apps. Eur J Inf Syst 2020 Jul 27;29(4):415-428 [FREE Full text] [doi: 10.1080/0960085x.2020.1784046]
9. Wiertz C, Banerjee A, Acar OA, Ghosh A. Predicted adoption rates of contact tracing app configurations - Insights from a choice-based conjoint study with a representative sample of the UK population. SSRN 2020 Apr 30:1-19 [FREE Full text] [doi: 10.2139/ssrn.3589199]
10. Jonker M, de Bekker-Grob E, Veldwijk J, Goossens L, Bour S, Rutten-Van Mölken M. COVID-19 contact tracing apps: Predicted uptake in the Netherlands based on a discrete choice experiment. JMIR Mhealth Uhealth 2020 Oct 09;8(10):e20741 [FREE Full text] [doi: 10.2196/20741] [Medline: 32795998]
11. Walrave M, Waeterloos C, Ponnet K. Ready or not for contact tracing? Investigating the adoption intention of COVID-19 contact-tracing technology using an extended unified theory of acceptance and use of technology model. Cyberpsychol Behav Soc Netw 2021 Jun;24(6):377-383. [doi: 10.1089/cyber.2020.0483] [Medline: 33017171]
12. Venkatesh V, Thong J, Xu X. Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. MIS Q 2012;36(1):157. [doi: 10.2307/41410412]
13. Altmann S, Milsom L, Zillessen H, Blasone R, Gerdon F, Bach R, et al. Acceptability of app-based contact tracing for COVID-19: Cross-country survey study. JMIR Mhealth Uhealth 2020 Aug 28;8(8):e19857 [FREE Full text] [doi: 10.2196/19857] [Medline: 32759102]
14. Shaman J, Galanti M. Will SARS-CoV-2 become endemic? Science 2020 Oct 30;370(6516):527-529. [doi: 10.1126/science.abe5960] [Medline: 33055131]
15. Wessels NJ, Hulshof L, Loohuis AMM, van Gemert-Pijnen L, Jellema P, van der Worp H, et al. User experiences and preferences regarding an app for the treatment of urinary incontinence in adult women: Qualitative study. JMIR Mhealth Uhealth 2020 Jun 12;8(6):e17114 [FREE Full text] [doi: 10.2196/17114] [Medline: 32530431]
16. Stanton J, Weiss E. Electronic monitoring in their own words: An exploratory study of employees' experiences with new types of surveillance. Comput Human Behav 2000 Jul;16(4):423-440. [doi: 10.1016/s0747-5632(00)00018-2]
17. Thomson K, Brouwers C, Damman OC, de Bruijne MC, Timmermans DR, Melles M. How health care professionals evaluate a digital intervention to improve medication adherence: Qualitative exploratory study. JMIR Hum Factors 2018 Feb 20;5(1):e7 [FREE Full text] [doi: 10.2196/humanfactors.8948] [Medline: 29463494]
18. NZ COVID Tracer app. New Zealand Ministry of Health. URL: https://www.health.govt.nz/our-work/diseases-and-conditions/covid-19-novel-coronavirus/covid-19-resources-and-tools/nz-covid-tracer-app [accessed 2020-11-18]
19. Robert A. Lessons from New Zealand's COVID-19 outbreak response. Lancet Public Health 2020 Nov;5(11):e569-e570 [FREE Full text] [doi: 10.1016/S2468-2667(20)30237-1] [Medline: 33065024]

20. NZ COVID Tracer app released to support contact tracing. New Zealand Ministry of Health. 2020 May 20. URL: https://www.health.govt.nz/news-media/media-releases/nz-covid-tracer-app-released-support-contact-tracing [accessed 2020-11-18]
21. NZ COVID Tracer. Apple App Store. URL: https://apps.apple.com/nz/app/nz-covid-tracer/id1511667597 [accessed 2020-11-18]
22. NZ COVID Tracer. Google Play. URL: https://play.google.com/store/apps/details?id=nz.govt.health.covidtracer [accessed 2020-11-18]
23. NZ COVID Tracer app and daily numbers update. New Zealand Ministry of Health. 2020 Jun 10. URL: https://www.health.govt.nz/news-media/media-releases/nz-covid-tracer-app-and-daily-numbers-update [accessed 2020-11-18]
24. Peck K. Early diagnosis and rapid isolation: Response to COVID-19 outbreak in Korea. Clin Microbiol Infect 2020 Jul;26(7):805-807 [FREE Full text] [doi: 10.1016/j.cmi.2020.04.025] [Medline: 32344168]
25. 4 cases of COVID-19 – None in the community. New Zealand Ministry of Health. 2020 Nov 13. URL: https://www.health.govt.nz/news-media/media-releases/4-cases-covid-19-none-community [accessed 2020-11-18]
26. Update to NZ COVID Tracer mobile app released. New Zealand Ministry of Health. 2020 Jul 30. URL: https://www.health.govt.nz/news-media/media-releases/update-nz-covid-tracer-mobile-app-released [accessed 2020-11-18]
27. Update - Monday, 24 August 2020. New Zealand Department of Internal Affairs. 2020 Aug 24. URL: https://www.dia.govt.nz/COVID-19-Response-Update-Monday-24-August-2020 [accessed 2020-11-18]
28. 5 new cases of COVID-19. New Zealand Ministry of Health. 2020 Nov 03. URL: https://www.health.govt.nz/news-media/media-releases/5-new-cases-covid-19-9 [accessed 2020-11-18]
29. 2 additional cases of COVID-19. New Zealand Ministry of Health. 2020 Nov 04. URL: https://www.health.govt.nz/news-media/media-releases/2-additional-cases-covid-19 [accessed 2020-11-18]
30. COVID-19 data portal. Stats NZ. URL: https://www.stats.govt.nz/experimental/covid-19-data-portal [accessed 2020-11-18]
31. COVID-19: Source of cases. New Zealand Ministry of Health. URL: https://www.health.govt.nz/our-work/diseases-and-conditions/covid-19-novel-coronavirus/covid-19-data-and-statistics/covid-19-source-cases [accessed 2020-11-18]
32. 2018 census place summaries: New Zealand. Stats NZ. URL: https://www.stats.govt.nz/tools/2018-census-place-summaries/new-zealand [accessed 2020-11-18]
33. History of the COVID-19 Alert System. New Zealand Government. URL: https://covid19.govt.nz/alert-system/history-of-the-covid-19-alert-system/ [accessed 2020-11-18]
34. Coronavirus Resource Center. John Hopkins University & Medicine. URL: https://coronavirus.jhu.edu/ [accessed 2020-11-18]
35. 1 case of COVID-19 in quarantine worker. New Zealand Ministry of Health. 2020 Nov 07. URL: https://www.health.govt.nz/news-media/media-releases/1-case-covid-19-quarantine-worker [accessed 2020-11-18]
36. Ling M, Kothe EJ, Mullan BA. Predicting intention to receive a seasonal influenza vaccination using protection motivation theory. Soc Sci Med 2019 Jul;233:87-92. [doi: 10.1016/j.socscimed.2019.06.002] [Medline: 31195194]
37. Maddux J, Rogers R. Protection motivation and self-efficacy: A revised theory of fear appeals and attitude change. J Exp Soc Psychol 1983 Sep;19(5):469-479. [doi: 10.1016/0022-1031(83)90023-9]
38. Social media stats in New Zealand - December 2020. StatCounter. 2020. URL: https://gs.statcounter.com/social-media-stats/all/new-zealand [accessed 2021-01-14]
39. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
40. Byambasuren O, Beller E, Hoffmann T, Glasziou P. Barriers to and facilitators of the prescription of mHealth apps in Australian general practice: Qualitative study. JMIR Mhealth Uhealth 2020 Jul 30;8(7):e17447 [FREE Full text] [doi: 10.2196/17447] [Medline: 32729839]
41. No new cases of COVID-19. New Zealand Ministry of Health. 2020 Aug 09. URL: https://www.health.govt.nz/news-media/media-releases/no-new-cases-covid-19-44 [accessed 2020-11-18]
42. Roy EA. Ardern thanks 'team of 5 million' as New Zealand reopens schools and offices. The Guardian. 2020 May 11. URL: https://www.theguardian.com/world/2020/may/11/ardern-announces-new-zealand-will-reopen-schools-offices-and-restaurants-this-week [accessed 2020-11-18]
43. Kokolakis S. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. Comput Secur 2017 Jan;64:122-134. [doi: 10.1016/j.cose.2015.07.002]
44. Riemer K, Ciriello R, Peter S, Schlagwein D. Digital contact-tracing adoption in the COVID-19 pandemic: IT governance for collective action at the societal level. Eur J Inf Syst 2020 Sep 23;29(6):731-745 [FREE Full text] [doi: 10.1080/0960085x.2020.1819898]
45. Pearce N, Lawlor D, Brickley E. Comparisons between countries are essential for the control of COVID-19. Int J Epidemiol 2020 Aug 01;49(4):1059-1062 [FREE Full text] [doi: 10.1093/ije/dyaa108] [Medline: 32601669]
46. Baker MG, Wilson N, Anglemyer A. Successful elimination of COVID-19 transmission in New Zealand. N Engl J Med 2020 Aug 20;383(8):e56 [FREE Full text] [doi: 10.1056/nejmc2025203]
47. Cousins S. New Zealand eliminates COVID-19. Lancet 2020 May 09;395(10235):1474 [FREE Full text] [doi: 10.1016/S0140-6736(20)31097-7] [Medline: 32386582]
48. Epstein LD. Political Parties in Western Democracies. New Brunswick, NJ: Transaction Publishers; 1980.

49. Huang QS, Wood T, Jelley L, Jennings T, Jefferies S, Daniells K, NPIsImpactOnFlu Consortium, et al. Impact of the COVID-19 nonpharmaceutical interventions on influenza and other respiratory viral infections in New Zealand. Nat Commun 2021 Feb 12;12(1):1001 [FREE Full text] [doi: 10.1038/s41467-021-21157-9] [Medline: 33579926]
50. Kung S, Doppen M, Black M, Hills T, Kearns N. Reduced mortality in New Zealand during the COVID-19 pandemic. Lancet 2021 Jan 02;397(10268):25 [FREE Full text] [doi: 10.1016/S0140-6736(20)32647-7] [Medline: 33333005]
51. Every-Palmer S, Jenkins M, Gendall P, Hoek J, Beaglehole B, Bell C, et al. Psychological distress, anxiety, family violence, suicidality, and wellbeing in New Zealand during the COVID-19 lockdown: A cross-sectional study. PLoS One 2020;15(11):e0241658 [FREE Full text] [doi: 10.1371/journal.pone.0241658] [Medline: 33147259]
52. Foley DA, Kirk M, Jepp C, Brophy-Williams S, Tong SYC, Davis JS, Study endorsed by ASID CRNANZPID Group. COVID-19 and paediatric health services: A survey of paediatric physicians in Australia and New Zealand. J Paediatr Child Health 2020 Aug;56(8):1219-1224 [FREE Full text] [doi: 10.1111/jpc.14903] [Medline: 32364301]
53. Fouda A, Mahmoudi N, Moy N, Paolucci F. The COVID-19 pandemic in Greece, Iceland, New Zealand, and Singapore: Health policies and lessons learned. Health Policy Technol 2020 Dec;9(4):510-524 [FREE Full text] [doi: 10.1016/j.hlpt.2020.08.015] [Medline: 32874858]
54. Jefferies S, French N, Gilkison C, Graham G, Hope V, Marshall J, et al. COVID-19 in New Zealand and the impact of the national response: A descriptive epidemiological study. Lancet Public Health 2020 Nov;5(11):e612-e623 [FREE Full text] [doi: 10.1016/S2468-2667(20)30225-5] [Medline: 33065023]
55. McGuire D, Cunningham J, Reynolds K, Matthews-Smith G. Beating the virus: An examination of the crisis communication approach taken by New Zealand Prime Minister Jacinda Ardern during the COVID-19 pandemic. Hum Resour Dev Int 2020 Jun 22;23(4):361-379 [FREE Full text] [doi: 10.1080/13678868.2020.1779543]
56. Summers J, Cheng H, Lin H, Barnard LT, Kvalsvig A, Wilson N, et al. Potential lessons from the Taiwan and New Zealand health responses to the COVID-19 pandemic. Lancet Reg Health West Pac 2020 Nov;4:100044 [FREE Full text] [doi: 10.1016/j.lanwpc.2020.100044] [Medline: 34013216]
57. Polit DF, Beck CT. Generalization in quantitative and qualitative research: Myths and strategies. Int J Nurs Stud 2010 Nov;47(11):1451-1458. [doi: 10.1016/j.ijnurstu.2010.06.004] [Medline: 20598692]

Abbreviations

PMT: protection motivation theory
QR: Quick Response
UTAUT: unified theory of acceptance and use of technology

Edited by L Buis; submitted 07.12.20; peer-reviewed by N Wessels, A Mahnke; comments to author 08.01.21; revised version received 26.01.21; accepted 15.07.21; published 08.09.21.

Please cite as:
Tretiakov A, Hunter I
User Experiences of the NZ COVID Tracer App in New Zealand: Thematic Analysis of Interviews
JMIR Mhealth Uhealth 2021;9(9):e26318
URL: https://mhealth.jmir.org/2021/9/e26318
doi:10.2196/26318
PMID:34292868

Original Paper

Diagnostic Accuracy of Smartphone-Based Audiometry for Hearing Loss Detection: Meta-analysis

Chih-Hao Chen[1], MD; Heng-Yu Haley Lin[2*], MPH, MD; Mao-Che Wang[1,3,4*], MD, PhD; Yuan-Chia Chu[5,6,7], PhD; Chun-Yu Chang[8], MD; Chii-Yuan Huang[1,3], MD, PhD; Yen-Fu Cheng[1,3,9,10], MD, PhD

[1]Department of Otolaryngology-Head and Neck Surgery, Taipei Veterans General Hospital, Taiwan, Taipei City, Taiwan
[2]Department of Medical Education, Taipei Veterans General Hospital, Taipei City, Taiwan
[3]Faculty of Medicine, National Yang Ming Chiao Tung University, Taipei City, Taiwan
[4]Institute of Hospital and Health Care Administration, National Yang Ming Chiao Tung University, Taipei City, Taiwan
[5]Information Management Office, Taipei Veterans General Hospital, Taipei City, Taiwan
[6]Medical AI Development Center, Taipei Veterans General Hospital, Taipei City, Taiwan
[7]Department of Information Management, National Taipei University of Nursing and Health Sciences, Taipei City, Taiwan
[8]Department of Anesthesiology, Taipei Tzu Chi Hospital, Buddhist Tzu Chi Medical Foundation, New Taipei City, Taiwan
[9]Department of Medical Research, Taipei Veterans General Hospital, Taipei City, Taiwan
[10]Institute of Brain Science, National Yang Ming Chiao Tung University, Taipei City, Taiwan
[*]these authors contributed equally

Corresponding Author:
Yen-Fu Cheng, MD, PhD
Department of Otolaryngology-Head and Neck Surgery
Taipei Veterans General Hospital
Taiwan
No.201, Sec. 2, Shipai Rd., Beitou District
Taipei City, 11217
Taiwan
Phone: 886 2 2871 2121 ext 1292
Email: yfcheng2@vghtpe.gov.tw

Abstract

Background: Hearing loss is one of the most common disabilities worldwide and affects both individual and public health. Pure tone audiometry (PTA) is the gold standard for hearing assessment, but it is often not available in many settings, given its high cost and demand for human resources. Smartphone-based audiometry may be equally effective and can improve access to adequate hearing evaluations.

Objective: The aim of this systematic review is to synthesize the current evidence of the role of smartphone-based audiometry in hearing assessments and further explore the factors that influence its diagnostic accuracy.

Methods: Five databases—PubMed, Embase, Cochrane Library, Web of Science, and Scopus—were queried to identify original studies that examined the diagnostic accuracy of hearing loss measurement using smartphone-based devices with conventional PTA as a reference test. A bivariate random-effects meta-analysis was performed to estimate the pooled sensitivity and specificity. The factors associated with diagnostic accuracy were identified using a bivariate meta-regression model. Study quality was assessed using the Quality Assessment of Diagnostic Accuracy Studies-2 tool.

Results: In all, 25 studies with a total of 4470 patients were included in the meta-analysis. The overall sensitivity, specificity, and area under the receiver operating characteristic curve for smartphone-based audiometry were 89% (95% CI 83%-93%), 93% (95% CI 87%-97%), and 0.96 (95% CI 0.93-0.97), respectively; the corresponding values for the smartphone-based speech recognition test were 91% (95% CI 86%-94%), 88% (95% CI 75%-94%), and 0.93 (95% CI 0.90-0.95), respectively. Meta-regression analysis revealed that patient age, equipment used, and the presence of soundproof booths were significantly related to diagnostic accuracy.

Conclusions: We have presented comprehensive evidence regarding the effectiveness of smartphone-based tests in diagnosing hearing loss. Smartphone-based audiometry may serve as an accurate and accessible approach to hearing evaluations, especially in settings where conventional PTA is unavailable.

(JMIR Mhealth Uhealth 2021;9(9):e28378) doi:10.2196/28378

KEYWORDS

audiometry; hearing loss; hearing test; mhealth; mobile health; digital health; meta-analysis; mobile phone; smartphone diagnostic test accuracy

Introduction

Background

Hearing loss is one of the most common disabilities affecting both individual and public health. Hearing loss has been linked to multiple physical [1,2], cognitive [3,4], and psychosocial [5,6] outcomes and is associated with problematic health care use and higher medical expenses [7]. According to previous studies and World Health Organization estimates, more than 5% of the world's population is affected by hearing impairment, especially older adults aged above 65 years [8-10]. Notably, the prevalence of hearing loss is 50% higher in low-income countries [11]. Within the disease spectrum of hearing impairment, a considerable number of cases, such as those involving idiopathic sudden sensorineural hearing loss (SSNHL) and noise-induced hearing loss, are preventable and can be treated effectively and in a timely manner [12-14].

Pure tone audiometry (PTA) is the gold standard for current hearing assessment batteries [15]. However, this measurement is often unavailable, given its demanding nature with regard to equipment, certified personnel, space, and expenses, particularly in settings such as primary care practices, urgent care, and in low- and middle-income countries [16-18]. As hearing loss has been identified as the single largest potentially modifiable risk factor for dementia in midlife [19] and most patients with hearing impairment can benefit from timely interventions, a more accessible and equally accurate approach to hearing assessment is warranted. Great efforts have been made to create more cost-effective devices and automate audiologic examinations, resulting in the rapid development of smartphone audiometry. Because of the universal availability of mobile technology and cellular networks, smartphone-based hearing tests may provide an adequate assessment of hearing as an alternative to conventional PTA and assist large-scale hearing screening [16,20,21].

Objective

A considerable number of smartphone apps have been introduced for hearing screening [22,23], evaluation [24-26], and even rehabilitation and care [27,28] in recent years, and previous research has compared the performance of these apps with standard audiometry [21,29]. However, these studies were heterogeneous in terms of study design, use of equipment, and baseline characteristics of the participants, which resulted in inconsistent data on the diagnostic performance of smartphone audiometry. The aim of this study is to synthesize the most updated and comprehensive evidence of the diagnostic value of smartphone-based hearing assessments for hearing loss. We performed a meta-analysis with meta-regression to summarize the diagnostic accuracy of smartphone audiometry and investigated the factors affecting the test results. We aim to provide more definitive evidence of the utility of smartphone audiometry in clinical application in the future.

Methods

Study Design

This meta-analysis followed the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) Diagnostic Test Accuracy Studies statement [30].

Search Strategy

In all, five databases—PubMed, Embase, Cochrane Library, Web of Science and Scopus—were searched from inception through January 15, 2021, by 2 authors (CHC and HYHL). The Boolean operator *OR* was used to cover similar concepts, whereas *AND* was used to intersect different concepts. We used a combination of Medical Subject Headings and text words to create three subsets of citations: the first included studies on hearing loss (*hearing loss*, *hypoacusis*, and *hearing impairment*), the second included studies on smartphones (*smartphone*, *cellular phone*, *mobile*, and *mobile phone*), and the third included studies on the concept of use (*diagnosis*, *audiometry*, and *self-examination*). The detailed search strategy is presented in Multimedia Appendix 1. The identified citations were imported into the reference software and screened by title, abstract, and keyword. Potentially eligible records were then subjected to a full-text review.

Eligibility Criteria

The included studies were selected based on the following criteria: (1) PTA was used as a reference test, (2) audiometry was used on smart devices (ie, PTA and speech recognition audiometry) as an index test, and (3) adequate information was reported on diagnostic accuracy (ie, prevalence, sensitivity, and specificity) to quantify the effect estimates for meta-analysis. Studies with outcomes that did not relate to the diagnostic accuracy of the index test or did not provide enough information for meta-analysis were excluded. We did not exclude studies based on country, language, or publication date.

Study Selection and Data Extraction

All studies were fully reviewed and selected by 2 authors (CHC and HYHL). If there were any disagreements in the study selection, they were resolved by a third author (YFC) through consensus or discussion. The extracted data included the author's name, publication year, country, test setting, number of patients, mean age of the study population, operating system of the smart device, equipment used during the examination, and use of a soundproof booth. The disease population was defined as comprising patients with abnormal reference test results in each study. The quantitative data were either extracted directly from raw data or converted from the diagnostic parameters (ie, sensitivity, specificity, and prevalence) in each study to construct

standard diagnostic test 2×2 tables containing true-positive, false-positive, false-negative, and true-negative samples for the index text.

Study Quality Assessment

The quality of the included studies was assessed by 2 authors (CHC and HYHL) using the Quality Assessment of Diagnostic Accuracy Studies-2 tool. A third reviewer (YFC) resolved disagreements regarding the methodological quality through consensus or discussion.

Statistical Analysis

Overview

Sensitivity and specificity were calculated for each extracted data set. A negative correlation between sensitivity and specificity caused by different thresholds was observed; therefore, we adopted a bivariate random-effects model to estimate the pooled sensitivity and specificity of the index test and to account for the heterogeneity that commonly exists in meta-analyses of diagnostic accuracy tests [31]. The bivariate random-effects model assumes logit-transformed sensitivity and specificity as bivariable distributions, and it also considers the threshold effect, which is an indication of the trade-off phenomenon in most diagnostic accuracy tests because the threshold differs among studies [32]. To investigate the covariate among the index studies, bivariate meta-regression analysis was performed [33], one at a time. For the covariate effect on age, we divided the studies into child, elderly, and adult groups. People aged below 18 years were considered to be in the child group, whereas people aged above 65 years were considered to be in the elderly group based on the World Health Organization criteria [34]. First, we examined whether the covariate caused variance in the sensitivity and specificity measures. The following likelihood-ratio chi-square test was used to determine whether the covariate served as a significant variable by testing the hypothesis that these covariates do not explain variance in the logit-transformed pairs of sensitivity and specificity. To further illustrate the diagnostic accuracy and compare the discriminatory properties, we constructed hierarchical summary receiver operating characteristic curves for the overall result as well as the subgroup results identified by the meta-regression analysis by accounting for the correlation in the data through a hierarchical approach. To deal with zero observations in the 2×2 contingency tables, 0.5 was added to each cell to reduce the influence of small studies. We calculated 95% CIs on the basis of the binominal distribution of the truly positive and truly negative samples. Publication bias was examined using the Deeks funnel plot using the natural logarithm of the diagnostic odds ratio against 1/(effective sample size)$^{1/2}$ to plot the asymmetry of the included studies. Effective sample size (ESS) was calculated by the number of examinees who were diseased (n1) and not diseased (n2) as:

$$ESS = (4n1\ n2) / (n1 + n2) \quad \textbf{(1)}$$

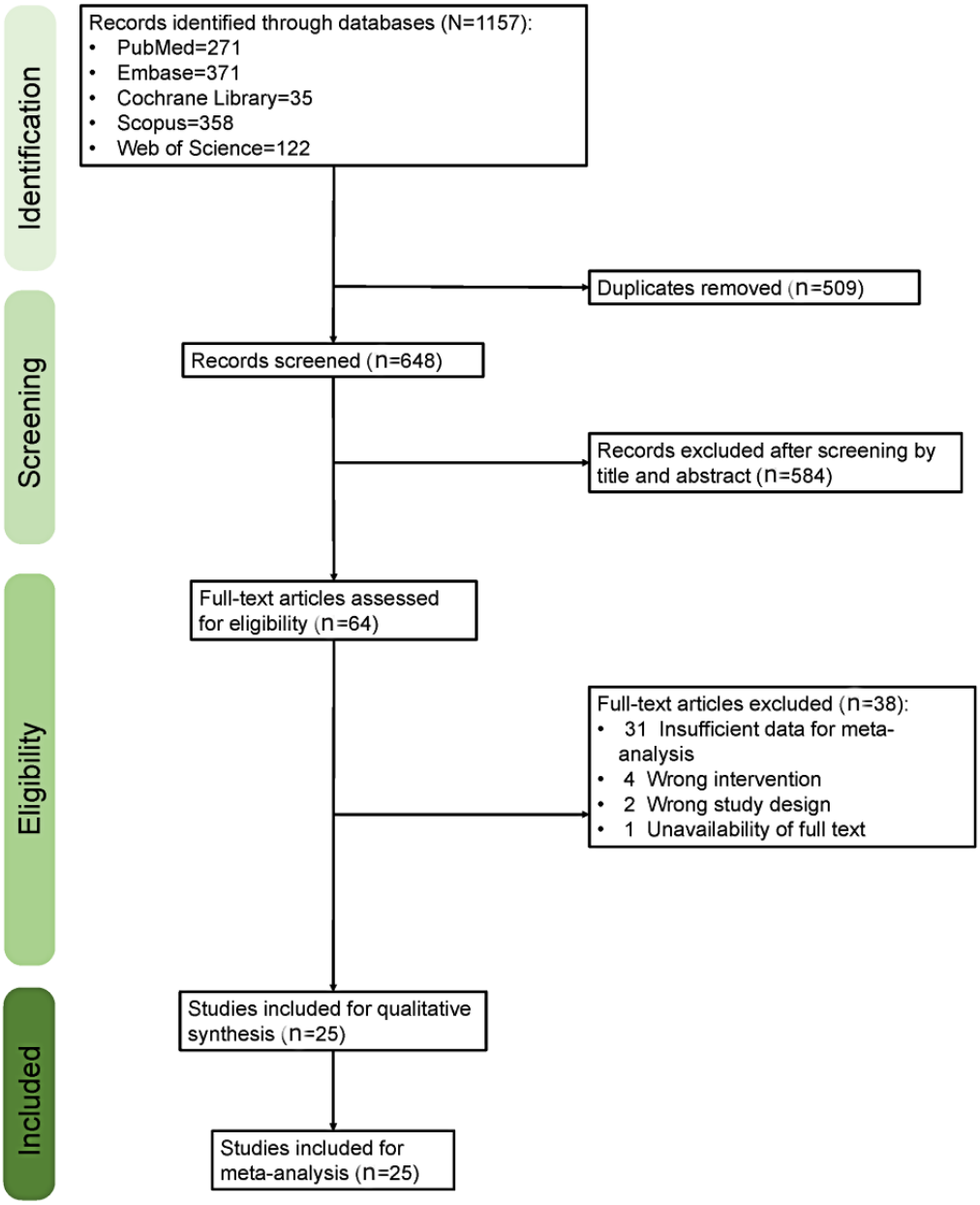
ESS considers that unequal numbers of individuals who are diseased and not diseased reduce the precision of test accuracy estimates [31,35]. A $P<.10$ for the regression tests suggests significant publication bias. Statistical analyses were conducted using Stata version 15.0 (StataCorp), with the midas and metandi commands. All statistical tests were two-sided, and $P<.05$ was considered statistically significant.

Results

Study Identification and Selection

A total of 1157 studies were identified through the databases. Of the 1157 studies, 648 (56%) remained in the preliminary search after the removal of 509 (44%) duplicates. Of the 648 studies, 584 (90.1%) were excluded after 2 authors (CHC and HYHL) screened the titles and abstracts; a total of 9.9% (64/648) of studies then underwent full-text review. Of the 64 studies, 39 (61%) were excluded because of the following reasons: insufficient data for meta-analysis, index tests not used, inappropriate study design, or unavailability of the full text. As a result, of the 64 studies, 25 (39%) studies with a total of 4470 patients were included in the meta-analysis. The detailed PRISMA flow diagram is presented in Figure 1.

Figure 1. The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analyses) flow diagram.



Study Characteristics

Of the 25 studies, 21 were prospective [10,21,22,29,36-52], 1 was retrospective [53], and the remaining 3 studies did not report the study design [23,54,55]. In all, 20 studies used PTA as the index test [10,21,22,29,36-42,44-46,48-52,54], whereas the remaining 5 studies applied a speech recognition test (SRT) as the index test [23,43,47,53,55]. A total of 4 studies enrolled elderly participants [10,36,37,39], whereas 7 studies included children [21-23,38,41,49,55], and 13 studies enrolled adult participants [29,40,42-44,46,48-54]. The remaining study did not report the age of the study population [45]. In all, 15 studies operated audiometry through an iPhone (Apple Inc) operating system–based app [10,22,29,36,37,39,40,45-48,50,52,54,55], whereas the remaining 10 used an Android (Google LLC) operating system–based audiometry app [21,23,38,41-44,49,51,53]. A total of 15 studies used headphones for testing [10,21,23,38,41-47,49,50,52,54], 9 studies used earphones for the examination [22,29,36,37,39,40,48,53,55], and 1 study did not mention the equipment used [51]. In all, 6 studies conducted the examination in a soundproof booth [10,49-51,53,55], 18 studies did not use a soundproof booth to conduct the examination [21-23,29,36-47,52,54], and 1 study did not report whether the test was conducted in a soundproof booth [48]. A total of 4 studies [45,46,49,50] conducted the index test among different independent populations, yielding a total of 30 study groups for the analysis. Further information regarding the included study populations and statistics is presented in Multimedia Appendices 2 and 3.

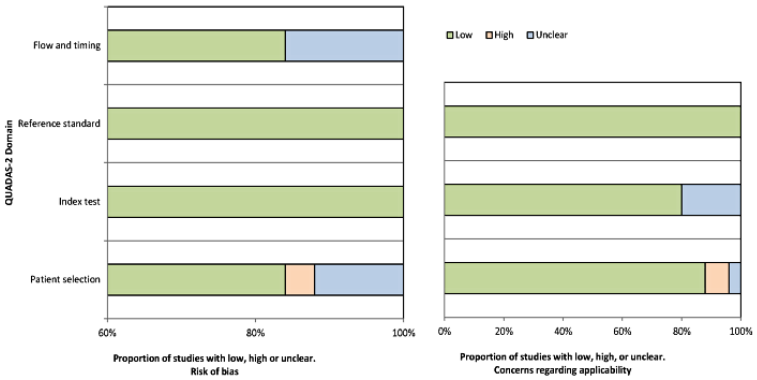
Quality and Risk-of-Bias Assessment

Quality Assessment of Diagnostic Accuracy Studies-2 scores were used to evaluate the quality of the included studies. Regarding the evaluation of the risk of bias, all the studies carried out index studies without knowing the results of the reference test in advance and set the threshold before testing. A total of 4 studies did not clearly describe the sequence

between the index and reference tests [47,53-55]. Regarding the evaluation of applicability, 1 study enrolled patients with underlying otitis media [38], and another 2 studies included patients with SSNHL [29,52]. In all, 5 studies used unmarketed apps as index tests. A detailed assessment and an overall picture of the methodological quality of the included studies are presented in Figure 2.

Figure 2. Quality assessment results based on the Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) guidelines.

	Risk of bias				Applicability		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Abu-Ghanem et al [36]	Low	Low	Low	Low	Low	Low	Low
Anuar et al [48]	Low	Low	Low	Low	Low	Low	Low
Aremu et al [51]	Low	Low	Low	Low	Low	Low	Low
Bauer et al [45]	Low	Low	Low	Low	Low	Low	Low
Chu et al [22]	Low	Low	Low	Low	Low	Low	Low
Corona et al [49]	Low	Low	Low	Low	Low	Low	Low
Derin et al [54]	Unclear	Low	Low	Unclear	Low	Low	Low
Durgut et al [38]	Low	Low	Low	Low	High	Low	Low
Handzel et al [29]	Low	Low	Low	Low	High	Low	Low
Kelly et al [46]	Low	Low	Low	Low	Low	Low	Low
Li et al [10]	Low	Low	Low	Low	Low	Low	Low
Lin et al [52]	Low	Low	Low	Low	Unclear	Low	Low
Livshirz et al [37]	Low	Low	Low	Low	Low	Low	Low
Louw et al [42]	Low	Low	Low	Low	Low	Low	Low
Lycke et al [39]	Low	Low	Low	Low	Low	Low	Low
Mahomed-Asmail et al [41]	Low	Low	Low	Low	Low	Low	Low
Peer et al [40]	Low	Low	Low	Low	Low	Low	Low
Potgieter et al [43]	Low	Low	Low	Low	Low	Unclear	Low
Potgieter et al [53]	High	Low	Low	Unclear	Low	Unclear	Low
Saliba et al [50]	Low	Low	Low	Low	Low	Low	Low
Samlli et al [47]	Unclear	Low	Low	Unclear	Low	Unclear	Low
Sandström et al [44]	Low	Low	Low	Low	Low	Low	Low
Swanepoel et al [21]	Low	Low	Low	Low	Low	Low	Low
Xiao et al [55]	Unclear	Low	Low	Unclear	Low	Unclear	Low
Yimtae et al [23]	Low	Low	Low	Low	Low	Unclear	Low



Overall Diagnostic Performance

Overall, the studies using a smartphone app with PTA showed a sensitivity of 89% (95% CI 83%-93%) and specificity of 93% (95% CI 87%-97%), whereas studies using an app involving SRT revealed a sensitivity of 91% (95% CI 86%-94%) and specificity of 88% (95% CI 75%-94%). The hierarchical summary receiver operating characteristic curves with summary points for both PTA and SRT are shown in Figures 3 and 4. The predicted values for the area under the receiver operating characteristic curve (AUC) for the PTA and SRT measures were 0.96 (95% CI 0.93-0.97) and 0.93 (95% CI 0.90-0.95), respectively.

Figure 3. The HSROC for pure tone audiometry. HSROC: hierarchical summary receiver operating characteristic.

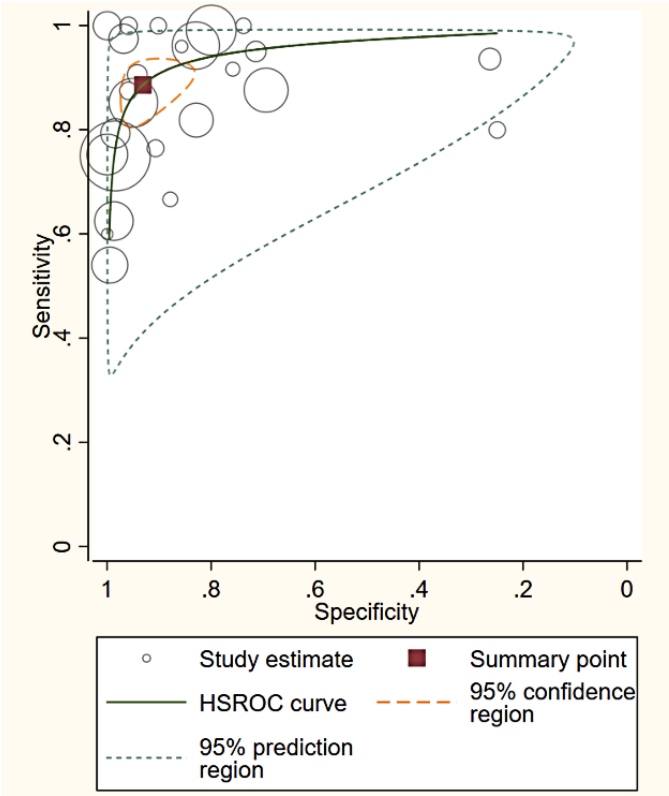
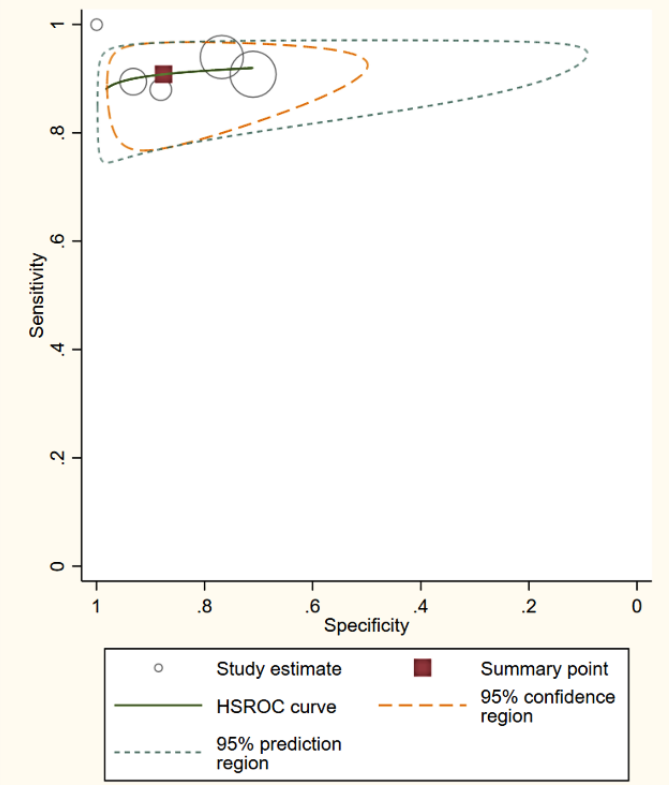


Figure 4. The HSROC for the speech recognition test. HSROC: hierarchical summary receiver operating characteristic.



Meta-Regression and Subgroup Analysis

The bivariate meta-regression analysis showed a significant influence of the operating system on sensitivity (88% vs 89%). The likelihood-ratio chi-square test revealed that elderly group (χ^2_1=85.9; *P*<.001), child group (χ^2_1=62.9; *P*<.001), headphone use (χ^2_1=17.8; *P*<.001), and soundproof booth use (χ^2_1=19.5; *P*<.001) were significant covariates causing variance between paired sensitivity and specificity, whereas the operating system did not reveal such a difference (χ^2_1=0.02; *P*=.99). The AUC values for the elderly group versus the adult group were 0.90

(95% CI 0.87-0.92) versus 0.96 (95% CI 0.94-0.97), respectively, whereas the AUC values for the child group versus the adult group were 0.90 (95% CI 0.88-0.93) versus 0.96 (95% CI 0.94-0.97), respectively. The AUC values for the headphone group versus the earphone group were 0.96 (95% CI 0.94-0.97) versus 0.92 (95% CI 0.89-0.94), respectively. The AUC values for the soundproof booth group versus the non–soundproof booth group were 0.99 (95% CI 0.97-0.99) versus 0.94 (95% CI 0.91-0.96), respectively. The AUC values for the iPhone operating system group versus the Android operating system group were 0.95 (95% CI 0.93-0.97) versus 0.96 (95% CI 0.94-0.97), respectively. The detailed results are presented in Table 1.

Table 1. Results of the bivariate meta-regression analysis (N=25).

Covariate	Number	Sensitivity (95% CI)	*P* value	Specificity (95% CI)	*P* value	Likelihood-ratio test	Chi-square (*df*)	Area under the curve (95% CI)
Age								
Elderly [10,36,37,39]	4	0.77 (0.55-0.99)	.04	0.92 (0.80-1.00)	.99	<.001[a]	85.9 (1)	0.90 (0.87-0.92)
Child [21,22,38,41,49]	5	0.85 (0.69-1.00)	.10	0.96 (0.89-1.00)	.34	<.001	62.9(1)	0.90 (0.88-0.93)
Adult [29,40,42,44,46,48-52,54]	14	0.90 (0.85-0.96)	—[b]	0.91 (0.82-1.00)	—	—	—	0.96 (0.94-0.97)
Operating system								
iPhone operating system [10,22,29,36,37,39,40,45,46,48,50,52,54]	17	0.88 (0.82-0.94)	.04	0.93 (0.87-0.99)	.51	.99	0.02(1)	0.95 (0.93-0.97)
Android [21,38,41,42,44,49,51]	8	0.89 (0.81-0.97)	—	0.93 (0.85-1.00)	—	—	—	0.96 (0.94-0.97)
Equipment								
Headphone [10,21,38,41,42,44-47,49,50,52,54]	17	0.91 (0.87-0.95)	.85	0.89 (0.82-0.97)	.05	<.001	17.8(1)	0.96 (0.94-0.97)
Earphone [22,29,36,37,39,40,48]	7	0.80 (0.65-0.95)	—	0.97 (0.92-1.00)	—	—	—	0.92 (0.89-0.94)
Soundproof booth								
Yes [10,49-51]	6	0.95 (0.90-1.00)	.72	0.95 (0.87-1.00)	.83	<.001	19.5(1)	0.99 (0.97-0.99)
No [21,22,29,36-42,44-46,52,54]	18	0.87 (0.82-0.93)	—	0.91 (0.85-0.98)	—	—	—	0.94 (0.91-0.96)

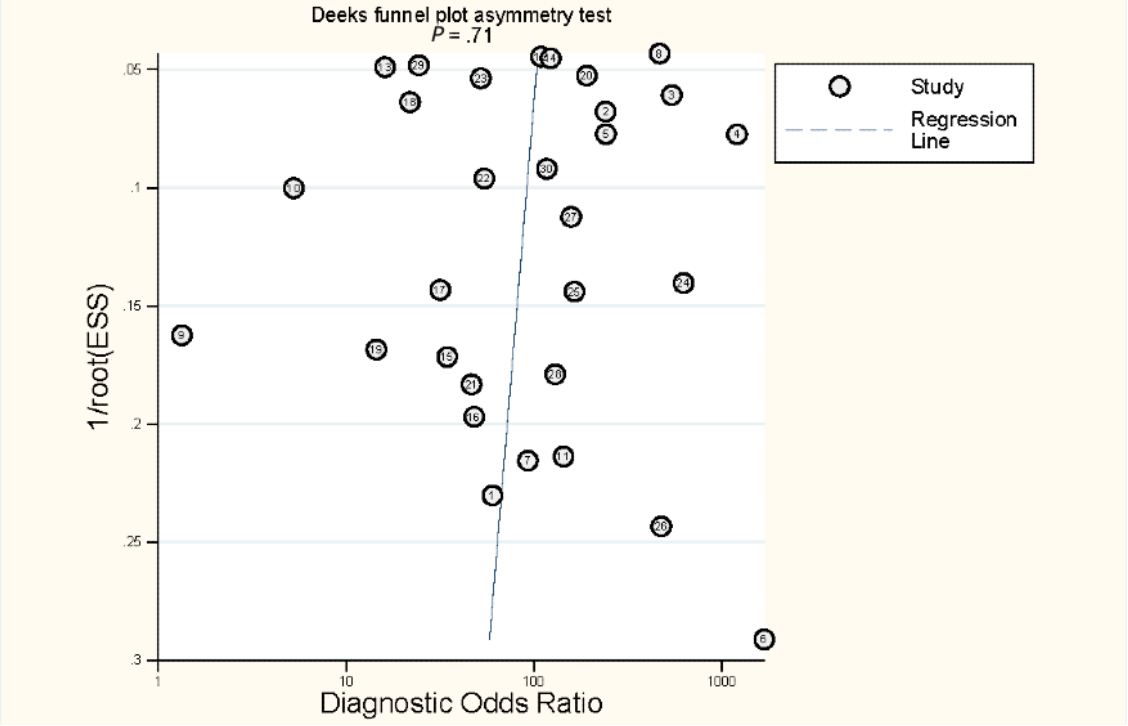
[a]Significant *P*<.05.

[b]Reference of likelihood-ratio chi-square test.

Publication Bias

The Deeks funnel plot revealed no asymmetrical distribution for the included studies, and the regression test did not show a significant publication bias (*P*=.71; Figure 5).

Figure 5. The Deeks funnel plot. ESS: effective sample size.



Discussion

Principal Findings

In this study, we performed a meta-analysis to estimate the pooled diagnostic accuracy of smartphone-based hearing tests using conventional PTA as the gold standard. The overall sensitivity of smartphone-based audiometry was 89%, the specificity was 93%, and the AUC was 0.95, which suggested outstanding diagnostic performance for identifying hearing loss using PTA as the gold standard test. When using the SRT as the gold standard test, our results showed a sensitivity of 91%, specificity of 88%, and AUC of 0.93, which also indicated excellent diagnostic accuracy. On the basis of the results of the bivariate meta-regression analysis, we found that participant age, equipment used, and the use of a soundproof booth significantly affected the diagnostic accuracy of smartphone-based audiometry, whereas the operating system of the smartphone did not. To our knowledge, this is the first meta-analysis that provides comprehensive evidence of the diagnostic performance of a smartphone-based approach to detecting hearing loss.

PTA assesses a person's lowest threshold response to pure tone stimuli at various frequencies [56]. It is still considered the gold standard test for audiologic examinations and provides information regarding the severity and type of hearing loss. According to the American National Standards Institute specifications, there are four types of PTA. Type 1 audiometry (advanced clinical or research) involves a completely equipped audiometer that can conduct both air and bone conduction tests. Type 2 (clinical) fits the same specifications as type 1, except for the requirement of loudspeaker equipment. Portable audiometers without speech-comprehension measurements are classified as type 3 (diagnostic), whereas type 4 (screening) consists of screening audiometers with the basic functions of a hearing test [56]. Although types 1 and 2 are considered the most informative and comprehensive audiometry, they are often not available in many settings, especially in resource-limited areas such as in low- and middle-income countries and rural regions. Even in resource-rich countries, standard PTA is not usually available at primary care practices [17]. Standard PTA tests require certified professionals to administer them, whereas audiologic training is generally lacking in resource-limited countries—there is less than one audiologist for every 1 million people according to previous studies [23,53,54]. Furthermore, the equipment for conventional PTA, including a soundproof booth and a calibrated audiometer, involves both cost and space. The demanding nature of conventional PTA may result in its low accessibility and further affect the generality of hearing screening and quality of hearing care [57,58].

In recent years, mobile health devices have evolved rapidly, as have smartphone-based hearing approaches. Smartphone-based audiometry is a cost-effective, convenient, and reliable tool for screening hearing loss. As smartphones are common in the modern society and the apps are very accessible, given their low cost or no cost, smartphone-based hearing tests could potentially bridge the gap between patients with hearing loss and adequate audiologic assessments and, potentially, hearing care. Previous studies have confirmed that such apps were able to provide basic hearing screening wherever the individual was located as long as the location met the required level of background noise, reducing the need to travel and pay for a hearing examination [59,60]. These smartphone-based hearing tests are usually designed to be user friendly because automated diagnostic audiometry simplifies complex audiologic protocols, allowing their use by nonprofessionals [61,62]. Studies have also described the use of smartphone-based audiometry in settings such as primary care practices and community health clinics for routine hearing screening to identify potentially

handicapping hearing loss [59]. The findings of this study confirm that the diagnostic performance of smartphone-based audiometry aligns perfectly with conventional PTA in identifying hearing loss and adds to previous research with a larger pooled sample size and systematic scope.

Although this study highlights the high diagnostic value of smartphone-based hearing tests and their promising role in hearing screening, we identified several possible variables that may influence diagnostic performance. First, the accuracy of smartphone-based audiometry was lower in elderly individuals and children. This may suggest technical barriers between smart devices and elderly individuals and children. Previous studies have found that factors such as prevalent vision impairments and slower learning curves in managing technological devices because of lack of experience and functional decline may contribute to the higher level of difficulties when using smartphone-based apps among elderly populations [60,63]. At the same time, a previous study also found that children achieved lower accuracy in PTA [64]. Our results also showed that headphone use during the hearing examination may improve the diagnostic accuracy of smartphone-based audiometry. Earphones are a required component in standard audiometry because they prevent the collapse of the external ear canal and reduce the level of ambient noise [65-67]. However, if the participant does not insert the earphone correctly in the automated examination, it could be a problem. A previous study showed that earphone positioning may affect audiologic assessment results and whether the earphone is positioned by a trained examiner or by the examinee may affect audiologic assessment [68]. The negative effects of background noise may further support our finding that examinations conducted in soundproof booths have better diagnostic accuracy. The influence of ambient noise, which results in erroneous test results of smartphone audiometry, has been reported in previous studies [69,70], leading to the conclusion that the use of soundproof booths may increase the diagnostic accuracy of smartphone-based hearing tests [71]. Although some of the included studies reported comparable results of hearing assessments outside of a soundproof booth with passive attenuation and simultaneous ambient noise monitoring [71-73], most of the studies did not provide information regarding the management of ambient noise. The diagnostic value of this subgroup, however, still appeared feasible, because their AUC values exceeded the cutoff point of 0.9 [74]. In summary, our findings suggest that adequate adjustment of the variables that significantly affect the accuracy of smartphone-based audiometry may improve its diagnostic performance in diagnosing hearing loss. Approaches such as adding instructions regarding the examination protocol and correct use of earphones, providing customized audiologist consultations for elderly individuals, improving the app's function in monitoring environmental noise, and regularly collecting feedback from users could be added to the current implementation methods.

Limitations

This study has several limitations. First, as in most studies of diagnostic test accuracy, different thresholds exist among the studies and may have caused the threshold effect. A prior test calculation of the correlation between sensitivity and specificity revealed a negative result, confirming the threshold effect in this study. Therefore, we adopted the bivariate random-effects model to account for the cross-study threshold difference as suggested by previous studies [31,32]. Second, there was heterogeneity regarding the study designs, test protocols, and reference PTA thresholds for diagnosing hearing loss across the included studies, which may have biased the results when pooling them into the meta-analyses. Future studies with homogenous gold standards and uniform protocols for smartphone-based hearing tests are needed. Third, ambient noise monitoring is a key factor influencing the accuracy of audiometry [75]. Although most of the included studies did monitor noise, no data on the accuracy without ambient noise monitoring were provided. As a result, we were not able to perform the meta-regression analysis according to this factor. Fourth, frequency may act as a confounder, but most of the included studies did not provide diagnostic accuracy for each frequency; therefore, we could evaluate the diagnostic performance of smartphone audiometry only with the average threshold calculated from the frequencies. Fifth, most of the included studies did not describe the masking procedure, possibly because the included studies sampled healthy people, and the threshold difference between bilateral ears could hardly exceed 40 dB. In addition, some smartphone audiometry methods did not provide an automasking procedure during the automated examination. We suggest that future studies describe the masking procedure in detail, regardless of whether it is used. Sixth, of the 25 included studies, most did not describe the calibration method, whereas 9 (36%) used reference equivalent threshold sound pressure levels. A previous study revealed that the differences in hearing thresholds among the device models were significant, which might directly result from the biological calibration method used to determine the reference sound level [75]. Calibration information was lacking, possibly because of the intrinsic lack of a calibration function in the app. We suggest that future studies address this issue. Finally, some included studies enrolled patients with underlying diseases such as otitis media and SSNHL. Although, ideally, subgroup analyses should have been performed for these unique studies for more accurate results, we were not able to implement this investigation because of the scarcity of relevant studies. We look forward to more studies that investigate the value of smartphone audiometry in identifying different types of hearing loss in the future because they can provide more solid and specific evidence for apps in different clinical settings.

Conclusions

In this meta-analysis, we have provided comprehensive evidence regarding the diagnostic performance of smartphone-based audiometry in diagnosing hearing loss. Given the high sensitivity and specificity of smartphone-based audiometry, along with its low cost and high accessibility, smartphone-based hearing assessments may serve as a cost-effective and equally accurate diagnostic tool, in comparison with conventional PTA, for assessing hearing loss, especially in resource-limited settings where conventional PTA is not feasible. Our findings also suggest that future improvements in smartphone-based audiometry should focus on adjusting the potential factors that may affect its diagnostic accuracy.

Acknowledgments

This work was supported by the Taipei Veterans General Hospital (V108C-145, V109C-135, V109E-008-5, and V110E-003-2) and the Ministry of Science and Technology (MOST109-2320-B075-00, MOST109-2314-B075-014-MY2, and MOST109-2320-B-075-006).

Authors' Contributions

Both YFC and CYH are the corresponding authors of this paper and have contributed equally to this work. CHC and YFC were responsible for data acquisition. CHC, HYHL, YCC, and CYH were responsible for the analysis and interpretation of the data. CHC, CYH, and HYHL drafted the manuscript. CHC, HYHL, and YCC performed the statistical analyses. YFC and CYH obtained the funding, and YCC, CYC, and MCW were responsible for administrative, technical, and material support. YFC and CYH supervised the study. All authors were responsible for the study concept and design, critical revision of the manuscript for important intellectual content, and approval of the final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1
The detailed search strategy.
[DOCX File , 22 KB - mhealth_v9i9e28378_app1.docx]

Multimedia Appendix 2
Table of study characteristics.
[DOCX File , 21 KB - mhealth_v9i9e28378_app2.docx]

Multimedia Appendix 3
Table of study diagnostic parameters.
[DOCX File , 21 KB - mhealth_v9i9e28378_app3.docx]

References

1. Deal J, Richey Sharrett A, Bandeen-Roche K, Kritchevsky S, Pompeii L, Gwen Windham B, et al. Hearing impairment and physical function and falls in the atherosclerosis risk in communities hearing pilot study. J Am Geriatr Soc 2016 Apr;64(4):906-908 [FREE Full text] [doi: 10.1111/jgs.14075] [Medline: 27100600]
2. Lin FR, Metter EJ, O'Brien RJ, Resnick SM, Zonderman AB, Ferrucci L. Hearing loss and incident dementia. Arch Neurol 2011 Feb;68(2):214-220 [FREE Full text] [doi: 10.1001/archneurol.2010.362] [Medline: 21320988]
3. Deal JA, Sharrett AR, Albert MS, Coresh J, Mosley TH, Knopman D, et al. Hearing impairment and cognitive decline: a pilot study conducted within the atherosclerosis risk in communities neurocognitive study. Am J Epidemiol 2015 May 01;181(9):680-690 [FREE Full text] [doi: 10.1093/aje/kwu333] [Medline: 25841870]
4. Shukla A, Reed N, Armstrong N, Lin F, Deal J, Goman A. Hearing loss, hearing aid use, and depressive symptoms in older adults-findings from the Atherosclerosis Risk in Communities Neurocognitive Study (ARIC-NCS). J Gerontol B Psychol Sci Soc Sci 2021 Feb 17;76(3):518-523 [FREE Full text] [doi: 10.1093/geronb/gbz128] [Medline: 31628485]
5. Brewster K, Ciarleglio A, Brown P, Chen C, Kim H, Roose S, et al. Age-related hearing loss and its association with depression in later life. Am J Geriatr Psychiatry 2018 Jul;26(7):788-796 [FREE Full text] [doi: 10.1016/j.jagp.2018.04.003] [Medline: 29752060]
6. Dalton DS, Cruickshanks KJ, Klein BE, Klein R, Wiley TL, Nondahl DM. The impact of hearing loss on quality of life in older adults. Gerontologist 2003 Oct;43(5):661-668. [doi: 10.1093/geront/43.5.661] [Medline: 14570962]
7. Reed NS, Altan A, Deal JA, Yeh C, Kravetz AD, Wallhagen M, et al. Trends in health care costs and utilization associated with untreated hearing loss over 10 years. JAMA Otolaryngol Head Neck Surg 2019 Jan 01;145(1):27-34 [FREE Full text] [doi: 10.1001/jamaoto.2018.2875] [Medline: 30419131]
8. WHO global estimates on prevalence of hearing loss. World Health Organization. 2012. URL: https://www.who.int/pbd/deafness/WHO_GE_HL.pdf [accessed 2021-08-09]
9. Stevens G, Flaxman S, Brunskill E, Mascarenhas M, Mathers C, Finucane M, Global Burden of Disease Hearing Loss Expert Group. Global and regional hearing impairment prevalence: an analysis of 42 studies in 29 countries. Eur J Public Health 2013 Feb;23(1):146-152. [doi: 10.1093/eurpub/ckr176] [Medline: 22197756]
10. Li LY, Wang S, Wu C, Tsai C, Wu T, Lin Y. Screening for hearing impairment in older adults by smartphone-based audiometry, self-perception, hhie screening questionnaire, and free-field voice test: comparative evaluation of the screening

accuracy with standard pure-tone audiometry. JMIR Mhealth Uhealth 2020 Oct 27;8(10):e17213 [FREE Full text] [doi: 10.2196/17213] [Medline: 33107828]
11. Prescott C, Omoding S, Fermor J, Ogilvy D. An evaluation of the 'voice test' as a method for assessing hearing in children with particular reference to the situation in developing countries. Int J Pediatr Otorhinolaryngol 1999 Dec 15;51(3):165-170 [FREE Full text] [doi: 10.1016/s0165-5876(99)00263-3] [Medline: 10628542]
12. Kuhn M, Heman-Ackah SE, Shaikh JA, Roehm PC. Sudden sensorineural hearing loss: a review of diagnosis, treatment, and prognosis. Trends Amplif 2011 Sep;15(3):91-105 [FREE Full text] [doi: 10.1177/1084713811408349] [Medline: 21606048]
13. Conlin AE, Parnes LS. Treatment of sudden sensorineural hearing loss: I. A systematic review. Arch Otolaryngol Head Neck Surg 2007 Jun;133(6):573-581. [doi: 10.1001/archotol.133.6.573] [Medline: 17576908]
14. Daniel E. Noise and hearing loss: a review. J Sch Health 2007 May;77(5):225-231. [doi: 10.1111/j.1746-1561.2007.00197.x] [Medline: 17430434]
15. Sindhusake D, Mitchell P, Smith W, Golding M, Newall P, Hartley D, et al. Validation of self-reported hearing loss. The Blue Mountains Hearing Study. Int J Epidemiol 2001 Dec;30(6):1371-1378. [doi: 10.1093/ije/30.6.1371] [Medline: 11821349]
16. Swanepoel DW, Clark JL, Koekemoer D, Hall JW, Krumm M, Ferrari DV, et al. Telehealth in audiology: the need and potential to reach underserved communities. Int J Audiol 2010 Mar;49(3):195-202. [doi: 10.3109/14992020903470783] [Medline: 20151929]
17. Flaegel K, Brandt B, Goetz K, Steinhaeuser J. Which procedures are performed by general internists practicing primary care in Germany? - a cross-sectional study. BMC Fam Pract 2020 Apr 29;21(1):73 [FREE Full text] [doi: 10.1186/s12875-020-01136-7] [Medline: 32349681]
18. Louw C, Swanepoel D, Eikelboom R. Self-reported hearing loss and pure tone audiometry for screening in primary health care clinics. J Prim Care Community Health 2018;9:2150132718803156 [FREE Full text] [doi: 10.1177/2150132718803156] [Medline: 30278815]
19. Livingston G, Huntley J, Sommerlad A, Ames D, Ballard C, Banerjee S, et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. Lancet 2020 Aug 08;396(10248):413-446 [FREE Full text] [doi: 10.1016/S0140-6736(20)30367-6] [Medline: 32738937]
20. Kelly T, Minges M. Executive summary. In: Information and Communications for Development 2012. Washington, DC: World Bank Group Publication; Aug 2012.
21. Swanepoel DW, Myburgh HC, Howe DM, Mahomed F, Eikelboom RH. Smartphone hearing screening with integrated quality control and data management. Int J Audiol 2014 Dec;53(12):841-849. [doi: 10.3109/14992027.2014.920965] [Medline: 24998412]
22. Chu Y, Cheng Y, Lai Y, Tsao Y, Tu T, Young ST, et al. A mobile phone-based approach for hearing screening of school-age children: cross-sectional validation study. JMIR Mhealth Uhealth 2019 Apr 01;7(4):e12033 [FREE Full text] [doi: 10.2196/12033] [Medline: 30932870]
23. Yimtae K, Israsena P, Thanawirattananit P, Seesutas S, Saibua S, Kasemsiri P, et al. A tablet-based mobile hearing screening system for preschoolers: design and validation study. JMIR Mhealth Uhealth 2018 Oct 23;6(10):e186 [FREE Full text] [doi: 10.2196/mhealth.9560] [Medline: 30355558]
24. Romero RL, Kates F, Hart M, Ojeda A, Meirom I, Hardy S. Quality of deaf and hard-of-hearing mobile apps: evaluation using the Mobile App Rating Scale (MARS) with additional criteria from a content expert. JMIR Mhealth Uhealth 2019 Oct 30;7(10):e14198 [FREE Full text] [doi: 10.2196/14198] [Medline: 31670695]
25. Hollander C, Joubert K, Schellack N. An Ototoxicity Grading System Within a Mobile App (OtoCalc) for a resource-limited setting to guide grading and management of drug-induced hearing loss in patients with drug-resistant tuberculosis: prospective, cross-sectional case series. JMIR Mhealth Uhealth 2020 Jan 14;8(1):e14036 [FREE Full text] [doi: 10.2196/14036] [Medline: 31934875]
26. Maidment DW, Heyes R, Gomez R, Coulson NS, Wharrad H, Ferguson MA. Evaluating a theoretically informed and cocreated mobile health educational intervention for first-time hearing aid users: qualitative interview study. JMIR Mhealth Uhealth 2020 Aug 05;8(8):e17193 [FREE Full text] [doi: 10.2196/17193] [Medline: 32755885]
27. Ratanjee-Vanmali H, Swanepoel DW, Laplante-Lévesque A. Patient uptake, experience, and satisfaction using web-based and face-to-face hearing health services: process evaluation study. J Med Internet Res 2020 Mar 20;22(3):e15875 [FREE Full text] [doi: 10.2196/15875] [Medline: 32196459]
28. Meijerink JF, Pronk M, Lissenberg-Witte BI, Jansen V, Kramer SE. Effectiveness of a web-based SUpport PRogram (SUPR) for hearing aid users aged 50+: two-arm, cluster randomized controlled trial. J Med Internet Res 2020 Sep 22;22(9):e17927 [FREE Full text] [doi: 10.2196/17927] [Medline: 32960175]
29. Handzel O, Ben-Ari O, Damian D, Priel MM, Cohen J, Himmelfarb M. Smartphone-based hearing test as an aid in the initial evaluation of unilateral sudden sensorineural hearing loss. Audiol Neurootol 2013;18(4):201-207. [doi: 10.1159/000349913] [Medline: 23689282]

30. McInnes M, Moher D, Thombs B, McGrath T, Bossuyt P, The PRISMA-DTA Group, et al. Preferred Reporting Items for a Systematic Review and Meta-analysis of diagnostic test accuracy studies: The PRISMA-DTA statement. J Am Med Assoc 2018 Jan 23;319(4):388-396. [doi: 10.1001/jama.2017.19163] [Medline: 29362800]
31. Leeflang MM, Deeks JJ, Takwoingi Y, Macaskill P. Cochrane diagnostic test accuracy reviews. Syst Rev 2013 Oct 07;2:82 [FREE Full text] [doi: 10.1186/2046-4053-2-82] [Medline: 24099098]
32. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 2005 Oct;58(10):982-990. [doi: 10.1016/j.jclinepi.2005.02.022] [Medline: 16168343]
33. Thompson SG, Higgins JP. How should meta-regression analyses be undertaken and interpreted? Stat Med 2002 Jun 15;21(11):1559-1573. [doi: 10.1002/sim.1187] [Medline: 12111920]
34. World Health Organization. Global Age-friendly Cities: A Guide. Geneva, Switzerland: World Health Organization; 2007.
35. van Enst WA, Ochodo E, Scholten RJ, Hooft L, Leeflang MM. Investigation of publication bias in meta-analyses of diagnostic test accuracy: a meta-epidemiological study. BMC Med Res Methodol 2014 May 23;14:70 [FREE Full text] [doi: 10.1186/1471-2288-14-70] [Medline: 24884381]
36. Abu-Ghanem S, Handzel O, Ness L, Ben-Artzi-Blima M, Fait-Ghelbendorf K, Himmelfarb M. Smartphone-based audiometric test for screening hearing loss in the elderly. Eur Arch Otorhinolaryngol 2016 Feb;273(2):333-339. [doi: 10.1007/s00405-015-3533-9] [Medline: 25655259]
37. Livshitz L, Ghanayim R, Kraus C, Farah R, Even-Tov E, Avraham Y, et al. Application-based hearing screening in the elderly population. Ann Otol Rhinol Laryngol 2017 Jan;126(1):36-41. [doi: 10.1177/0003489416672873] [Medline: 27913721]
38. Durgut O, Ekim B, Dikici O, Solmaz F, Ağırgöl B, Özbakan A. Evaluation of hearing thresholds by using a mobile application in children with otitis media with effusion. Audiol Neurootol 2020;25(3):120-124. [doi: 10.1159/000505309] [Medline: 31962328]
39. Lycke M, Debruyne PR, Lefebvre T, Martens E, Ketelaars L, Pottel H, et al. The use of uHear™ to screen for hearing loss in older patients with cancer as part of a comprehensive geriatric assessment. Acta Clin Belg 2018 Apr;73(2):132-138. [doi: 10.1080/17843286.2017.1392070] [Medline: 29063810]
40. Peer S, Fagan JJ. Hearing loss in the developing world: evaluating the iPhone mobile device as a screening tool. S Afr Med J 2015 Jan;105(1):35-39. [doi: 10.7196/samj.8338] [Medline: 26046161]
41. Mahomed-Asmail F, Swanepoel DW, Eikelboom RH, Myburgh HC, Hall J. Clinical validity of hearScreen™ smartphone hearing screening for school children. Ear Hear 2016;37(1):11-17. [doi: 10.1097/AUD.0000000000000223] [Medline: 26372265]
42. Louw C, Swanepoel DW, Eikelboom RH, Myburgh HC. Smartphone-based hearing screening at primary health care clinics. Ear Hear 2017;38(2):93-100. [doi: 10.1097/AUD.0000000000000378] [Medline: 27764002]
43. Potgieter J, Swanepoel DW, Myburgh HC, Smits C. The South African English smartphone digits-in-noise hearing test: effect of age, hearing loss, and speaking competence. Ear Hear 2018;39(4):656-663. [doi: 10.1097/AUD.0000000000000522] [Medline: 29189432]
44. Sandström J, Swanepoel D, Laurent C, Umefjord G, Lundberg T. Accuracy and reliability of smartphone self-test audiometry in community clinics in low income settings: a comparative study. Ann Otol Rhinol Laryngol 2020 Jun;129(6):578-584. [doi: 10.1177/0003489420902162] [Medline: 31965808]
45. Bauer MA, Sales A, Teixeira AR, Morsch P, Lessa AH, Bós Â. Development and accuracy of a hearing screening application. Braz J Otorhinolaryngol 2020 May 05:A (forthcoming) [FREE Full text] [doi: 10.1016/j.bjorl.2020.03.009] [Medline: 32620319]
46. Kelly EA, Stadler ME, Nelson S, Runge CL, Friedland DR. Tablet-based screening for hearing loss: feasibility of testing in nonspecialty locations. Otol Neurotol 2018 Apr;39(4):410-416. [doi: 10.1097/MAO.0000000000001752] [Medline: 29494473]
47. Samelli AG, Rabelo CM, Sanches SG, Aquino CP, Gonzaga D. Tablet-based hearing screening test. Telemed J E Health 2017 Sep;23(9):747-752. [doi: 10.1089/tmj.2016.0253] [Medline: 28328389]
48. Anuar K, Rani M, Hitam S, Noh A. Exploring Uhear Hearing Application as a mobile screening tool for the underprivileged. Rawal Med J 2018;43(4):717-720 [FREE Full text]
49. Corona AP, Ferrite S, Bright T, Polack S. Validity of hearing screening using hearTest smartphone-based audiometry: performance evaluation of different response modes. Int J Audiol 2020 Sep;59(9):666-673. [doi: 10.1080/14992027.2020.1731767] [Medline: 32134341]
50. Saliba J, Al-Reefi M, Carriere JS, Verma N, Provencal C, Rappaport JM. Accuracy of mobile-based audiometry in the evaluation of hearing loss in quiet and noisy environments. Otolaryngol Head Neck Surg 2017 Apr;156(4):706-711. [doi: 10.1177/0194599816683663] [Medline: 28025906]
51. Aremu SK. Evaluation of the hearing test pro application as a screening tool for hearing loss assessment. Niger Med J 2018;59(5):55-58 [FREE Full text] [doi: 10.4103/nmj.NMJ_160_18] [Medline: 31293289]

52. Lin HH, Chu Y, Lai Y, Cheng H, Lai F, Cheng Y, et al. A smartphone-based approach to screening for sudden sensorineural hearing loss: cross-sectional validity study. JMIR Mhealth Uhealth 2020 Nov 11;8(11):e23047 [FREE Full text] [doi: 10.2196/23047] [Medline: 33174845]
53. Potgieter J, Swanepoel DW, Smits C. Evaluating a smartphone digits-in-noise test as part of the audiometric test battery. S Afr J Commun Disord 2018 May 21;65(1):1-6 [FREE Full text] [doi: 10.4102/sajcd.v65i1.574] [Medline: 29781704]
54. Derin S, Cam OH, Beydilli H, Acar E, Elicora SS, Sahan M. Initial assessment of hearing loss using a mobile application for audiological evaluation. J Laryngol Otol 2016 Mar;130(3):248-251. [doi: 10.1017/S0022215116000062] [Medline: 26781699]
55. Xiao L, Zou B, Gao L, Weng M, Lando M, Smith AE, et al. A novel tablet-based approach for hearing screening of the pediatric population, 516-patient study. Laryngoscope 2020 Sep;130(9):2245-2251. [doi: 10.1002/lary.28329] [Medline: 31661567]
56. Burkard R. Development of ANSI standards: an audiologist's view. Ashawire. URL: https://leader.pubs.asha.org/doi/10.1044/leader.FTR1.09112004.2 [accessed 2021-08-09]
57. Fagan J, Jacobs M. Survey of ENT services in Africa: need for a comprehensive intervention. Glob Health Action 2009 Mar 19;2(1):1932 [FREE Full text] [doi: 10.3402/gha.v2i0.1932] [Medline: 20027268]
58. van Tonder J, Swanepoel DW, Mahomed-Asmail F, Myburgh H, Eikelboom RH. Automated smartphone threshold audiometry: validity and time efficiency. J Am Acad Audiol 2017 Mar;28(3):200-208. [doi: 10.3766/jaaa.16002] [Medline: 28277211]
59. Corry M, Sanders M, Searchfield GD. The accuracy and reliability of an app-based audiometer using consumer headphones: pure tone audiometry in a normal hearing group. Int J Audiol 2017 Sep;56(9):706-710. [doi: 10.1080/14992027.2017.1321791] [Medline: 28485673]
60. Mohadisdudis H, Ali N. A study of smartphone usage and barriers among the elderly. In: Proceedings of the 3rd International Conference on User Science and Engineering (i-USEr). 2014 Presented at: 3rd International Conference on User Science and Engineering (i-USEr); Sep 2-5, 2014; Shah Alam, Malaysia. [doi: 10.1109/iuser.2014.7002686]
61. Swanepoel D, Mngemane S, Molemong S, Mkwanazi H, Tutshini S. Hearing assessment-reliability, accuracy, and efficiency of automated audiometry. Telemed J E Health 2010 Jun;16(5):557-563. [doi: 10.1089/tmj.2009.0143] [Medline: 20575723]
62. Clark J, Swanepoel D. Technology for hearing loss--as we know it, and as we dream it. Disabil Rehabil Assist Technol 2014 Sep;9(5):408-413. [doi: 10.3109/17483107.2014.905642] [Medline: 24712413]
63. Andone I, Błaszkiewicz K, Eibes M, Trendafilov B, Montag C, Markowetz A. How age and gender affect smartphone usage. In: Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct. 2016 Presented at: UbiComp '16: The 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing; September 12-16, 2016; Heidelberg, Germany p. 9-12. [doi: 10.1145/2968219.2971451]
64. Kamil RJ, Genther DJ, Lin FR. Factors associated with the accuracy of subjective assessments of hearing impairment. Ear Hear 2015 Jan;36(1):164-167 [FREE Full text] [doi: 10.1097/AUD.0000000000000075] [Medline: 25158982]
65. Erber NP. Variables that influence sound pressures generated in the ear canal by an audiometric earphone. J Acoust Soc Am 1968 Aug;44(2):555-562. [doi: 10.1121/1.1911120] [Medline: 5665526]
66. Marangoni A, Scharlach R, Silveira M, Calais L, Gil D. "Insert earphones: application to avoid collapse of the external auditory canal". Rev Soc Bras Fonoaudiol 2012 Mar;17(1):61-65 [FREE Full text] [doi: 10.1590/s1516-80342012000100012]
67. Borton TE, Nolen BL, Luks SB, Meline NC. Clinical applicability of insert earphones for audiometry. Audiology 1989;28(2):61-70. [doi: 10.3109/00206098909081611] [Medline: 2930364]
68. Almeida B, Menezes P, Andrade KC, Teixeira C. Positioning of earphones and variations in auditory thresholds. Braz J Otorhinolaryngol 2015;81(6):642-646 [FREE Full text] [doi: 10.1016/j.bjorl.2015.08.016] [Medline: 26454532]
69. McPherson B, Law M, Wong M. Hearing screening for school children: comparison of low-cost, computer-based and conventional audiometry. Child Care Health Dev 2010 May;36(3):323-331. [doi: 10.1111/j.1365-2214.2010.01079.x] [Medline: 20507326]
70. Kam A, Gao H, Li L, Zhao H, Qiu S, Tong M. Automated hearing screening for children: a pilot study in China. Int J Audiol 2013 Dec;52(12):855-860. [doi: 10.3109/14992027.2013.832419] [Medline: 24059596]
71. Swanepoel D, Matthysen C, Eikelboom R, Clark J, Hall J. Pure-tone audiometry outside a sound booth using earphone attentuation, integrated noise monitoring, and automation. Int J Audiol 2015;54(11):777-785. [doi: 10.3109/14992027.2015.1072647] [Medline: 26514954]
72. Swanepoel D, Olusanya B, Mars M. Hearing health-care delivery in sub-Saharan Africa--a role for tele-audiology. J Telemed Telecare 2010;16(2):53-56. [doi: 10.1258/jtt.2009.009003] [Medline: 20008052]
73. Storey K, Muñoz K, Nelson L, Larsen J, White K. Ambient noise impact on accuracy of automated hearing assessment. Int J Audiol 2014 Oct;53(10):730-736. [doi: 10.3109/14992027.2014.920110] [Medline: 24909592]
74. Hosmer JD, Lemeshow S, Sturdivant R. Applied Logistic Regression. Hoboken, New Jersey: John Wiley & Sons; Feb 26, 2013.
75. Masalski M, Morawski K. Worldwide prevalence of hearing loss among smartphone users: cross-sectional study using a mobile-based app. J Med Internet Res 2020 Jul 23;22(7):e17238 [FREE Full text] [doi: 10.2196/17238] [Medline: 32706700]

Abbreviations

AUC: area under the receiver operating characteristic curve
ESS: effective sample size
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PTA: pure tone audiometry
SRT: speech recognition test

Edited by L Buis; submitted 04.03.21; peer-reviewed by P Israsena, F Lai; comments to author 23.04.21; revised version received 22.05.21; accepted 17.06.21; published 10.09.21.

Please cite as:
Chen CH, Lin HYH, Wang MC, Chu YC, Chang CY, Huang CY, Cheng YF
Diagnostic Accuracy of Smartphone-Based Audiometry for Hearing Loss Detection: Meta-analysis
JMIR Mhealth Uhealth 2021;9(9):e28378
URL: https://mhealth.jmir.org/2021/9/e28378/
doi:10.2196/28378
PMID:34515644

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