JMIR mHealth and uHealth

Impact Factor (2022): 5.0
Volume 7 (2019), Issue 1 ISSN 2291-5222 Editor in Chief: Laurie Buis, PhD, MSI

Contents

Reviews

Original Papers

Viewpoint

Short Paper

Corrigenda and Addendas

Review

User Models for Personalized Physical Activity Interventions: Scoping Review

Suparna Ghanvatkar[1], MTech; Atreyi Kankanhalli[1], PhD; Vaibhav Rajan[1], PhD

Department of Information Systems and Analytics, School of Computing, National University of Singapore, Singapore, Singapore

Corresponding Author:
Vaibhav Rajan, PhD
Department of Information Systems and Analytics
School of Computing
National University of Singapore
Computing 1, 13 Computing Drive
Singapore, 117417
Singapore
Phone: 65 65166737
Email: vaibhav.rajan@nus.edu.sg

Abstract

Background: Fitness devices have spurred the development of apps that aim to motivate users, through interventions, to increase their physical activity (PA). Personalization in the interventions is essential as the target users are diverse with respect to their activity levels, requirements, preferences, and behavior.

Objective: This review aimed to (1) identify different kinds of personalization in interventions for promoting PA among any type of user group, (2) identify user models used for providing personalization, and (3) identify gaps in the current literature and suggest future research directions.

Methods: A scoping review was undertaken by searching the databases PsycINFO, PubMed, Scopus, and Web of Science. The main inclusion criteria were (1) studies that aimed to promote PA; (2) studies that had personalization, with the intention of promoting PA through technology-based interventions; and (3) studies that described user models for personalization.

Results: The literature search resulted in 49 eligible studies. Of these, 67% (33/49) studies focused solely on increasing PA, whereas the remaining studies had other objectives, such as maintaining healthy lifestyle (8 studies), weight loss management (6 studies), and rehabilitation (2 studies). The reviewed studies provide personalization in 6 categories: goal recommendation, activity recommendation, fitness partner recommendation, educational content, motivational content, and intervention timing. With respect to the mode of generation, interventions were found to be semiautomated or automatic. Of these, the automatic interventions were either knowledge-based or data-driven or both. User models in the studies were constructed with parameters from 5 categories: PA profile, demographics, medical data, behavior change technique (BCT) parameters, and contextual information. Only 27 of the eligible studies evaluated the interventions for improvement in PA, and 16 of these concluded that the interventions to increase PA are more effective when they are personalized.

Conclusions: This review investigates personalization in the form of recommendations or feedback for increasing PA. On the basis of the review and gaps identified, research directions for improving the efficacy of personalized interventions are proposed. First, data-driven prediction techniques can facilitate effective personalization. Second, use of BCTs in automated interventions, and in combination with PA guidelines, are yet to be explored, and preliminary studies in this direction are promising. Third, systems with automated interventions also need to be suitably adapted to serve specific needs of patients with clinical conditions. Fourth, previous user models focus on single metric evaluations of PA instead of a potentially more effective, holistic, and multidimensional view. Fifth, with the widespread adoption of activity monitoring devices and mobile phones, personalized and dynamic user models can be created using available user data, including users' social profile. Finally, the long-term effects of such interventions as well as the technology medium used for the interventions need to be evaluated rigorously.

(JMIR Mhealth Uhealth 2019;7(1):e11098) doi:10.2196/11098

KEYWORDS

review; exercise; physical fitness; automation; mobile apps; web browser; health communication; health promotion

Introduction

Background

Insufficient physical activity (PA) is a worldwide concern as it is a major cause of obesity and the fourth leading risk factor for mortality, accounting for an estimated 3.2 million deaths globally [1]. Maintaining or increasing PA of patients is also an important goal in the treatment for various chronic diseases such as diabetes and cardiovascular illnesses.

Fitness trackers, such as Fitbit and Jawbone, are increasingly being used to monitor personal PA. Activity data collected by associated smartphone apps are being utilized, along with other user-specific or contextual data, to design interventions with the aim of motivating users to increase their PA [2,3]. These interventions take varied forms ranging from activity status reports to personalized fitness-buddy recommendations.

Increasing PA often requires a change in lifestyle or behavior of the user. Feedback based on activity status reports is a common strategy that is often augmented with educational information on the benefits of increased PA. The key limitation of such interventions is the reliance on the self-motivation of users to increase their PA [4,5]. Users may not be motivated for various reasons, for example, they may be inactive by habit or the presented activity goal may be too intimidating for them. Other factors may also play a role in determining the efficacy of interventions. For example, some users may not have the time to perform a recommended PA [6] or there may be constraints imposed by the users' location, weather, or working environments. Providing information on the benefits of increased PA rarely suffices; effecting behavior change to increase PA additionally requires motivational interventions [7].

The aims, behavior, preferences, context, and lifestyle of users have to be taken into account by apps to design effective interventions [8,9]. A "one size fits all" approach is unable to effectively serve a diverse set of users. Even simple activity recommendations, such as "60 minutes of moderately vigorous physical activity (MVPA)" may be too daunting for a sedentary user or for a cardiac patient. Thus, there is a need for personalization of interventions for promoting PA among users. Personalization implies a modification in the intervention generation or delivery aimed at a specific user. A status feedback does not indicate personalization; personalization implies customized content or advice to help the targeted user in increasing PA.

Previous reviews [10-13] on interventions for increasing PA have studied internet-based or Web-based interventions without focusing exclusively on personalization. They evaluated the success of included studies with respect to intervention delivery (eg, email and website-based) and discussed the utility of theory-based interventions [10]. Other reviews have specifically studied target groups, such as stroke patients [14] or cardiovascular disease (CVD) patients [15]. Another recent review [16] has analyzed the decision support systems used in PA interventions but does not focus on personalization. A survey of tailoring techniques used in real-time PA coaching systems published before August 2013 is presented in the study by op den Akker et al [17].

The term "personalization" has multiple definitions in different domains [18]. We follow the commonly accepted definition in the study by Fan et al [18], which defines personalization as "a process which changes the functionality, interface, content or distinctiveness of system to increase its personal relevance." According to this definition, if the system is not altered in any of the dimensions mentioned to increase personal relevance, it is not considered as personalization. An earlier study by Hawkins et al [19], defines "tailoring" as a generic term for providing feedback, personalization, and content matching. It uses the term personalization to encompass the tactics of identification, raising expectation, and contextualization. However, following our adopted definition, from the study by Fan et al [18], we also include the category of "content matching" within "personalization." The review by Akker et al [17] identified 7 categories based on tailoring techniques for activity coaching—feedback, inter-human interaction, adaptation, user targeting, goal setting, context awareness, and self-learning—and discussed relevant studies in these categories. Thus, tailoring has been used as a broad term in the literature and does not necessarily provide the "modification" required for personalization in our adopted definition. In this study, we use the term "personalization" to denote a user-specific modification of an intervention.

Objectives

The purpose of this review was to identify recent literature where the technology-based intervention is personalized with the aim of increasing PA of users. The feedback or recommendation is not just a presentation of the users' activity status. It is either a personalized feedback based on the history and status of the user to motivate or educate the user or a recommendation to potentially increase PA. A key aspect of such studies, which we focus on, is the user model created, which in turn helps generate personalized recommendations. Findings from this review provide important insights into the current literature and identify significant gaps in the literature. Addressing these gaps could lead to more effective, personalized, and technology-based interventions for promoting PA.

Methods

The Scoping Review

This review aims to identify various interventions, customizations, and user models generated for personalization of technology-based interventions to increase PA of users. We employed a rigorous literature search and chose to conduct a scoping review to analyze our research questions. The research questions we have focused on are as follows: (1) what are the ways of providing technology-based personalized interventions for increasing PA among users and (2) what are the user models used to provide such personalization? For ensuring quality of the included studies, we have used only peer-reviewed articles, including research-in-progress articles, which had full text available. We do not perform additional quality analysis of the studies, as quality assessment does not form part of the scoping

study remit [20]. This paper follows the methodology and directions given in the study by Arksey and O'Malley [20].

Search Strategy

PubMed, PsycINFO, Scopus, and Web of Science databases were used to select relevant studies. A comprehensive search was conducted till August 23, 2018, in which articles published since 2013 were targeted. The search string was constructed by considering the criteria required to be satisfied by the studies to be considered: {physical activity} {interventions} having {personalization} provided through some {technology} and identifying or creating a {user model} for the same. The following search string was used: ((fitness OR exercise OR "physical activity" OR "activity level" OR "active living") AND (intervention OR recommend* OR prescribe OR prescription OR feedback OR message) AND (tailor* OR personaliz* OR personalis*) AND (mobile OR internet OR computer OR device OR "fitness trackers" OR website OR online) AND (profil* OR model)). The search was restricted to papers published in English. This search string ensured our condition of the technology-based intervention having a user model or profile identified for providing the personalization.

In addition to the database searches, we also performed hand searches for additional relevant studies. These studies were found by identifying relevant references from the studies selected. These references were also analyzed for the selection criteria and included in the review if they met the criteria. In addition, a hand search of *Journal of Medical Internet Research* results for "physical activity interventions" was done to identify several other relevant studies.

Data Extraction

We selected the articles in 2 phases and used Mendeley reference manager to organize them. The first phase involved title, abstract, and keyword review as obtained from the databases searched. This phase was applied to all results obtained from the databases after merging duplicates (a feature provided by Mendeley). The second phase included reviewing the full text of the articles. This was done by obtaining PDF documents for each of the articles that met the inclusion criteria. The full texts were analyzed using the inclusion and exclusion criteria, and studies that were deemed relevant after this phase were included in the scoping review.

Selection Criteria

Studies were eligible for this review if all the following were true: (1) there was an attempt to increase or regulate PA among the target users; (2) the studies had some form of personalized intervention, as recommendations or feedback intended to promote PA of the users; (3) a user model was generated and used for providing the personalized intervention described; (4) the intervention was provided through usage of technology; (5) studies were in English and published in or after the year 2013; and (6) they were not review papers, dissertations, or letters and were published through a peer-reviewed process.

Studies published before 2013 were not included as the popularity of fitness devices and attempts to create trackers and coaches have increased in the last 4 to 5 years, that is, older literature may be less relevant to today's apps. Moreover, relevant literature until then has already been reviewed in the study by Akker et al [17]. There was no restriction on the study objectives, type of users, or the type of intervention or feedback, other than the focus on personalization for PA promotion. The focus of the review is on methods of personalization and user model generation for technology-based interventions. The interventions where personalization was provided manually were excluded as the user model used for delivering the personalization cannot be identified in a manual process. A comprehensive review of different ways to model users and provide personalized, technology-based interventions for increasing PA in different settings was desired.

The exclusion criteria for this review were as follows: (1) personalization not aimed at increasing PA (eg, personalization in activity tracking or gait detection); (2) personalization provided only in terms of using name or activity status in message, these parameters were filled into standard messages; (3) no user model identified during the intervention; (4) personalization generated manually, even though may be delivered using technology through a website; (5) gender- or culture-based standard tailoring for intervention; and (6) only reports provided without any personalized content for encouraging or educating user, or without any advice.

The inclusion criterion entailed that the technology-based, personalized intervention had to be necessarily aimed at increasing PA. The criterion of increasing PA was not necessarily the main objective of included studies but had to be one of the objectives. For example, in some studies, medication adherence [21] or weight loss [2] was the other objective.

Results

Screening and Study Selection

The screening procedure and study selection was undertaken by 1 researcher and then independently verified by 2 other researchers for adherence to the selection criteria. The initial results were screened for the inclusion criteria and the full-text articles were analyzed using the exclusion criteria. Initial results were obtained by setting the filters of language and duration for all the databases (536 results) and were then searched for duplicates, which resulted in 355 unique studies. The abstracts of these studies were then screened for the criteria of whether the paper tried to increase or regulate PA. In addition, 15 relevant studies were identified by hand searching and cross-references. This led to a selection of 181 papers, which were screened on full text for the remaining selection criteria, resulting in 57 studies.

We found several groups of studies that studied the same system, that is, they were parts or improvements of the same intervention. We also found additional studies through hand searches that belonged to these groups, which helped us understand details of the interventions. We grouped these related studies together and used only 1 representative publication for each of the 23 groups. The groups and the representative studies are listed in Table 1. This step reduced the final number of studies to 49.

Table 1. Studies grouped by the intervention developed or investigated.

Intervention	Related studies	Representative study
ACKTUS	Janols et al and Lindgren et al [23,24]	Janols et al [23]
Active Plus	Boekhout et al, Peels et al, Peels et al, and van Stralen et al [25-28]	Peels et al [27]
Active2Gether	Klein et al and Klein et al [3,29]	Klein et al 2017 [3]
Active-O-Meter	Cook et al and De Bourdeaudhuji et al [30,31]	Cook et al [30]
ATHENA	Ali et al and Fahim et al [32,33]	Fahim et al [32]
Food4Me	Marsaux et al, Morales et al, and Marsaux et al [34-36]	Marsaux et al [35]
I Move	Friederichs et al and Friederichs et al [37,38]	Friederichs et al [38]
MOPO	Ahola et al, Jauho et al, and Pyky et al [39-41]	Pyky et al [41]
My Activity Coach	Alley et al and Alley et al [42,43]	Alley et al [43]
MyBehavior	Rabbi et al [44,45]	Rabbi et al [45]
myHealthyBehavior	Schulz et al and Schulz et al [46,47]	Schulz et al [47]
PATH-In	Brooks et al and Williams et al [48,49]	Williams et al [48]
PATHway	Chatzitofis et al, Claes et al, and Triantafyllidis et al [50-52]	Triantafyllidis et al [52]
Personalized Coaching System	Cabrita et al, Hermens et al, and Op den Akker et al [53-55]	Hermens et al [54]
PRO-Fit	Dharia et al [56-58]	Dharia et al [56]
REACH	Mitchell et al and Mitchell et al [59,60]	Mitchell et al [60]
RENATA	Reinwad et al and Storm et al [7,61]	Storm et al [7]
SmartLoss	Martin et al and Martin et al [2,62]	Martin et al [2]
Start to Stand	De Cocker and De Cocker [63,64]	De Cocker [64]
TaylorActive	Soetens et al and Vandelanotte et al [65,66]	Vandelanotte et al [66]
TXT2Bfit	Hebden et al and Partridge et al [67,68]	Partridge et al [68]
Weight in Balance	Walthouwer et al, Walthouwer et al, and Walthouwer et al [69-71]	Walthouwer et al [71]
YEAH	Kattelmann et al [72,73]	Kattelmann et al [72]

Among the 181 studies assessed for eligibility, most of the studies could be screened using our exclusion criteria. However, a few studies, such as the study by Liu and Chan [22], were identified through the search but were excluded because the definition of personalization used was different. It focused on whether or not to prompt the user based on current and predicted activity status, which differs from the conceptualization adopted.

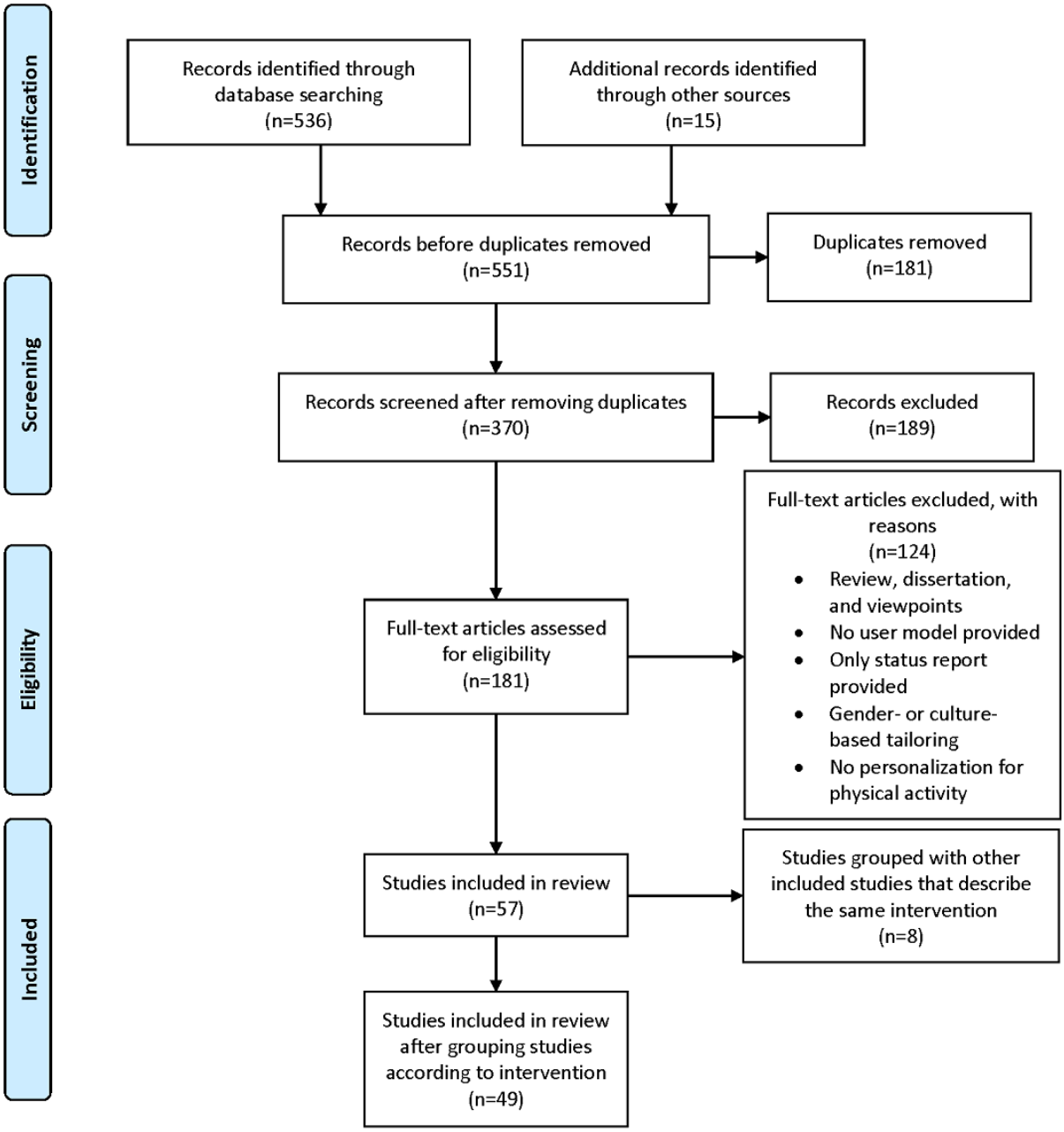
Figure 1 illustrates the flowchart representing the study selection process.

Overview of Studies

We placed no restrictions on the research objective or methodology of the studies to be included in the review, other than following our study criteria. As a result, the studies differ considerably with respect to their research objectives, interventions, data collection methods, and target users. We summarize these diverse settings below before examining the personalized interventions and user models employed in more detail. Moreover, as mentioned earlier, we identified 23 interventions, each of which has been described in more than 1 study. These groups are listed in Table 1. In this review, we study all the articles listed in the table (column 2) but represent each group by a representative study (column 3).

Increasing PA was the research objective in 33 studies. In the remaining studies, the objectives were weight loss, weight management, or obesity prevention (6 studies); maintaining healthy lifestyle that included diet, smoking, alcohol, or exercise management (8 studies); and rehabilitation (2 studies), with PA increase being an auxiliary goal. Furthermore, 1 study had a combined goal of weight management and healthy lifestyle. Among the 33 studies on PA, 31 directly aimed to increase PA of users, 1 study aimed to reduce workplace sitting time [64], and the last study aimed to encourage medical adherence in addition to increasing PA [21]. These studies not only tried to monitor and increase PA but some also focused on helping users overcome barriers to increasing PA and improving their self-efficacy, for example in Oosterom-Calo et al [21]. The 6 studies aimed at weight loss attempted to increase the PA of users to achieve the desired results [eg, 2,74]. One study aimed at providing rehabilitation to patients provided recommendations in consultation with health care experts [75]. The last study aimed at rehabilitation provided real-time as well as weekly adaptation of exercises for the patients [52]. For the 8 studies on healthy lifestyle, their objectives included wellness services [32], exercise and diet recommendation [6,7,47,76-78], and a personalized coaching system [54], which was illustrated through 3 use cases, that is, neck coach, activity coach, and stress coach.

Figure 1. Flowchart for study selection process.



The interventions were presented or delivered to users in multiple ways, for example, through Web apps [7,66,79], mobile phone apps [2,3,74], Kinect devices [52,80], specific activity monitors [81], PDF report [35], text message (short message service, SMS) [9,68,82], printout [27,83,84], or telephone call [60]. In some cases the users were actively pushed by the system toward their PA goal by automatic delivery of interventions periodically [7,9,54]. In other cases, the interventions were relatively more passive and expected higher levels of motivation from the users. They required the user to access the app [45] or answer questionnaires [21,64,79] before they could obtain the personalized intervention.

Data were collected for the intervention (for monitoring users and generating user models) in various ways, for example, through questionnaires [3,21,80], mobile phone sensors [9,32,45,85], specific activity monitors [41,64,66,81], or fitness trackers [2,3,86]. The target population of the studies varied from specific to general users, that is, people with chronic disease [76], elderly adults [79,80], diabetes patients [9,87], cancer patients [78,82,84], people with CVD or at risk of CVD or heart failure [7,21], osteoarthritis patients [48], young adults [3,72,88], and general users [2,32,45,86].

Apart from these differences, the included studies differed in the intervention generation techniques, type of personalization, and user models. In this review, we systematically study these 3 aspects in detail. Note that some studies included in the review may also have incorporated personalization of diet or models for activity detection, but our focus is restricted to only the part of the study concerning PA. The overview presented in Multimedia Appendix 1 highlights the objectives, interventions,

personalization, user models, and theoretical models used in the included studies. These are ordered by their research objectives and by the intervention generation mode within each objective.

Types of Interventions

The interventions included in this review provide various forms of feedback or recommendation. We distinguish between feedback and recommendation, where recommendations are prescriptive in nature, whereas feedback is an informative response to the users' actions. For example, feedback can be information regarding tips to increase PA such as "exercising with a partner can be a fun and motivating experience," whereas recommendations are prescriptive suggestions of an activity or goal such as "30 minutes of brisk walk along with your mother" provided to the user. We only consider feedback that is personalized in some aspect (eg, content or timing).

Personalization is achieved in different ways, that is, by personalizing goal, activity, or fitness partner recommendations or by personalizing messages and their timings, as discussed in the following section. On the basis of how this personalized intervention is generated, interventions in our review could be classified into 2 categories: semiautomated and automated. We excluded those studies that had only manual interventions.

Semiautomated Interventions

Semiautomated interventions are those where personalization is not completely automated but includes manual effort from the health care provider. There were 9 studies with semiautomated interventions, and the combinations of manual and automatic elements in them varied.

In the study by Tseng et al [76], automated activity and goal recommendations were provided, which could be modified by a medical expert. Similarly, for the case of SmartLoss [2], the system required a goal to be set in consultation with the nurse, but the platform also automatically provided a set of "SmartTips" in case the user was predicted to be deviating from the weight loss program. Another semiautomated system [35] provided automated educational content to the users using a machine learning method along with manual personalized advice and intervention from an expert. My Activity Coach study [43] used automatic advice recommendation as well as a one-on-one video calling interaction with coach. Similarly, 3 other studies [60,68,82] used telephonic conversations for motivational interviewing, but the participants had website access or automatically delivered messages. The remaining 2 studies [75,80] generated automatic personalized activity or game levels within the limits defined by a health care expert.

Automated Interventions

Automated interventions present in 40 papers in our review used either knowledge-based or data-driven approaches or both to automate the personalization. All the knowledge-based systems relied on either behavior change techniques (BCTs) or PA or clinical guidelines. All the data-driven systems used machine learning techniques to learn user models from historical data.

Knowledge-Based Systems

There were 30 studies using knowledge-based approaches. These systems were rule-based and provided feedback and recommendations based on reasoning modules or rules.

Of these studies, 22 of them attempted to encode knowledge into the system derived from behavior change theories. They provided the personalization intervention by inferring the most suitable user category, where the categories were theory-based. Thus, they provided category-level personalization instead of an individual level of personalization. In addition, 2 studies added an individual level personalization by considering the preference of intervention time [89,90] or by providing walking and cycling routes based on location [27].

There were 8 studies that used PA or clinical guidelines. The study by Ali et al [85] used a hybrid rule- and case-based reasoning model but tried to identify similar cases using the K-nearest neighbor algorithm. Their system was based on the Center for Disease Control and Prevention (CDC) guidelines [91]. The study by Coolbaugh et al [81] used a specific activity monitor for providing the intervention and performed goal adaptation in accordance with American College of Sports Medicine training progression guidelines [92]. This intervention was time-bound and progressed according to the rules laid out in the process flowchart. Moreover, 2 studies [48,93] also used various PA guidelines to recommend step goals. The last 3 studies [52,78,87] used PA guidelines for their specific types of users, that is, cancer patients and diabetes patients. Few studies among these [6,52,78,87] use knowledge-based system constructed using a BCT but provide recommendations based on clinical guidelines.

Data-Driven Systems

Data-driven systems using machine learning approaches were described in 7 studies. The class of machine learning techniques used falls under 2 broad categories, that is, reinforcement learning [9,45] and supervised learning [35,56,74,86,94].

In reinforcement learning, an automated agent learns a policy to optimize a cumulative reward function while sequentially interacting with the environment. At each step, the agent performs an action, obtains a reward, and decides its next action based on the reward with the aim of optimizing the reward function. Thus, in the study by Yom-Tov et al [9], at each step, the agent sent a message, obtained information about the user's PA, and determined which message to send in the next step. The messaging policy was personalized for the user to maximize cumulative PA. For the other study, multi-armed bandits, a form of reinforcement learning, was used for suggestion generation in the MyBehavior app [45].

Supervised machine learning techniques learn a model from historical data to predict dependent variables from independent variables. The model may be static (such as support vector machine [SVM] or decision trees in the study Marsaux et al [35]) or temporal (such as recurrent neural networks [RNN] in the study by Lim et al [86])—the former does not explicitly model temporal correlations, whereas the latter does. In another study, PRO-fit recommended a fitness partner using geolocation, activity preference, and calendar-based availability on a

smartphone [56]. It also provided activity recommendation using collaborative filtering [57] and activity prediction from raw accelerometer data. An Internet of Things–based app [94] proposed a context-aware recommendation system to generate a suitable activity for the user based on current fatigue and fitness level. Finally, in the study by Lim et al [86], a lifestyle model parameterized by heart rate (HR), step count, and burned calories was constructed using RNN, and recommendations for healthy behavior were based on forecasts of these variables.

Combined Knowledge-Based and Data-Driven System

There were 3 studies that combined knowledge-based and data-driven approaches. The ATHENA system [32] defined a rule-based recommendation, in which only ranking and validation were done using machine learning. Ensemble-based supervised learning was used for recommendation of food, physical, and mental therapy in study. In the study by Hermens et al [54], a rule-based system was used for message content along with machine learning to appropriately time the message—an SVM was trained based on historical data to predict effective timing. Another personalized health care system [77] proposed an ontology-based knowledge base, which uses decision trees for providing relevant recommendation to the user.

Overview of Personalization

Personalization was found in both recommendations and feedback. In the case of recommendations, personalization was seen with respect to goal setting, activity suggestion, and selection of fitness partners. Feedback was found to be personalized with respect to the content, which could be motivational or educational, or with respect to the timing of its delivery. Status comparative feedback was also considered to be personalized as it was provided to only those for whom it was considered beneficial. Thus, we classify personalized interventions into 6 categories, that is, goal recommendation, activity recommendation, fitness partner recommendation, educational content, motivational content, and intervention timing. These categories are not mutually exclusive, as several studies had more than 1 type of personalization.

Goal Recommendation

The category of goal recommendation refers to the prescription of a quantified target goal. This target is in terms of an activity evaluation metric, such as duration of activity, step count, or calorie expenditure. Note that if an activity is prescribed without quantification, then we classify it as an activity recommendation and not a goal recommendation.

Of the studies in the review, 20 of them provided personalized goal recommendation. The specification of the goals differed across the systems and apps. The goals could be specified in terms of game level [80], training zone and HR [52,75,81], activity duration [3,43,76,81,82,89], step data [2,54,60,93,95], or activity level prescription by an expert [23]. These goals were adapted according to the person's status and did not follow standardized fixed goals (such as "30 minutes of MVPA"). For the case of the REACH intervention [60], it has been mentioned that personalized step goals are generated based on rate of perceived exertion. However, it is not stated if this is done automatically or by researchers and delivered manually.

In 4 of these studies [2,45,66,85], personalized goals were indirectly defined or altered after obtaining fixed goals from the user or a guideline. The SmartLoss app [2] aimed to make the user follow their regular exercise program of 7000 to 8000 steps per day. It defined a "zone of adherence," which is a weight range indicating that the weight loss of the user is as expected. Goal adaptation occurred when a user was repeatedly outside this "zone of adherence" and was provided other options for increasing PA. In the multimodal reasoning system [85], the example goal was in terms of kilograms to lose, but the personalized goals in terms of target metabolic equivalents (METs) and calories were also calculated and recommended by the system. The MyBehavior app [45] used the weekly weight loss goal entered by the user to obtain a personalized target calorie goal using the Harris-Benedict equation [96]. The TaylorActive system [66] also provided goal recommendations and suggestions during a session, but the actual goal was set by the user.

All the above-mentioned studies set a goal for the user before the user activity began. However, in the personalized PA prescription intervention study [81], the goal was not explicitly known by the user before the activity, although a Web interface allowed the user to check the goal recommendation. It also defined a user goal in terms of target HR and duration of activity, which was sent to the activity monitor. The activity monitor provided visual feedback (blinks on the monitor) to the user when the target goal was achieved or if their HR exceeded the target.

In another study [35], the feedback was whether the user must increase, strongly increase, or maintain their PA. However, the feedback was not quantified, and thus, this study was not classified as a goal recommendation.

Activity Recommendation

The category of activity recommendation includes studies where 1 or more appropriate activities (eg, running and cycling) or behaviors (eg, sleep for X hours) are prescribed to the user. The 22 studies in our review that provided activity recommendation also retained the monitored PA as part of the user model.

Of these, 4 studies offered semiautomated interventions. The activity recommendations were in the form of health care experts' treatment advice (where the treatment included PA) [80] or activity suggestions [35,76]. Furthermore, 1 study [82] did not use a health care expert for the activity plan but suggested activity in the messages sent to the user.

The remaining 18 studies generated activity recommendations based on automated systems. These studies generated the activity or behavior recommendation by considering contextual information such as location [45,56,86] or preferences [32,66,85].

In another study [64], the recommended activities were restricted to standing or walking. It is also important to note that this system encouraged the user to create a goal and activity plan with the aid of the system, which shows the importance of user's

motivation and involvement in planning. In the study by Klein et al [3], the activity and goal recommendations were provided, but the final choice was left to the user. Similarly, in the study by Williams et al [48], the activity plan was generated by the system and adjusted to user level, but the user could rerequest the plan generation. For the personalized coaching system in the study by Hermens et al [54], the activity recommendation was provided in the content of the message [54].

Fitness Partner Recommendation

The aim of fitness partner recommendation is to match users of a system who are similar, to motivate them and help them maintain their PA. Of the 3 studies of this type, 2 studies [56,74] used recommender systems for finding a suitable partner. The other study [95] attempted to find a similar user by matching all users who crossed the particular user during a running activity.

Educational Content

In the educational content category (21 studies), personalized feedback aimed to increase the knowledge of the users about the importance of or techniques for improving PA.

There is a vast amount of information available on the internet, and providing the user with the most relevant content is the aim of such personalization. A direct way to do this was to provide appropriate links to website content (eg, in Food4Me [35], SmartLoss [2], and the multimodal system in the study by Ali et al [85]). The "My Activity Coach" system [43] and Ninas Saludables intervention [88] provide tailored Web content to users, for example, obese users receive additional content not provided to users with normal weight. The Active2Gether system [3] had an educational phase, where messages that put a user's insufficient performance into perspective were sent along with the need and benefits of PA. The "start to stand" [64] app provided feedback messages, which also helped impress the harmful effects of too much sitting or sedentary lifestyle, based on the decision rules. Some studies such as those by Storm et al and Short et al [7,84] provided tailored example plans to the user to aid in goal setting. Some studies provided content in terms of tips to increase PA if found to be relevant [30,38,71,97,98]. The I-Move for Life study [84] provided information on the benefits of PA tailored according to the expected outcomes. The Active Plus system [27] provided the user with information on sports opportunities tailored to the location, along with walking and cycling routes. This information is also educational as it provides a feasible method for improving PA of the user.

Educational content may be motivational as well, for example, if the content is provided to help users overcome their specific barriers to performing PA [21,32,66,79,83]. This was in the form of tailored video or textual content.

Motivational Content

This category (29 studies) contains personalized feedback that aims to motivate users to improve their PA. What may motivate a user can be inferred from specific rules or BCTs. Note that motivational messages that were not personalized (eg, "Good!" [2]) were not included. Messages in this category targeted users specifically to elicit an action by also utilizing techniques including the users' name or providing users' current PA status [3]. However, as mentioned in the exclusion criteria, using only statistics or name in a standard template message is not considered personalization.

A reinforcement learning based study [9] aimed to learn which type of message (negative feedback, positive relative to self, positive relative to others, or no message) best motivates a user. A few studies had both motivational and educational messages as they target the beliefs of users [43,83,84,88]. The studies targeting "stage of change" of the user generally provided personalized motivation [3,41,54,68,72] by determining the stage the user is currently in, for example, precontemplation, preparation, or maintenance. The TaylorActive system [66] provided personally relevant feedback in various categories including what they called the "boosting your confidence" category.

The multimodal system in the study by Ali et al [85] offered motivating content that was not personalized. The Social POD app [74] provided personalized fitness partner recommendation. Personalization of motivational content was done through the fitness partner, who selected a motivational message to be sent to the user.

Intervention Timing

This form of personalization takes the context into account and finds the right time to send a feedback or recommendation to the user. Timing of feedback is known to play an important role, for example, a notification reminder sent when the user is busy is likely to be ignored and forgotten.

In our review, 7 studies provided this kind of personalization. Of these, 2 studies [54,86] learned the most appropriate time for intervention from past data using machine learning. The neural network–based model [86] used a greedy policy to determine the best time, which learns from user feedback, after predicting the users' activity. The personalized coaching system in the study by Hermens et al [54] trained an SVM to determine the appropriateness of a given time to send a message. Moreover, 2 studies (PRO-Fit [56] and Step Up Life [89]) used the calendar context to determine if a given time was suitable for recommendation. Step Up Life intervention [89] additionally used location to determine home and "friendly" locations for providing intervention or reminders. Furthermore, 2 studies [82,90] used the preference of timing obtained from user for providing personalization. The last study [6] mentioned sending reminders at opportune moments, but the exact methodology is unclear.

Theoretical Background for Personalization

In our review, 41 studies used a theoretical framework or foundation for providing personalization. Apps with a theoretical background for their personalization either followed guidelines from sports or health care bodies or used BCTs.

Activity training guidelines from the American College of Sports Medicine [92] were followed for recommending activity increments to avoid injuries in the studies by Lee et al and Coolbaugh et al [78,81]. Guidelines from the CDC [91] were

used to generate hybrid rule-based techniques to recommend a suitable activity to users in the studies by Fahim et al and Ali et al [32,85].

BCTs are theory-based methods for changing 1 or several psychological determinants of behavior, such as a person's attitude or self-efficacy. They aim to create a change among users through appropriate persuasion. Several studies used knowledge-based approaches to incorporate BCTs in providing personalization. The BCTs used were based on Fogg's Behavior Model [99], Social Cognitive Theory (SCT) [100], Transtheoretical Model (TTM) [101], Theory of Planned Behavior (TPB) [102], I-Change Model [103], Behavioral Change Wheel [104], Activity Theory [104], Protection Motivation Theory [105], Motivational Interviewing [106], and health action process approach (HAPA) [107]. The study by Mukhtar [6] used Fogg's Behavior Model to create what is termed as a "persuasion strategy" for the user, which takes into account motivation, ability, and trigger as parameters for appropriate recommendation. The Step Up Life intervention also utilizes the Fogg's Behavior Theory for designing the model [89]. HAPA [107] was used to target different user stages and provide information on behavior risk and intention formation to the user in the study by Storm et al [7]. TTM defines stages of change in users and was used to determine the feedback given to the user, a direct rule-based implementation of the underlying BCT in the study by Pyky et al [41]. A system utilizing TPB and the Stage of Change Model [101] represented the constructs through questions as psychosocial correlates with PA. I-Change Model was used to design the system and questionnaires to effectively motivate users in a few studies [47,71,97]. The SCT has also been used to design the intervention considering that including and addressing social mediators such as family and peer would elicit a positive and sustained response from the user [60,88]. The TaylorActive study [66] used TPB, SCT, and self-determination theory to assess various constructs such as self-efficacy, intrinsic/extrinsic motivation, and action planning during different sessions designed for the user.

Theoretical frameworks were not present in studies using machine learning algorithms for recommending a goal or activity, as the algorithm was used to model user activity and suggest or recommend a better alternative. Some studies [45,74] used BCTs to make design decisions and choices but did not use BCT parameters for user modeling. For instance, the MyBehavior app [44,45] followed BCT to provide low-effort suggestions and used a form of reinforcement learning for activity recommendation. Another study [74] used SCT to design messages and used machine learning to recommend a fitness partner. Table 2 shows the different types personalization provided by the studies in our review.

User Models

Each study in this review created a different user model and defined the user through various attributes. We classify user-related attributes into 5 categories, that is, PA profile, demographics, medical data, BCT parameters, and contextual information.

User models can have a static and/or a dynamic component. The static component is collected only once, typically at the start of the intervention, for example, demographics and preferences. The dynamic component gets updated regularly and includes the monitored quantity describing PA. Some user models also used the personalized quantity as part of the model. All the collected information may not be part of the user model; here, only the data required and used to provide personalization are described under the user model. In cases where it could not be determined how the measured quantity was used, it has been mentioned as part of the profile descriptions.

Physical Activity Profile

The user model nearly always included the quantity being monitored—weight, diet, or PA—either recorded automatically or logged by the user. The monitored quantity differs in the included studies because of differing research objectives, intervention systems, and evaluation metrics. PA profile consisted of this monitored quantity along with the historical data of feedback, goals, or activity.

Evaluation of PA was necessary in almost all cases as personalized advice to users would need to consider current PA status of the users. Thus, PA profile was used as part of the user model in 47 studies. However, PA profile data were not used in 2 studies that provided behavior advice to its users based on the assessment and identified problematic beliefs and barriers [21,68].

Most of the studies evaluated PA by calorie or energy expenditure in terms of METs [45,85,86]. Some others estimated it by the time spent at different PA levels, such as vigorous or MVPA [41,87]. There were studies which set the target HR and used specific HR monitors for data collection [80,81,95], whereas 1 used a smartwatch [75]. Step count was another measure used to evaluate PA, obtained directly from fitness tracking devices [2,3,9]. PA was also evaluated by the time spent in performing an activity [30,74] or the duration. Another study [35] used metrics such as PA level and activity energy expenditure to estimate the level and energy expenditure in performing the PA. Stairs climbed was also used as a measure of PA in the study by Klein et al [3]. The activity level was a common metric used by studies which collected PA-related data through questionnaires [38,43,79,97].

The parameters listed in the PA profile of the studies (see Multimedia Appendix 2) are self-explanatory except for 2 of them. The "start and stand" app [64] had a data attribute named "level of sitting time in 5 domains." This was obtained through the Workforce Sitting Questionnaire and included time spent in (1) traveling, (2) at work, (3) watching television, (4) using computer at home, and (5) other leisure activities. In the study by Martin et al [2], the "zone of adherence" was a quantity calculated by their mathematical model to predict whether the user needs to be provided special interventions. Furthermore, 1 study [77] used the term "lifestyle" for personalizing the exercise recommendation to a person. This has been categorized as a PA profile metric as lifestyle can be used to deduce the current level of PA of the person. In addition, 1 study [60] used a metric termed as Signal Vector Magnitude to calculate the vector magnitude of acceleration corrected for gravity.

Table 2. Personalization provided.

Serial #	Paper reference	Goal recommendation	Activity recommendation	Fitness partner	Educational content	Motivational content	Intervention timing
1	Vandelanotte et al [66]	Y[a]	Y	—[b]	Y	Y	—
2	Ahire et al [77]	—	Y	—	—	—	—
3	Mukhtar [6]	U[c]	Y	—	—	—	Y
4	Tseng et al [76]	Y	Y	—	—	—	—
5	Storm et al [7]	—	—	—	Y	Y	—
6	Schulz et al [47]	—	—	—	—	Y	—
7	Hermens et al [54]	Y	Y	—	—	Y	Y
8	Lee et al [78]	Y	Y	—	Y	Y	—
9	Fahim et al [32]	—	Y	—	—	—	—
10	Dharia et al [56]	—	Y	Y	—	—	Y
11	Rabbi et al [45]	Y	Y	—	—	—	—
12	Twardowski et al [94]	Y	Y	—	—	—	—
13	Yom-Tov et al [9]	—	—	—	—	Y	—
14	Lim et al [86]	—	Y	—	—	—	Y
15	Cook et al [30]	—	—	—	Y	Y	—
16	Larsen et al [88]	—	—	—	Y	—	—
17	Short et al [84]	—	—	—	Y	Y	—
18	Boudreau et al [97]	—	—	—	Y	Y	—
19	Moreau et al [87]	—	—	—	Y	Y	—
20	Rajanna et al [89]	Y, U	Y	—	—	—	Y
21	Irvine et al [83]	—	Y	—	Y	U	—
22	Friederichs et al [38]	—	—	—	Y	Y	—
23	Blake et al [108]	—	—	—	—	Y	—
24	Coolbaugh et al [81]	Y	—	—	—	—	—
25	Hargreaves et al [93]	Y	—	—	—	Y, U	—
26	Williams et al [48]	Y-initially	Y	—	—	—	—
27	Kwasnicka et al [98]	—	—	—	Y	Y	—
28	Janols et al [23]	U	—	—	—	Y	—
29	Ali et al [85]	Y	Y	—	Y	—	—
30	Mistry et al [90]	—	—	—	—	Y	Y
31	Peels et al [27]	—	Y	—	Y	Y	—
32	Klein et al [3]	Y	Y	—	Y	Y	—
33	Ammann et al [79]	—	—	—	Y	Y	—
34	Pyky et al [41]	—	U	—	U	Y	—
35	Varadharajan et al [95]	Y	—	Y	—	Y	—
36	Codreanu et al [80]	Y	Y	—	—	—	—
37	Marsaux et al [35]	—	Y	—	Y	—	—
38	Alley et al [43]	Y	—	—	Y	Y	—
39	Mitchell et al [60]	Y, U	—	—	—	—	—
40	Oosterom-Calo et al [21]	—	—	—	Y	Y	—
41	De Cocker et al [64]	—	Y	—	Y	Y	—

Serial #	Paper reference	Goal recommendation	Activity recommendation	Fitness partner	Educational content	Motivational content	Intervention timing
42	Triantafyllidis et al [52]	Y	Y	—	—	Y (if required)	—
43	Dobrican et al [75]	Y	—	—	—	—	—
44	Hales et al [74]	—	—	Y	—	Indirect	—
45	Martin et al [2]	Y	—	—	Y	—	—
46	Spark et al [82]	Y, U	Y, U	—	—	Y	Y
47	Kattelmann et al [72]	U	U	—	—	Y	—
48	Partridge et al [68]	—	—	—	—	Y	—
49	Walthouwer et al [71]	—	—	—	Y	U	—

[a]Y: personalization present.

[b]Personalization absent.

[c]U: unclear.

Demographics

Demographics formed a part of the user model for 39 studies. Demographic data collected included age and gender [75], body mass index [7,32], employment [64,79], nationality [7], weight [79,95], marital status [47,88,97], and education [64,79]. Several studies collected demographic information but did not use it for providing the personalized intervention. User demographics formed an important part of the user model in 16 papers. Among these, some studies [7,41,54,74,79] did not explicitly state whether demographics was used for personalization or not.

Medical Data

In our review, 16 studies (aimed at rehabilitation, healthy lifestyle, and increasing PA) used medical data as part of their user model. Personalization was based on clinical symptoms [77,80], cholesterol levels [35,93], medical records [6,76,94], pain [23,48,52], and anaerobic threshold (the point between aerobic and anaerobic training of the user) along with HR and HR at rest [75]. It is unclear if the study by Mitchell et al [60] used the medical data for providing personalization. Sleep data were also collected by 2 studies [23,66]; however, whether it was directly used for providing personalization is not clear.

Behavior Change Technique Parameters

In our review, 30 studies used various BCT parameters such as stage of change [79]; subjective PA [3]; motivation [66]; skills, barriers, goals, and outcome expectations [3]; habit strength [7]; and rate of perceived exertion [60]. The MOPO study [41] based its personalization on a data attribute termed “life satisfaction,” which is a self-reported scale on happiness, interest in life, feeling of loneliness, and the ease of living. Various psychosocial parameters such as attitudes, intention, motivation, and confidence are also used along with stage of change [87,97].

Such BCT parameters were inferred using questionnaires such as the 20-item Weight Efficacy-Lifestyle Questionnaire and the 44-item Big Five Inventory Questionnaire that sought answers from users. Studies using BCT parameters had interventions that were knowledge-based, except in the studies by Hermens et al and Hales et al [54,74].

Contextual Information

Contextual information in the user models refers to any additional information that provided cues to the context and/or behavior of the user. The context of the user varied considerably across the 14 studies that used this type of information in our review. This category included user preferences, social media profile, location, time, mood, and energy levels among others.

Activity preferences of a user were generally obtained from the user to recommend a suitable activity to the user. All the 12 studies utilizing preferences also used PA profile in their user models. Preferences were also inferred based on users’ history and adherence to recommendations in the studies by Yom-Tov et al and Lim et al [9,86]. Location and time information was used to determine the feasibility of certain activity recommendations in the studies by Klein et al and Short et al [3,84]. For example, jogging may not be feasible during rainy weather. A study by Codreanu and Florea [80] used the estimated energy level (rested, fatigued, or energetic), defined by the “mood temperature factor.”

In our review, 4 papers used the social media profile to motivate users through activity status posts on websites or by inspiration from friends. Of these, 2 studies [32,95] used the profile to provide better recommendations and persuasion to users. On the other hand, the Active2Gether and PRO-Fit systems used social media in a direct way to generate social comparison [3] and recommend a fitness buddy to the user [56].

The TaylorActive app [66] used various indicators to gauge quality of life, perceived neighborhood environment, learning style, and delivery mode preference. All of these were measured using questionnaires provided to the user. Multimedia Appendix 2 summarizes the parameters across the 5 categories for the user models of the studies in our review.

Results of Individual Studies

The studies included in this review have diverse aims and, thus, different evaluation metrics. For our review, we have considered only the results relevant to PA of users. Not all studies in the review presented evaluations of their proposed interventions and not all of them evaluated a PA metric. In our review, only 27 studies presented evaluations of their proposed interventions

for PA. Out of these, 15 studies have reported positive statistically significant outcomes. The remaining 12 studies have not shown statistically significant results or shown no improvement at all. The impact and extent of the results vary in the studies as all are not randomized controlled trials (RCTs) and do not try to address similar questions. Table 3 shows the evaluation and results for the RCTs included in this review.

There are 20 RCTs (listed in Table 3), which evaluated their proposed interventions. Of these, some studies were evaluated on the basis of self-reported PA [7,47,83,84] and others used objectively measured PA through devices [35,41,66,93]. A metric used in these studies is MET-minutes or MET hours, which is the metabolic equivalent unit for energy expenditure. These MET minutes have been observed from self-reported data collected through questionnaires. We observe that not all studies report an improvement in PA after intervention as compared with the control group. In the study by De Cocker et al [64], the objectively measured sitting time has no significant difference; however, the self-reported data show significant difference between the intervention and control groups. The MOPO study [41] also reports significant change in self-rated fitness but no significant change in self-reported daily sitting time. On the other hand, studies such as the ones by Vandelanotte et al, Partridge et al, and Irvine et al [66,68,83] reported a significant improvement in PA of users after the intervention. The Reinforcement Learning(RL)-based messaging intervention [9] observed a significant improvement in the messages sent through the learned policy for the user in comparison with the initial random policy.

Some of the studies evaluated the difference in intervention delivery mediums. In the study by Peels et al [27], 2 kinds of personalized interventions were used—basic and environmental—where environmental intervention provided users with more contextual information, such as walking and cycling routes. In the study by Van Stralen et al [28], it was found that the printed interventions—basic as well as environmental—were significantly effective; however, the Web-based interventions were not. However, in the study by Walthouwer et al [71], no significant difference was observed when participants were provided interventions through the medium of their choice (text or video). Similarly, the study by Blake et al [108] observed no significant difference between delivery modes, email, and SMS. In addition, the study by Schulz et al [47] observed no statistical difference among the sequential intervention module delivery versus the simultaneous module delivery. Another study, the TaylorActive system [65], reported an increase in PA for all groups of intervention delivery—text, video, and combination.

Table 4 shows the evaluation of other studies (which are not RCTs), along with the methodology used for evaluation.

There are 7 studies in the review, which are not RCTs, but present some feasibility or usability analyses [74,81] or are observational [79,88] or single group studies [45,54,82]. Of these, some studies such as the studies by Hermens et al [54] and Coolbaugh et al [81] have very low sample size (8 and 2, respectively). The Personalized Coaching System study [54] conducted many different experiments. We consider the one mentioned in the paper, which aims to improve long-term activity behavior of chronic obstructive pulmonary disease patients. Moreover, 5 out of 8 patients had an improvement in activity level, although exercise capacity and health status show clinical improvement in 3 of these 5 patients. The feasibility of personalized PA prescription intervention [81] was tested on 2 users. Of these, 1 subject showed excellent adherence until week 10, but the other subject had inconsistent participation. These studies do not demonstrate the effectiveness of the interventions due to their low sample sizes. However, they provide directions toward potential feasible interventions for increasing PA.

Other studies such as the studies by Rabbi et al, Ammann et al, and Spark et al [45,79,82] show significant improvements in PA for their users. A different evaluation metric is used by Short et al [84], which evaluated habit strength of performing PA. The self-reported habit strength for PA increased, which has been considered as an effective improvement for PA intervention.

In both RCT and other studies, several studies have shown significant improvement in the PA of the participants due to personalized interventions. The study by Cook et al [30] showed a significant intervention effect with an increase in active commute and leisure time PA as well as PA in schools for the adolescents. The MyBehavior app evaluation study [45] also stated an increase in walking minutes and calories burnt in nonwalking exercises as compared with the baseline. A study for older adults [83] reported a positive impact on PA, with improvements in endurance, strengthening, stretching, and balance improvements. Similarly, I-Move [38] achieved a small but significant improvement in weekly minutes of MVPA. The study by Partridge et al [68] also reported a statistically significant increase in mean MET-minutes per week. In addition, the total walking days increased in the intervention group as compared with the control group. An increase in weekly minutes of MVPA was reported by Larsen et al [88]. They also reported an increase in the diversity of activities undertaken by participants as compared with the baseline.

A total of 2 studies have reported an improvement in self-reported values but have not observed the same for the objectively measured PA values [35,64]. Several of the studies do observe improvements in PA in the intervention groups; however, these are not statistically significant [9,41,47,74]. Furthermore, 1 study [71] tried to analyze the matched delivery preference and reported no intervention effect with a delivery method (video-only, text-only, or combined) of choice. It also reports that the video-only intervention did not see any improvements.

Table 3. Results of individual studies—randomized controlled trials.

Serial #	Paper reference	Dataset size	Variables evaluated	Results
1	Soetens et al [65]	803	Effect of time over increase in PA[a]	PA increases in all groups, time has no significant effect on all completers though has significant effect on those who had low baseline scores for total PA minutes (*P*<.001)
5	Storm et al [7]	790	Strength of habit for PA measured with abbreviated version of Self-Reported Habit Index, self-efficacy, and planning	Self-efficacy (*P*=.1), planning (*P*=.2), and habit strength (*P*=.006) improved in the intervention group
6	Schulz et al [47]	5055	Minutes of PA per day in control, sequential intervention module delivery, and simultaneous module delivery	No statistical difference in sequential and simultaneous delivery for PA or with respect to control group. Sequential delivery could be more effective than simultaneous module delivery after 12 months (*P*=.7)
13	Yom-Tov et al [9]	27	PA minutes per week, change in activity with message policy, change from initial to RL[b]-based learned policy	No statistical difference in treatment and control arm (*P*=.30) for PA minutes per week. Difference in change of activity between initial and learned message policy statistically significant (*P*=.004)
15	Cook et al [30]	555	PA (minutes per week) behavior difference at baseline and postmeasurement for 3 parameters: commuting, leisure time PA, and PA in school	Improvement found in leisure time MVPA[c] (*P*<.05), for increase in commute by bicycle (around 30 min) (*P*<.01) and total MVPA (*P*<.05)
17	Short et al [84]	724	Minutes per week of MVPA and resistance training score for all 3 arms–3 module interventions delivered monthly, weekly, or single-module	Significant improvement of MVPA across all groups (*P*<.05). Significant improvement in resistance score from monthly 3-module intervention to single module (*P*=.01)
21	Irvine et al [83]	368	Cardiovascular exercises, stretching exercises, strength exercises, balance exercises (all measured in minutes per week), and number of activities	Improvement in intervention group as compared with control in all (*P*<.001)
22	Friederichs et al [38]	4302	Minutes of MVPA per week and number of days ≥30 min activity in I-Move intervention, Active Plus intervention, and control group	I-Move had small but more significant effect than Active Plus in minutes of MVPA per week (*P*=.03 and *P*=.07). I-Move had medium sized effect and Active Plus had large size effect for number of days ≥30 min
23	Blake et al [108]	296	Active travel, moderate activity at work and recreation and vigorous activity at work and recreation in 2 arms for different delivery modes, both with tailored content, one with SMS[d] and another with email	No significant difference between email and SMS, but significant difference in moderate activity at work (hours per day), with email more effective than SMS (*P*=.24).
25	Hargreaves et al [93]	97	Step count	No difference at baseline and 12 weeks. Significant increase in step count of intervention group between week 12 and week 24 (*P*=.055) but not so significant in comparison group (*P*=.15)
30	Mistry et al [90]	337	PA between the 3 groups–standard care, generic message, and intervention group after 4 weeks	No significant difference between groups for change in PA (*P*>.05)
31	Peels et al [27]	1729	Number of MET[e] hours in 4 kinds of tailoring: printed, and Web-based (basic and environment-based in each) and control group	Printed (both basic and environmental) had statistically significant increase in MET hours (*P*=.025 and *P*=.31, respectively). No significant increase in both Web-based interventions (*P*=.59 and *P*=.887, respectively)
34	Pyky et al [41]	496	Self-rated health and fitness and leisure time PA	Changes in self-rated fitness and leisure time PA are associated with improved self-rated health (*P*<.026 and *P*<.04, respectively). No significant difference between intervention and control for self-reported daily sitting (*P*=.32) and light housework (but no other leisure time) PA (*P*=.43)
37	Marsaux et al [35]	1607	Objective PA in control group, group with personalized advice on diet and PA (L1 group), L1+phenotype (L2 group) and L2+genotype (L3 group)	No significant difference between control and any of the 3 groups in objective PA level measured (*P*=.73)

Serial #	Paper reference	Dataset size	Variables evaluated	Results
38	Alley et al [43]	154	PA (min per week) for 3 groups: control, tailoring only, and tailoring+video coaching group	Significant difference in PA between tailoring+video coaching versus control group (*P*=.01) but no significant difference in PA between the 2 intervention groups (*P*=.54)
39	Mitchell et al [60]	171	Sedentary time, LPA[f], and MVPA for intervention group with personalized step goals versus control group with generic advice	Decrease in sedentary time, Improvement in LPA and MVPA for both groups (*P*<.005).
41	De Cocker et al [64]	312	Sitting time in 3 groups: control, generic intervention, and tailored intervention	Self-reported total sitting time decreased more in tailored group compared with both generic group (*P*=.002) and control group (*P*=.002). But no significant difference in objectively measured data
47	Kattelmann et al [72]	1639	Total MET-minutes per week estimated from self-reported data	No difference between control and intervention for total MET-minutes per week (*P*=.90). Significant time effect for moderate MET-minutes per week (*P*=.002) and significant time × group × gender effect for vigorous MET-min per week (*P*=.05)
48	Partridge et al [68]	214	Self-reported PA data analyzed as MET-minutes per week	Significant effect of intervention on average MET minutes per week at 12 weeks (*P*=.05). Total PA days (*P*=.003) and number of walking days (*P*=.02) increased in intervention group
49	Walthouwer et al [71]	1419	PA duration in text-tailored, video-tailored, and control arm. In the tailoring group, 2 groups were compared, 1 where preference of user to video/text was matched and another without the matching	No significant difference in condition match/mismatch for PA (*P*=.33). Also, no significant difference for video-tailoring × intervention used (*P*=.83) and text-tailoring × intervention used (*P*=.81)

[a]PA: physical activity.

[b]RL: reinforcement learning

[c]MVPA: moderately vigorous physical activity.

[d]SMS: short messaging service.

[e]MET: metabolic equivalent.

[f]LPA: light physical activity.

Table 4. Results of individual studies—nonrandomized controlled trials.

Serial #	Paper reference	Method of study design	Dataset size	Variables evaluated	Result
7	Hermens et al [54]	Single-case experimental study	8	Objectively measured activity behavior (activity level)	5 patients had increased PA[a] level
11	Rabbi et al [45]	Single case experiment with multiple baseline	16	Minutes of walking per day and calories burnt in nonwalking exercise per day	Intervention had significant effect for walking (P<.005) and exercise (P<.05)
16	Larsen et al [88]	Observational study	21	Change in minutes of MVPA[b] using a semistructured interview among adolescent girls after 12 weeks	Statistically significant increase in weekly minutes of MVPA (P<.001). Also reported activity types had larger variation than baseline
24	Coolbaugh et al [81]	Feasibility study	2	12 weeks of personalized intervention	Feasibility could not be ascertained
33	Ammann et al [79]	Observational study	803	Weekly total PA minutes across young, middle age, and old age groups	Significant increase in MVPA from baseline for older adults (P<.5). All age groups increased weekly PA significantly (P<.05) and walking minutes (P<.01) over time in intention-to-treat analysis
44	Hales et al [74]	Pilot study and iterative usability study	9	Calories spent during intentional activity for users as compared with baseline	Calories expended increased from baseline but not statistically significant (P=.57)
46	Spark et al [82]	Single group, pre- and post-test study	29	Duration of MVPA for participants in initial intervention (6 months), followed by extended contact information (6-12 months) and no contact follow up (12-18 months)	Significant improvement in minutes/day MVPA to 6 months from baseline (P=.006) and to 18 months from baseline (P=.003)

[a]PA: physical activity.

[b]MVPA: moderately vigorous physical activity.

Discussion

Principal Findings

This study provides a review of studies on personalized technology–based interventions for increasing PA. This review adds to the PA literature in several ways. It provides an overview of personalization provided to users in the context of apps that aim at increasing PA. It examines various attributes, which can be personalized for encouraging the user, and identifies the theoretical frameworks used in these studies. This review included all research designs and, thus, provides a comprehensive view of ideas for effectively encouraging PA by means of personalization. We now discuss the review implications with respect to interventions, personalization, user models, theory and guidelines, and results.

Interventions

The widespread adoption of activity monitoring devices, increasing accuracy of data-driven prediction techniques, and ease of automation all facilitate the use of automated interventions. However, PA changes in patients who are under clinical observation may need to be assessed by a health care expert, leading to manual interventions.

Semiautomated systems combine and thereby aim to provide the best of both worlds—automated and manual interventions. Though these are often specialized for patients [75,80], they can also be available for the general user [2,35]. Having a health care expert–based intervention is less scalable but often necessary for patients under specific medical treatments. An interesting case of semiautomation is seen in the study by Dobrican and Zampunieris [75], where the targets were cardiac patients and the aim was rehabilitation. The doctor was involved for medical advice, but adaptive goals were set based on the European Society of Cardiology guidelines [109]. Note that there are arguments suggested against completely automated systems, for example, they have not been effective in weight loss [2].

Commercial fitness apps designed for the general user could take into account specific requirements of users with clinical conditions, including chronic diseases such as diabetes, who may benefit from such interventions. Current systems would need to include adaptive goal recommendation [54] to offer personalization in light of medical constraints and not just preferences of the users (eg, no swimming for elderly patients). From this review, we observe that user models for patients with chronic diseases are similar. PA guidelines, such as European Society of Cardiology's guidelines [109] for cardiac patients or by Canadian Diabetes Association [110] and BCT–based design could be incorporated to enable effective behavior change.

Personalization

Interventions in the included studies were personalized in one or more ways. Recommendations were personalized with respect to goals, activity, or fitness partners. Feedback was personalized

with respect to its educational or motivational content and, in some cases, its timing.

Personalization was done either individually or in a category-based manner. The former includes individual models, for example, based on a user's lifestyle [86], rate of progression [81], and preferences [32] or determined by a health care specialist [80]. In the latter case, category-specific personalization was provided after identifying the most appropriate category for the user. The categories were defined based on BCT [eg, 3] or activity status [eg, 2].

User Models

User models were created using a variety of different measurements, that is, PA profiles, demographics, medical data, BCT parameters, and contextual information.

Various parameters were used to evaluate PA, and all the profiles aimed to measure 1 or more "dimensions" of PA. An interesting visualization of multidimensional PA was proposed in 1 study [111]. The premise is that PA cannot be judged only on 1 criterion, for example, number of steps or time of vigorous activity, and has multiple dimensions including sedentary time. All the interventions for PA were restricted in their dimensions, and a multidimensional profile would be useful to obtain a holistic view of the user.

User models based on social profiles used the least amount of other contextual parameters. They promoted behavior change through social influence and are promising for both effective persuasion and user modeling. Among the included studies, social profiles were used for buddy matching [74] and also to post status data on social media to promote PA.

Personalized and dynamic user models can be created using the wealth of multimodal user data available from smartphones. Most of the existing apps do not use all the available data. PRO-fit utilized some of the available data sources—the phones' geolocation, the users' social network, and the users' calendar—effectively [56]. By integrating all the available data, a richer profile can be created, and when combined with reinforcement learning techniques, the most effective interventional policies for each user can be learnt. As user behavior may change over time, it is important to employ online learning algorithms that can continuously monitor user models, adapt to their changing lifestyle patterns, and accordingly modify interventions as well.

Theory and Guidelines

Theory-based studies used BCTs to only make design decisions. Furthermore, 1 study did not completely define all the phases of TTM during the design process but utilized the readiness parameter defined by the model [45]. In addition, incorporation of BCTs was usually done via questionnaires in these studies, which may be infeasible or obtrusive to the user. Thus, automated learning of BCT parameters may be worth exploring. There is preliminary work in this direction. A user's awareness depends on both the actual and perceived behavior [3]. A study that personalized messages using reinforcement learning concluded that the difference in users' exercise on a given day could be learnt by the learning algorithm, thus making user behavior predictable [9]. The methodology of utilizing activities of daily life for profiling users and their behavior [86] is another approach for estimating user behavior. User preferences could also be learnt through greedy approaches [86] or through inherent model design [45].

Another problem with methodologies based on BCTs is that they generally set a fixed ideal goal for a user. In contrast, PA guidelines suggest PA progression to prevent fatigue or muscular injuries. The generic goals of 60 min of PA or 10,000 steps may be too difficult and hence demotivating to a user who is sedentary or has clinical complications. Such users often require help, in the form of intermediate goals, to reach the final goal. PA guidelines can be utilized in such cases. There are attempts in studies [87] to use PA guidelines while using BCT for motivating users. Another study [78] also encourages its users, that is, cancer patients, to follow guidelines set by American Cancer Association [110] while planning their PA.

As identified across the PA literature, an "intention-behavior" gap exists among users. This poses the classic problem that although users are motivated and have intentions to increase their PA, they are not sufficiently active. Many studies were based on BCTs. However, healthy lifestyle induced during the intervention does not ensure that the user does not go back to a sedentary lifestyle after the intervention [54]. The sustained effects of interventions were not evaluated by all the studies but only by a few studies (e.g. [45,60,82]). Habit strength and formation has been addressed and evaluated in the study by Storm et al [7]. It is important that the sustained long-term effects of intervention are analyzed, as it would help to identify effective methods of promoting PA.

Results of Individual Studies

Direct positive results demonstrating the effectiveness of personalized interventions have been observed in a diverse set of studies—there are studies implementing data-driven automated systems [45], which recommend activities, whereas there are also studies which provide only personalized educational and motivational content [30]. These results indicate that activity or definite goal recommendation is not required for an effective personalized PA intervention. Effectively personalized motivational and educational content can help induce behavior change among the participants as well. It is also interesting to note that most of the studies with significant improvements are based on theoretical models (e.g.[7,30,54,88]). However, most of these studies also use self-reported values and collect data through questionnaires (e.g.[79,84,108]).

The self-reported PA values need to be considered with caution. As observed in the studies by Marsaux et al and De Cocker et al [35,64], there can be different results when self-reported and objectively measured data are compared. Thus, positive results obtained by interventions based on self-reported data need to be evaluated with objectively measured data through accelerometers and sensors. However, it can be inferred by the positive results obtained through BCT that their incorporation could help users, even if the users' perceived PA level is incorrect. This makes it worthwhile to find ways of incorporating BCTs and theoretical guidelines in other data-driven–based interventions. However, it also needs to be

noted that even the studies that do not show significant improvements use BCTs [47,74,90].

Studies have evaluated not just the PA metrics but also the intervention delivery mediums. The intervention delivery was found to not matter in the cases of video versus text [71] and SMS versus email [108]. However, a difference was found in the case of print versus Web-based intervention [27]. It is possible that the print medium was found to be effective as the participants were adults over the age of 50 years. However, further studies need to be performed to analyze the differences between intervention delivery mediums and their effects on the users.

The sample sizes of the studies reviewed vary considerably. The studies have also not been analyzed for quality to recommend future directions. However, our review indicates that there is scope for more rigorous evaluation in terms of intervention delivery, personalization, and intervention method. Many studies in the review perform pilot studies or feasibility studies or identify RCT protocols, which are yet to be completely evaluated. Evaluation of various systems to identify the effectiveness of intervention medium (along with the personalization aspect) in motivating users could be useful.

Limitations

This review was restricted to specific databases and an appropriate search query. It is possible that some studies may have been left out due to their journal or indexing bias. In addition, the search was restricted to a time frame that was considered relevant for the personalization aspect of the study and could again have led to studies being left out of the review. Moreover, as this is a scoping review, we have included studies without quality analysis and also studies without any evaluation. Though it helps identify the breadth of research, as the quality of studies is not assessed, the gaps identified may not be completely accurate.

Conclusions

This study provides a comprehensive review of personalized technology–based interventions, as recommendations or feedback, for promoting PA. Overall, the studies show that these interventions for increasing PA are more effective when they are personalized, compared with a “one size fits all” generic advice. Gaps have been identified in several aspects, such as in the development of a multidimensional user model and the use of behavioral theory in automated personalization. On the basis of these gaps, research directions for improving the efficacy of personalized technology–based interventions have been suggested.

Acknowledgments

This work was supported by the Ministry of Education, Singapore (grants R-253-000-129-114 and R-253-000-138-133).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Overview of studies in the review.

[PDF File (Adobe PDF File), 97KB - mhealth_v7i1e11098_app1.pdf]

Multimedia Appendix 2

User model parameters for studies in the review.

[PDF File (Adobe PDF File), 83KB - mhealth_v7i1e11098_app2.pdf]

References

1. World Health Organization. Physical activity URL: http://www.who.int/topics/physical_activity/en/ [accessed 2018-10-10] [WebCite Cache ID 6yMEoVUW3]
2. Martin CK, Gilmore LA, Apolzan JW, Myers CA, Thomas DM, Redman LM. Smartloss: a personalized mobile health intervention for weight management and health promotion. JMIR Mhealth Uhealth 2016 Mar 16;4(1):e18 [FREE Full text] [doi: 10.2196/mhealth.5027] [Medline: 26983937]
3. Klein MC, Manzoor A, Mollee JS. Active2Gether: A personalized m-health intervention to encourage physical activity. Sensors (Basel) 2017 Jun 19;17(6) [FREE Full text] [doi: 10.3390/s17061436] [Medline: 28629178]
4. Forastiere M, De Pietro G, Sannino G. An mHealth Application for a Personalized Monitoring of One’s Own Wellness: Design and Development. In: Smart Innovation, Systems and Technologies.: Springer, Cham; 2016 Presented at: International Conference on Innovation in Medicine and Healthcare 2016; 15-17 June; Puerto de la Cruz, Spain p. 269-278. [doi: 10.1007/978-3-319-39687-3_26]
5. Kerner C, Goodyear V. The motivational impact of wearable healthy lifestyle technologies: a self-determination perspective on Fitbits with adolescents. Am J Heal Educ 2017 Jul 28;48(5):287-297. [doi: 10.1080/19325037.2017.1343161]

6. Mukhtar H. Using persuasive recommendations in wellness applications based upon user activities. Int J Adv Comput Sci Appl 2016 Jan;7(8):349-356. [doi: 10.14569/IJACSA.2016.070847]
7. Storm V, Dörenkämper J, Reinwand DA, Wienert J, De Vries H, Lippke S. Effectiveness of a web-based computer-tailored multiple-lifestyle intervention for people interested in reducing their cardiovascular risk: a randomized controlled trial. J Med Internet Res 2016 Apr 11;18(4):e78 [FREE Full text] [doi: 10.2196/jmir.5147] [Medline: 27068880]
8. Kroeze W, Werkman A, Brug J. A systematic review of randomized trials on the effectiveness of computer-tailored education on physical activity and dietary behaviors. Ann Behav Med 2006 Jun;31(3):205-223. [doi: 10.1207/s15324796abm3103_2] [Medline: 16700634]
9. Yom-Tov E, Feraru G, Kozdoba M, Mannor S, Tennenholtz M, Hochberg I. Encouraging physical activity in patients with diabetes: intervention using a reinforcement learning system. J Med Internet Res 2017 Oct 10;19(10):e338 [FREE Full text] [doi: 10.2196/jmir.7994] [Medline: 29017988]
10. Joseph RP, Durant NH, Benitez TJ, Pekmezi DW. Internet-based physical activity interventions. Am J Lifestyle Med 2014 Dec;8(1):42-68 [FREE Full text] [doi: 10.1177/1559827613498059] [Medline: 25045343]
11. Baker PR, Francis DP, Soares J, Weightman AL, Foster C. Community wide interventions for increasing physical activity. Cochrane Database Syst Rev 2015;1:CD008366. [doi: 10.1002/14651858.CD008366.pub3] [Medline: 25556970]
12. Rogers MA, Lemmen K, Kramer R, Mann J, Chopra V. Internet-delivered health interventions that work: systematic review of meta-analyses and evaluation of website availability. J Med Internet Res 2017 Mar 24;19(3):e90 [FREE Full text] [doi: 10.2196/jmir.7111] [Medline: 28341617]
13. Baker P, Dobbins M, Soares J, Francis D, Weightman A, Costello J. Public health interventions for increasing physical activity in children, adolescents and adults: an overview of systematic reviews. Cochrane Database Syst Rev 2015;8(1):57-58. [doi: 10.1002/14651858.CD011454]
14. Saunders DH, Sanderson M, Brazzelli M, Greig CA, Mead GE. Physical fitness training for stroke patients. Cochrane Database Syst Rev 2013 Oct 21(10):CD003316. [doi: 10.1002/14651858.CD003316.pub5] [Medline: 24142492]
15. Duff OM, Walsh DM, Furlong BA, O'Connor NE, Moran KA, Woods CB. Behavior change techniques in physical activity eHealth interventions for people with cardiovascular disease: systematic review. J Med Internet Res 2017 Aug 02;19(8):e281 [FREE Full text] [doi: 10.2196/jmir.7782] [Medline: 28768610]
16. Claes J, Buys R, Woods C, Briggs A, Geue C, Aitken M, et al. PATHway I: design and rationale for the investigation of the feasibility, clinical effectiveness and cost-effectiveness of a technology-enabled cardiac rehabilitation platform. BMJ Open 2017 Jun 30;7(6):e016781 [FREE Full text] [doi: 10.1136/bmjopen-2017-016781] [Medline: 28667228]
17. op den Akker H, Jones VM, Hermens HJ. Tailoring real-time physical activity coaching systems: a literature survey and model. User Model User-Adap Inter 2014 Jun 25;24(5):351-392. [doi: 10.1007/s11257-014-9146-y]
18. Fan H, Poole MS. What is personalization? Perspectives on the design and implementation of personalization in information systems. J Org Comp Elect Com 2006 Jan;16(3-4):179-202. [doi: 10.1080/10919392.2006.9681199]
19. Hawkins RP, Kreuter M, Resnicow K, Fishbein M, Dijkstra A. Understanding tailoring in communicating about health. Health Educ Res 2008 Jun;23(3):454-466 [FREE Full text] [doi: 10.1093/her/cyn004] [Medline: 18349033]
20. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
21. Oosterom-Calo R, Te Velde SJ, Stut W, Brug J. Development of Motivate4Change using the intervention mapping protocol: an interactive technology physical activity and medication adherence promotion program for hospitalized heart failure patients. JMIR Res Protoc 2015;4(3):e88 [FREE Full text] [doi: 10.2196/resprot.4282] [Medline: 26195072]
22. Liu C, Chan C. A fuzzy logic prompting mechanism based on pattern recognition and accumulated activity effective index using a smartphone embedded sensor. Sensors (Basel) 2016 Aug 19;16(8):1322 [FREE Full text] [doi: 10.3390/s16081322] [Medline: 27548184]
23. Janols R, Guerrero E, Lindgren H. A pilot study on personalised coaching to increase older adults' physical and social activities. In: Advances in Intelligent Systems and Computing.: Springer, Cham; 2017 Presented at: Ambient Intelligence–Software and Applications – 8th International Symposium on Ambient Intelligence (ISAmI 2017); 21-23 June; Porto, Portugal p. 140-148. [doi: 10.1007/978-3-319-61118-1_18]
24. Lindgren H, Guerrero E, Janols R. Personalised Persuasive Coaching to Increase Older Adults' Physical and Social Activities: A Motivational Model. In: Advances in Practical Applications of Cyber-Physical Multi-Agent Systems: The PAAMS Collection.: Springer, Cham; 2017 Presented at: 15th International Conference on Practical Applications of Agents and Multi-Agent Systems; June 21-23, 2017; Porto, Portugal p. 182. [doi: 10.1007/978-3-319-59930-4_14]
25. Boekhout JM, Berendsen BA, Peels DA, Bolman CA, Lechner L. Evaluation of a computer-tailored healthy ageing intervention to promote physical activity among single older adults with a chronic disease. Int J Environ Res Public Health 2018 Dec 15;15(2):346 [FREE Full text] [doi: 10.3390/ijerph15020346] [Medline: 29462862]
26. Peels DA, de Vries H, Bolman C, Golsteijn RH, van Stralen MM, Mudde AN, et al. Differences in the use and appreciation of a web-based or printed computer-tailored physical activity intervention for people aged over 50 years. Health Educ Res 2013 Aug;28(4):715-731 [FREE Full text] [doi: 10.1093/her/cyt065] [Medline: 23784076]
27. Peels DA, Hoogenveen RR, Feenstra TL, Golsteijn RH, Bolman C, Mudde AN, et al. Long-term health outcomes and cost-effectiveness of a computer-tailored physical activity intervention among people aged over fifty: modelling the results

of a randomized controlled trial. BMC Public Health 2014 Oct 23;14:1099 [FREE Full text] [doi: 10.1186/1471-2458-14-1099] [Medline: 25342517]
28. Van Stralen MM, Kok G, de Vries H, Mudde AN, Bolman C, Lechner L. The Active plus protocol: systematic development of two theory- and evidence-based tailored physical activity interventions for the over-fifties. BMC Public Health 2008 Dec 04;8:399 [FREE Full text] [doi: 10.1186/1471-2458-8-399] [Medline: 19055806]
29. Klein MC, Manzoor A, Middelweerd A, Mollee JS, te Velde SJ. Encouraging physical activity via a personalized mobile system. IEEE Internet Comput 2015 Jul;19(4):20-27. [doi: 10.1109/MIC.2015.51]
30. Cook TL, De Bourdeaudhuij I, Maes L, Haerens L, Grammatikaki E, Widhalm K, et al. Moderators of the effectiveness of a web-based tailored intervention promoting physical activity in adolescents: the HELENA Activ-O-Meter. J Sch Health 2014 Apr;84(4):256-266. [doi: 10.1111/josh.12140] [Medline: 24617909]
31. De Bourdeaudhuij I, Maes L, De Henauw S, De Vriendt T, Moreno LA, Kersting M, HELENA Study Group. Evaluation of a computer-tailored physical activity intervention in adolescents in six European countries: the Activ-O-Meter in the HELENA intervention study. J Adolesc Health 2010 May;46(5):458-466. [doi: 10.1016/j.jadohealth.2009.10.006] [Medline: 20413082]
32. Fahim M, Idris M, Ali R, Nugent C, Kang B, Huh E, et al. ATHENA: a personalized platform to promote an active lifestyle and wellbeing based on physical, mental and social health primitives. Sensors (Basel) 2014 May 23;14(5):9313-9329 [FREE Full text] [doi: 10.3390/s140509313] [Medline: 24859031]
33. Ali R, Siddiqi M, Kang B, Lee S. KARE: A Hybrid Reasoning Approach for Promoting Active Lifestyle. In: Proceedings of the 9th International Conference on Ubiquitous Information Management and Communication. 2015 Presented at: 9th International Conference on Ubiquitous Information Management and Communication; January 08 - 10, 2015; Bali, Indonesia. [doi: 10.1145/2701126.2701156]
34. Marsaux CF, Celis-Morales C, Fallaize R, Macready AL, Kolossa S, Woolhead C, et al. Effects of a web-based personalized intervention on physical activity in European adults: a randomized controlled trial. J Med Internet Res 2015 Oct 14;17(10):e231 [FREE Full text] [doi: 10.2196/jmir.4660] [Medline: 26467573]
35. Marsaux CF, Celis-Morales C, Livingstone KM, Fallaize R, Kolossa S, Hallmann J, et al. Changes in physical activity following a genetic-based internet-delivered personalized intervention: randomized controlled trial (Food4Me). J Med Internet Res 2016 Feb 05;18(2):e30 [FREE Full text] [doi: 10.2196/jmir.5198] [Medline: 26851191]
36. Celis-Morales C, Livingstone KM, Marsaux CF, Forster H, O'Donovan CB, Woolhead C, et al. Design and baseline characteristics of the Food4Me study: a web-based randomised controlled trial of personalised nutrition in seven European countries. Genes Nutr 2015 Jan;10(1):450 [FREE Full text] [doi: 10.1007/s12263-014-0450-2] [Medline: 25491748]
37. Friederichs SA, Oenema A, Bolman C, Guyaux J, Van Keulen HM, Lechner L. I Move: systematic development of a web-based computer tailored physical activity intervention, based on motivational interviewing and self-determination theory. BMC Public Health 2014;14:212 [FREE Full text] [doi: 10.1186/1471-2458-14-212] [Medline: 24580802]
38. Friederichs SA, Oenema A, Bolman C, Lechner L. Motivational interviewing and self-determination theory in a web-based computer tailored physical activity intervention: a randomized controlled trial. Psychol Health 2016 Aug;31(8):907-930. [doi: 10.1080/08870446.2016.1151018] [Medline: 26849996]
39. Ahola R, Pyky R, Jämsä T, Mäntysaari M, Koskimäki H, Ikäheimo TM, et al. Gamified physical activation of young men--a Multidisciplinary Population-Based Randomized Controlled Trial (MOPO study). BMC Public Health 2013;13:32 [FREE Full text] [doi: 10.1186/1471-2458-13-32] [Medline: 23311678]
40. Jauho A, Pyky R, Ahola R, Kangas M, Virtanen P, Korpelainen R, et al. Effect of wrist-worn activity monitor feedback on physical activity behavior: a randomized controlled trial in Finnish young men. Prev Med Rep 2015;2:628-634 [FREE Full text] [doi: 10.1016/j.pmedr.2015.07.005] [Medline: 26844128]
41. Pyky R, Koivumaa-Honkanen H, Leinonen A, Ahola R, Hirvonen N, Enwald H, et al. Effect of tailored, gamified, mobile physical activity intervention on life satisfaction and self-rated health in young adolescent men: a population-based, randomized controlled trial (MOPO study). Comput Human Behav 2017 Jul;72:13-22. [doi: 10.1016/j.chb.2017.02.032]
42. Alley S, Jennings C, Plotnikoff RC, Vandelanotte C. My Activity Coach - using video-coaching to assist a web-based computer-tailored physical activity intervention: a randomised controlled trial protocol. BMC Public Health 2014;14:738 [FREE Full text] [doi: 10.1186/1471-2458-14-738] [Medline: 25047900]
43. Alley S, Jennings C, Plotnikoff RC, Vandelanotte C. Web-based video-coaching to assist an automated computer-tailored physical activity intervention for inactive adults: a randomized controlled trial. J Med Internet Res 2016 Aug 12;18(8):e223 [FREE Full text] [doi: 10.2196/jmir.5664] [Medline: 27520283]
44. Rabbi M, Pfammatter A, Zhang M, Spring B, Choudhury T. Automated personalized feedback for physical activity and dietary behavior change with mobile phones: a randomized controlled trial on adults. JMIR Mhealth Uhealth 2015;3(2):e42 [FREE Full text] [doi: 10.2196/mhealth.4160] [Medline: 25977197]
45. Rabbi M, Aung MH, Zhang M, Choudhury T. MyBehavior: automatic personalized health feedback from user behaviors and preferences using smartphones. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2015 Presented at: International Joint Conference on Pervasive and Ubiquitous Computing; September 07 - 11, 2015; Osaka, Japan p. 707-718. [doi: 10.1145/2750858.2805840]

46. Schulz DN, Kremers SP, van Osch LA, Schneider F, van Adrichem MJ, de Vries VH. Testing a Dutch web-based tailored lifestyle programme among adults: a study protocol. BMC Public Health 2011;11:108 [FREE Full text] [doi: 10.1186/1471-2458-11-108] [Medline: 21324181]
47. Schulz DN, Kremers SP, Vandelanotte C, van Adrichem MJ, Schneider F, Candel MJ, et al. Effects of a web-based tailored multiple-lifestyle intervention for adults: a two-year randomized controlled trial comparing sequential and simultaneous delivery modes. J Med Internet Res 2014;16(1):e26 [FREE Full text] [doi: 10.2196/jmir.3094] [Medline: 24472854]
48. Williams QI, Gunn AH, Beaulieu JE, Benas BC, Buley B, Callahan LF, et al. Physical therapy vs. internet-based exercise training (PATH-IN) for patients with knee osteoarthritis: study protocol of a randomized controlled trial. BMC Musculoskelet Disord 2015 Sep 28;16:264 [FREE Full text] [doi: 10.1186/s12891-015-0725-9] [Medline: 26416025]
49. Brooks MA, Beaulieu JE, Severson HH, Wille CM, Cooper D, Gau JM, et al. Web-based therapeutic exercise resource center as a treatment for knee osteoarthritis: a prospective cohort pilot study. BMC Musculoskelet Disord 2014 May 17;15:158 [FREE Full text] [doi: 10.1186/1471-2474-15-158] [Medline: 24884547]
50. Chatzitofis A, Zarpalas D, Filos D, Triantafyllidis A, Chouvarda I, Maglaveras N. Technological Module for Unsupervised, Personalized Cardiac Rehabilitation Exercising. : IEEE; 2017 Presented at: 2017 IEEE 41st Annual Computer Software and Applications Conference (COMPSAC); 4-8 July 2017; Turin, Italy p. 125. [doi: 10.1109/COMPSAC.2017.230]
51. Claes J, Buys R, Woods C, Briggs A, Geue C, Aitken M, et al. PATHway I: design and rationale for the investigation of the feasibility, clinical effectiveness and cost-effectiveness of a technology-enabled cardiac rehabilitation platform. BMJ Open 2017 Jun 30;7(6):e016781 [FREE Full text] [doi: 10.1136/bmjopen-2017-016781] [Medline: 28667228]
52. Triantafyllidis A, Filos D, Buys R, Claes J, Cornelissen V, Kouidi E, et al. Computerized decision support for beneficial home-based exercise rehabilitation in patients with cardiovascular disease. Comput Methods Programs Biomed 2018 Aug;162:1-10. [doi: 10.1016/j.cmpb.2018.04.030] [Medline: 29903475]
53. Cabrita M, Akker HD, Achterkamp R, Hermens H, Vollenbroek-Hutten M. Automated Personalized Goal-setting in an Activity Coaching Application. In: Proceedings of the 3rd International Conference on Sensor Networks - Volume 1: SENSORNETS. 2014 Presented at: SensorNets 2014; Jan 2014; Lisbon p. 389-396. [doi: 10.5220/0004878703890396]
54. Hermens H, Op den Akker H, Tabak M, Wijsman J, Vollenbroek M. Personalized Coaching Systems to support healthy behavior in people with chronic conditions. J Electromyogr Kinesiol 2014 Dec;24(6):815-826. [doi: 10.1016/j.jelekin.2014.10.003] [Medline: 25455254]
55. Op den Akker H, Cabrita M, Op den Akker R, Jones VM, Hermens HJ. Tailored motivational message generation: a model and practical framework for real-time physical activity coaching. J Biomed Inform 2015 Jun;55:104-115 [FREE Full text] [doi: 10.1016/j.jbi.2015.03.005] [Medline: 25843359]
56. Dharia S, Jain V, Patel J, Vora J, Yamauchi R, Eirinaki M. PRO-Fitxercise with friends. In: Proceedings of the 2016 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. 2016 Presented at: ASONAM '16; August 18 - 21, 2016; Davis, California. [doi: 10.1109/ASONAM.2016.7752437]
57. Dharia S, Eirinaki M, Jain V, Patel J, Varlamis I, Vora J, et al. Social recommendations for personalized fitness assistance. Pers Ubiquit Comput 2017 Jun 29;22(2):245-257. [doi: 10.1007/s00779-017-1039-8]
58. Dharia S, Jain V, Patel J, Vora J, Chawla S, Eirinaki M. PRO-Fit: A personalized fitness assistant framework. 2016 Presented at: The 28th International Conference on Software Engineering & Knowledge Engineering; July 1 - July 3, 2016; San Francisco Bay, California, USA. [doi: 10.18293/SEKE2016-174]
59. Mitchell BL, Lewis NR, Smith AE, Rowlands AV, Parfitt G, Dollman J. Rural Environments and Community Health (REACH): a randomised controlled trial protocol for an online walking intervention in rural adults. BMC Public Health 2014 Sep 18;14:969 [FREE Full text] [doi: 10.1186/1471-2458-14-969] [Medline: 25236776]
60. Mitchell BL, Smith AE, Rowlands AV, Fraysse F, Parfitt G, Lewis NR, et al. Promoting physical activity in rural Australian adults using an online intervention. J Sci Med Sport 2018 Jul 10 Epub ahead of the print (forthcoming) [FREE Full text] [doi: 10.1016/j.jsams.2018.07.002] [Medline: 30031747]
61. Reinwand D, Kuhlmann T, Wienert J, de Vries H, Lippke S. Designing a theory- and evidence-based tailored eHealth rehabilitation aftercare program in Germany and the Netherlands: study protocol. BMC Public Health 2013;13:1081 [FREE Full text] [doi: 10.1186/1471-2458-13-1081] [Medline: 24245493]
62. Martin CK, Miller AC, Thomas DM, Champagne CM, Han H, Church T. Efficacy of SmartLoss, a smartphone-based weight loss intervention: results from a randomized controlled trial. Obesity (Silver Spring) 2015 May;23(5):935-942. [doi: 10.1002/oby.21063] [Medline: 25919921]
63. De Cocker K, De Bourdeaudhuij I, Cardon G, Vandelanotte C. Theory-driven, web-based, computer-tailored advice to reduce and interrupt sitting at work: development, feasibility and acceptability testing among employees. BMC Public Health 2015;15(1):959 [FREE Full text] [doi: 10.1186/s12889-015-2288-y] [Medline: 26404915]
64. De Cocker K, De Bourdeaudhuij I, Cardon G, Vandelanotte C. The effectiveness of a web-based computer-tailored intervention on workplace sitting: a randomized controlled trial. J Med Internet Res 2016 May 31;18(5):e96 [FREE Full text] [doi: 10.2196/jmir.5266] [Medline: 27245789]
65. Soetens KC, Vandelanotte C, de Vries VH, Mummery KW. Using online computer tailoring to promote physical activity: a randomized trial of text, video, and combined intervention delivery modes. J Health Commun 2014 Dec;19(12):1377-1392. [doi: 10.1080/10810730.2014.894597] [Medline: 24749983]

66. Vandelanotte C, Short C, Plotnikoff RC, Hooker C, Canoy D, Rebar A, et al. TaylorActive--Examining the effectiveness of web-based personally-tailored videos to increase physical activity: a randomised controlled trial protocol. BMC Public Health 2015 Oct 05;15:1020 [FREE Full text] [doi: 10.1186/s12889-015-2363-4] [Medline: 26438225]
67. Hebden L, Balestracci K, McGeechan K, Denney-Wilson E, Harris M, Bauman A, et al. 'TXT2BFiT' a mobile phone-based healthy lifestyle program for preventing unhealthy weight gain in young adults: study protocol for a randomized controlled trial. Trials 2013;14:75 [FREE Full text] [doi: 10.1186/1745-6215-14-75] [Medline: 23506013]
68. Partridge SR, McGeechan K, Hebden L, Balestracci K, Wong AT, Denney-Wilson E, et al. Effectiveness of a mHealth lifestyle program with telephone support (TXT2BFiT) to prevent unhealthy weight gain in young adults: randomized controlled trial. JMIR Mhealth Uhealth 2015;3(2):e66 [FREE Full text] [doi: 10.2196/mhealth.4530] [Medline: 26076688]
69. Walthouwer MJ, Oenema A, Lechner L, de Vries H. Use and effectiveness of a video- and text-driven web-based computer-tailored intervention: randomized controlled trial. J Med Internet Res 2015 Sep 25;17(9):e222 [FREE Full text] [doi: 10.2196/jmir.4496] [Medline: 26408488]
70. Walthouwer MJ, Oenema A, Soetens K, Lechner L, de Vries H. Implementation of web-based interventions by Dutch occupational health centers. Health Promot Int 2017 Oct 01;32(5):818-830. [doi: 10.1093/heapro/daw014] [Medline: 27020573]
71. Walthouwer MJ, Oenema A, Soetens K, Lechner L, De Vries H. Systematic development of a text-driven and a video-driven web-based computer-tailored obesity prevention intervention. BMC Public Health 2013;13:978 [FREE Full text] [doi: 10.1186/1471-2458-13-978] [Medline: 24138937]
72. Kattelmann KK, White AA, Greene GW, Byrd-Bredbenner C, Hoerr SL, Horacek TM, et al. Development of Young Adults Eating and Active for Health (YEAH) internet-based intervention via a community-based participatory research model. J Nutr Educ Behav 2014;46(2):S10-S25. [doi: 10.1016/j.jneb.2013.11.006] [Medline: 24456834]
73. Kattelmann KK, Bredbenner CB, White AA, Greene GW, Hoerr SL, Kidd T, et al. The effects of Young Adults Eating and Active for Health (YEAH): a theory-based Web-delivered intervention. J Nutr Educ Behav 2014;46(6):S27-S41. [doi: 10.1016/j.jneb.2014.08.007] [Medline: 25457733]
74. Hales S, Turner-McGrievy G, Fahim A, Freix A, Wilcox S, Davis RE, et al. A mixed-methods approach to the development, refinement, and pilot testing of social networks for improving healthy behaviors. JMIR Hum Factors 2016 Feb 12;3(1):e8 [FREE Full text] [doi: 10.2196/humanfactors.4512] [Medline: 27026535]
75. Dobrican R, Zampunieris D. A Proactive Solution, using Wearable and Mobile Applications, for Closing the Gap between the Rehabilitation Team and Cardiac Patients. : IEEE; 2016 Presented at: IEEE International Conference on Healthcare Informatics (ICHI); 4-7 Oct, 2016; Chicago, IL, USA p. 146. [doi: 10.1109/ICHI.2016.23]
76. Tseng J, Lin BH, Lin YF, Tseng V, Day ML, Wang SH, et al. An interactive healthcare system with personalized diet and exercise guideline recommendation. : IEEE; 2015 Presented at: Conference on Technologies and Applications of Artificial Intelligence (TAAI); 20-22 Nov, 2015; Tainan, Taiwan. [doi: 10.1109/TAAI.2015.7407106]
77. Ahire S, Khanuja H. A Personalized Framework for Health Care Recommendation. : IEEE; 2015 Presented at: International Conference on Computing Communication Control and Automation; 26-27 Feb, 2015; Pune, India. [doi: 10.1109/ICCUBEA.2015.92]
78. Lee MK, Park H, Yun YH, Chang YJ. Development and formative evaluation of a web-based self-management exercise and diet intervention program with tailored motivation and action planning for cancer survivors. JMIR Res Protoc 2013;2(1):e11 [FREE Full text] [doi: 10.2196/resprot.2331] [Medline: 23612029]
79. Ammann R, Vandelanotte C, de Vries H, Mummery WK. Can a website-delivered computer-tailored physical activity intervention be acceptable, usable, and effective for older people? Health Educ Behav 2013 Apr;40(2):160-170. [doi: 10.1177/1090198112461791] [Medline: 23077157]
80. Codreanu I, Florea A. A Proposed Serious Game Architecture to Self-Management HealthCare for Older Adults. : IEEE; 2015 Presented at: 17th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC); 21-24 Sept, 2015; Timisoara, Romania. [doi: 10.1109/SYNASC.2015.71]
81. Coolbaugh CL, Raymond SC, Hawkins DA. Feasibility of a dynamic web guidance approach for personalized physical activity prescription based on daily information from wearable technology. JMIR Res Protoc 2015 Jun 04;4(2):e67 [FREE Full text] [doi: 10.2196/resprot.3966] [Medline: 26043793]
82. Spark LC, Fjeldsoe BS, Eakin EG, Reeves MM. Efficacy of a text message-delivered extended contact intervention on maintenance of weight loss, physical activity, and dietary behavior change. JMIR Mhealth Uhealth 2015;3(3):e88 [FREE Full text] [doi: 10.2196/mhealth.4114] [Medline: 26373696]
83. Irvine AB, Gelatt VA, Seeley JR, Macfarlane P, Gau JM. Web-based intervention to promote physical activity by sedentary older adults: randomized controlled trial. J Med Internet Res 2013;15(2):e19 [FREE Full text] [doi: 10.2196/jmir.2158] [Medline: 23470322]
84. Short CE, Rebar A, James EL, Duncan MJ, Courneya KS, Plotnikoff RC, et al. How do different delivery schedules of tailored web-based physical activity advice for breast cancer survivors influence intervention use and efficacy? J Cancer Surviv 2017 Feb;11(1):80-91. [doi: 10.1007/s11764-016-0565-0] [Medline: 27498099]
85. Ali R, Afzal M, Hussain M, Ali M, Siddiqi MH, Lee S, et al. Multimodal hybrid reasoning methodology for personalized wellbeing services. Comput Biol Med 2016 Feb 01;69:10-28. [doi: 10.1016/j.compbiomed.2015.11.013] [Medline: 26705863]

86. Lim C, Kim ZM, Choi H. Developing a mobile wellness management system for healthy lifestyle by analyzing daily living activities. Stud Health Technol Inform 2017;245:146-150. [Medline: 29295070]
87. Moreau M, Gagnon M, Boudreau F. Development of a fully automated, web-based, tailored intervention promoting regular physical activity among insufficiently active adults with type 2 diabetes: integrating the I-change model, self-determination theory, and motivational interviewing components. JMIR Res Protoc 2015;4(1):e25 [FREE Full text] [doi: 10.2196/resprot.4099] [Medline: 25691346]
88. Larsen B, Benitez T, Cano M, Dunsiger SS, Marcus BH, Mendoza-Vasconez A, et al. Web-based physical activity intervention for Latina adolescents: feasibility, acceptability, and potential efficacy of the Niñas Saludables study. J Med Internet Res 2018 May 09;20(5):e170 [FREE Full text] [doi: 10.2196/jmir.9206] [Medline: 29743151]
89. Rajanna V, Behera D, Goldberg D, Hammond T. Step Up Life: A Context Aware Health Assistant. In: Proceedings of the Third ACM SIGSPATIAL International Workshop on the Use of GIS in Public Health.: ACM; 2014 Presented at: HealthGIS '14; November 04, 2014; Dallas, Texas p. 21-30. [doi: 10.1145/2676629.2676636]
90. Mistry CD, Sweet SN, Rhodes RE, Latimer-Cheung AE. Text2Plan: exploring changes in the quantity and quality of action plans and physical activity in a text messaging intervention. Psychol Health 2015;30(7):839-856. [doi: 10.1080/08870446.2014.997731] [Medline: 25659571]
91. Centers for Disease Control and Prevention. National Center for Chronic Disease Prevention and Health Promotion URL: https://www.cdc.gov/chronicdisease/index.htm [accessed 2018-05-15] [WebCite Cache ID 6zQgfZMRs]
92. Haskell WL, Lee I, Pate RR, Powell KE, Blair SN, Franklin BA, et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. Med Sci Sports Exerc 2007 Aug;39(8):1423-1434. [doi: 10.1249/mss.0b013e3180616b27] [Medline: 17762377]
93. Hargreaves EA, Mutrie N, Fleming JD. A web-based intervention to encourage walking (StepWise): pilot randomized controlled trial. JMIR Res Protoc 2016 Jan 25;5(1):e14 [FREE Full text] [doi: 10.2196/resprot.4288] [Medline: 26810251]
94. Twardowski B, Ryzko D. IoT and Context-Aware Mobile Recommendations Using Multi-agent Systems. : IEEE; 2015 Presented at: IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT); 6-9 Dec. 2015; Singapore, Singapore. [doi: 10.1109/WI-IAT.2015.120]
95. Varadharajan V, Kannav V, Pasala A. Sensor based coaching for a fit and healthy society. : IEEE; 2016 Presented at: IEEE International Symposium on Technology and Society (ISTAS); 20-22 Oct. 2016; Kollam, India. [doi: 10.1109/ISTAS.2016.7764275]
96. Harris JA, Benedict FG. A biometric study of human basal metabolism. Proc Natl Acad Sci U S A 1918 Dec;4(12):370-373 [FREE Full text] [Medline: 16576330]
97. Boudreau F, Walthouwer MJL, De Vries H, Dagenais GR, Turbide G, Bourlaud A, et al. Rationale, design and baseline characteristics of a randomized controlled trial of a web-based computer-tailored physical activity intervention for adults from Quebec City. BMC Public Health 2015 Oct 09;15:1038 [FREE Full text] [doi: 10.1186/s12889-015-2364-3] [Medline: 26453041]
98. Kwasnicka D, Vandelanotte C, Rebar A, Gardner B, Short C, Duncan M, et al. Comparing motivational, self-regulatory and habitual processes in a computer-tailored physical activity intervention in hospital employees - protocol for the PATHS randomised controlled trial. BMC Public Health 2017 Dec 26;17(1):518 [FREE Full text] [doi: 10.1186/s12889-017-4415-4] [Medline: 28549476]
99. Fogg BJ. A behavior model for persuasive design. In: Proceedings of the 4th International Conference on Persuasive Technology. New York, New York, USA: ACM; 2009 Presented at: International Conference on Persuasive Technology; April 26 - 29, 2009; Claremont, California, USA. [doi: 10.1145/1541948.1541999]
100. Bandura A. Social cognitive theory of self-regulation. Organ Behav Hum Decis Process 1991 Dec;50(2):248-287. [doi: 10.1016/0749-5978(91)90022-L]
101. Prochaska JO. Transtheoretical Model of Behavior Change. In: Gellman MD, Turner JR, editors. Encyclopedia of Behavioral Medicine. New York, NY: Springer; 2013.
102. Ajzen I. The theory of planned behavior. Organ Behav Hum Decis Process 1991 Dec;50(2):179-211. [doi: 10.1016/0749-5978(91)90020-T]
103. de Vries H, Mudde A, Leijs I, Charlton A, Vartiainen E, Buijs G, et al. The European Smoking Prevention Framework Approach (EFSA): an example of integral prevention. Health Educ Res 2003 Oct;18(5):611-626 [FREE Full text] [Medline: 14572020]
104. Mitchie S, Atkins L, West R. The Behaviour Change Wheel: A Guide To Designing Interventions. United Kingdom: Silverback Publishing; 2014.
105. Rosenstock IM. Adoption and maintenance of lifestyle modifications. Am J Prev Med 1988;4(6):349-352. [Medline: 3228532]
106. Rollnick S, Miller W, Butler C, Aloia M. Motivational interviewing in health care: helping patients change behavior. J Chronic Obstr Pulm Dis 2009 Jul 02;5(3):203-203. [doi: 10.1080/15412550802093108]
107. Schwarzer R, Lippke S, Luszczynska A. Mechanisms of health behavior change in persons with chronic illness or disability: the Health Action Process Approach (HAPA). Rehabil Psychol 2011 Aug;56(3):161-170. [doi: 10.1037/a0024509] [Medline: 21767036]

108. Blake H, Suggs LS, Coman E, Aguirre L, Batt ME. Active8! Technology-based intervention to promote physical activity in hospital employees. Am J Health Promot 2017 Dec;31(2):109-118. [doi: 10.4278/ajhp.140415-QUAN-143] [Medline: 26559712]
109. McMurray JJ, Adamopoulos S, Anker SD, Auricchio A, Böhm M, Dickstein K, Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology, et al. ESC guidelines for the diagnosis and treatment of acute and chronic heart failure 2012: The Task Force for the Diagnosis and Treatment of Acute and Chronic Heart Failure 2012 of the European Society of Cardiology. Developed in collaboration with the Heart Failure Association (HFA) of the ESC. Eur J Heart Fail 2012 Aug;14(8):803-869. [doi: 10.1093/eurjhf/hfs105] [Medline: 22828712]
110. Canadian Diabetes Association Clinical Practice Guidelines Expert Committee, Booth G, Cheng AYY. Canadian Diabetes Association 2013 clinical practice guidelines for the prevention and management of diabetes in Canada. Methods. Can J Diabetes 2013 Apr;37(Suppl 1):S4-S7. [doi: 10.1016/j.jcjd.2013.01.010] [Medline: 24070961]
111. Peacock OJ, Western MJ, Batterham AM, Stathi A, Standage M, Tapp A, et al. Multidimensional individualised Physical ACTivity (Mi-PACT)--a technology-enabled intervention to promote physical activity in primary care: study protocol for a randomised controlled trial. Trials 2015 Aug 28;16:381 [FREE Full text] [doi: 10.1186/s13063-015-0892-x] [Medline: 26314577]

Abbreviations

BCT: behavior change technique
CDC: Centers for Disease Control and Prevention
CVD: cardiovascular disease
HAPA: health action process approach
HR: heart rate
LPA: light physical activity
MET: metabolic equivalent
MVPA: moderately vigorous physical activity
PA: physical activity
RCT: randomized controlled trial
RL: reinforcement learning
RNN: recurrent neural networks
SCT: social cognitive theory
SMS: short messaging service
SVM: support vector machine
TPB: Theory of Planned Behavior
TTM: Trans-Theoretical Model

Edited by G Eysenbach; submitted 21.05.18; peer-reviewed by S Lippke, M Cabrita; comments to author 07.08.18; revised version received 01.10.18; accepted 26.10.18; published 16.01.19.

Please cite as:
Ghanvatkar S, Kankanhalli A, Rajan V
User Models for Personalized Physical Activity Interventions: Scoping Review
JMIR Mhealth Uhealth 2019;7(1):e11098
URL: http://mhealth.jmir.org/2019/1/e11098/
doi:10.2196/11098
PMID:30664474

Original Paper

A Library of Analytic Indicators to Evaluate Effective Engagement with Consumer mHealth Apps for Chronic Conditions: Scoping Review

Quynh Pham[1,2], MSc; Gary Graham[2], MSc; Carme Carrion[3,4], PhD; Plinio P Morita[1,5], PEng, PhD; Emily Seto[1,2], PEng, PhD; Jennifer N Stinson[1,6,7,8], RN, PhD, CPNP; Joseph A Cafazzo[1,2,9], PEng, PhD

[1]Institute of Health Policy, Management and Evaluation, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[2]Centre for Global eHealth Innovation, Techna Institute, University Health Network, Toronto, ON, Canada
[3]eHealth Center, Universitat Oberta de Catalunya, Catalonia, Spain
[4]eHealth Lab Research Group, School of Health Sciences, Universitat Oberta de Catalunya, Catalonia, Spain
[5]School of Public Health and Health Systems, Faculty of Applied Health Sciences, University of Waterloo, Toronto, ON, Canada
[6]Child Health Evaluative Sciences Research Institute, The Hospital for Sick Children, Toronto, ON, Canada
[7]Department of Anesthesia and Pain Medicine, The Hospital for Sick Children, Toronto, ON, Canada
[8]Lawrence S Bloomberg Faculty of Nursing, University of Toronto, Toronto, ON, Canada
[9]Institute of Biomaterials and Biomedical Engineering, Faculty of Applied Science and Engineering, University of Toronto, Toronto, ON, Canada

Corresponding Author:
Quynh Pham, MSc
Institute of Health Policy, Management and Evaluation
Dalla Lana School of Public Health
University of Toronto
Health Sciences Building, Suite 425
155 College Street
Toronto, ON, M5T 3M6
Canada
Phone: 1 416 340 4800 ext 4765
Email: q.pham@mail.utoronto.ca

Abstract

Background: There is mixed evidence to support current ambitions for mobile health (mHealth) apps to improve chronic health and well-being. One proposed explanation for this variable effect is that users do not engage with apps as intended. The application of analytics, defined as the use of data to generate new insights, is an emerging approach to study and interpret engagement with mHealth interventions.

Objective: This study aimed to consolidate how analytic indicators of engagement have previously been applied across clinical and technological contexts, to inform how they might be optimally applied in future evaluations.

Methods: We conducted a scoping review to catalog the range of analytic indicators being used in evaluations of consumer mHealth apps for chronic conditions. We categorized studies according to app structure and application of engagement data and calculated descriptive data for each category. Chi-square and Fisher exact tests of independence were applied to calculate differences between coded variables.

Results: A total of 41 studies met our inclusion criteria. The average mHealth evaluation included for review was a two-group pretest-posttest randomized controlled trial of a hybrid-structured app for mental health self-management, had 103 participants, lasted 5 months, did not provide access to health care provider services, measured 3 analytic indicators of engagement, segmented users based on engagement data, applied engagement data for descriptive analyses, and did not report on attrition. Across the reviewed studies, engagement was measured using the following 7 analytic indicators: the number of measures recorded (76%, 31/41), the frequency of interactions logged (73%, 30/41), the number of features accessed (49%, 20/41), the number of log-ins or sessions logged (46%, 19/41), the number of modules or lessons started or completed (29%, 12/41), time spent engaging with the app (27%, 11/41), and the number or content of pages accessed (17%, 7/41). Engagement with unstructured apps was mostly measured by the number of features accessed (8/10, *P*=.04), and engagement with hybrid apps was mostly measured by the

number of measures recorded (21/24, *P*=.03). A total of 24 studies presented, described, or summarized the data generated from applying analytic indicators to measure engagement. The remaining 17 studies used or planned to use these data to infer a relationship between engagement patterns and intended outcomes.

Conclusions: Although researchers measured on average 3 indicators in a single study, the majority reported findings descriptively and did not further investigate how engagement with an app contributed to its impact on health and well-being. Researchers are gaining nuanced insights into engagement but are not yet characterizing effective engagement for improved outcomes. Raising the standard of mHealth app efficacy through measuring analytic indicators of engagement may enable greater confidence in the causal impact of apps on improved chronic health and well-being.

(JMIR Mhealth Uhealth 2019;7(1):e11941) doi:10.2196/11941

KEYWORDS

analytics; effective engagement; engagement; adherence; log data; mobile health; mobile applications; chronic disease; scoping review

Introduction

Background

There is mixed evidence to support current ambitions for mobile health (mHealth) apps to improve chronic health and well-being [1]. While some apps have demonstrated efficacy in definitive trials [2-5], others have performed poorly [6-9]. One proposed explanation for this variable effect is that users do not engage with apps as intended [10]. The construct of engagement has been quantitatively conceptualized as the amount, duration, breadth, and depth of intervention usage [11,12]. For many mHealth app evaluations, users can be segmented along a continuum of engagement; some will never use the app, some will use it but quickly abandon it, and some will use it in unexpected ways. Complex patterns of engagement with mHealth apps are emerging and challenge current conceptual paradigms for interpreting their impact on chronic health outcomes. These digitally mediated mechanisms of action require more granular evaluations capable of analyzing multilevel, temporally dense engagement data [13]. Evaluating engagement is therefore a priority and calls for the integration of nonintrusive measures of this construct in mHealth evaluation methodology [14].

Recently, scholars sought to further the conceptualization of engagement by proposing that it may be more valuable to identify the mechanisms that underlie *effective engagement*, defined as *sufficient engagement with an intervention to achieve intended outcomes* [14,15]. The construct of effective engagement differs conceptually from both engagement and adherence, which have historically been used interchangeably [16]. Sieverink et al reason that the following 3 elements are necessary to determine adherence to a digital health intervention: (1) the ability to measure usage behaviors, (2) an operationalization of intended use, and (3) an empirical, theoretical, or rational justification of intended use [17]. We propose that effective engagement is more intentional than engagement but less justified than adherence. It sits between both constructs and bridges the transition from identifying patterns of engagement toward evidencing their capacity to achieve intended outcomes.

There has been recognition that the definition of engagement has evolved to include *offline* interactions with the behavior change mediated by a digital health intervention. Yardley et al have been instrumental in furthering this conceptualization of engagement by suggesting that there are 2 levels of engagement: (1) the *micro* level of immediate engagement with the digital health intervention and (2) the *macro* level of engagement with the wider intervention-mediated behavior change [14]. They posit that engagement is a dynamic process marked by shifts in both micro and macroengagement, which will vary depending on the intervention, the user, and their context. Users may be macroengaging and experiencing positive behavior change, but this may not necessarily be reflected in their microengagement analytics data. In acknowledgment of this distinction between engagement with the technological and behavioral aspects of an intervention, Yardley et al critically posit that microengagement alone cannot be taken as a valid indicator of effective engagement. We do not dispute Yardley et al's arguments and recognize the limitations of relying solely on microengagement data to infer effective engagement. However, we posit that measuring and reporting on microengagement is fundamental to understanding how people actually use an app to improve their health and well-being. In turn, these analytic insights can be coupled with measures of macroengagement to identify the mediating mechanisms that motivate effective engagement.

The application of analytics, defined as the use of data to generate new insights [18], is an emerging approach to study and interpret engagement with mHealth interventions [19]. Van Gemert-Pijnen et al have advanced the application of log data analysis to inform how an intervention works in practice and which components should be improved to yield greater benefit [20-22]. Arden-Close et al have developed and implemented a novel R-based tool to visually explore patterns of engagement [23]. Heckler et al have called for the adoption of a continuous optimization model of evaluation that leverages simulated computational models to predict how users might engage with an intervention before data collection [24]. Scherer et al have demonstrated the value of joint models in the analysis of longitudinal engagement data. In fact, Scherer et al recently participated in a workshop sponsored by the National Institutes of Health on emerging technology and data analytics for behavioral health, and espoused the need for new analytic methods that can scale to thousands of individuals and billions of data points [19]. Short et al recently published a viewpoint

on engagement measurement options that can be employed in electronic health (eHealth) and mHealth behavior change intervention evaluations [25]. They found that system engagement data are the most commonly collected and reported measures of engagement in eHealth and mHealth interventions. From this, they recommend having shared ways of conceptualizing these data as the field progresses to consolidate categorization.

Objectives

Motivated by the proven value of analytics to study engagement with mHealth apps, we sought to compile and catalog a library of analytic indicators of engagement with consumer mHealth apps for self-managing chronic conditions. We defined analytic indicators as proxy measures of engagement with an mHealth app based on objective usage that generates log data [14,22]. When positioned alongside other measures suitable for evaluating the subjective experience of mHealth app engagement, they may provide complementary data-driven insights into the objective extent of engagement. We propose that analytic indicators of engagement do exactly this: they *indicate* that users may be engaging effectively with a digital health intervention but do not definitively confirm a relationship between engagement and intended outcomes. Establishing this relationship requires adopting a mixed-methods multidimensional approach to measure effective engagement using multiple assessment strategies [14,25].

While many researchers have included analytic indicators as a study measure when evaluating apps, they are not consistent or systematic in their selection [26]. We propose that there is benefit to understanding how engagement with mHealth apps for chronic conditions has been defined, measured, and analyzed across evaluations. The aim of this scoping review was therefore to consolidate how analytic indicators of engagement have previously been applied across clinical and technological contexts to inform how they might be optimally applied in future evaluations.

Methods

Review Framework

This scoping review was guided by the methodological framework developed by Arksey and O'Malley [27] and advanced by Levac et al [28]. They endorse an iterative review process with 5 distinct steps: (1) identifying the research question, (2) searching for relevant studies, (3) selecting studies, (4) charting the data, and (5) collating, summarizing, and reporting results. This framework is particularly relevant to disciplines with emerging evidence, such as mHealth, in which the paucity of definitive research makes it difficult for researchers to undertake systematic reviews [28]. In this context, conducting a scoping review allowed us to incorporate a range of study designs beyond those accepted for inclusion in systematic reviews, to generate broad findings on how researchers are measuring engagement with consumer mHealth apps for chronic conditions. We made efforts to adhere to recommendations for each step, starting with the selection of a research question that was sufficiently broad to map the extent, range, and nature of mHealth engagement research activity. We conducted this review to explore the following research question: *what analytic indicators of engagement are being used in evaluations of consumer mHealth apps for chronic conditions?*

Search Strategy

A literature search was conducted in the MEDLINE, PsycINFO, CINAHL, and EMBASE databases. In addition, the *Journal of Medical Internet Research* and its sister journals were independently searched given their frequent and high-impact publication of mHealth research. A combination of different keywords for the constructs "engagement" and "mHealth" was used. No search terms for chronic conditions were defined a priori to broaden search results. We adopted the World Health Organization's definition of a chronic condition as a "non-communicable disease of long duration and slow progression [29]." Multimedia Appendix 1 presents our search strategy for MEDLINE on the Ovid platform.

Eligibility Criteria

Titles and abstracts retrieved from the search strategy were screened for inclusion against the following criteria: (1) the article described an evaluation or a protocol for an evaluation of a consumer mHealth app for self-managing a chronic condition; (2) the study included operationalization of an engagement-related construct—Multimedia Appendix 1 provides the full list of screened constructs; (3) the study included objective, quantifiable measurements using log data analytics; (4) the app was intended to be used more than once; (5) the article was published between November 1, 2015, and November 1, 2017; and (6) the article was published in English.

Studies were excluded if (1) the mHealth app was solely an appointment reminder service; (2) the primary app technology was short message service or interactive voice response; (3) the app was for an acute condition or preventive health purposes; (4) the app was a support tool for a patient's circle of care; (5) the app did not require user input through active or passive (sensor) data entry; (6) the app only delivered educational content; and (7) the article primarily described the design, development, or usability testing of the app.

Data Collection and Analysis

The first author conducted the electronic searches with support from a faculty-affiliated librarian and reviewed the reference lists of relevant articles. All identified titles and abstracts were downloaded and merged using Mendeley (Elsevier) [30] and duplicated records were removed. The first author independently screened all titles and abstracts against eligibility criteria. Any articles that caused the author uncertainty were retained until data extraction when more information was available to make an informed decision for inclusion in the review. Following title and abstract review, full papers of included abstracts were assessed for final selection by all study authors.

Textbox 1. Codes extracted from included articles.

1. General information regarding the study title, authors, journal, year, and country.
2. App information, specifically the public name, chronic condition addressed, and accessibility of health care provider services.
3. Study information, specifically the purpose, duration, sample size, and design.
4. App structure (structured, hybrid, or unstructured): "Structured" apps contained locked, sequential components (eg, modules, lessons, and features) that users had to complete before moving forward. "Hybrid" apps contained both fixed core components and variable components for free use. "Unstructured" apps contained variable components that users could access and use at will.
5. Analytic indicators used to measure engagement, specifically the number of log-ins or sessions logged, the number of modules or lessons started or completed, the number of features accessed, the number of measures recorded, the number or content of pages accessed, the frequency of interactions logged, and total time spent engaging with the app.
6. Engagement-based segmentation: studies that segmented users based on engagement data (eg, "of the users who logged in at least five times…") were assigned this code.
7. Application of engagement data (descriptive or inferential): a "descriptive" code was assigned to studies that presented, described, or summarized engagement data. An "inferential" code was assigned to studies that used engagement data to predict the intended outcome. Outcome types were coded for studies that applied engagement data inferentially.
8. Attrition type (dropout or nonusage) and statistical method of analysis: dropout attrition is the phenomenon of users not returning to complete follow-up study activities. Nonusage attrition is the phenomenon of users losing interest in a digital health intervention and ceasing to use it [10].

A data extraction form was developed by the first author to extract relevant study information. We referenced work by Sieverink [17] and Kelders [31] on analytic indicators of adherence to eHealth technologies to establish preliminary codes. The form was piloted on a sample of included articles to validate proposed codes and add emergent codes. The codes extracted from each study are presented in Textbox 1. All study data were entered into SPSS version 24 (IBM) [32]. Each study along with its corresponding data was treated as a separate case. We categorized studies according to app structure and application of engagement data and calculated descriptive data for each category. Chi-square and Fisher exact tests of independence were applied to calculate differences between coded variables. A Monte Carlo correction was applied when observed counts were below expected counts.

Results

Study Selection

A total of 1873 articles were identified through the database search. Of the 60 full texts screened, 19 were excluded, 8 of which did not include objective, quantifiable measurements using log data analytics. In total, 41 articles comprising 33 studies and 8 protocols met the eligibility criteria and were included for review. Figure 1 presents the Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram of the study selection progress [33].

Methodological Characteristics

The first authors of reviewed studies were affiliated with institutions in the United States (46%, 19/41), Canada (20%, 8/41), the United Kingdom (10%, 4/41), Australia (5%, 2/41), Germany (5%, 2/41), the Netherlands (5%, 2/41), France (2%, 1/41), India (2%, 1/41), Singapore (2%, 1/41), Spain (2%, 1/41), Sweden (2%, 1/41), and Switzerland (2%, 1/41).

Researchers reported log data analytics across 14 different engagement-related constructs: engagement (27%, 11/41), adherence (17%, 7/41), usage (15%, 6/41), use (15%, 6/41), feasibility (10%, 4/41), acceptability (7%, 3/41), utilization (5%, 2/41), attrition (5%, 2/41), participation (5%, 2/41), activity (2%, 1/41), adoption (2%, 1/41), compliance (2%, 1/41), fidelity (2%, 1/41), and retention (2%, 1/41). There was significant variation in how constructs were defined across studies, which limited our ability to (1) extract reliable definitions for each construct, (2) map analytic indicators to specific constructs, and (3) conduct cross-construct comparisons of analytic indicators.

The majority of reviewed studies were experimental (51%, 21/41), with the two-group pretest-posttest randomized controlled trial (RCT) as the most prevalent experimental study design (48%, 10/21), followed by the one-group pretest-posttest design (43%, 9/21). Quasi-experimental design selection (17%, 7/41) was more diverse and included cohort (29%, 2/7), interrupted time-series (14%, 1/7), and single case (14%, 1/7) studies. The remaining 13 studies included for review were observational in design (32%, 13/41). Studies were on average 5 months long (median 152 days, interquartile range, IQR 106), with a sample size of over 100 participants (median 103, IQR 252). The longest reviewed observational study conducted by Serrano et al was 7 years long with over 1 million participants [34]. A total of 19 studies applied engagement-based segmentation and reported results for separate user cohorts (58%, 19/33). In total, 14 of the reviewed studies were published in the *Journal of Medical Internet Research* or its sister journals (34%, 14/41).

Intervention Characteristics

A wide range of chronic conditions were targeted through the apps under study, with mental health (29%, 12/41), chronic pain (12%, 5/41), asthma (10%, 4/41), cardiovascular disease (7%, 3/41), and diabetes (type 1 and 2; 15%, 6/41) leading the clinical charge. Researchers also evaluated apps for cancer (5%, 2/41), hypertension (5%, 2/41), obesity (5%, 2/41), chronic kidney disease (2%, 1/41), chronic obstructive pulmonary disease (2%, 1/41), cystic fibrosis and inflammatory bowel disease (2%, 1/41), Parkinson disease (2%, 1/41), and sleep apnea (2%, 1/41). Over half of the apps had a hybrid structure (59%, 24/41), 10

apps were unstructured (24%), and 7 apps were structured (17%). Nearly half of all structured apps were aimed at improving mental health (40%, 4/10). Health care provider services were accessible to users to support managing their condition in nearly half of all reviewed apps (44%, 18/41). Characteristics of the included studies are presented in Multimedia Appendix 2 alongside the full dataset of coded analytic indicators for each study, which are summarized below.

Analytic Indicators

Across the reviewed studies, engagement was measured using the following 7 analytic indicators in order of prevalence: the number of measures recorded (76%, 31/41), the frequency of interactions logged (73%, 30/41), the number of features accessed (49%, 20/41), the number of log-ins or sessions logged (46%, 19/41), the number of modules or lessons started or completed (29%, 12/41), time spent engaging with the app (27%, 11/41), and the number or content of pages accessed (17%, 7/41). Table 1 presents a tally of the analytic indicators measured in each included study. On average, researchers applied 3 different analytic indicators to measure their engagement data (mean 3.20, SD 1.42; median 3, IQR 2). The Fisher exact test of independence indicated that engagement with unstructured apps was mostly measured by the number of features accessed (8/10, P=.04), and engagement with hybrid apps was mostly measured by the number of measures recorded (21/24, P=.03). Table 2 provides a descriptive overview of structured, hybrid, and unstructured apps across study characteristics and analytic indicators.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flow diagram. mHealth: mobile health; SMS: short message service.

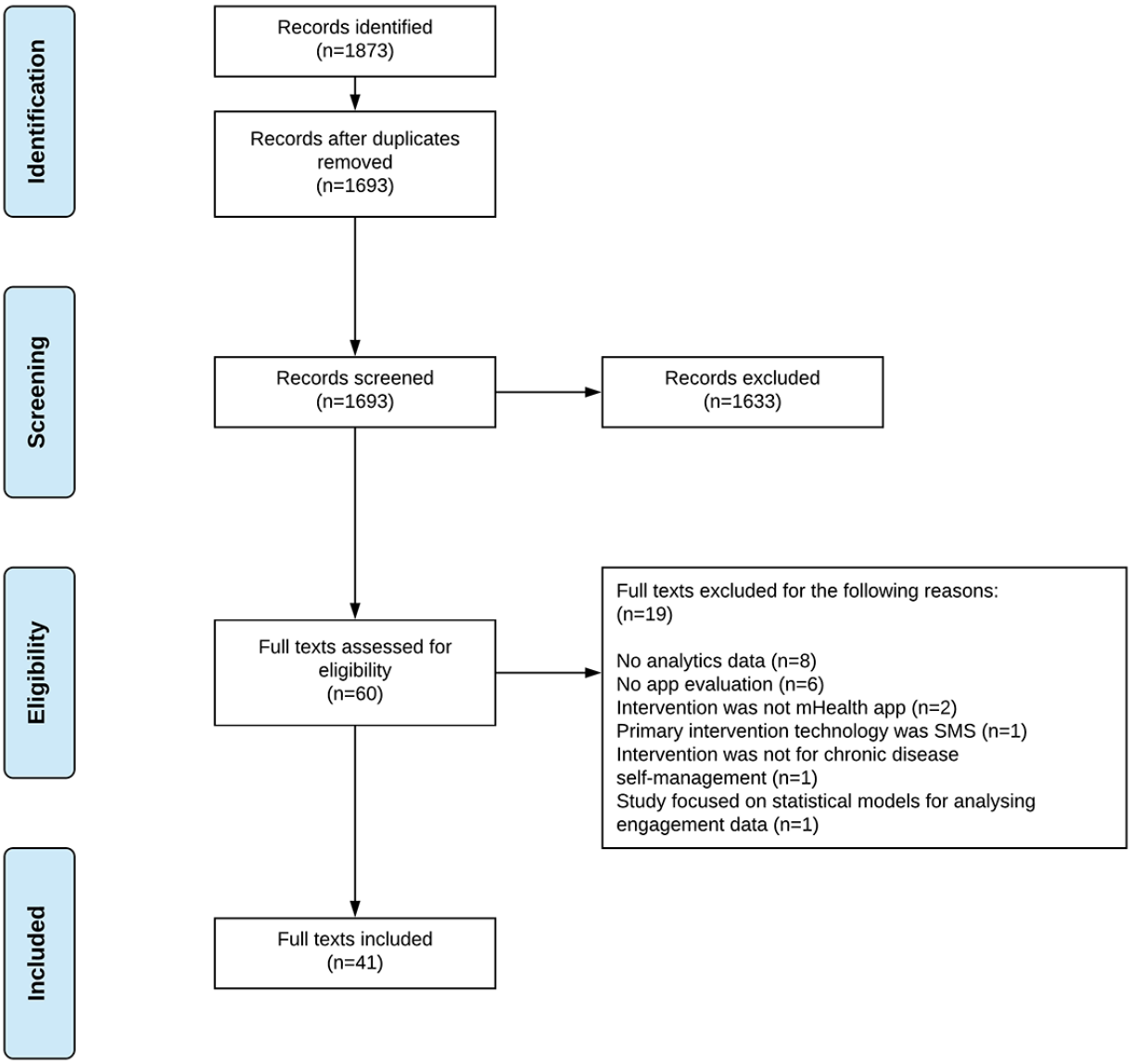


Table 1. Tally of analytic indicators used in reviewed studies.

Author	Measures	Interactions	Features	Log-ins	Modules	Time spent	Pages
Mental health (n=12)							
Beiwinkel et al [35]	✓[a]	✓	—[b]	✓	✓	✓	—
Ben-Zeev et al [36]	✓	✓	✓	✓	✓	—	—
Ben-Zeev et al [37]	✓	✓	✓	—	✓	—	—
Davies et al [38]	✓	✓	✓	✓	—	✓	✓
Frisbee et al [39]	—	—	✓	✓	—	—	—
Kinderman et al [40]	✓	—	—	✓	—	—	—
Kuhn et al [41]	—	✓	—	—	—	—	✓
Owen et al [42]	—	✓	✓	✓	—	✓	✓
Pham et al [43]	—	✓	—	✓	—	✓	—
Torous et al [44]	✓	✓	✓	—	—	—	—
Vansimaeys et al [45]	✓	✓	—	—	✓	—	—
Wahle et al [46]	✓	✓	—	✓	✓	—	—
Chronic pain (n=5)							
Fortier et al [47]	✓	—	—	—	—	—	—
Jamison et al [48]	✓	✓	—	✓	—	—	—
Jibb et al [49]	✓	✓	—	—	—	—	—
Reade et al [50]	✓	✓	—	✓	✓	—	—
Skrepnik et al [51]	✓	✓	—	—	—	—	—
Asthma (n=4)							
Chan et al [52]	✓	✓	✓	—	✓	—	—
Cook et al [53]	—	—	—	✓	—	—	—
Fedele et al [54]	—	✓	✓	✓	—	—	—
Kosse et al [55]	✓	—	—	—	✓	—	—
Cardiovascular disease (n=3)							
Agboola et al [56]	✓	—	—	✓	—	✓	✓
Goyal et al [57]	✓	✓	✓	✓	✓	—	—
Sakakibara et al [58]	✓	—	✓	—	—	—	—
Type 1 diabetes (n=3)							
Goyal et al [59]	—	✓	✓	—	—	—	—
Ryan et al [60]	✓	✓	✓	—	—	✓	—
Sieber et al [61]	✓	—	—	—	—	—	—
Type 2 diabetes (n=3)							
Desveaux et al [62]	—	✓	✓	✓	—	✓	—
Goh et al [63]	—	✓	—	—	—	—	—
Kleinman et al [64]	✓	—	—	✓	—	—	—
Other (n=11)							
Bot et al [65]	✓	✓	✓	—	✓	—	—
Hardinge et al [66]	✓	✓	✓	—	—	✓	—
Isetta et al [67]	—	✓	—	—	—	—	—
Kaplan et al [68]	✓	✓	✓	—	✓	✓	—
Langius-Eklof et al [69]	✓	—	—	—	✓	—	—

Author	Measures	Interactions	Features	Log-ins	Modules	Time spent	Pages
Ong et al [70]	✓	—	✓	✓	—	—	—
Pham et al [71]	✓	✓	✓	✓	—	—	✓
Serrano et al [34]	✓	✓	✓	—	—	—	✓
Taki et al [72]	✓	✓	—	✓	—	✓	✓
Thies et al [73]	✓	✓	✓	—	—	✓	—
Toro-Ramos et al [74]	✓	✓	—	—	✓	—	—

[a]Analytic indicators of engagement used in reviewed studies.

[b]Not applicable.

Number of Measures

Of the analytic indicators identified in this review, the number of measures recorded by users on an app was the most commonly used indicator of engagement with mHealth apps for chronic conditions. Researchers evaluated a range of measures that aligned with their target chronic condition, such as blood glucose [56,60,61,64,73], weight [56,73,74], symptoms [66,68,69], patient-reported outcomes [38,46,52,65,71], diary entries [47,66], and steps [51]. There was some overlap in the types of measures being collected across apps targeting the same chronic conditions, such as the number of blood glucose readings recorded as an indicator of engagement with diabetes apps. Overall, the target chronic condition and functionality of the app under study ultimately determined which measures would be collected and subsequently reported as an analytic indicator of engagement.

Frequency of Interactions

The frequency of interactions logged was the second most prevalent analytic indicator of engagement. Researchers often chose to complement assessing the number of measures recorded on an app with the frequency by which the measures were recorded. Stratifying frequency of interactions by specific date ranges was also common; Davies et al measured the number of users who used a mental health app at least once after 1 week, 4 weeks, and 20 weeks [38]. They also applied within-date range indicators such as the number of users who used the app once, 2 to 3 times, 4 to 6 times, or 6 or more times per week. Some researchers assigned a benchmark number of days to signify engagement, such as Isetta et al who measured the number of users who engaged with an app for sleep apnea on at least 66% of all days in the study [67]. Others assigned significance to a specific day and considered reaching it as an indicator of engagement, such as Jamison et al who measured the number of users who continued to submit daily assessments of their chronic pain after 90 and 180 days [48]. Layering this analytic indicator over other indicators added temporal context to better understand how users were engaging over time.

Number of Features

The range of features accessed by users in an app was frequently measured as an analytic indicator of engagement. Researchers primarily logged (1) the number of features accessed and (2) the number of times each feature was accessed. In their trial of the Veterans Affairs' Comprehensive Assistance for Family Caregivers Program where users were provided with access to a suite of 6 apps for posttraumatic stress disorder (PTSD) self-management, Frisbee et al measured the number of unique apps used in the suite [39]. To better understand user preferences between 2 features of their app for schizophrenia self-management, Ben-Zeev et al measured the number of times users chose the video feature over the written content feature [36]. Our research group proposed exploring whether users would access all the features made available in their app for prostate cancer survivorship care, particularly whether users would enable caregiver permissions or write notes to document changes in their care [71]. Overall, researchers applied this analytic indicator to explore the breadth of app engagement and inform feature popularity and relevance for the target population.

Number of Log-Ins

The number of log-ins or sessions logged by users continues to be a commonly used analytic indicator of engagement. This indicator was often coupled with the frequency of interactions logged to standardize counts. Researchers also frequently measured the number of users who opened an app at least once to segment them from users who had downloaded the app but never logged any subsequent activity. Owen et al made both these associations by measuring the number of sessions logged by users on their PTSD self-management app, as well as the number of users who logged at least one session on the first day, week, and month post download [42]. Researchers used this analytic indicator to reflect the shift from adoption to habituation, with a greater number of log-ins or sessions denoting greater engagement.

Table 2. Descriptive overview of app structures across study characteristics and analytic indicators.

Characteristics	Structured (N=7), n (%)	Hybrid (N=24), n (%)	Unstructured (N=10), n (%)
Chronic condition			
Mental health (n=12)	2 (29)	6 (25)	4 (40)
Chronic pain (n=5)	2 (29)	3 (13)	0 (0)
Asthma (n=4)	1 (14)	3 (13)	0 (0)
Cardiovascular disease (n=3)	0 (0)	2 (8)	1 (10)
Type 1 diabetes (n=3)	1 (14)	1 (4)	1 (10)
Type 2 diabetes (n=3)	0 (0)	1 (4)	2 (20)
Other (n=11)	1 (14)	8 (33)	2 (20)
Segmentation			
Yes (n=19)	1 (14)	12 (50)	6 (60)
No (n=14)	4 (47)	7 (29)	3 (30)
Analytic indicators			
Number of measures (n=31)[a]	6 (86)	21 (88)	4 (40)
Frequency of interactions (n=30)	4 (57)	18 (75)	8 (80)
Number of features (n=20)[a]	2 (29)	10 (42)	8 (80)
Number of log-ins (n=19)	4 (57)	12 (50)	3 (30)
Number of modules (n=12)	2 (29)	10 (42)	0 (0)
Time spent (n=11)	0 (0)	8 (33)	3 (30)
Number of pages (n=7)	0 (0)	4 (17)	3 (30)
Application of engagement data			
Descriptive (n=24)	7 (100)	13 (54)	4 (40)
Inferential (n=17)	0 (0)	11 (46)	6 (60)
Study design			
Experimental (n=21)	3 (43)	13 (54)	5 (50)
Quasi-experimental (n=7)	0 (0)	6 (25)	1 (10)
Observational (n=13)	4 (57)	5 (21)	4 (40)
Number of indicators			
1 (n=5)	1 (14)	3 (13)	1 (1)
2 (n=10)	2 (29)	4 (17)	4 (40)
3 (n=8)	3 (43)	4 (17)	1 (10)
4 (n=10)	1 (14)	6 (25)	3 (30)
5 (n=7)	0 (0)	6 (25)	1 (10)
6 (n=1)	0 (0)	1 (4)	0 (0)

[a]$P<.05$.

Number of Modules

When defining analytic indicators for categorization, we differentiated between unrestricted and restricted data collection. Unrestricted data collection was defined as data that could be entered into an app at a frequency or volume dictated by the user, such as the number of blood glucose readings or medications recorded [64]. Restricted data collection was defined as requiring the user to enter data according to a set frequency or volume, such as a list of assigned articles to be read [74] or challenges to be completed [57]. We coded studies reporting unrestricted data collection as *number of measures* and coded studies reporting restricted data collection as *number of modules*. A range of studies measured the number of outcome surveys completed from those assigned [45,68,75]. Others assessed the number of videos watched from a playlist [36,55], educational modules completed [52], or self-care advice accessed [69]. Overall, researchers studying apps with modular content considered module completion to be indicative of

engagement and consequently, tracked module progression and completion rates.

Time Spent

The amount of time that users engaged with an app was considered by a subset of researchers to be an analytic indicator of engagement. Researchers measured the time spent on unique sections of an app [66], the time spent on unique pages [56], the length of a unique session [38,42,43,71], the length between unique sessions [72], and the total time spent on an app [62,68,73]. Davies et al also segmented sessions by those that were in the 30- to 60-second range [38]. Measuring time spent engaging with an app helped researchers to distinguish between exploratory and purposeful engagement; a rapid succession of short page views was indicative of scanning through content, whereas prolonged viewing suggested greater intention and interest in content. Overall, this analytic indicator informed defining accurate session duration parameters to track session-based analytics.

Number of Pages

The number of pages accessed by users was logged by researchers to reflect overall patterns of app engagement and discoverability of specific content. Kuhn et al measured the number and content of pages visited by users in their app for PTSD self-management, as did other researchers [38,41,71]. Taki et al combined session analytics with page analytics and measured the number of pages viewed per session in their app for obesity self-management [72]. Owen et al recorded click stream data documenting their users' navigation through page content [42]. Insights gleaned from this analytic indicator provided researchers with a broader understanding of the user journey through an app and drew attention to specific content that might drive engagement.

Conceptual Categories of Analytic Indicators

We sought to conceptually clarify the 7 identified analytic indicators by grouping them according to the 4 categories that constitute the quantitative conceptualization of engagement: amount, duration, breadth, and depth [11,12]. Table 3 presents an overview of the categories, their comprised analytic indicators, and the number of reviewed studies that fall into each category. The focus of most reviewed studies was on the depth (76%, 31/41) and amount of engagement (73%, 30/41). There was less attention on the breadth (49%, 20/41) and duration (27%, 11/41) of engagement. TThese findings suggest that a subset of researchers are either not measuring the breadth and duration of engagement in their mHealth evaluations or underreporting the findings.

Application of Engagement Data

Of the 41 studies included for review, 24 presented, described, or summarized the data generated from applying analytic indicators to measure engagement. The remaining 17 studies used or planned to use these data to infer a relationship between engagement patterns and intended outcomes.

Clinical Outcomes

Over half of all researchers assessed the relationship between engagement and clinical outcomes (53%, 9/17). Toro-Ramos et al measured the number of weeks users engaged with their hypertension self-management app and found that users with sustained usage across 19 weeks experienced significant reductions in systolic blood pressure and weight [74]. In their trial of an app for PTSD self-management, Kuhn et al applied the number of days and weeks users engaged with the app as a predictor variable for changes in PTSD symptoms but did not find a significant relationship [41]. Goyal et al segmented all users who reported 5 or more blood glucose readings a day into a subgroup for secondary analyses and found a significant relationship between increased readings and improved glycated hemoglobin after 6 months [59]. They also identified a significant interaction between users who entered a reading on at least three days a week, and improved daily blood glucose self-monitoring. Overall, there was evidence of predictive validity across reviewed studies, with engagement correlating with improved clinical outcomes. However, the majority of analyses conducted to establish this predictive validity relied on nonexperimental variations in engagement due to nonadherence or implementation infidelity. Future evaluations assessing the relationship between engagement and clinical outcomes should consider alternative trial designs with multiple randomizations to ensure that findings are not biased by confounding [76-78].

Table 3. Conceptual categories of analytic indicators.

Category and analytic indicators	Studies, n (%)
Amount	
Frequency of interactions	30 (73)
Number of log-ins	30 (73)
Duration: Time spent	11 (27)
Breadth	
Number of features	20 (49)
Number of pages	20 (49)
Depth	
Number of modules	31 (76)
Number of measures	31 (76)

Engagement Outcomes

Many researchers sought to investigate the effect of engagement behaviors on other engagement outcomes (53%, 9/17). In their study examining engagement with a weight loss app, Serrano et al applied classification and regression tree methods to identify subgroups with unique engagement behaviors [79]. They were able to distinguish highly engaged subgroups by the number of customizations made to the diet and exercise features of the app. Ben-Zeev et al found that participants who engaged with their schizophrenia self-management app for a period of 5 to 6 months also had a higher frequency of interactions and engaged 4.3 days per week on average [37]. Torous et al also characterized engagement for a schizophrenia self-management app through fitting frequency of interaction data to a piecewise power law distribution [44]. They found that future use with the app is directly related to prior app use, suggesting that those who engage with the app more often will have a higher probability of app engagement in the future. In their trial of a caloric-monitoring app for type 2 diabetes self-management, Goh et al applied latent-class growth modeling to delineate 8-week trajectories of app engagement [63]. They were able to identify 3 distinct app trajectories based on the frequency of interactions and also associate patient characteristics with these trajectories. In summary, there were strong predictive relationships between numerous engagement domains. This finding motivates establishing complementary domains across multiple contexts to optimize data triangulation.

Utilization Outcomes

Two studies proposed to evaluate the impact of engagement patterns on health care utilization outcomes (12%, 2/17). Kaplan et al plan to examine the impact of sustained engagement over time with an app for pediatric cystic fibrosis and inflammatory bowel disease self-management on the number of hospitalizations and emergency department visits [68]. However, they anticipate that changes in these outcomes may not be realized in a 6-month intervention period. Our research group is evaluating a prostate cancer survivorship app [71] and aims to investigate the relationship between (1) the number of patient-reported outcome measures completed and (2) the frequency of interactions logged on the number of in-clinic visits for prostate cancer–related concerns. Altogether, the limited sample of reviewed studies suggests that the relationship between engagement and utilization outcomes is underdeveloped and warrants further study.

The Fisher exact test of independence indicated that studies of structured apps were more likely to only report descriptive statistics on engagement data (7/7, P=.04). In addition, most studies that applied inferential statistics also measured the frequency of interactions logged (16/17, P=.014). Most researchers who did not segment users into cohorts based on engagement data only reported descriptive statistics on their engagement data (13/14, P<.001), while researchers who segmented their users into cohorts were more likely to conduct subgroup analyses and infer properties of the larger clinical population (14/19, P<.001). Table 4 provides a descriptive overview of studies applying descriptive or inferential analyses on engagement data.

Attrition Type and Analyses

The majority of reviewed studies did not report on attrition (70%, 23/33). Of the 10 studies that did, 5 reported on dropout attrition (50%), 4 reported on nonusage attrition (40%), and 1 reported on both phenomena (10%). Researchers were more likely to descriptively summarize raw attrition proportions than statistically analyze them (70%, 7/10). Those that conducted comparisons across attrition curves used Kaplan-Meier survival curves (10%, 1/10), Cox regression models (10%, 2/10), and latent class growth models (10%, 1/10).

Table 4. Descriptive overview of descriptive and inferential engagement data application across study characteristics and analytic indicators.

Characteristics	Descriptive (N=24), n (%)	Inferential (N=17), n (%)
Chronic condition		
Mental health (n=12)	6 (25)	6 (35)
Chronic pain (n=5)	4 (17)	1 (6)
Asthma (n=4)	3 (13)	1 (6)
Cardiovascular disease (n=3)	2 (8)	1 (6)
Type 1 diabetes (n=3)	2 (8)	1 (6)
Type 2 diabetes (n=3)	2 (8)	1 (6)
Other (n=11)	5 (21)	6 (35)
Segmentation		
Yes (n=19)[a]	5 (21)	14 (82)
No (n=14)[a]	13 (54)	1 (6)
Analytic indicators		
Number of measures (n=31)	20 (83)	11 (65)
Frequency of interactions (n=30)[a]	14 (58)	16 (94)
Number of features (n=20)	11 (46)	9 (53)
Number of log-ins (n=19)	12 (50)	7 (41)
Number of modules (n=12)	7 (29)	5 (29)
Time spent (n=11)	8 (33)	3 (18)
Number of pages (n=7)	3 (13)	4 (24)
Structure		
Structured (n=7)[a]	7 (29)	0 (0)
Hybrid (n=24)	13 (54)	11 (65)
Unstructured (n=10)	4 (17)	6 (35)
Study design		
Experimental (n=21)	13 (54)	8 (47)
Quasi-experimental (n=7)	4 (17)	3 (18)
Observational (n=13)	7 (29)	6 (35)
Number of indicators		
1 (n=5)	3 (13)	2 (12)
2 (n=10)	7 (29)	3 (18)
3 (n=8)	3 (13)	5 (29)
4 (n=10)	7 (29)	3 (18)
5 (n=7)	3 (13)	4 (24)
6 (n=1)	1 (4)	0 (0)

[a]$P<.05$.

Discussion

Principal Findings

In conducting this scoping review, we sought to catalog the range of analytic indicators being used in evaluations of consumer mHealth apps for chronic conditions. We applied Arksey and O'Malley's methods of reporting and provided a descriptive analysis of the extent, nature, and distribution of analytic indicators across 41 studies, as well as a narrative and thematic summary of collected data [27]. The average mHealth evaluation included for review was a two-group pretest-posttest RCT of a hybrid-structured app for mental health self-management, had 103 participants, lasted 5 months, did not provide access to health care provider services, measured 3 analytic indicators of engagement, segmented users based on

engagement data, applied engagement data for descriptive analyses, and did not report on attrition.

Analytic Indicators

Our results indicate that researchers are measuring engagement across 7 analytic indicators, specifically: (1) the number of measures recorded, (2) the frequency of interactions logged, (3) the number of features accessed, (4) the number of log-ins or sessions logged, (5) the number of modules or lessons started or completed, (6) time spent engaging with the app, and (7) the number or content of pages accessed. We found that the researchers favored evaluating the number of measures recorded on an app as an indicator of engagement, closely followed by the frequency of interactions logged. We also found that both these indicators were most often used to assess hybrid and unstructured apps; these 2 app structures also made up the majority of apps under review.

We noted that researchers were least likely to measure the number of pages accessed and time spent engaging with the app; the latter indicator was mostly reported descriptively (73%, 8/11). This finding was surprising given the historical popularity of these indicators for measuring engagement with Web-based interventions [17,23,80]. The breadth and duration categories that conceptually comprise these analytic indicators were also deprioritized. We propose that these indicators are falling out of favor because of the growing recognition that users engage differently with apps. Users perceive apps to be a short-term commitment [81] and access app-based content sporadically for shorter periods of time compared with Web-based interventions [82]. Recent research by Morrison et al comparing patterns of engagement with a stress management intervention delivered via website versus app mitigated these differences by significantly reducing the number of pages on the app version of the intervention compared with the website [83]. They subsequently found that app users logged in twice as often but spent half as much time engaging compared with website users. They did not report the number of pages accessed or time spent engaging with the app as indicators of engagement. This body of research, in conjunction with our own findings, suggests that researchers evaluating mHealth apps for self-managing chronic conditions should refrain from measuring and reporting these 2 analytic indicators of engagement unless they are expressly relevant to the app under study.

Our identification of the number of measures recorded on an app as an analytic indicator of engagement deviates from previous research by Sieverink et al on usage and adherence to eHealth interventions [17], which found no evidence that researchers were operationalizing constructs in this way. Our focus on reviewing studies of mHealth apps for self-managing chronic conditions may explain this finding, as these interventions encourage users to systematically record data and capture the variability of their disease state over time [84]. In thinking of the frequency of interactions logged as a common analytic indicator of engagement, we note that there has been a shift toward on-demand apps with features and functionality that users can engage with at their own discretion. Benchmarking engagement by time range provides more context on a user's intentions and needs than just the total amount of engagement.

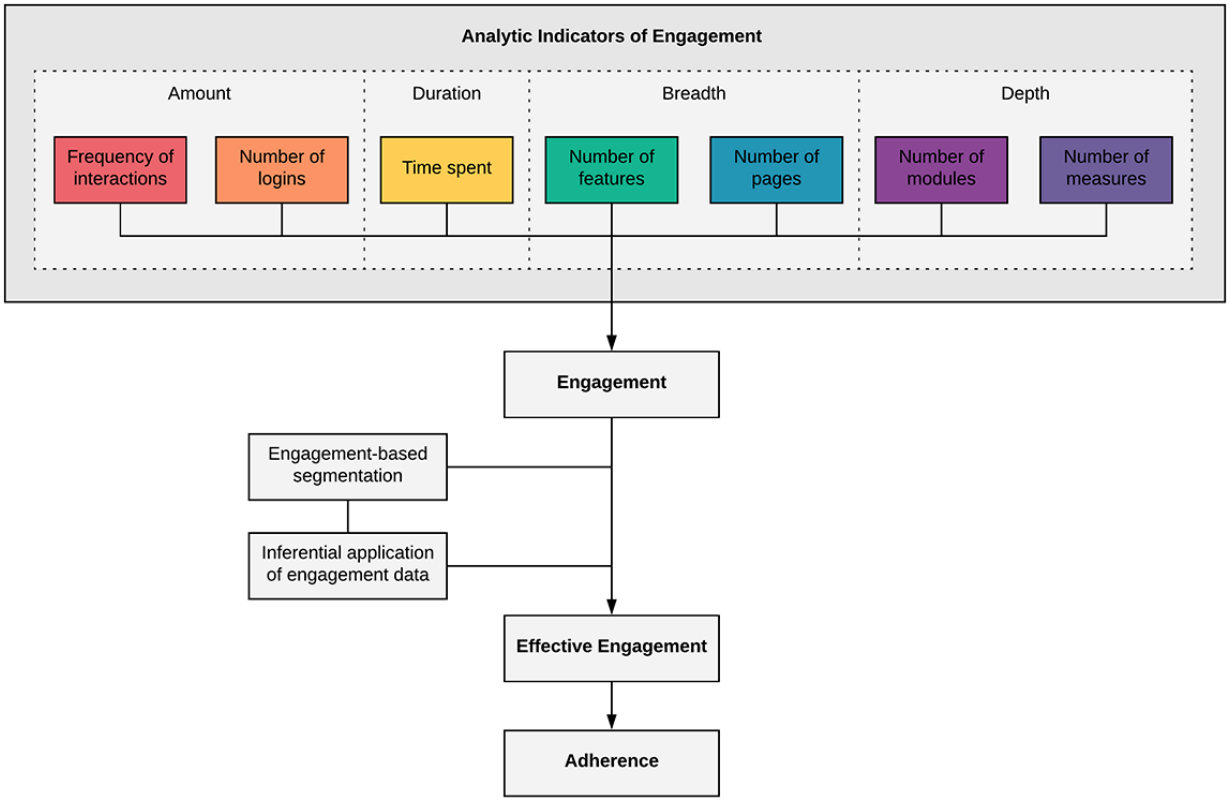
We did not observe any significant differences between the number or type of analytic indicators used to measure engagement across chronic conditions. Researchers applied indicators that were relevant to the features and functionality of their app. For example, studies of apps for diabetes self-management often measured the number of blood glucose readings due to the popularity of this feature but never measured the number of modules or lessons because these features were not offered to users. In a recent review on the barriers and facilitators of engagement with remote measurement technology for managing health, Simblett et al found that studies were reporting idiosyncratic measures of engagement and adherence that were not comparable across studies [26]. Their findings align with our own, and support Yardley et al's assertion that effective engagement is defined in relation to the purpose of a specific intervention and can only be established empirically in the context of that intervention [14]. Although Simblett et al call for less variation in how engagement is quantitatively measured across studies, we propose that researchers continue to apply context-specific analytic indicators but report them more systematically to enable cross-study comparison. Researchers might consider categorizing indicators according to the 7 domains identified in this research and providing detailed specifications on the analytic tags required to implement each indicator. When reporting on indicators, researchers should specify that they are measuring the construct of engagement and then catalog each domain. This practice may contribute to greater taxonomic consensus by curbing the arbitrary reporting of engagement-related constructs identified in this review.

Application of Engagement Data

Although researchers measured, on average, 3 indicators in a single study, the majority reported findings descriptively and did not further investigate how engagement with an app contributed to its impact on health and well-being. This finding suggests that researchers are gaining nuanced insights into how users are engaging with their apps but are not conducting inferential analyses to characterize effective engagement for improved outcomes. Relating analytic engagement patterns to behavior change and intended outcomes has been advocated across the behavioral and computational sciences [14,15,24,85,86], with recent efforts made to equip researchers with strategies for performing inferential analyses on engagement data [22,87,88]. Our analyses indicated that studies of structured apps were more likely to only report descriptive statistics on engagement data. Given that structured apps primarily require users to follow a predetermined engagement pathway and complete a series of milestones, it is reasonable for researchers to report on completion rates and identify drop-off points. However, it may be helpful to conduct inferential analyses to understand if completion of an app-mediated program is required to achieve intended outcomes, or whether users may derive proportional benefits from progressing through stages of the program. Of the studies that applied inferential statistics, most measured the number of days, week, or months users engaged with an app. This finding suggests that researchers consider a temporal understanding of

engagement to be important in determining a predictive effect on intended outcomes.

Figure 2. Process model of methodological continuum for evaluating mobile health engagement to adherence.



Recommendations

In their systematic review, Sieverink et al found that over half of all reviewed studies measured adherence to eHealth interventions using a single analytic indicator, and a quarter used 2 indicators [17]. The authors conclude that a limited but deliberate set of only one of 2 different indicators in accordance with the goal of the technology is sufficient to operationalize adherence. On reviewing how researchers were operationalizing adherence, they found that the majority reported adherence only in terms of how an intervention was used. The absence of a comparison to a threshold for intended use renders this operationalization incongruent with the definition of adherence. Instead, we propose that it aligns with the current understanding of engagement, which is more exploratory in nature and thus supports applying a greater number of analytic indicators.

In contrast to Sieverink et al's findings, the majority of our reviewed studies applied between 2 and 4 analytic indicators to measure engagement. This variance suggests that researchers are starting to recognize a conceptual and methodological distinction between the constructs of engagement and adherence. From these findings, we make the following recommendation: researchers seeking to gain a preliminary understanding of how users are engaging with their app are encouraged to apply all relevant analytic indicators from those identified in this review. Multimedia Appendix 2 presents data that may support researchers to select indicators that have previously been measured for their target chronic condition or for an app with similar features and functionality. Upon generation of analytic findings, researchers might consider segmenting users by engagement behaviors to interrogate the data and refine their engagement models. Conducting inferential subgroup analyses with engagement as a predictor of observed health outcomes might uncover potential patterns of effective engagement and inform an operationalization of intended use. In this way, measuring engagement can be positioned on a methodological continuum toward determining adherence. Figure 2 presents a process model of our recommendations.

During our full-text review, we excluded a large number of studies because they did not include objective, quantifiable measurements using log data analytics. Some studies had users self-report their engagement, whereas others omitted reporting engagement altogether and solely related findings on app efficacy. One possible explanation for this gap might be that researchers are unfamiliar with how to derive analytic insights from their app. From our experience, the process of tagging interaction data to enable analytic insights requires deliberate foresight. A shared understanding between a researcher and a software developer of the research questions being answered is critical to determine how analytics data should be modeled. Multimedia Appendix 3 presents a use case for applying analytic tags to evaluate effective engagement.

Our final recommendation concerns the reporting of attrition in data-driven mHealth evaluations. In 2005, Eysenbach published landmark work on the *law of attrition* [10], which was his observation that a substantial portion of participants in

eHealth trials stop using the intervention before study end. He posits that attrition is a fundamental characteristic and methodological challenge in the evaluation of eHealth interventions and recommends that "usage metrics and determinants of attrition should be highlighted, measured, analyzed, and discussed" [10]. Our findings suggest that this counsel has not fully translated into practice in the mHealth field. There is less inclination to log and report on analytic indicators of disengagement. We encourage researchers to attribute the same value to attrition data as they currently do to engagement data, as both constructs provide consequential insights into the viability of an app in the real world.

Limitations

Some methodological limitations of our scoping review warrant discussion, the most significant being that we only reviewed articles published over a 2-year period. This sampling frame may not have captured a representative sample of mHealth literature. As such, we may have missed relevant studies published before November 2015 and after November 2017 that would have met our eligibility criteria. While we acknowledge that our sampling frame is limited in scoping the entire field of mHealth, we believe it captures the application of analytics within the field of mHealth. From our review of the literature before conducting our search, we identified a paucity of papers that focused on mHealth log data analyses. The systematic review on usage-based adherence to eHealth interventions conducted by Sieverink et al reviewed 62 papers, of which 7 were on smartphone-based interventions [17]. Of those 7 papers, 5 were published after 2016, and the other 2 were both published in 2013. Perski et al conducted a systematic review on engagement with digital behavior change interventions that comprised all studies up to November 2015 [11]. They reviewed 113 studies, of which 13 were on mobile phone–based interventions. Only 4 of those studies applied log data analyses to study engagement with the intervention. These insights confirm that our scoping review did not include all studies that applied log data analyses to study engagement with mHealth apps. However, they also suggest that the number of studies we omitted is small. Our sampling frame of November 2015 to November 2017 directly follows Perski et al's review and includes 41 studies to address our specific research questions. For these reasons, we posit that our sample is sufficiently robust to provide a representative understanding of how analytics are being applied to study engagement with mHealth apps. Due to limited resources, only 1 reviewer conducted the electronic searches and screened all titles and abstracts against eligibility criteria, thereby potentially introducing bias. We did not assess the quality of included articles; however, this is in line with our review framework, which does not mandate this methodological practice. Finally, we did not map analytic indicators to the 14 identified engagement-related constructs for analysis. We acknowledge that conceptual differences exist between some of these constructs (eg, usage, feasibility, and adherence), and it is possible to use multiple constructs in the same study. However, we reviewed each construct and its analytic operationalizations separately during our data extraction process and could not discern significant differences. As such, we feel that we have included a homogenous body of research in this review and provided accurate insights into how researchers have used analytic indicators to measure engagement.

Conclusions

To date, the potential for mHealth apps to positively impact chronic health outcomes has not yet been realized [89]. This is, in part, due to the difficulties of generating a solid evidence base to guide clinical, policy, and regulatory decision making [90]. Indeed, the mHealth field has been reproached for arguing that apps warrant *digital exceptionalism* given the iterative nature of their design and the prohibitive cost of trials compared with their perceived level of risk [91]. We propose that our review supports researchers to harness these natural attributes for conducting data-driven evaluations of digitally mediated behavior change. Without objective knowledge of how users engage with an app to care for themselves, the mechanisms of action that underlie complex models of digitally mediated behavior change cannot be identified.

Our proposed library of analytic indicators to evaluate effective engagement with consumer mHealth apps for chronic conditions may be of value to researchers as a resource to support their evaluative practice. Researchers can systematically incorporate these analytic indicators into their study measures by adding analytic tags to their app's source code, allowing them to measure engagement without creating user burden or reactivity. Once generated, these data can be used in inferential analyses to delineate relationships with observed health outcomes. Researchers can further interrogate these data by conducting rapid cycles of research and development to validate hypothesized models of effective engagement. On the basis of these insights, researchers can (1) build a cumulative body of evidence for how users should engage with their app to achieve intended outcomes, (2) incrementally improve their app to optimize effective engagement, and (3) determine the optimal digital dose of effective engagement with their app for validation in a definitive trial to meet required levels of evidence for procurement and distribution [92]. Successful implementation of these practices may elevate the discourse of these apps beyond the coarse evaluations and monolithic policy recommendations against their value in health care.

Raising the standard of mHealth app efficacy through measuring analytic indicators of engagement may enable greater confidence in the causal impact of apps on improved chronic health and well-being. It is this opportunity afforded by data-driven research to close the gap between promised and realized health benefits that is most meaningful.

Acknowledgments

The authors wish to thank Ms Vincci Lui, Faculty Liaison and Instruction Librarian for the Institute of Health Policy, Management and Evaluation, for guiding our search strategy. The authors are grateful to Dr Chitra Lalloo for reviewing an early draft of this manuscript and providing comments.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Scoping review search strategy.

[PDF File (Adobe PDF File), 24KB - mhealth_v7i1e11941_app1.pdf]

Multimedia Appendix 2

Full dataset of coded analytic indicators.

[XLSX File (Microsoft Excel File), 27KB - mhealth_v7i1e11941_app2.xlsx]

Multimedia Appendix 3

Practical considerations for applying analytic indicators to evaluate effective engagement.

[PDF File (Adobe PDF File), 27KB - mhealth_v7i1e11941_app3.pdf]

References

1. 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):1089-1098 [FREE Full text] [doi: 10.1038/s41746-018-0021-9]
2. Proudfoot J, Clarke J, Birch MR, Whitton AE, Parker G, Manicavasagar V, et al. Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: a randomised controlled trial. BMC Psychiatry 2013 Nov 18;13:312 [FREE Full text] [doi: 10.1186/1471-244X-13-312] [Medline: 24237617]
3. Clarke J, Proudfoot J, Birch MR, Whitton AE, Parker G, Manicavasagar V, et al. Effects of mental health self-efficacy on outcomes of a mobile phone and web intervention for mild-to-moderate depression, anxiety and stress: secondary analysis of a randomised controlled trial. BMC Psychiatry 2014 Sep 26;14:272 [FREE Full text] [doi: 10.1186/s12888-014-0272-1] [Medline: 25252853]
4. Ivanova E, Lindner P, Ly KH, Dahlin M, Vernmark K, Andersson G, et al. Guided and unguided acceptance and commitment therapy for social anxiety disorder and/or panic disorder provided via the internet and a smartphone application: a randomized controlled trial. J Anxiety Disord 2016 Dec;44:27-35 [FREE Full text] [doi: 10.1016/j.janxdis.2016.09.012] [Medline: 27721123]
5. Roepke AM, Jaffee SR, Riffle OM, McGonigal J, Broome R, Maxwell B. Randomized controlled trial of SuperBetter, a smartphone-based/internet-based self-help tool to reduce depressive symptoms. Games Health J 2015 Jun;4(3):235-246 [FREE Full text] [doi: 10.1089/g4h.2014.0046] [Medline: 26182069]
6. Laing BY, Mangione CM, Tseng CH, Leng M, Vaisberg E, Mahida M, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. Ann Intern Med 2014 Nov 18;161(10 Suppl):S5-12 [FREE Full text] [doi: 10.7326/M13-3005] [Medline: 25402403]
7. Direito A, Jiang Y, Whittaker R, Maddison R. Smartphone apps to improve fitness and increase physical activity among young people: protocol of the Apps for IMproving FITness (AIMFIT) randomized controlled trial. BMC Public Health 2015 Jul 11;15:635 [FREE Full text] [doi: 10.1186/s12889-015-1968-y] [Medline: 26159834]
8. Turner-McGrievy G, Tate D. Tweets, apps, and pods: results of the 6-month Mobile Pounds Off Digitally (Mobile POD) randomized weight-loss intervention among adults. J Med Internet Res 2011 Dec 20;13(4):e120 [FREE Full text] [doi: 10.2196/jmir.1841] [Medline: 22186428]
9. Holmen H, Torbjørnsen A, Wahl AK, Jenum AK, Småstuen MC, Arsand E, et al. A mobile health intervention for self-management and lifestyle change for persons with type 2 diabetes, part 2: one-year results from the Norwegian randomized controlled trial RENEWING HEALTH. JMIR Mhealth Uhealth 2014 Dec 11;2(4):e57 [FREE Full text] [doi: 10.2196/mhealth.3882] [Medline: 25499872]
10. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]

11. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Dec;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
12. O'Brien H, Toms EG. What is user engagement? A conceptual framework for defining user engagement with technology. J Am Soc Inf Sci 2008 Apr;59(6):938-955. [doi: 10.1002/asi.20801]
13. Pham Q, Graham G, Lalloo C, Morita PP, Seto E, Stinson JN, et al. An Analytics Platform to Evaluate Effective Engagement With Pediatric Mobile Health Apps: Design, Development, and Formative Evaluation. JMIR Mhealth Uhealth 2018 Dec 21;6(12):e11447 [FREE Full text] [doi: 10.2196/11447] [Medline: 30578179]
14. Yardley L, Spring BJ, Riper H, Morrison LG, Crane DH, Curtis K, et al. Understanding and promoting effective engagement with digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):833-842. [doi: 10.1016/j.amepre.2016.06.015] [Medline: 27745683]
15. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Dec 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
16. Barello S, Triberti S, Graffigna G, Libreri C, Serino S, Hibbard J, et al. eHealth for patient engagement: a systematic review. Front Psychol 2016 Jan 08;6:2013 [FREE Full text] [doi: 10.3389/fpsyg.2015.02013] [Medline: 26779108]
17. Sieverink F, Kelders SM, van Gemert-Pijnen JE. Clarifying the concept of adherence to eHealth technology: systematic review on when usage becomes adherence. J Med Internet Res 2017 Dec 06;19(12):e402 [FREE Full text] [doi: 10.2196/jmir.8578] [Medline: 29212630]
18. Hervatis V, Loe A, Barman L, O'Donoghue J, Zary N. A conceptual analytics model for an outcome-driven quality management framework as part of professional healthcare education. JMIR Med Educ 2015 Oct 06;1(2):e11 [FREE Full text] [doi: 10.2196/mededu.4789] [Medline: 27731840]
19. Kotz D, Lord SE, O'Malley AJ, Stark L, Marsch LA. Workshop on emerging technology and data analytics for behavioral health. JMIR Res Protoc 2018 Jun 20;7(6):e158 [FREE Full text] [doi: 10.2196/resprot.9589] [Medline: 29925493]
20. Van Gemert-Pijnen JE, Kelders SM, Bohlmeijer ET. Understanding the usage of content in a mental health intervention for depression: an analysis of log data. J Med Internet Res 2014 Jan 31;16(1):e27 [FREE Full text] [doi: 10.2196/jmir.2991] [Medline: 24486914]
21. Sieverink F, Kelders SM, Braakman-Jansen LM, van Gemert-Pijnen JE. The added value of log file analyses of the use of a personal health record for patients with type 2 diabetes mellitus: preliminary results. J Diabetes Sci Technol 2014 Mar;8(2):247-255 [FREE Full text] [doi: 10.1177/1932296814525696] [Medline: 24876574]
22. Sieverink F, Kelders S, Poel M, van Gemert-Pijnen L. Opening the black box of electronic health: collecting, analyzing, and interpreting log data. JMIR Res Protoc 2017 Aug 07;6(8):e156 [FREE Full text] [doi: 10.2196/resprot.6452] [Medline: 28784592]
23. Arden-Close EJ, Smith E, Bradbury K, Morrison L, Dennison L, Michaelides D, et al. A visualization tool to analyse usage of web-based interventions: the example of Positive Online Weight Reduction (POWeR). JMIR Hum Factors 2015 May 19;2(1):e8 [FREE Full text] [doi: 10.2196/humanfactors.4310] [Medline: 27026372]
24. Hekler EB, Klasnja P, Riley WT, Buman MP, Huberty J, Rivera DE, et al. Agile science: creating useful products for behavior change in the real world. Transl Behav Med 2016 Dec;6(2):317-328 [FREE Full text] [doi: 10.1007/s13142-016-0395-7] [Medline: 27357001]
25. Short C, DeSmet A, Woods C, Williams SL, Maher C, Middelweerd A, et al. Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. J Med Internet Res 2018 Nov 16;20(11):e292 [FREE Full text] [doi: 10.2196/jmir.9397] [Medline: 30446482]
26. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and facilitators of engagement with remote measurement technology for managing health: systematic review and content analysis of findings. J Med Internet Res 2018 Jul 12;20(7):e10480 [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]
27. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb 01;8(1):19-32 [FREE Full text] [doi: 10.1080/1364557032000119616]
28. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010 Sep 20;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
29. Global Status Report on Noncommunicable Diseases. In: World Health Organization. Geneva: World Health Organization; 2010.
30. Elsevier. Mendeley Internet URL: https://www.mendeley.com/?interaction_required=true [accessed 2018-12-22] [WebCite Cache ID 74qvREYJ5]
31. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012 Nov 14;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
32. IBM Corporation. IBM SPSS Statistics URL: https://www.ibm.com/products/spss-statistics [accessed 2018-12-22] [WebCite Cache ID 74qvec8Se]

33. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. 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]
34. Serrano KJ, Coa KI, Yu M, Wolff-Hughes DL, Atienza AA. Characterizing user engagement with health app data: a data mining approach. Transl Behav Med Internet 2017 Jun;7(2):277-285 [FREE Full text] [doi: 10.1007/s13142-017-0508-y] [Medline: 28616846]
35. Beiwinkel T, Kindermann S, Maier A, Kerl C, Moock J, Barbian G, et al. Using smartphones to monitor bipolar disorder a symptoms: a pilot study. JMIR Ment Health 2016 Jan 06;3(1):e2 [FREE Full text] [doi: 10.2196/mental.4560] [Medline: 26740354]
36. Ben-Zeev D, Brian R, Aschbrenner KA, Jonathan G, Steingard S. Video-based mobile health interventions for people with schizophrenia: bringing the "pocket therapist" to life. Psychiatr Rehabil J 2018 Mar;41(1):39-45 [FREE Full text] [doi: 10.1037/prj0000197] [Medline: 27295133]
37. Ben-Zeev D, Scherer EA, Gottlieb JD, Rotondi AJ, Brunette MF, Achtyes ED, et al. mHealth for schizophrenia: patient engagement with a mobile phone intervention following hospital discharge. JMIR Ment Health 2016 Jul 27;3(3):e34 [FREE Full text] [doi: 10.2196/mental.6348] [Medline: 27465803]
38. Davies EB, Craven MP, Martin JL, Simons L. Proportionate methods for evaluating a simple digital mental health tool. Evid Based Ment Health 2017 Nov;20(4):112-117 [FREE Full text] [doi: 10.1136/eb-2017-102755] [Medline: 28993317]
39. Frisbee KL. Variations in the use of mHealth tools: the VA mobile health Study. JMIR Mhealth Uhealth 2016 Jul 19;4(3):e89 [FREE Full text] [doi: 10.2196/mhealth.3726] [Medline: 27436165]
40. Kinderman P, Hagan P, King S, Bowman J, Chahal J, Gan L, et al. The feasibility and effectiveness of Catch It, an innovative CBT smartphone app. BJPsych Open 2016 May;2(3):204-209 [FREE Full text] [doi: 10.1192/bjpo.bp.115.002436] [Medline: 27703777]
41. Kuhn E, Kanuri N, Hoffman JE, Garvert DW, Ruzek JI, Taylor CB. A randomized controlled trial of a smartphone app for posttraumatic stress disorder symptoms. J Consult Clin Psychol 2017 Mar;85(3):267-273 [FREE Full text] [doi: 10.1037/ccp0000163] [Medline: 28221061]
42. Owen JE, Jaworski BK, Kuhn E, Makin-Byrd KN, Ramsey KM, Hoffman JE. mHealth in the wild: using novel data to examine the reach, use, and impact of PTSD coach. JMIR Ment Health 2015 Mar 25;2(1):e7 [FREE Full text] [doi: 10.2196/mental.3935] [Medline: 26543913]
43. Pham Q, Khatib Y, Stansfeld S, Fox S, Green T. Feasibility and efficacy of an mHealth game for managing anxiety: "Flowy" randomized controlled pilot trial and design evaluation. Games Health J 2016 Feb;5(1):50-67 [FREE Full text] [doi: 10.1089/g4h.2015.0033] [Medline: 26536488]
44. Torous J, Staples P, Slaters L, Adams J, Sandoval L, Onnela JP, et al. Characterizing smartphone engagement for schizophrenia: results of a naturalist mobile health study. Clin Schizophr Relat Psychoses 2017 Aug 04 Epub ahead of print [FREE Full text] [doi: 10.3371/CSRP.JTPS.071317] [Medline: 28777029]
45. Vansimaeys C, Zuber M, Pitrat B, Join-Lambert C, Tamazyan R, Farhat W, et al. Combining standard conventional measures and ecological momentary assessment of depression, anxiety and coping using smartphone application in minor stroke population: a longitudinal study protocol. Front Psychol 2017 Jul 12;8:1172 [FREE Full text] [doi: 10.3389/fpsyg.2017.01172] [Medline: 28747895]
46. Wahle F, Kowatsch T, Fleisch E, Rufer M, Weidt S. Mobile sensing and support for people with depression: a pilot trial in the wild. JMIR Mhealth Uhealth 2016 Sep 21;4(3):e111 [FREE Full text] [doi: 10.2196/mhealth.5960] [Medline: 27655245]
47. Fortier MA, Chung WW, Martinez A, Gago-Masague S, Sender L. Pain buddy: a novel use of m-health in the management of children's cancer pain. Comput Biol Med 2016 Dec 01;76:202-214 [FREE Full text] [doi: 10.1016/j.compbiomed.2016.07.012] [Medline: 27479493]
48. Jamison RN, Jurcik DC, Edwards RR, Huang CC, Ross EL. A pilot comparison of a smartphone app with or without 2-way messaging among chronic pain patients: who benefits from a pain app? Clin J Pain 2017 Aug;33(8):676-686 [FREE Full text] [doi: 10.1097/AJP.0000000000000455] [Medline: 27898460]
49. Jibb LA, Stevens BJ, Nathan PC, Seto E, Cafazzo JA, Johnston DL, et al. Implementation and preliminary effectiveness of a real-time pain management smartphone app for adolescents with cancer: a multicenter pilot clinical study. Pediatr Blood Cancer 2017 Oct;64(10) [FREE Full text] [doi: 10.1002/pbc.26554] [Medline: 28423223]
50. Reade S, Spencer K, Sergeant JC, Sperrin M, Schultz DM, Ainsworth J, et al. Cloudy with a chance of pain: engagement and subsequent attrition of daily data entry in a smartphone pilot study tracking weather, disease severity, and physical activity in patients with rheumatoid arthritis. JMIR Mhealth Uhealth 2017 Mar 24;5(3):e37 [FREE Full text] [doi: 10.2196/mhealth.6496] [Medline: 28341616]
51. Skrepnik N, Spitzer A, Altman R, Hoekstra J, Stewart J, Toselli R. Assessing the impact of a novel smartphone application compared with standard follow-up on mobility of patients with knee osteoarthritis following treatment with hylan G-F 20: a randomized controlled trial. JMIR Mhealth Uhealth 2017 May 09;5(5):e64 [FREE Full text] [doi: 10.2196/mhealth.7179] [Medline: 28487266]

52. 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 Apr;35(4):354-362 [FREE Full text] [doi: 10.1038/nbt.3826] [Medline: 28288104]
53. Cook KA, Modena BD, Simon RA. Improvement in asthma control using a minimally burdensome and proactive smartphone application. J Allergy Clin Immunol Pract 2016;4(4):730-7.e1 [FREE Full text] [doi: 10.1016/j.jaip.2016.03.005] [Medline: 27107690]
54. Fedele DA, McConville A, Graham Thomas J, McQuaid EL, Janicke DM, Turner EM, et al. Applying interactive mobile health to asthma care in teens (AIM2ACT): development and design of a randomized controlled trial. Contemp Clin Trials 2018 Dec;64:230-237. [doi: 10.1016/j.cct.2017.09.007] [Medline: 28986245]
55. Kosse RC, Bouvy ML, de Vries TW, Kaptein AA, Geers HC, van Dijk L, et al. mHealth intervention to support asthma self-management in adolescents: the ADAPT study. Patient Prefer Adherence 2017 Mar 16;11:571-577 [FREE Full text] [doi: 10.2147/PPA.S124615] [Medline: 28356720]
56. Agboola S, Palacholla RS, Centi A, Kvedar J, Jethwani K. A multimodal mHealth intervention (FeatForward) to improve physical activity behavior in patients with high cardiometabolic risk factors: rationale and protocol for a randomized controlled trial. JMIR Res Protoc 2016 May 12;5(2):e84 [FREE Full text] [doi: 10.2196/resprot.5489] [Medline: 27174783]
57. Goyal S, Morita PP, Picton P, Seto E, Zbib A, Cafazzo JA. Uptake of a consumer-focused mHealth application for the assessment and prevention of heart disease: the <30 days study. JMIR Mhealth Uhealth 2016 Mar 24;4(1):e32 [FREE Full text] [doi: 10.2196/mhealth.4730] [Medline: 27012937]
58. Sakakibara BM, Ross E, Arthur G, Brown-Ganzert L, Petrin S, Sedlak T, et al. Using mobile-health to connect women with cardiovascular disease and improve self-management. Telemed J E Health 2017 Mar;23(3):233-239 [FREE Full text] [doi: 10.1089/tmj.2016.0133] [Medline: 27623231]
59. Goyal S, Nunn CA, Rotondi M, Couperthwaite AB, Reiser S, Simone A, et al. A mobile app for the self-management of type 1 diabetes among adolescents: a randomized controlled trial. JMIR Mhealth Uhealth 2017 Jun 19;5(6):e82 [FREE Full text] [doi: 10.2196/mhealth.7336] [Medline: 28630037]
60. Ryan EA, Holland J, Stroulia E, Bazelli B, Babwik SA, Li H, et al. Improved A1C levels in type 1 diabetes with smartphone App use. Can J Diabetes 2017 Feb;41(1):33-40. [doi: 10.1016/j.jcjd.2016.06.001] [Medline: 27570203]
61. Sieber J, Flacke F, Link M, Haug C, Freckmann G. Improved glycemic control in a patient group performing 7-point profile self-monitoring of blood glucose and intensive data documentation: an open-label, multicenter, observational study. Diabetes Ther 2017 Oct;8(5):1079-1085 [FREE Full text] [doi: 10.1007/s13300-017-0306-z] [Medline: 28913822]
62. Desveaux L, Agarwal P, Shaw J, Hensel JM, Mukerji G, Onabajo N, et al. A randomized wait-list control trial to evaluate the impact of a mobile application to improve self-management of individuals with type 2 diabetes: a study protocol. BMC Med Inform Decis Mak 2016 Nov 15;16(1):144 [FREE Full text] [doi: 10.1186/s12911-016-0381-5] [Medline: 27842539]
63. Goh G, Tan NC, Malhotra R, Padmanabhan U, Barbier S, Allen JC, et al. Short-term trajectories of use of a caloric-monitoring mobile phone app among patients with type 2 diabetes mellitus in a primary care setting. J Med Internet Res 2015 Feb 03;17(2):e33 [FREE Full text] [doi: 10.2196/jmir.3938] [Medline: 25648130]
64. Kleinman NJ, Shah A, Shah S, Phatak S, Viswanathan V. Improved medication adherence and frequency of blood glucose self-testing using an m-Health platform versus usual care in a multisite randomized clinical trial among people with type 2 diabetes in India. Telemed J E Health 2017 Sep;23(9):733-740 [FREE Full text] [doi: 10.1089/tmj.2016.0265] [Medline: 28328396]
65. Bot BM, Suver C, Neto EC, Kellen M, Klein A, Bare C, et al. The mPower study, Parkinson disease mobile data collected using ResearchKit. Sci Data 2016 Mar 03;3:160011 [FREE Full text] [doi: 10.1038/sdata.2016.11] [Medline: 26938265]
66. Hardinge M, Rutter H, Velardo C, Shah SA, Williams V, Tarassenko L, et al. Using a mobile health application to support self-management in chronic obstructive pulmonary disease: a six-month cohort study. BMC Med Inform Decis Mak 2015 Jun 18;15:46 [FREE Full text] [doi: 10.1186/s12911-015-0171-5] [Medline: 26084626]
67. Isetta V, Torres M, González K, Ruiz C, Dalmases M, Embid C, et al. A new mHealth application to support treatment of sleep apnoea patients. J Telemed Telecare 2017 Jan;23(1):14-18 [FREE Full text] [doi: 10.1177/1357633X15621848] [Medline: 26672606]
68. Kaplan HC, Thakkar SN, Burns L, Chini B, Dykes DM, McPhail GL, et al. Protocol of a pilot study of technology-enabled coproduction in pediatric chronic illness care. JMIR Res Protoc 2017 Apr 28;6(4):e71 [FREE Full text] [doi: 10.2196/resprot.7074] [Medline: 28455274]
69. Langius-Eklöf A, Crafoord MT, Christiansen M, Fjell M, Sundberg K. Effects of an interactive mHealth innovation for early detection of patient-reported symptom distress with focus on participatory care: protocol for a study based on prospective, randomised, controlled trials in patients with prostate and breast cancer. BMC Cancer 2017 Jul 04;17(1):466 [FREE Full text] [doi: 10.1186/s12885-017-3450-y] [Medline: 28676102]
70. Ong S, Jassal SV, Miller JA, Porter EC, Cafazzo JA, Seto E, et al. Integrating a smartphone-based self-management system into usual care of advanced CKD. Clin J Am Soc Nephrol 2016 Dec 06;11(6):1054-1062 [FREE Full text] [doi: 10.2215/CJN.10681015] [Medline: 27173169]

71. Pham Q, Cafazzo JA, Feifer A. Adoption, acceptability, and effectiveness of a mobile health app for personalized prostate cancer survivorship care: protocol for a realist case study of the Ned App. JMIR Res Protoc 2017 Oct 12;6(10):e197 [FREE Full text] [doi: 10.2196/resprot.8051] [Medline: 29025699]
72. Taki S, Lymer S, Russell CG, Campbell K, Laws R, Ong KL, et al. Assessing user engagement of an mHealth intervention: development and implementation of the growing healthy app engagement index. JMIR Mhealth Uhealth 2017 Jun 29;5(6):e89 [FREE Full text] [doi: 10.2196/mhealth.7236] [Medline: 28663164]
73. 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 [FREE Full text] [doi: 10.2196/humanfactors.7709] [Medline: 28974481]
74. Toro-Ramos T, Kim Y, Wood M, Rajda J, Niejadlik K, Honcz J, et al. Efficacy of a mobile hypertension prevention delivery platform with human coaching. J Hum Hypertens 2017 Dec;31(12):795-800 [FREE Full text] [doi: 10.1038/jhh.2017.69] [Medline: 28972573]
75. Beiwinkel T, Hey S, Bock O, Rössler W. Supportive mental health self-monitoring among smartphone users with psychological distress: protocol for a fully mobile randomized controlled trial. Front Public Health 2017;5:249 [FREE Full text] [doi: 10.3389/fpubh.2017.00249] [Medline: 28983477]
76. Lei H, Nahum-Shani I, Lynch K, Oslin D, Murphy SA. A "SMART" design for building individualized treatment sequences. Annu Rev Clin Psychol 2012;8:21-48 [FREE Full text] [doi: 10.1146/annurev-clinpsy-032511-143152] [Medline: 22224838]
77. Almirall D, Nahum-Shani I, Sherwood NE, Murphy SA. Introduction to SMART designs for the development of adaptive interventions: with application to weight loss research. Transl Behav Med 2014 Sep;4(3):260-274 [FREE Full text] [doi: 10.1007/s13142-014-0265-0] [Medline: 25264466]
78. Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. Am J Prev Med 2007 May;32(5 Suppl):S112-S118 [FREE Full text] [doi: 10.1016/j.amepre.2007.01.022] [Medline: 17466815]
79. Serrano KJ, Yu M, Coa KI, Collins LM, Atienza AA. Mining health app data to find more and less successful weight loss subgroups. J Med Internet Res 2016 Dec 14;18(6):e154 [FREE Full text] [doi: 10.2196/jmir.5473] [Medline: 27301853]
80. Morrison C, Doherty G. Analyzing engagement in a web-based intervention platform through visualizing log-data. J Med Internet Res 2014 Nov 13;16(11):e252 [FREE Full text] [doi: 10.2196/jmir.3575] [Medline: 25406097]
81. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013 Apr 18;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
82. Morrison LG, Hargood C, Lin SX, Dennison L, Joseph J, Hughes S, et al. Understanding usage of a hybrid website and smartphone app for weight management: a mixed-methods study. J Med Internet Res 2014 Oct 22;16(10):e201 [FREE Full text] [doi: 10.2196/jmir.3579] [Medline: 25355131]
83. Morrison LG, Geraghty AW, Lloyd S, Goodman N, Michaelides DT, Hargood C, et al. Comparing usage of a web and app stress management intervention: an observational study. Internet Interv 2018 Jun 21;12:74-82 [FREE Full text] [doi: 10.1016/j.invent.2018.03.006] [Medline: 30135771]
84. Stinson JN, Lalloo C, Harris L, Isaac L, Campbell F, Brown S, et al. iCanCope with Pain™: user-centred design of a web- and mobile-based self-management program for youth with chronic pain based on identified health care needs. Pain Res Manag 2014;19(5):257-265 [FREE Full text] [Medline: 25000507]
85. Yardley L, Choudhury T, Patrick K, Michie S. Current issues and future directions for research into digital behavior change interventions. Am J Prev Med 2016 Nov;51(5):814-815. [doi: 10.1016/j.amepre.2016.07.019] [Medline: 27745680]
86. Patrick K, Hekler EB, Estrin D, Mohr DC, Riper H, Crane D, et al. The pace of technologic change: implications for digital health behavior intervention research. Am J Prev Med 2016 Nov;51(5):816-824. [doi: 10.1016/j.amepre.2016.05.001] [Medline: 27745681]
87. Scherer E, Ben-Zeev D, Li Z, Kane J. Analyzing mHealth engagement: joint models for intensively collected user engagement data. JMIR Mhealth Uhealth 2017 Jan 12;5(1):e1 [FREE Full text] [doi: 10.2196/mhealth.6474] [Medline: 28082257]
88. Tignor N, Wang P, Genes N, Rogers L, Hershman SG, Scott ER, et al. Methods for clustering time series data acquired from mobile health apps. Pac Symp Biocomput 2017;22:300-311 [FREE Full text] [doi: 10.1142/9789813207813_0029] [Medline: 27896984]
89. Nilsen W. American Association for the Advancement of Science. 2015. mHealth's Revolution: Balancing Help and Harm URL: https://www.aaas.org/sites/default/files/Nilsen%20mHealths%20Revolution%20Balancing%20Help%20and%20Harm.pdf [accessed 2018-12-24] [WebCite Cache ID 74u134YRR]
90. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
91. The Lancet. Is digital medicine different? Lancet 2018 Dec 14;392(10142):95. [doi: 10.1016/S0140-6736(18)31562-9] [Medline: 30017135]

92. The National Institute for Health and Care Excellence. NICE IAPT assessment programm: Eligibility requirements of digitally enabled therapy technologies URL: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-advice/IAPT/eligibility-and-prioritisation-criteria.pdf [accessed 2018-12-22] [WebCite Cache ID 74qy3qHlR]

Abbreviations

eHealth: electronic health
IQR: interquartile range
mHealth: mobile health
PTSD: posttraumatic stress disorder
RCT: randomized controlled trial

Edited by S Kitsiou; submitted 14.08.18; peer-reviewed by S Kelders, H Oh, E Meinert; comments to author 28.09.18; revised version received 04.10.18; accepted 10.12.18; published 18.01.19.

Please cite as:
Pham Q, Graham G, Carrion C, Morita PP, Seto E, Stinson JN, Cafazzo JA
A Library of Analytic Indicators to Evaluate Effective Engagement with Consumer mHealth Apps for Chronic Conditions: Scoping Review
JMIR Mhealth Uhealth 2019;7(1):e11941
URL: http://mhealth.jmir.org/2019/1/e11941/
doi:10.2196/11941
PMID:30664463

Review

Mobile-Based Interventions for Dietary Behavior Change and Health Outcomes: Scoping Review

Atreyi Kankanhalli[1], PhD; Jieun Shin[1], MS, MBA; Hyelim Oh[1], PhD

Department of Information Systems and Analytics, National University of Singapore, Singapore, Singapore

Corresponding Author:
Hyelim Oh, PhD
Department of Information Systems and Analytics
National University of Singapore
13 Computing Drive
Singapore,
Singapore
Phone: 65 6516 6732
Fax: 65 6779 4580
Email: hyelim.oh@nus.edu.sg

Abstract

Background: Mobile apps are being widely used for delivering health interventions, with their ubiquitous access and sensing capabilities. One such use is the delivery of interventions for healthy eating behavior.

Objective: The aim of this study was to provide a comprehensive view of the literature on the use of mobile interventions for eating behavior change. We synthesized the studies with such interventions and mapped out their input methods, interventions, and outcomes.

Methods: We conducted a scoping literature search in PubMed/MEDLINE, Association for Computing Machinery Digital Library, and PsycINFO databases to identify relevant papers published between January 2013 and April 2018. We also hand-searched relevant themes of journals in the *Journal of Medical Internet Research* and registered protocols. Studies were included if they provided and assessed mobile-based interventions for dietary behavior changes and/or health outcomes.

Results: The search resulted in 30 studies that we classified by 3 main aspects: input methods, mobile-based interventions, and dietary behavior changes and health outcomes. First, regarding input methods, 5 studies allowed photo/voice/video inputs of diet information, whereas text input methods were used in the remaining studies. Other than diet information, the content of the input data in the mobile apps included user's demographics, medication, health behaviors, and goals. Second, we identified 6 categories of intervention contents, that is, self-monitoring, feedback, gamification, goal reviews, social support, and educational information. Although all 30 studies included self-monitoring as a key component of their intervention, personalized feedback was a component in 18 studies, gamification was used in 10 studies, goal reviews in 5 studies, social support in 3 studies, and educational information in 2 studies. Finally, we found that 13 studies directly examined the effects of interventions on health outcomes and 12 studies examined the effects on dietary behavior changes, whereas only 5 studies observed the effects both on dietary behavior changes and health outcomes. Regarding the type of studies, although two-thirds of the included studies conducted diverse forms of randomized control trials, the other 10 studies used field studies, surveys, protocols, qualitative interviews, propensity score matching method, and test and reference method.

Conclusions: This scoping review identified and classified studies on mobile-based interventions for dietary behavior change as per the input methods, nature of intervention, and outcomes examined. Our findings indicated that dietary behavior changes, although playing a mediating role in improving health outcomes, have not been adequately examined in the literature. Dietary behavior change as a mechanism for the relationship between mobile-based intervention and health outcomes needs to be further investigated. Our review provides guidance for future research in this promising mobile health area.

(JMIR Mhealth Uhealth 2019;7(1):e11312) doi:10.2196/11312

KEYWORDS

mHealth; mobile-based intervention; dietary behavior; food intake; behavior change; health outcomes

Introduction

Technology and Healthy Eating Promotion

Changes in lifestyle have resulted in dietary problems, such as consumption of high-calorie and low-nutrient foods. Consumption of such foods is associated with obesity and chronic diseases, such as cardiovascular disease and diabetes [1-3]. The problem of widespread obesity is a serious public health concern for individuals, care providers, and policy makers [2]. To mitigate these issues, technological interventions are being developed to encourage people to consume a diversified, balanced, and healthy diet depending on individual needs (eg, age, gender, and lifestyle), cultural context, locally available foods, and dietary customs. According to the Food and Agriculture Organization, most people make their dietary choices for personal reasons, for example, based on time constraints and convenience, personal preferences, and everyday habits [4], rather than basing them on good nutrition and health. Although individuals often realize that their eating behaviors are not ideal for their health and endeavor to change them, failure to maintain a long-lasting, healthy lifestyle is very common [5]. In this regard, technological interventions have the potential to assist in encouraging users to consume a healthy diet.

Mobile-Based Health Interventions

Particularly, mobile-based interventions have become a popular means for the promotion and continuity of self-management of users' health [6]. The distinct features that facilitate the adoption of mobile health (mHealth) apps are the ubiquity and sensing capabilities of mobile devices. Compared with Web-based interventions, mobile apps enable users to log their food intake behaviors and other activities throughout the day. Moreover, the technical capabilities of mobile devices equipped with sensors and computing power have enabled mobile app designers to develop interactive interventions that facilitate monitoring and self-management of health behaviors, such as physical activity [7,8].

At the same time, health psychology literature identifies various behavior change techniques that can be used as interventions to promote healthy behavior [9,10]. The change techniques cited include providing general health information, instructions, prompt reviews of behavior goals, and self-monitoring of users' behavior. Physical activity and eating behaviors are 2 of the most targeted health behaviors, because of their importance for health outcomes. In this regard, there is considerable literature reviewing the interventions for physical activity, for example, exercise and fitness [7,8].

With diverse behavior change techniques, mobile interventions on users' diets can lead to weight loss, diabetes management, or health promotion, in general [11-14]. However, despite the need for a systematic investigation of such interventions for dietary behavior changes, there are limited related reviews [15,16]. This could be partly because of the erstwhile difficulties of measuring a person's dietary behavior and then improving on it [15-17]. As one approach, Roy et al [18] assessed healthy eating behavior by measuring if patients' diets are healthy through allocating scores for consuming a variety of foods or recommended foods and nutrients. Further, in prior research on mobile interventions for dietary behavior changes, healthy eating behavior is suggested as a mediator for health outcomes because it is a crucial step for attaining the outcomes [13,19-32]. Yet, some studies only examined the dietary behavior changes, for example, fruit and vegetable intake change, after mobile-based interventions [19-24,27-30]. We found only a few studies assessing both dietary behavior changes and health outcomes [13,25,26,31,32]. However, other researchers have directly studied the effects of mobile-based interventions on health outcomes (eg, weight loss) [11,12,14,33-42].

Research Gap

To the best of our knowledge, previous literature reviews have not incorporated a comprehensive list of studies on mobile interventions for dietary behavior change. Bardus et al [43], DiFilippo et al [15], and Nour et al [16] are the closest to our scoping review, yet their focus differs from our review. Bardus et al [43] examined both the use of mobile phones and websites for weight management, and mainly focused on the comparison between these 2 technologies. DiFilippo et al [15] included only 4 studies in their review, which evaluated weight loss as the outcome of better nutrition, while excluding mobile interventions using text messaging or digital photography. Nour et al [16] focused on mobile interventions whose objective was limited to increasing vegetable intake in young adults. In contrast, we aimed to review the mobile interventions for healthy eating more broadly by synthesizing all studies that focus on mobile interventions for dietary behavior change and health outcomes.

Objectives

The objectives of this scoping review were to identify and synthesize the existing literature on mobile-based interventions for dietary behavior changes and health outcomes. This enables us to better understand the mobile interventions that affect user's eventual health outcomes, which can lead to more efficient and effective promotion of diet guidelines, and improved consumer health and lifestyle policies. Specifically, our review sought to identify and categorize 3 main aspects: (1) the diet input methods of these mobile apps; (2) the mobile-based interventions; and (3) the dietary behavior changes and health outcomes. In addition, we coded the study sample characteristics and methods. Our general research question that guided this scoping review is as follows: "How do mobile interventions influence dietary behavior changes and health outcomes?"

Methods

Scoping Review Methodology

Given the rapid evolution of mHealth apps, we chose a scoping review methodology to obtain an overview of the extant literature on mobile interventions for dietary behavior changes and health outcomes. A scoping review is a literature review technique that is useful to *map* relevant literature in a field of interest [44,45]. At a general level, a scoping review 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" [46]. Therefore, a scoping review addresses broader topics where many different study designs might be applicable [47]. On the contrary, a systematic review answers a well-defined question from studies with appropriate designs, which typically focus on randomized controlled trials (RCTs) or quality-assessed studies with a relatively narrow range to synthesize evidence from them [48].

As described earlier, the literature on dietary behavior changes through mobile interventions has yet to be comprehensively reviewed, thus motivating our scoping review. Furthermore, although 20 of the studies in our review [11-13,19-21,23,24,28,30-37,40,42,49] conducted RCTs, they differed in research questions and objectives, and had diverse outcome variables. Thus, the effectiveness of different interventions in these studies becomes incomparable through a systematic review. In this sense, it was more meaningful to conduct a scoping review for mapping out diverse literature on mobile interventions for dietary behavior changes and health outcomes.

Identifying Relevant Studies and Study Selection

Search Strategy

A scoping literature search was performed on the PubMed/MEDLINE, Association for Computing Machinery Digital Library, and PsycINFO databases. Additionally, we did hand-searches through all relevant themes of journals in *Journal ofMedical Internet Research* (JMIR) and through the registered and published protocols in PROSPERO. The search was restricted to publications from January 2013 to April 2018. The reason that we chose to start the search from 2013 is that the rise in popularity of mobile apps began then. These databases were searched for relevant publications in fields of the title, abstract and keywords using the following search terms: "([food OR diet OR nutrition OR intake] and [mHealth OR mobile OR smartphone OR mobile application])."

Eligibility and Exclusion Criteria

Our aim was to include papers that describe mobile-based interventions for dietary behavior changes and/or health outcomes. Studies were included if they (1) were an original paper published in peer-reviewed journals (except review papers); (2) included mobile-based interventions to influence users' dietary behavior; and (3) reported dietary behavior changes or health outcomes from the mobile-based interventions. With respect to dietary behavior changes (eg, eating more vegetables and consuming food with fewer calories), biochemical outcomes (eg, blood glucose and urinary sodium changes), and health status changes (eg, weight loss), studies targeting multiple health behavior changes/outcomes (eg, changes in both dietary behavior and health status) were also included, as long as at least one change or outcome was related to diet.

Regarding our exclusion criteria, studies dealing with consumption of drugs, toxic substances, chemicals, or pharmaceutical elements were not included. Studies targeting only people with specific diseases or disorders, such as AIDS, cancer, or mental disorder, were also excluded. However, we included studies targeting people who are obese or diabetic as these diseases are directly related to dietary behavior and are more common in the general public. Furthermore, studies focusing on eating disorders such as anorexia and binge eating behavior were not included as the results would not be applicable to a large population. Similarly, studies targeting very specific groups of people, such as pregnant women, children, or athletes, were excluded as the necessary components of their diet, such as minerals (eg, zinc), are not applicable to the broader population. Additionally, studies about the design, development, usability, acceptability, or feasibility of mHealth apps are not within the scope of this review.

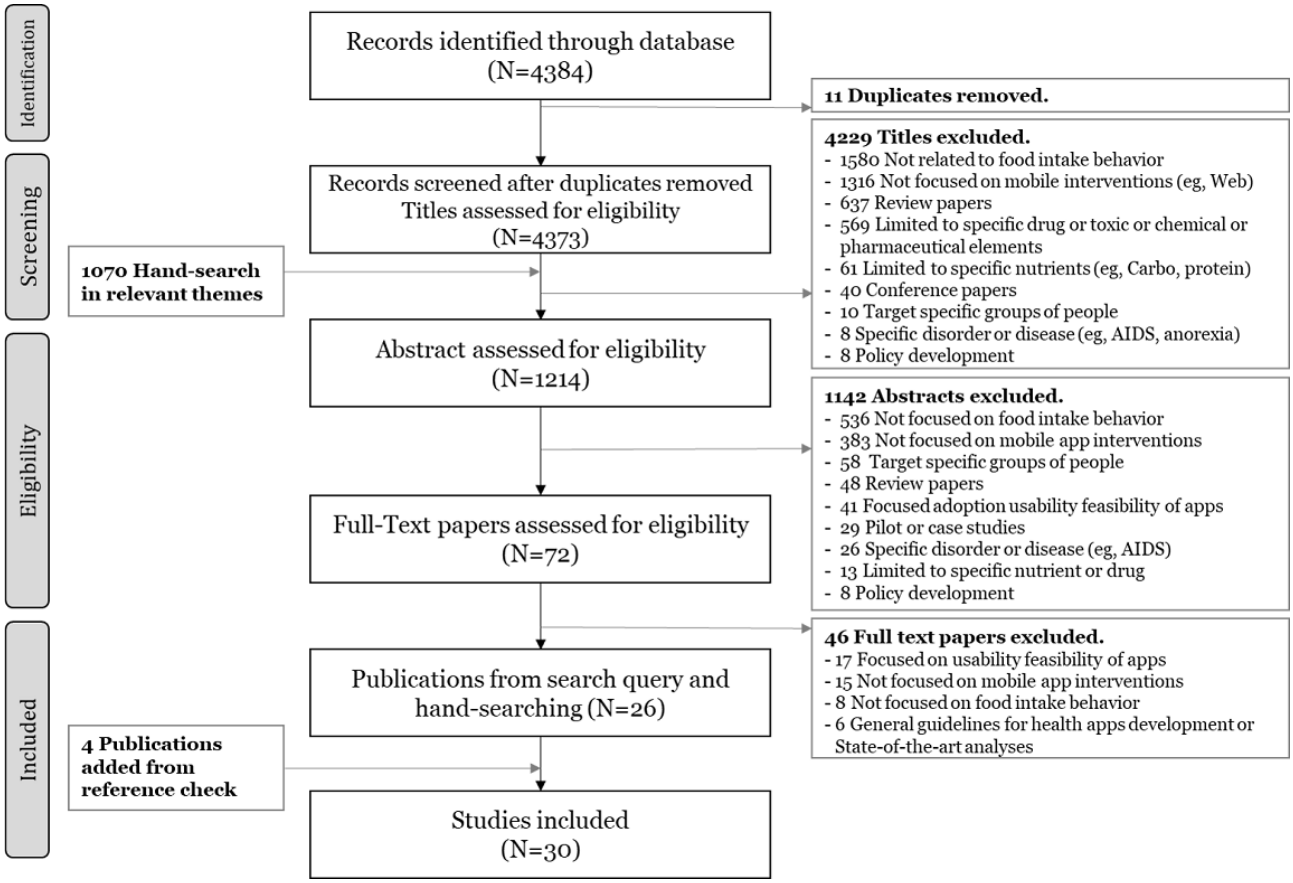
Study Selection

We downloaded the titles and abstracts of all screened studies and used EndNote X8 (Thomson Reuters) for citation management. Duplicates were removed, and the titles and abstracts were reviewed by grouping papers into 4 categories: (1) studies meeting our selection criteria; (2) studies requiring further examination; (3) excluded studies; and (4) other review papers. The selection of studies for our research was conducted and reported according to the guidelines for conducting scoping reviews [47]. In the searching stage, the reference list of all identified reports and papers was searched to include additional studies. Furthermore, 3 reviewers discussed the inclusion and exclusion criteria, and 2 reviewers independently reviewed abstracts for inclusion. Subsequently, papers that were determined to be potentially relevant to our review were downloaded in entirety and reviewed for eligibility. In sum, our search identified 5607 papers, of which 26 studies met the inclusion criteria. In addition, 4 studies were added from reference checks, as the scoping review methodology allows to refine the search strategy during the selection process [43,49,50]. As a result, 30 studies were selected for our final list. The complete selection process is illustrated in Figure 1.

Extraction and Charting of Results

Following study selection, data extraction was completed according to the standard practice for high-quality scoping reviews. A data-charting form was created by the research team to include study characteristics, mobile app input characteristics, mobile-based intervention characteristics, and outcomes of the study. The components of the mobile interventions were classified by a set of behavior change techniques that were seen to have an effect on health behavior [9,10,51]. These included self-monitoring, feedback, gamification, goal reviews, social support, and educational information (Textbox 1). All relevant data from the studies were coded using the data-charting form, and short summaries were obtained to provide an overview of the included studies presented in Multimedia Appendix 1.

Figure 1. Flowchart of the selection process for the scoping review.



Textbox 1. Components of the data-charting form.

Study characteristics:

- Title
- Author (year)
- Participants
- Country of the study
- Study method(s)
- Duration of the study

Mobile app input characteristics:

- Modes of input data
  - Audio/video/photo recognition
  - Text
- Contents of input data
  - Demographic information
  - Medication
  - Health measures
  - Goals
  - Food intake (diet)

Mobile intervention characteristics:

- Modes of mobile interventions
  - In-app log
  - In-app feedback
  - Notification from mobile app
  - Other notification (short message service and email)
- Content of mobile interventions
  - Self-monitoring
  - Feedback
  - Gamification
  - Goal-setting and review
  - Social support
  - Educational information
- Related theories
  - Learning theory (LT)
  - Theory of planned behavior (TPB)
  - Social cognitive theory (SCT)
  - Self-regulation theory (SRT)
  - Theory of behavior changes (TBC)
  - Control theory (CT)
  - Self-determination theory (SDT)
  - Social network theory (SNT)
  - Processes-of-change theory (PCT)

Outcomes of the study:

- Dietary behavior changes (DBC)
- Biochemical outcomes (BO)
- Health status (HS)

Results

Overview

Our initial search identified 4384 papers from the databases. We also included 1201 hand-searched papers from relevant themes of journals in the *JMIR* journals, specifically from "mHealth for Wellness, Behavior Change and Prevention," "Instruments and Questionnaires for Nutrition and Food Intake," and "Mobile Health (mHealth)." Of these, we found that 131 studies had already been included through our previous searches. Thus, as per the procedure in Woodward et al's study [52], we included only 1070 hand-searched papers in the selection process after removing duplicates. The details of the hand-search process and outcomes are illustrated in Figures A1-A3 in Multimedia Appendix 2. Our search in PROSPERO did not yield any additional protocols that met the inclusion criteria. Most of the papers (4229) were excluded during the title screening stage itself because there were a limited number of previous studies where a nutrition app had been used to change dietary behavior [15]. When reviewing the abstracts, we further excluded 1142 papers that did not focus on mobile interventions or dietary behavior or were limited to examining the adoption, usability, and feasibility of mobile apps. Of the remaining 72 papers assessed for full-text eligibility, 15 did not focus on mobile app interventions. Another 17 studies focused on the usability and feasibility of mobile apps, 8 did not focus on the food intake behavior, and 6 studies provided either general guidelines for app development or state-of-the-art analyses of mobile apps. Finally, a check of the reference lists of the included papers resulted in the addition of 4 more publications, bringing our total to 30 papers.

Characteristics of Studies

More than a third of the studies (ie, 11) were conducted in the United States [11-13,26,27,30,32,35,41,49,50], followed by 7 studies in Australia [19,22,24,31,36-38]. In addition, 2 studies each were conducted in China [14,39], Canada [21,25], and Spain [28,40]. Furthermore, 1 study each was conducted in Austria [33], Finland [23], Japan [42], the Netherlands [20], Portugal [29], and Singapore [34]. In terms of the methods of the included studies, 20 studies [11-13,19-21,23,24,28, 30-37,41,42,49] conducted RCTs and 3 conducted field studies [22,27,50]. Among the field studies, Wharton et al [50] conducted a field study that randomized 3 groups, but did not have a control group. Gilson et al [22] performed a field study with Australian truck drivers, but they also did not have a control group. Pirolli et al [27] designed an experiment with 2×2 conditions but also worked without a control group. The remaining 7 studies did not use either an RCT or a field study. Specifically, He et al [14] collected secondary data from 15,310 WeChat group users to study weight loss, and used the propensity score method. Rodrigues et al [29] and Lieffers et al [25] used survey methods, whereas Bejar et al conducted a cross-sectional survey spanning 28 days [40]. Mummah et al [26] used a qualitative interview method, whereas Rollo et al [38] observed 10 patients with type 2 diabetes mellitus. Smith et al [39] studied beverage intake behaviors of 110 young adults over a 3-day period, and checked the volume of urine thereafter. In terms of duration, about 53% (16/30) of the studies covered a period of 3 months or more [11-14,19,20,22-24,28, 30-32,36,41,42], whereas 14 studies were conducted over less than 3 months [21,25-27,29,33-35,37-40,49,50].

Characteristics of Inputs

Most of the studies (ie, 24) used only the text-based input mode [11-13,19-22,24-36,39-41,42]. For example, users logged their dietary behaviors by entering text about what they ate on a mobile app during the study period. On the contrary, 6 studies used more advanced input modes, including photo recognition [23,37,38,42,49], voice logging [37], and social network service message logging [14]. Although all the studies in our review acquired data on diet intake, some studies also collected other types of data, such as weight [11-14,19,26,31-34,37,41,50]. Studies also captured other health measures such as body mass index (BMI; by using weight and height) [11,12,34,37] and waist circumference [12], as well as medication [36]. Moreover, 6 of the studies [11,19,27,34,49,50] included goal settings in their inputs.

Characteristics of Interventions

Mode and Theory of Mobile Interventions

On the basis of input data, the studies in our review provided the various modes of interventions as shown in Multimedia Appendix 1. We found that 12 studies [19,22,23,26,27,30,35-40] showed only the logged history of the user, which we refer to as *in-app log* mode. On the contrary, 18 studies [11-14,20,21,24,25,28,29,31-34,41,42,49,50] also included the feedback function in addition to the logged history, coded as "in-app feedback." For example, the mobile app could send in-app feedback messages to users either as a tailored feedback on their progress or as a notification to keep users updated. Among these 18 studies, 5 studies used multiple modes of interventions. Specifically, 1 study used a push notification from the app [33] which was coded as "notification from mobile app" and 4 studies used other forms which was coded as "other notification," that is, short message service notification [12,24,31], telephone calls [31], emails [11,12,31], Facebook messages [12].

Furthermore, some interventions in the included studies applied various theories for the design of their behavior change techniques. Specifically, 11 studies [11-13,19,21,23,24,26,31-33] included interventions based on theories, including learning theory (LT), theory of planned

behavior (TPB), social cognitive theory (SCT), self-regulation theory (SRT), theory of behavior change (TBC), control theory (CT), self-determination theory (SDT), social network theory (SNT), and the processes-of-change theory (PCT). The most frequently used theory was the SCT (used in 5 studies) [11,12,19,26,32], followed by LT [11,13,26], SRT [11,19,26], and TPB [11,26,33], which were used in 3 studies each. In addition, 2 studies each used TBC [11,21], CT [12,23], and SDT [23,24], whereas 1 study each was based on PCT [31] and SNT [12].

Content of Mobile Interventions

The contents of the interventions in our review studies were categorized into 6 types (ie, self-monitoring, feedback, gamification, goal reviews, social support, and educational information), as seen in Table 1 and Multimedia Appendix 1. In Table 1, they were further divided into studies evaluating the effects of the interventions on (1) only dietary behavior changes, (2) only health outcomes, and (3) both.

In the *self-monitoring* category, all 30 studies contained diet-logging functions as part of their mobile interventions. Further, 12 studies [11-13,20,21,25,28,29,31,33,41,50] provided more detailed results including progress, of which 4 studies [20,21,28,29] assessed their effects on dietary behavior change, 3 studies [13,25,31] on both dietary behavior change and health outcomes, and the remaining 5 studies on health outcomes only. Interestingly, 2 studies [12,13] also provided functions for improving users' adherence to healthier diets and recommended calorie prescriptions. Furthermore, 1 study used an intervention of sending messages about the status of users' weight to increase user engagement [12].

Feedback-based interventions were seen in 18 studies [11-14,20,21,24,25,28,29,31-34,41,42,49,50]. Feedback is differentiated from self-monitoring in terms of its scale of interactions. Feedback intends to change users' beliefs by providing a high level of interaction [53]. With respect to their effects, 6 studies [20,21,24,28,29,49] assessed the impacts of these interventions on dietary behavior change, 4 studies [13,25,31,32] on both dietary behavior change and health outcomes, and the remaining 8 studies on health outcomes only. As for intervention content, all 18 studies provided both progress reviews and recommendations, whereas 4 studies provided reminders [27,30,33,39].

In the *gamification* category, we identified and coded for 7 key elements of gamification: points, leaderboard, levels, quests and challenges, progression, viral loop [54], and trading [55].

We found 10 studies [11,14,19,23,27,31,32,34,49,50] that include gamification elements. Among them, 9 used progression elements [11,14,19,27,31,32,34,49,50], whereas 2 studies also provided quest and challenge elements [14,50]. Furthermore, 2 studies provided a leaderboard in their interventions, on the basis of points given to users [23,32]. There were no studies with gamification elements of level, viral loop, and trading in our review. With respect to their effects, 4 studies [19,23,27,49] assessed the impacts of these interventions on dietary behavior change, 3 studies [11,31,32] on both dietary behavior change and health outcomes, and the remaining 3 studies on health outcomes only.

In the category of *goal review*, 5 studies [11,19,27,34,50] with goal setting inputs provided a review of goals. Among these, 1 study [11] supported users by providing goal progress reports. In sum, 2 studies [19,27] assessed intervention impacts on dietary behavior change, whereas the remaining 3 studies evaluated effects on health outcomes.

In the *social support* category, 3 studies [14,23,32] used interventions that provide such support. Compared with self-monitoring, feedback, and gamification categories, the goal review and social support categories of interventions were less used. With respect to their effect, 1 study [23] assessed the impact of social support on dietary behavior change, another study [32] on both dietary behavior change and health outcomes, and the remaining study on health outcomes only, that is, weight loss. In terms of content, all 3 studies provided general social support, whereas 1 study [23] provided comparison functions among peers in its intervention.

The last category of mobile interventions consisted of those providing *educational information*. Here, 2 studies [19,35] provided educational materials on diets and the challenges to adhere to the prescribed diets. Furthermore, 1 study [19] assessed the effect of the interventions on dietary behavior change, whereas the other study examined health outcomes only.

Table 1. Content of interventions and effects.

Content of interventions	Studies with interventions on dietary behavior changes only (N=12), n (%)	Studies with interventions on both dietary behavior changes and health outcomes (N=5), n (%)	Studies with interventions on health outcomes only (N=13), n (%)
Self-monitoring	12 (100)	5 (100)	13 (100)
Logging	12 (100)	5 (100)	13 (100)
Progress	4 (33)	3 (60)	5 (38)
Adherence	0 (0)	1 (20)	1 (8)
Engagement	0 (0)	0 (0)	1 (8)
Feedback	6 (50)	4 (80)	8 (62)
Reviews on their progress	6 (50)	4 (80)	8 (62)
Recommendations	6 (50)	4 (80)	8 (62)
Reminders	2 (17)	0 (0)	2 (15)
Gamification	4 (33)	3 (60)	3 (23)
Progression	3 (25)	3 (60)	3 (23)
Quest and challenge	0 (0)	0 (0)	2 (15)
Leaderboard	1 (8)	1 (20)	0 (0)
Points	1 (8)	1 (20)	0 (0)
Goal reviews	2 (16)	0 (0)	3 (23)
Reviews of goal logs	2 (16)	0 (0)	3 (23)
Goal progress support	0 (0)	0 (0)	1 (8)
Social support	1 (8)	1 (20)	1 (8)
Social support	1 (8)	1 (20)	1 (8)
Social comparisons	1 (8)	0 (0)	0 (0)
Educational information	1 (8)	0 (0)	1 (8)

Characteristics of Outcomes

The mode of measurement of outcomes in the included studies (24 studies) was mostly self-reported by participants themselves [12-14,19-22,24-34,36,37,40,41,49,50]. Among the 30 studies, 2 studies measured the outcome variable by a blood test [11,42] and 2 studies measured it by a urine volume or sodium test [35,39]. Furthermore, 2 studies measured the participants' weight by a scale in a lab [11,38], and 2 studies provided a photo recognition mode for users to report the outcome [23,49].

Regarding the outcomes assessed by the 30 papers, 12 studies [19-24,27-30,40,49] evaluated dietary behavior change as their only outcome. On the contrary, 13 studies [11,12,14,33-39,41,42,50] directly examined the effects of mobile interventions on health outcomes, whereas 5 studies [13,25,26,31,32] assessed the effects on both dietary behavior change and health outcomes.

Dietary Behavior Change

As mentioned above, 12 studies aimed to change users' dietary behavior as their main outcome. Indeed, we found that the dietary behaviors examined in the studies were quite diverse. Several studies focused on the intake of specific food types, including high-fiber bread and low-fat milk [19], vegetables [27], fats [23], as well as low-calorie foods [29,49]. However, most of the studies [20,21,22,24,28,30,40] focused on healthy food intake defined by different combinations of fruit, vegetable, processed food, sugar, fat, salt, sugar-sweetened drinks, calories, and, lastly, a Mediterranean diet.

Health Outcomes

On the contrary, 13 studies [11,12,14,33-39,41,42,50] assessed the direct effect of mobile interventions on users' health outcomes, that is, health status changes and biochemical outcomes. Of these, 9 studies [11,12,14,33,34,37,38,41,50] measured health status outcomes in terms of weight loss and BMI. Furthermore, 6 studies measured biochemical outcomes, including blood glucose level [11,36,42], urine volume [39], urinary sodium [35], and blood pressure/hemoglobin [34].

Dietary Behavior Change and Health Outcomes

In our review, 5 studies [13,25,26,31,32] aimed at dietary behavior change as well as achieving better health outcomes. These studies are rare, but important, because they show the relationship between dietary behavior changes and health outcomes. For this reason, we investigated each study in detail. First, *Martin et al* [13] examined the effect of a weight loss intervention that delivers personalized recommendations and educational materials via the multimedia capabilities of participants' smartphones. They found that the participants successfully adhered to their calorie intake prescriptions

provided by the intervention, which resulted in weight loss at the end of 12 weeks. *Lieffers et al* [25] examined the effects of self-monitoring by using logs of food and calorie intakes, as well as recipes, exercise, and restaurant nutrition information. They surveyed dietitians to evaluate the effectiveness of the recommendations on nutrition and food apps. They reported that 41% of dietitians felt the studied apps would help users in managing their body weight and result in healthier body composition. *Mummah et al* [26] found that mobile-based self-monitoring resulted in changes in users' vegetable intake behavior, thereby achieving weight loss. The RCT tested the effect of theory-driven mobile interventions using 18 behavior change techniques. *Hales et al* [32] reported that social support using a social network function resulted in users' dietary behavior change, that is, consuming fewer calories, and health outcomes, that is, weight loss. *Hebden et al* [31] developed a custom program to provide personalized coaching as their intervention. Their study found that the intervention changes the types of food consumed, thereby improving health outcomes, such as weight loss.

Discussion

Our scoping review aimed at identifying and synthesizing prior studies of mobile-based interventions for dietary behavior change and health outcomes. The implications of the findings of this review are discussed below along with the strengths and limitations.

Principal Findings

From our review, we identified the most common input mode as text-based input. However, using more advanced methods, such as photo recognition, can ease the burden of diet input and logging. Furthermore, although all the reviewed studies captured users' diet intake, future apps could benefit from collecting other types of data, including demographics (eg, height and weight), medication (eg, insulin dosage), health measures (eg, BMI), and goal setting.

Moreover, we found that self-monitoring, followed by personalized feedback type of mobile intervention, was most common. Although both self-monitoring and personalized feedback were found to help in achieving the desired dietary behavior changes, the other (less common) content categories of gamification, goal reviews, social support, and educational information were also helpful in this regard. Thus, mobile apps in future can make better use of these other intervention categories.

In our review, all the 11 studies [11-13,19,21,23,24,26,31-33] stating that their interventions were based on behavior change theories made little explicit reference to theory. Of these, 10 studies merely mentioned the theories but did not describe how the theoretical constructs were used to derive their interventions. As an exception, only 1 study [26] explained how behavioral theories were used to derive their intervention approach, but did not examine the underlying mechanisms of the behavior change. Moreover, the remaining 19 studies in our review made no reference to theory. This suggests the need for more theory-based interventions for dietary behavior change. This is because theoretical models provide links between intervention content and mediating processes implied by theory. They can enable the identification of features that systematically influence the effectiveness of interventions and, hence, help build a cumulative understanding of what works and how [56]. Without understanding the underlying mechanisms of behavior change techniques, decision makers lack information to make choices about what interventions are likely to be effective in their own settings.

The studies we reviewed investigated how their mobile interventions affected dietary behavior changes and/or health outcomes. However, there was little consistency among the dietary behavior changes examined, suggesting that more comprehensive and consistent measures can be developed for this purpose. Furthermore, in terms of health outcomes, most studies have focused on assessing weight loss. Although weight loss is an important measure of health improvement, other outcomes may also need to be examined.

We also found that most of the studies did not focus on dietary behavior change as a mediator for health outcomes. Although weight loss or blood glucose control is crucial for some user groups, such as obese people or diabetics, a healthy diet helps to improve the overall health for most people. Thus, we need to understand the mechanisms behind mobile-based interventions' effects on dietary behavior changes and health outcomes. This limitation of existing literature is somewhat related to the current underutilization of theory-based mobile interventions.

Strengths and Limitations

One of the strengths of this research is that by following the objective of scoping reviews, this study synthesizes the extant literature and highlights potential gaps in it. We provide a comprehensive map of the literature on the underexplored topic of mobile interventions for dietary behavior change. The review covers aspects related to study characteristics, input mode and contents, mode and content of mobile interventions, related behavioral theories, and outcomes. A number of gaps in this area are identified in the Discussion section above based on our review.

Compared with prior scoping reviews in this domain, another strength of our review is that it identifies the underinvestigated mediating process of dietary behavior change in the relationship between mobile interventions and health outcomes. Furthermore, previous reviews have not incorporated a wide range of studies concerning the effect of mobile interventions on dietary behavior change, as we do. Specifically, they did not provide a descriptive overview by synthesizing studies [15,43], did not include a broad range of study designs and methodologies [43], or focused on a limited scope, that is, vegetable intake [16]. Thus, our review adds to the literature by providing a more comprehensive view of mobile-based interventions for dietary behavior changes and their outcomes.

One limitation of this scoping review is that potential biases might have influenced the results. *First*, publication bias could be present, indicated by the absence of negative effects of reported interventions included in this review. There was only

1 ineffective study [38], which failed to show the weight loss of users from using the mobile app. Although measuring the effectiveness of the interventions is not within the scope of our study, we mainly focused on identifying the relationship between mobile interventions, dietary behavior changes, and health outcomes. *Second*, we observed that the interventions described in earlier studies [29,30,31,33,36,41] in 2013, differ from the interventions described in newer studies, [14,27,35,37] in 2013. Compared with the recent apps with advanced technologies such as self-tracking sensors, food photo recognition, and customized real-time feedback, the apps with older interventions might not prove to be as effective. *Finally*, the search criteria we used for retrieving the studies were very broad and initially started with a large number (5607) of studies. As there is no consistent terminology for dietary behavior changes and outcomes, we tried to cover all the aspects that have been studied in this regard.

Conclusions

To the best of our knowledge, our scoping review provides the first overview of the relationships among mobile-based interventions, dietary behavior change, and health outcomes. In contrast to the general belief of the importance of dietary behavior changes, not many studies have examined dietary behavior changes as a mediator for health outcomes. Future research needs to be conducted to understand the effects of mobile interventions for dietary behavior changes on health outcomes.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Summary of included studies.

[PDF File (Adobe PDF File), 225KB - mhealth_v7i1e11312_app1.pdf]

Multimedia Appendix 2

Supplementary information of hand-search process.

[PDF File (Adobe PDF File), 223KB - mhealth_v7i1e11312_app2.pdf]

References

1. Anders S, Schroeter C. Diabetes, diet-health behavior, and obesity. Front Endocrinol (Lausanne) 2015 Mar 16;6:33 [FREE Full text] [doi: 10.3389/fendo.2015.00033] [Medline: 25852643]
2. Jiao J, Moudon AV, Kim SY, Hurvitz PM, Drewnowski A. Health implications of adults' eating at and living near fast food or quick service restaurants. Nutr Diabetes 2015 Jul 20;5(7):e171 [FREE Full text] [doi: 10.1038/nutd.2015.18] [Medline: 26192449]
3. Gittelsohn J, Trude A. Diabetes and obesity prevention: changing the food environment in low-income settings. Nutr Rev 2017 Jan;75(suppl 1):62-69 [FREE Full text] [doi: 10.1093/nutrit/nuw038] [Medline: 28049750]
4. Menza V, Probart C. Eating Well for Good Health Lessons on Nutrition & Healthy Diets. Rome, Italy: Food & Agriculture Organization of the UN; 2019:361.
5. Dalle Grave R, Centis E, Marzocchi R, El Ghoch M, Marchesini G. Major factors for facilitating change in behavioral strategies to reduce obesity. Psychol Res Behav Manag 2013 Oct 3;6:101-110 [FREE Full text] [doi: 10.2147/PRBM.S40460] [Medline: 24124398]
6. Rathbone AL, Prescott J. The use of mobile apps and SMS messaging as physical and mental health interventions: systematic review. J Med Internet Res 2017 Aug 24;19(8):e295 [FREE Full text] [doi: 10.2196/jmir.7740] [Medline: 28838887]
7. Bort-Roig J, Gilson ND, Puig-Ribera A, Contreras RS, Trost SG. Measuring and influencing physical activity with smartphone technology: a systematic review. Sports Med 2014 May;44(5):671-686. [doi: 10.1007/s40279-014-0142-5] [Medline: 24497157]
8. Lewis BA, Napolitano MA, Buman MP, Williams DM, Nigg CR. Future directions in physical activity intervention research: expanding our focus to sedentary behaviors, technology, and dissemination. J Behav Med 2017 Feb;40(1):112-126 [FREE Full text] [doi: 10.1007/s10865-016-9797-8] [Medline: 27722907]
9. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
10. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
11. Block G, Azar KM, Romanelli RJ, Block TJ, Hopkins D, Carpenter HA, et al. Diabetes prevention and weight loss with a fully automated behavioral intervention by email, web, and mobile phone: a randomized controlled trial among persons

with prediabetes. J Med Internet Res 2015 Oct 23;17(10):e240 [FREE Full text] [doi: 10.2196/jmir.4897] [Medline: 26499966]
12. Patrick K, Marshall S, Davila E, Kolodziejczyk J, Fowler J, Calfas K, et al. Design and implementation of a randomized controlled social and mobile weight loss trial for young adults (project SMART). Contemp Clin Trials 2014 Jan;37(1):10-18 [FREE Full text] [doi: 10.1016/j.cct.2013.11.001] [Medline: 24215774]
13. Martin CK, Miller AC, Thomas DM, Champagne CM, Han H, Church T. Efficacy of SmartLoss, a smartphone-based weight loss intervention: results from a randomized controlled trial. Obesity (Silver Spring) 2015 May;23(5):935-942 [FREE Full text] [doi: 10.1002/oby.21063] [Medline: 25919921]
14. He C, Wu S, Zhao Y, Li Z, Zhang Y, Le J, et al. Social media-promoted weight loss among an occupational population: cohort study using a WeChat mobile phone app-based campaign. J Med Internet Res 2017 Oct 23;19(10):e357 [FREE Full text] [doi: 10.2196/jmir.7861] [Medline: 29061555]
15. DiFilippo KN, Huang WH, Andrade JE, Chapman-Novakofski KM. The use of mobile apps to improve nutrition outcomes: a systematic literature review. J Telemed Telecare 2015 Jul;21(5):243-253. [doi: 10.1177/1357633X15572203] [Medline: 25680388]
16. Nour M, Chen J, Allman-Farinelli M. Efficacy and external validity of electronic and mobile phone-based interventions promoting vegetable intake in young adults: systematic review and meta-analysis. J Med Internet Res 2016 Apr 8;18(4):e58 [FREE Full text] [doi: 10.2196/jmir.5082] [Medline: 27059765]
17. Teixeira PJ, Carraça EV, Marques MM, Rutter H, Oppert JM, De Bourdeaudhuij I, et al. Successful behavior change in obesity interventions in adults: a systematic review of self-regulation mediators. BMC Med 2015 Apr 16;13:84 [FREE Full text] [doi: 10.1186/s12916-015-0323-6] [Medline: 25907778]
18. Roy R, Hebden L, Rangan A, Allman-Farinelli M. The development, application, and validation of a Healthy eating index for Australian Adults (HEIFA-2013). Nutrition 2016 Apr;32(4):432-440. [doi: 10.1016/j.nut.2015.10.006] [Medline: 26740257]
19. Duncan M, Vandelanotte C, Kolt GS, Rosenkranz RR, Caperchione CM, George ES, et al. Effectiveness of a web- and mobile phone-based intervention to promote physical activity and healthy eating in middle-aged males: randomized controlled trial of the ManUp study. J Med Internet Res 2014 Jun 12;16(6):e136 [FREE Full text] [doi: 10.2196/jmir.3107] [Medline: 24927299]
20. Elbert SP, Dijkstra A, Oenema A. A mobile phone app intervention targeting fruit and vegetable consumption: the efficacy of textual and auditory tailored health information tested in a randomized controlled trial. J Med Internet Res 2016 Jun 10;18(6):e147 [FREE Full text] [doi: 10.2196/jmir.5056] [Medline: 27287823]
21. Gilliland JA, Sadler RC, Clark AF, O'Connor CA, Milczarek M, Doherty ST. Using a smartphone application to promote healthy dietary behaviours and local food consumption. Biomed Res Int 2015 Sep;9(Suppl 3):841368 [FREE Full text] [doi: 10.1155/2015/841368] [Medline: 26380298]
22. Gilson NG, Pavey TG, Vandelanotte C, Duncan MJ, Gomersall SR, Trost SG, et al. Chronic disease risks and use of a smartphone application during a physical activity and dietary intervention in Australian truck drivers. Aust N Z J Public Health 2016 Feb;40(1):91-93. [doi: 10.1111/1753-6405.12501] [Medline: 26713400]
23. Helander E, Kaipainen K, Korhonen I, Wansink B. Factors related to sustained use of a free mobile app for dietary self-monitoring with photography and peer feedback: retrospective cohort study. J Med Internet Res 2014 Apr 15;16(4):e109 [FREE Full text] [doi: 10.2196/jmir.3084] [Medline: 24735567]
24. Kerr DA, Harray AJ, Pollard CM, Dhaliwal SS, Delp EJ, Howat PA, et al. The connecting health and technology study: a 6-month randomized controlled trial to improve nutrition behaviours using a mobile food record and text messaging support in young adults. Int J Behav Nutr Phys Act 2016 Apr 21;13:52 [FREE Full text] [doi: 10.1186/s12966-016-0376-8] [Medline: 27098449]
25. Lieffers JR, Vance VA, Hanning RM. Use of mobile device applications in Canadian dietetic practice. Can J Diet Pract Res 2014;75(1):41-47. [doi: 10.3148/75.1.2014.41] [Medline: 24606959]
26. Mummah SA, King AC, Gardner CD, Sutton SR. Iterative development of Vegethon: a theory-based mobile app intervention to increase vegetable consumption. Int J Behav Nutr Phys Act 2016 Aug 8;13(1):90 [FREE Full text] [doi: 10.1186/s12966-016-0400-z] [Medline: 27501724]
27. Pirolli P, Mohan S, Venkatakrishnan A, Nelson L, Silva M, Springer A. Implementation intention and reminder effects on behavior change in a mobile health system: a predictive cognitive model. J Med Internet Res 2017 Nov 30;19(11):e397 [FREE Full text] [doi: 10.2196/jmir.8217] [Medline: 29191800]
28. Recio-Rodriguez JI, Agudo-Conde C, Martin-Cantera C, González-Viejo MN, Fernandez-Alonso MD, Arietaleanizbeaskoa MS, EVIDENT Investigators. Short-term effectiveness of a mobile phone app for increasing physical activity and adherence to the Mediterranean diet in primary care: a randomized controlled trial (EVIDENT II study). J Med Internet Res 2016 Dec 19;18(12):e331 [FREE Full text] [doi: 10.2196/jmir.6814] [Medline: 27993759]
29. Rodrigues JJ, Lopes IM, Silva BM, Torre IL. A new mobile ubiquitous computing application to control obesity: SapoFit. Inform Health Soc Care 2013 Jan;38(1):37-53. [doi: 10.3109/17538157.2012.674586] [Medline: 22657250]
30. Turner-McGrievy GM, Beets MW, Moore JB, Kaczynski AT, Barr-Anderson DJ, Tate DF. Comparison of traditional versus mobile app self-monitoring of physical activity and dietary intake among overweight adults participating in an

mHealth weight loss program. J Am Med Inform Assoc 2013 May 1;20(3):513-518 [FREE Full text] [doi: 10.1136/amiajnl-2012-001510] [Medline: 23429637]
31. Hebden L, Cook A, van der Ploeg HP, King L, Bauman A, Allman-Farinelli M. A mobile health intervention for weight management among young adults: a pilot randomised controlled trial. J Hum Nutr Diet 2014 Aug;27(4):322-332. [doi: 10.1111/jhn.12155] [Medline: 23992038]
32. Hales S, Turner-McGrievy GM, Wilcox S, Fahim A, Davis RE, Huhns M, et al. Social networks for improving healthy weight loss behaviors for overweight and obese adults: a randomized clinical trial of the social pounds off digitally (Social POD) mobile app. Int J Med Inform 2016 Oct;94:81-90. [doi: 10.1016/j.ijmedinf.2016.07.003] [Medline: 27573315]
33. Brindal E, Hendrie G, Freyne J, Coombe M, Berkovsky S, Noakes M. Design and pilot results of a mobile phone weight-loss application for women starting a meal replacement programme. J Telemed Telecare 2013 Apr;19(3):166-174. [doi: 10.1177/1357633X13479702] [Medline: 23520213]
34. Goh G, Tan NC, Malhotra R, Padmanabhan U, Barbier S, Allen Jr JC, et al. Short-term trajectories of use of a caloric-monitoring mobile phone app among patients with type 2 diabetes mellitus in a primary care setting. J Med Internet Res 2015 Feb 3;17(2):e33 [FREE Full text] [doi: 10.2196/jmir.3938] [Medline: 25648130]
35. Ipjian ML, Johnston CS. Smartphone technology facilitates dietary change in healthy adults. Nutrition 2017 Jan;33:343-347. [doi: 10.1016/j.nut.2016.08.003] [Medline: 27742102]
36. Kirwan M, Vandelanotte C, Fenning A, Duncan MJ. Diabetes self-management smartphone application for adults with type 1 diabetes: randomized controlled trial. J Med Internet Res 2013 Nov 13;15(11):e235 [FREE Full text] [doi: 10.2196/jmir.2588] [Medline: 24225149]
37. Pendergast FJ, Ridgers ND, Worsley A, McNaughton SA. Evaluation of a smartphone food diary application using objectively measured energy expenditure. Int J Behav Nutr Phys Act 2017 Mar 14;14(1):30 [FREE Full text] [doi: 10.1186/s12966-017-0488-9] [Medline: 28288657]
38. Rollo ME, Ash S, Lyons-Wall P, Russell AW. Evaluation of a mobile phone image-based dietary assessment method in adults with type 2 diabetes. Nutrients 2015 Jun 17;7(6):4897-4910 [FREE Full text] [doi: 10.3390/nu7064897] [Medline: 26091234]
39. Smith LP, Hua J, Seto E, Du S, Zang J, Zou S, et al. Development and validity of a 3-day smartphone assisted 24-hour recall to assess beverage consumption in a Chinese population: a randomized cross-over study. Asia Pac J Clin Nutr 2014;23(4):678-690 [FREE Full text] [doi: 10.6133/apjcn.2014.23.4.10] [Medline: 25516327]
40. Bejar LM, Sharp BN, García-Perea MD. The e-EPIDEMIOLOGY mobile phone app for dietary intake assessment: comparison with a food frequency questionnaire. JMIR Res Protoc 2016 Nov 2;5(4):e208 [FREE Full text] [doi: 10.2196/resprot.5782] [Medline: 27806922]
41. Spring B, Duncan JM, Janke EM, Kozak AT, McFadden HG, DeMott A, et al. Integrating technology into standard weight loss treatment: a randomized controlled trial. JAMA Intern Med 2013 Jan 28;173(2):105-111 [FREE Full text] [doi: 10.1001/jamainternmed.2013.1221] [Medline: 23229890]
42. Waki K, Aizawa K, Kato S, Fujita H, Lee H, Kobayashi H, et al. DialBetics with a multimedia food recording tool, FoodLog: smartphone-based self-management for type 2 diabetes. J Diabetes Sci Technol 2015 May;9(3):534-540 [FREE Full text] [doi: 10.1177/1932296815579690] [Medline: 25883164]
43. Bardus M, Smith JR, Samaha L, Abraham C. Mobile phone and web 2.0 technologies for weight management: a systematic scoping review. J Med Internet Res 2015 Nov 16;17(11):e259 [FREE Full text] [doi: 10.2196/jmir.5129] [Medline: 26573984]
44. Dijkers M. What is a scoping review. KT Update 2015 Dec;4(1):1-4 [FREE Full text]
45. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
46. Mays N, Roberts E, Popay J. Synthesising research evidence. In: Fulop N, Allen P, Clarke A, Black N, editors. Studying the Organisation and Delivery of Health Services: Research Methods. London, United Kingdom: Routledge; 2001:188-220.
47. Peters MD, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]
48. Aromataris E, Riitano D. Constructing a search strategy and searching for evidence. A guide to the literature search for a systematic review. Am J Nurs 2014 May;114(5):49-56. [doi: 10.1097/01.NAJ.0000446779.99522.f6] [Medline: 24759479]
49. Rabbi M, Pfammatter A, Zhang M, Spring B, Choudhury T. Automated personalized feedback for physical activity and dietary behavior change with mobile phones: a randomized controlled trial on adults. JMIR Mhealth Uhealth 2015 May 14;3(2):e42 [FREE Full text] [doi: 10.2196/mhealth.4160] [Medline: 25977197]
50. Wharton CM, Johnston CS, Cunningham BK, Sterner D. Dietary self-monitoring, but not dietary quality, improves with use of smartphone app technology in an 8-week weight loss trial. J Nutr Educ Behav 2014;46(5):440-444. [doi: 10.1016/j.jneb.2014.04.291] [Medline: 25220777]
51. Michie S, Hyder N, Walia A, West R. Development of a taxonomy of behaviour change techniques used in individual behavioural support for smoking cessation. Addict Behav 2011 Apr;36(4):315-319. [doi: 10.1016/j.addbeh.2010.11.016] [Medline: 21215528]

52. Woodward A, Howard N, Wolffers I. Health and access to care for undocumented migrants living in the European Union: a scoping review. Health Policy Plan 2014 Oct;29(7):818-830 [FREE Full text] [doi: 10.1093/heapol/czt061] [Medline: 23955607]
53. Snyder M. Self-monitoring of expressive behavior. J Pers Soc Psychol 1974;30(4):526-537. [doi: 10.1037/h0037039]
54. Kankanhalli A, Taher M, Cavusoglu H, Kim SH. Gamification: A New Paradigm for Online User Engagement. Gamification: A New Paradigm for Online User Engagement, Proceedings of the 33rd International Conference on Information Systems; 2012 Presented at: International Conference on Information Systems; 2012; Orlando p. 16-19 URL: http://scholarbank.nus.edu.sg/handle/10635/78159
55. Liu D, Santhanam R, Webster J. Toward meaningful engagement: a framework for design and research of gamified information systems. MIS Q 2017 Apr 4;41(4):1011-1034. [doi: 10.25300/MISQ/2017/41.4.01]
56. Foy R, Ovretveit J, Shekelle PG, Pronovost PJ, Taylor SL, Dy S, et al. The role of theory in research to develop and evaluate the implementation of patient safety practices. BMJ Qual Saf 2011 May;20(5):453-459. [doi: 10.1136/bmjqs.2010.047993] [Medline: 21317181]

Abbreviations

BO: biochemical outcome
BMI: body mass index
CT: control theory
DBC: dietary behavior change
HS: health status
JMIR: *Journal of Medical Internet Research*
LT: learning theory
mHealth: mobile health
PCT: processes-of-change theory
RCT: randomized controlled trial
TBC: theory of behavior change
TPB: theory of planned behavior
SCT: social cognitive theory
SDT: self-determination theory
SNT: social network theory
SRT: self-regulation theory

Edited by G Eysenbach; submitted 16.06.18; peer-reviewed by C Carrion; comments to author 21.09.18; revised version received 16.11.18; accepted 22.11.18; published 21.01.19.

Please cite as:
Kankanhalli A, Shin J, Oh H
Mobile-Based Interventions for Dietary Behavior Change and Health Outcomes: Scoping Review
JMIR Mhealth Uhealth 2019;7(1):e11312
URL: http://mhealth.jmir.org/2019/1/e11312/
doi:10.2196/11312
PMID:30664461

Review

Technology-Based Interventions, Assessments, and Solutions for Safe Driving Training for Adolescents: Rapid Review

Emre Sezgin[1], PhD; Simon Lin[1], MBA, MD

Research Information Solutions and Innovation, The Research Institute, Nationwide Children's Hospital, Columbus, OH, United States

Corresponding Author:
Emre Sezgin, PhD
Research Information Solutions and Innovation, The Research Institute
Nationwide Children's Hospital
RB3 Building, 3rd Floor
700 Children's Drive
Columbus, OH, 43205
United States
Phone: 1 6143556814
Email: emre.sezgin@nationwidechildrens.org

Abstract

Background: Safe driving training for adolescents aims to prevent injury and promote their well-being. In that regard, information and communication technologies have been used to understand adolescent driving behavior and develop interventions.

Objective: The purpose of this review is to explore and discuss existing approaches to technology-based driving interventions, driving assessments, and solutions in the literature.

Methods: We searched the Web of Science and PubMed databases following a review protocol to collect relevant peer-reviewed journal articles. Inclusion criteria were (1) being published in the English language, (2) being published in a peer-reviewed journal, (3) testing the driving behavior of teens with technology-based intervention methods, and (4) being published between January 2000 and March 2018. We appraised the articles by reading their abstracts to select studies matching the inclusion criteria and reading the full text of articles for final refinement.

Results: Initial keyword searches on technology-based solutions resulted in 828 publications that we refined further by title screening (n=131) and abstract evaluation against inclusion criteria (n=29). Finally, we selected 16 articles that met the inclusion criteria and examined them regarding the use of technology-based interventions, assessments, and solutions. Use of built-in tracking devices and installation of black box devices were widely used methods for capturing driving events. Smartphones were increasingly adapted for data collection, and use of gamification for intervention design was an emerging concept. Visual and audio feedback also were used for intervention.

Conclusions: Our findings suggest that social influence is effective in technology-based interventions; parental involvement for promoting safe driving behavior is highly effective. However, the use of smartphones and gamification needs more study regarding their implementation and sustainability. Further developments in technology for predicting teen behavior and programs for behavioral change are needed.

(JMIR Mhealth Uhealth 2019;7(1):e11942) doi:10.2196/11942

KEYWORDS

adolescent health; assessment; driving safety; teen driving; technology-based intervention

Introduction

Background

The US National Center for Health Statistics reported that 73% of unintentional injury deaths among teenagers in the United States were caused by motor vehicle traffic incidents over the years 1999-2006 [1]. Motor vehicle crashes continue to be one of the leading causes of deaths among teenagers, and most incidents were attributed to risky behavior established during childhood [2]. Teen drivers have crash rates almost 3 times higher per mile driven than drivers 20 years and older [3]. Immaturity leads to speeding and other risky habits, and inexperience means teen drivers often do not recognize or know how to respond to hazards [4]. This issue highlights the question that has been raised by the US National Research Council,

Institute of Medicine, and Transportation Research Board [5]: "What are the best ways to influence teens' behavior?"

Current Practices

To answer the question, several solutions have been proposed by national organizations and associations in the United States. One major approach that is widely implemented is to improve driver education and training programs [6,7]; however, the effectiveness of such programs needs further evidence [6]. Another recent approach is the use of technology to monitor driving. The US Insurance Institute for Highway Safety reported that teens who used in-vehicle monitoring devices showed less risky behavior than unsupervised teens [8]. However, the technology-based intervention is only effective if parents are able to review the feedback and talk about it with their teen. In another practice by the US Governors Highway Safety Association, dashboard cameras for monitoring driving activity are used for postdriving training [4], but it was found not feasible due to the costs of equipment purchase and installation. The Minnesota Department of Transportation implemented a smartphone-based driving support system and tested it with the participation of teen-parent pairs [9]. Teens found the system helpful for complying with the rules and reducing risky driving behavior. However, the ability of teens to adapt to the system warnings and parents' concerns about their teens' privacy were major issues.

In-Vehicle Technologies

In-vehicle technologies include dedicated information system tools to understand driving conditions, environment, and behavior. These technologies can be stand-alone systems (black box) or integrated with other technologies such as mobile devices. The purpose of in-vehicle technologies is to create a real-time digital footprint of a driving event. The components of an in-vehicle system can be smartphones, communication tools (eg, short message service, email, or external information channels), and vehicle diagnostics to collect relevant information. For instance, the in-vehicle technology Foot-LITE uses real-time information on road conditions and vehicle operation, and collects data using a camera, a 3-axis accelerometer, and a global positioning system [10]. It connects to the vehicle with an onboard diagnostics (OBD-II standard) port and processes the data using an onboard processing unit (TRW Limited engine control unit). The system provides feedback on driving behavior, such as a lane departure warning, via a smartphone app. The advanced driver-assistance system (ADAS) is another widely used in-vehicle technology. ADAS has been used to adjust vehicle operation to improve safety and driving. It is an integrated system designed to understand real-time events and alert drivers to avoid collisions. The extent of the capabilities of ADAS may vary, but some examples are lane departure warning, blind-spot warning, adaptive lighting, and adaptive cruise control that autoaccelerates and autobrakes in traffic. Table 1 presents the information that can be potentially collected using in-vehicle technologies.

Table 1. Information collection and required hardware for in-vehicle technologies.

Data type	Source	Hardware
Weather conditions	National weather API[a] service	Wireless internet connection, modem
Road type (residential, city, rural)	Map API service	Wireless internet connection, modem
Traffic light status	Camera	Smartphone, external hardware
Traffic sign detection	Camera	Smartphone, external hardware
Lane-marking detection	Camera	Smartphone, external hardware
Traffic condition	Web source via traffic API service	Wireless internet connection, modem
Traveling distances	GPS[b]	Smartphone, black box
Changes in velocity	GPS, accelerometer	Smartphone, black box
Changes in acceleration	Accelerometer, gyrometer	Smartphone, black box
Changes in geolocation	GPS	Smartphone, black box
Heart rate, electrocardiogram	Monitoring sensor	Smartwatch
Seatbelt	Built-in sensor	External hardware
Light exposure	Camera, light sensor	Smartphone, external hardware
Accident detection (rollover and impact)	Accelerometer, gyrometer, magnetometer	Smartphone, black box
Acceleration, braking, and cornering behavior	Accelerometer, GPS, gyrometer	Smartphone, black box
Following distance	Camera, infrared sensor	Smartphone, external hardware
Driver identification	Camera	Smartphone, black box
Traveling pattern	GPS, magnetometer	Smartphone, black box

[a]API: application programming interface.

[b]GPS: global positioning system.

Objectives

Even though there is no "gold standard" for solving the safety issues of young drivers, the scientific quest for seeking solutions with technology-based interventions has been advancing. In that regard, some studies have revealed novel technologies and methods for increasing driving safety and awareness among teens. We believe that understanding these approaches would be helpful for identifying effective implementations of technology-based interventions. We sought to review the literature on the effect of technology-based interventions, assessments, and solutions on adolescent driving behavior. Therefore, we aimed to (1) explore the technology-based approaches reported in the literature, (2) discuss their methods and findings, and (3) suggest alternative approaches in the light of the findings.

Methods

We limited the scope of the literature search to peer-reviewed journal articles indexed in the Web of Science and PubMed databases, which provide access to scientifically rigorous studies in reputable and indexed journals. Inclusion criteria were (1) being published in the English language, (2) being published in a peer-reviewed journal, (3) testing the driving behavior of teens with technology-based intervention methods, and (4) being published between January 2000 and March 2018. Our search strategy was to (1) identify search keywords, (2) refine the selection of journal articles, (3) read abstracts to select studies matching the inclusion criteria, and (4) read the full text of articles for final refinement. We searched the databases using the following combinations of keywords: "teen" OR "adolescent" OR "young" AND "driving" OR "driver" AND "technology" OR "smartphone" OR "phone" AND "vehicle" AND "prevention" OR "intervention".

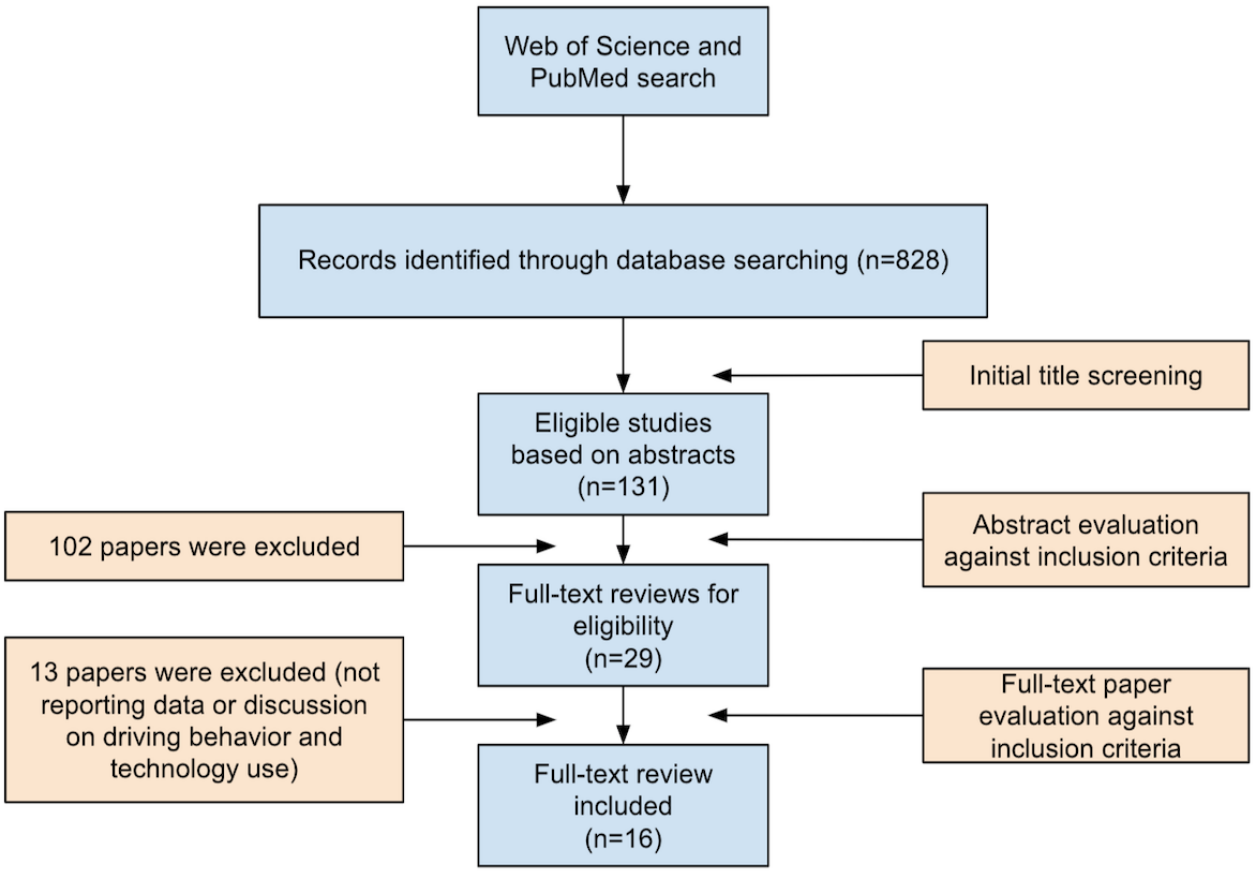
We extracted data into a predesigned Excel spreadsheet form (Microsoft Office 2016; Microsoft Corporation). The form included study title, year, journal, scope of the study, method, sample characteristics and size, and study findings. We performed a qualitative synthesis to descriptively synthesize the data.

Results

Search Results

We completed the search by March 2018. We initially identified 828 records (147 from Web of Science and 681 from PubMed), which we further refined by title screening (n=131) and abstract evaluation against the inclusion criteria (n=29). After full-text review, we identified 16 studies [11-26] that focused on the driving behavior of teens and promoting behavior change with technology-based interventions (Figure 1).

Figure 1. Review flow diagram.



Technologies to Improve Teen Driving Safety

Our research results fell into 3 categories of technology used for teen driving safety: (1) in-vehicle technologies (using built-in tracking devices in the car and an installed black box), (2) smartphones (using apps and the sensors of a smartphone), and (3) gamification (an extension of smartphones used to increase compliance and sustainability of safe driving).

In-Vehicle Technologies

Technology-Based Interventions and Assessments

The effectiveness of in-vehicle technologies for teen driving safety has been evaluated through longitudinal tests and randomized trials. In accordance with the reports mentioned in the introduction, parental involvement in combination with in-vehicle technology was identified as a highly effective intervention with the in-vehicle technologies. Farah et al [11] conducted a longitudinal study with the participation of 242 families of young male drivers. In-vehicle data recorders were placed in the cars for 12 months during a teen's first year of driving. The first 3 months consisted of the teen driving with a parent or family member, and the remaining 9 months consisted of the teen driving solo. There were 4 study groups based on family feedback and guidance on parental involvement. Drivers in the groups who received family feedback with or without guidance on parental involvement had lower event occurrence rates than did the control (no-feedback) group during the solo-driving phase. This finding indicated that early intervention using the combination of in-vehicle technology and family/parental involvement had a lasting effect on teen driving behavior. Similarly, Musicant and Lampel [12] recruited 32 young drivers to test the effectiveness of an intervention using an in-vehicle recording mechanism for capturing driving behavior from 113 to 237 days. The study had 2 phases, with and without feedback to families and teens. During the feedback phase, the occurrence of risky events was reduced. Farmer et al [13] tested 4 parameters of teen driving behavior using a driving detection system (acceleration, braking, speeding, and seat belt use). The groups had different interventions, such as receiving alert sounds for speeding and not using a seat belt, and parental access to Web-based feedback. This study found that the use of seat belts increased as the violations were reported to parents, compared with in-vehicle alert sounds. Speeding violations were decreased with in-vehicle alerts. Finally, technology-based monitoring of teen driving was able to reduce the incidence of risky behavior. However, the child-parent relationship and dynamics influenced the effectiveness of the intervention.

Impact of Social Influence and Parental Involvement

Simons-Morton et al reported in 2 studies [14,15] the implications family or parental involvement on teen driving behavior. The first study [14] tested the impact of an in-vehicle safety monitoring system on 2 groups of teens. One group received immediate feedback from the system about risky driving, while the other group received weekly reports and family access to driving scores in addition to the immediate feedback. The intervention that included parental involvement was more effective than the use of the technology alone. The follow up study [15] identified risk factors in teen driving. The participants were observed during the first 18 months after licensure, and data were collected on driving behaviors (eg, acceleration, braking, and location) via video, images, and periodic surveys. Saliva swabs were collected and tested for stress-induced compounds, and distraction and driving skills were scored. The findings suggested that crash and near-crash risks were almost 4 times higher in teens during the first 18 months after licensure than among adults. The authors argued that social norms strongly influence driving behavior: the risk of crash was higher for teens driving alone than while driving with passengers.

McGehee et al [16] used an event-triggered device to capture driving data, and audio and video feeds from the driver to measure risky behavior. Participants were observed using a 3-phase design to measure changes in driving behavior between the no-intervention and intervention (immediate feedback and feedback by parent and teen mentoring sessions) phases. Use of the event-triggered video system with weekly feedback and video review involving parents reduced the unsafe driving behavior of teens. A previous study by Carney et al [17] supported this finding. A similar 3-phase design with in-vehicle monitoring technology was implemented with the participation of 18 young drivers. Intervention with visual feedback and weekly event reports to teens and parents reduced risky driving behavior by 61%. To understand parental guidance for newly licensed young drivers, Prato et al [18] investigated the behavior of teen-parent pairs. They recruited 62 families; vehicle monitoring systems were installed in their cars, and driving behaviors were monitored (the first 3 months was the accompanied-driving setting, and the next 9 months was the solo-driving setting). Findings suggested that risk-taking behavior could be influenced by the driver's sex, observations of parental driving, sensation-seeking tendencies, duration of supervised driving, and the level of parental involvement in monitoring the teens.

Uptake Challenges

Interviews also revealed latent facts about integrated technology use. Gesser-Edelsburg and Guttman [19] interviewed 2 groups of teens: 1 group used an in-vehicle driver monitoring system (n=26) and 1 did not use a monitoring system (n=111). Findings suggested that the system may have had adverse effects on perceiving technology as a solution because it replaced parental accompaniment (as a tool for monitoring, punishment, and violation of privacy). However, the teens had a positive attitude toward the system for being an objective and credible source of driver behavior and for helping to improve driving skills. The authors argued that there is a need to create a support system of professionals for teens and parents, and that the technology should have the role of facilitator of the intervention. Weiss and colleagues' [20] focus group interviews demonstrated that teens were comfortable with the technology and familiar with its limitations. Thus, they were not willing to have interference by the system (ADAS) while driving to more naturalistically develop their skills.

To understand the perspective of parents, Guttman et al [21] interviewed parents of young drivers regarding the use of in-vehicle monitoring technologies. The participants addressed

slightly different issues, in that monetary cost and security concerns were the main factors discouraging installation of the technology. Young drivers receiving feedback and monitoring of driving were the main motivating factors to install the technology. Furthermore, the parents expected to take this step in the early stages of driving. Promotion of the technology would require more incentives and lower cost. Moreover, developing a clear policy on security and privacy about driving data and legal implications, addressing young drivers' privacy concerns, and providing resources for parents to guide their kids for safe driving were identified as critical aspects for implementing integrated technology approaches. In terms of policy implications to promote the uptake of these technologies, McGehee et al [16], Simons-Morton et al [14], and Carney et al [17] reported the significance of parental involvement in supervision in technology and use of graduated driver licensing.

Smartphone Use

Use of smartphones for teen driving safety is a relatively new concept. The increasing capability of smartphones, low cost of access, and higher accuracy in capturing events have promoted the use of smartphones for quantifying driver behavior. To understand the effect of safe driving apps, Creaser et al [22] tested the effect of a phone blocking app in 3 groups. The first group used the blocking app. The second group used the same app and parents received reports of risky events. For the third group, the control group, driving behaviors were just observed. Even though the study data revealed low use of the phone during driving, self-reported data showed that teens were able to find a way around blocking, and use of a mandatory setting would not have been helpful in the long run.

The use of safe driving apps, as well as the influence of the social environment, were investigated by Musicant et al [23]. They conducted a longitudinal study with young scouts and cadets to investigate the effect of a safe driving app, which was promoted with the use of a group incentive scheme. The app recorded events on each trip and provided feedback and a score based on driving behavior (speed, acceleration, braking, and cornering). The study demonstrated that young people may act for the benefit of the group. Low-cost and group incentive schemes could motivate young drivers to use the safety app. However, the effect may only have been temporary; lack of incentives, short trips, battery consumption, and forgetting to enable the app were the reasons given for not using the app. Similarly, Kervick et al [24] investigated the willingness of young drivers to use a smartphone-based driver support system. The perception of what the teen would gain from using the system and the influence of social environment on using the system were the factors determining the intention to use versus actual use.

Adaptation with Gamification

With the integration of smartphones in driving safety, gamification emerged as an intervention for behavioral change. In that regard, Steinberger et al [25] investigated several game concepts with young male drivers to encourage safe driving. Drivers tested the smartphone games (mounted on the dashboard) while driving via a drive simulator. Results showed that engagement was associated with economic concerns (fuel consumption) and anticipatory driving (what was ahead). In addition, participants expected a degree of challenge from the game to make driving fun, interactive with others, and personalizable (based on different characteristics and patterns). The authors also studied the effect of gamification on reducing driving boredom. Steinberger et al [26] tested a mobile game concept to encourage anticipatory driving by detecting speed limits and changes. They recruited 2 groups of teens as control and intervention groups to use a drive simulator. Driver data (eg, lane position, speed, video, and physiological measures) and subjective experience data (eg, surveys about boredom intensity, arousal, and perceived driving performance) were collected. Results showed that gamified intervention may reduce unsafe driving by reducing driving boredom. However, visual cues can increase cognitive workload and thus cause slower reaction times to driving events. The authors also indicated that physiological measures can help to identify driving boredom events.

Table 2 [11-26] summarizes the literature findings and provides a broader look at the study methods, significant findings, and barriers to technology use.

Table 2. Literature summary grouped by the type of technology used.

Study	Country	Method	Sample and size	Significant findings for technology use	Identified barriers to technology use
In-vehicle technologies					
McGehee, 2007 [16]	United States	Driving data analysis (technology used: DriveCam)	26 teens (16-17 years old)	Technology with periodic feedback and parental involvement were effective in reducing unsafe driving.	N/A[a]
Musicant, 2010 [12]	Israel	Driving data analysis	32 young drivers (17-24 years old)	Availability of feedback reduced event frequency by 50%,	N/A
Carney, 2010 [17]	United States	Driving data analysis (technology used: DriveCam)	18 teens (16 years old)	Intervention with visual feedback and weekly reports and videos to teens and parents increased safe driving.	N/A
Prato, 2010 [18]	Israel	Driving data analysis and survey	62 teen-parent pairs	Different sexes exhibited different risky behaviors; Tendency to seek sensation affects risky driving; Driving behavior of parents, duration of supervised driving, and level of parental monitoring influenced risky behavior.	N/A
Farmer, 2010 [13]	United States	Driving data analysis	85 teens (16-17 years old)	Reinforcement from parents was necessary for sustainable safe driving; Push notifications (emailing report cards and personalized feedback) were more effective than pull notifications (website access).	Alerts can be annoying; Too much information provided could be discouraging for parents
Guttman, 2011 [21]	Israel	Interview	906 parents of young drivers (17-24 years old)	Early stages of driving were considered a better time for installing the technology; Financial benefits and environmental considerations were perceived as incentives; Security of data and privacy of teens were common concerns; Technology may promote parent-teen driver communication; Parents should have access to monitoring data.	Cost; Security and privacy concerns; Confronting the young driver
Simons-Morton, 2013 [14]	United States	Driving data analysis and survey (technology used: DriveCam)	90 parent-teen couples (~16 years old)	Parental involvement increases effectiveness.	N/A
Simons-Morton, 2015 [15]	United States	Driving data analysis and survey	42 teens (~16 years old)	Social norms were important in risky behavior; Driving alone was riskier than with passengers.	N/A
Gesser-Edelsburg, 2013 [19]	Israel	Interview	137 teens (15-18 years old)	In-vehicle technology was an objective and credible source for driving; Replaced the role model of parents with objective feedback from the device.	Trust issues within parent-teen relationship; Invasion of privacy; Stress from parental punishment based on feedback; Doubts about the technology improving driving skills
Farah, 2013 [11]	Israel	Event frequency analysis (technology used: GreenRoad Tech)	212 teen-parent pairs	Periodic driving feedback, parental involvement, and guidance were effective in reducing risky driving.	N/A
Weiss, 2018 [20]	United States	Interview (technology used: advanced driver-assistance system)	24 teens (16-19 years old) and 12 parents	Teens were knowledgeable about and comfortable with the technology; Teen and parents preferred using a non–advanced driver-assistance system car to improve driving skills.	Teens are skeptical about abilities of the technology, knowing its limitations; The idea of giving control to a "machine" is not positively perceived
Smartphone					
Musicant, 2015 [23]	Israel	Interview and survey	24 scouts and 22 cadets (17-19 years old)	Group incentives and low cost improved uptake of in-vehicle technology.	Forgetfulness; Battery consumption; Lack of incentives

Study	Country	Method	Sample and size	Significant findings for technology use	Identified barriers to technology use
Creaser, 2015 [22]	United States	Survey	274 teens and 272 parents	The blocking app could be effective for new drivers; Parental involvement with the app increased the effectiveness.	Bypassing the app or using a friend's phone
Kervick, 2015 [24]	Ireland	Survey	333 teens (18-24 years old)	Perceived gains from use of the app and social influence affected acceptance of the driving support app.	N/A
Gamification with smartphone					
Steinberger, 2017 [25]	Australia	Design analysis and interview	24 young men (~20 years old)	Economic and anticipatory driving were engaging; Drivers expected a challenge from the game; Interaction with others was important; Personalization was desired	N/A
Steinberger, 2017 [26]	Australia	Driving data analysis and interview	32 young men (18-25 years old)	Ambient feedback with colors was useful.	Instant visual feedback can be distracting; Screen positioning can be distracting

[a]N/A: not available.

Discussion

In the light of the findings about the effects of in-vehicle technologies on teen driving behavior, we propose several implications and suggestions that could provide a basis for future development of interventions.

Parental Involvement in Technology Effectiveness

The findings indicate that in-vehicle technologies are useful for assisting teens with safe driving behaviors; parental involvement along with technology-based feedback has an even stronger influence on development of safe driving skills. Technology is viewed as a double-edged sword because it creates both an opportunity for and a barrier to the parent-teen relationship. On the one hand, it can be used to develop trust between parents and teens by providing an objective source for driving and feedback; however, teens can also perceive in-vehicle technologies as a sign of parental distrust. Therefore, when promoting the technology, the purpose of its use should emphasize the benefit to teens.

As reported in most of the studies, implementation of the technology should consider parents' attitudes toward the use of technology as well. In terms of parenting style, teens with highly involved parents (providing a level of support, but highly involved with rules and monitoring; ie, they are authoritative) are less likely to demonstrate risky (deliberate risk-taking) behavior and be more compliant with rules [27]. Similarly, involving adult passengers as well as receiving postdriving feedback with parental involvement could be effective in teens' development of safe driving skills [28]. While studies reported that parents have a major influential role in driving, the unwillingness of a teen's parent to implement in-vehicle technology or to be involved with the teen may be a barrier to the teen's development of safe driving skills [29]. In that regard, parents' personal traits and parent-teen dynamics, as well as environmental and living conditions, should be regarded as determining factors for the effectiveness of technology-based interventions. In addition, timely interventions, providing continuous feedback for parents and teens, and providing education resources and incentive mechanisms for parents and teens should be considered as key success factors for implementing technology-based interventions.

Extending Research and Gaining Evidence via Smartphone Use

Smartphone use is a promising means of delivering technology-based intervention for teens, as 2016 statistics from the Pew Research Center reported that 92% of young adults had smartphones [30]. Target audiences can be reached via smartphones at low cost, without interference with daily life, and without the need for high user involvement in data collection. Some studies reported that the use of smartphones has the potential to be attractive to teens because of the availability of timely feedback, interaction via apps, integration with other apps, and timely connection with users, as well as by offering incentives and the potential for social connection.

However, in terms of smartphone-based interventions and assessments to prevent risky teen driving, more evidence-based findings are needed, especially under real-world conditions. Because only a few studies about smartphone-based interventions for teen drivers have been conducted, the relevant literature about the general population can be used to design the methodology for developing interventions for teens. For example, the most desirable smartphone features were text blocking, collision warning, voice control, and driving data recorders for the general population [31]. Thus, the apps being used by adult drivers might be adapted and tested with teen users. In another case, a collision warning app effectively reduced event occurrence for adults [32], and feedback helped to improve driving efficiency and driving behavior for safe following distance [10]. The extension of these studies to teen driving safety and behavior change would help fill the knowledge gap. However, there could be challenges in adapting these existing methods for teens in terms of deciding on an intervention method, such as use of unobtrusive technology (which is more acceptable) versus intrusive technology (which is less acceptable) [28]. Battery consumption and incentivization are other challenges to overcome for ensuring teens' continuous use of smartphone-based solutions.

Gamification in Play

Gaming stimulates self-efficacy and reward mechanisms to promote particular behaviors. In the health care literature, gamification was reported to have positive effects on behavioral and cognitive outcomes [33]. Therefore, gamification could promote healthy driving behavior in adolescents via intrinsic motivation. The method also is highly accessible on mobile devices, is cost effective, and fits into the current lifestyle.

Based on the findings, use of a gaming approach could be effective in promoting behavioral change for safe driving. However, there is a lack of evidence of the effectiveness of design and in real-world driving implementations. Users expect a challenge, interaction, social connection, and personalization from a game, and it is challenging to fulfill these needs without causing a distraction. To further implement gamification without distracting the driver, postdriving feedback is suggested instead of feedback during driving. In that regard, drive scoring, leaderboards, and achievement badges could be used as feedback mechanisms. To assess the effectiveness of gamification in real-world implementation, wearables and biometric measures may provide feedback and observations on driving behavior change [34]. However, gamification has its own risks. Designers should note that increasing competition, design, and task evaluation issues may have adverse effects on behavior [35].

Regarding the extent of the research on in-vehicle technologies, smartphones, and gamification, the literature has presented more evidence of in-vehicle technology use with parental involvement. Studies of smartphones and gamification as safe driving interventions have been limited because they are relatively new concepts, and their effectiveness has been tested mostly under controlled environments.

Further Suggestions on Technology-Based Intervention Developments

The literature discussed in this study suggests that the capability of current technologies and their adaptation for effective use are increasing. However, the element that has not been discussed but is significant for long-term impact is the use of big data. In-vehicle technologies and smartphones have been extended to collect aggregated driving data for better quantifying driving behavior, such as understanding driver behavior with pattern recognition for identifying aggressive driving [36], and identifying driver behavior features for better feedback via machine learning [37,38]. Specifically, the ability of smartphones to collect data is as good as that of advanced in-vehicle technologies [39], and smartphones have the potential to provide further evidence of effectiveness of interventions and assessment in the long term.

The literature also lacks evidence of the long-term impact of technology-based interventions. Thus, in addition to the technical capabilities, a deeper look into multilevel influences (eg, the sociotechnical perspective, social determinants of health) on teen driving behavior would also contribute to the design of interventions. Furthermore, security of driving data, privacy, stress (teens being punished for bad driving and parents wanting to avoid confrontation), trust issues, cost of implementation, and lack of incentives were observed as the major barriers to use of the technology. Therefore, the design of digital driving behavior change intervention programs may benefit from considering the engaging factors, risk factors, and protective factors for teens; developing communication methods; evaluating teen driver behavior; monitoring progress; and ensuring compliance with ethics, regulations, and information governance [40,41].

Limitations

This review was limited to providing insight on in-vehicle technologies and intervention for teens based on the literature available in Web of Science and PubMed within our selection criteria. In addition, the study did not include research on driving distraction but focused on technology-based detection and intervention for injury prevention purposes. We listed the findings based on the technology being used, but we did not break down the findings to present method categorizations or the level of teens' learning progress (eg, graduated driver licensing level, early-period or late-period novice learners). Similarly, the review did not address regional or national policies and regulations for transportation and driving. Thus, readers should consider regional differences while interpreting the findings.

Conclusions

We reviewed the effects of technology-based interventions on adolescent driving behavior. We discussed in-vehicle technology and smartphone-based approaches and reported significant findings and observations. Finally, we provide suggestions for implementations and implications for further research. To our knowledge, there have been no literature reviews on teens and smartphone use and gamification of on-road driving. However, teen crash risks [42], distraction from mobile technology [43], effect of distraction on driving [44,45], and prevention of cell phone–based distractions [46] have been reviewed. This review extends the literature by filling in the gap in knowledge of technology-based intervention methods.

The study can be expanded with inclusion of other languages and databases. In that regard, we suggest including meta-analysis of trial studies with in-vehicle technologies in future work. Additional experimental studies on smartphones and gamification approaches would be useful to identify intervention methods, design requirements, and effectiveness of these new methods.

Conflicts of Interest

None declared.

References

1. Miniño A. Mortality among teenagers aged 12-19 years: United States, 1999-2006. NCHS Data Brief 2010 May(37):1-8 [FREE Full text] [Medline: 20450538]
2. Kann L, McManus T, Harris WA, Shanklin SL, Flint KH, Hawkins J, et al. Youth risk behavior surveillance - United States, 2015. MMWR Surveill Summ 2016 Dec 10;65(6):1-174. [doi: 10.15585/mmwr.ss6506a1] [Medline: 27280474]
3. Teenagers 2016. Ruckersville, VA: Insurance Institute for Highway Safety, Highway Loss Data Institute; 2018. URL: http://www.iihs.org/iihs/topics/t/teenagers/fatalityfacts/teenagers/2016 [accessed 2018-12-05] [WebCite Cache ID 74ROYG9IL]
4. Fischer P. Promoting Parent Involvement in Teen Driving: An In-Depth Look at the Importance and the Initiatives. Washington, DC: Governors Highway Safety Association; 2013. URL: https://www.ghsa.org/sites/default/files/2016-11/TeenDrivingParentReport%20low.pdf [accessed 2018-12-10] [WebCite Cache ID 74Z1IYYnb]
5. Institute of Medicine, Transportation Research Board, National Research Council. Preventing Teen Motor Crashes: Contributions from the Behavioral and Social Sciences: Workshop Report. Washington, DC: National Academies Press; 2007.
6. Hassan HM, Abdel-Aty MA. Exploring the safety implications of young drivers' behavior, attitudes and perceptions. Accid Anal Prev 2013 Jan;50:361-370. [doi: 10.1016/j.aap.2012.05.003] [Medline: 22658949]
7. Ferguson SA. Speeding-Related Fatal Crashes Among Teen Drivers and Opportunities for Reducing the Risks. Washington, DC: Governors Highway Safety Association; 2013. URL: https://www.ghsa.org/sites/default/files/2016-11/GHSA%20Teen%20SpeedingFinal.pdf [accessed 2018-12-05] [WebCite Cache ID 74RP5ERGQ]
8. Insurance Institute for Highway Safety (IIHS). Monitoring devices let parents supervise new drivers more closely. IIHS Status Report 2009 May 7;44(5) [FREE Full text]
9. Creaser J, Gorjestani A, Manser M, Donath M. Usability evaluation of a smart phone-based teen driver support system (TDSS). St Paul, MN: Minnesota Department of Transportation; 2011. URL: http://conservancy.umn.edu/bitstream/11299/150085/1/MnDOT2011-13.pdf [accessed 2018-12-05] [WebCite Cache ID 74RPsj33N]
10. Birrell SA, Fowkes M, Jennings PA. Effect of using an in-vehicle smart driving aid on real-world driver performance. IEEE Trans Intell Transport Syst 2014 Aug;15(4):1801-1810. [doi: 10.1109/TITS.2014.2328357]
11. Farah H, Musicant O, Shimshoni Y, Toledo T, Grimberg E, Omer H, et al. The first year of criving. Transport Res Rec 2018 May 22;2327(1):26-33. [doi: 10.3141/2327-04]
12. Musicant O, Lampel L. When technology tells novice drivers how to drive. Transport Res Rec 2018 May 22;2182(1):8-15. [doi: 10.3141/2182-02]
13. Farmer CM, Kirley BB, McCartt AT. Effects of in-vehicle monitoring on the driving behavior of teenagers. J Safety Res 2010 Feb;41(1):39-45. [doi: 10.1016/j.jsr.2009.12.002] [Medline: 20226949]
14. Simons-Morton BG, Bingham CR, Ouimet MC, Pradhan AK, Chen R, Barretto A, et al. The effect on teenage risky driving of feedback from a safety monitoring system: a randomized controlled trial. J Adolesc Health 2013 Jul;53(1):21-26 [FREE Full text] [doi: 10.1016/j.jadohealth.2012.11.008] [Medline: 23375825]
15. Simons-Morton BG, Klauer SG, Ouimet MC, Guo F, Albert PS, Lee SE, et al. Naturalistic teenage driving study: findings and lessons learned. J Safety Res 2015 Sep;54:41-44 [FREE Full text] [doi: 10.1016/j.jsr.2015.06.010] [Medline: 26403899]
16. McGehee DV, Raby M, Carney C, Lee JD, Reyes ML. Extending parental mentoring using an event-triggered video intervention in rural teen drivers. J Safety Res 2007;38(2):215-227. [doi: 10.1016/j.jsr.2007.02.009] [Medline: 17478192]
17. Carney C, McGehee DV, Lee JD, Reyes ML, Raby M. Using an event-triggered video intervention system to expand the supervised learning of newly licensed adolescent drivers. Am J Public Health 2010 Jun;100(6):1101-1106. [doi: 10.2105/AJPH.2009.165829] [Medline: 20395588]
18. Prato CG, Toledo T, Lotan T, Taubman-Ben-Ari O. Modeling the behavior of novice young drivers during the first year after licensure. Accid Anal Prev 2010 Mar;42(2):480-486. [doi: 10.1016/j.aap.2009.09.011] [Medline: 20159070]
19. Gesser-Edelsburg A, Guttman N. “Virtual” versus “actual” parental accompaniment of teen drivers: a qualitative study of teens’ views of in-vehicle driver monitoring technologies. Transport Res Part F Traffic Psychol Behav 2013 Feb;17:114-124. [doi: 10.1016/j.trf.2012.09.002]
20. Weiss E, Fisher Thiel M, Sultana N, Hannan C, Seacrist T. Advanced driver assistance systems for teen drivers: teen and parent impressions, perceived need, and intervention preferences. Traffic Inj Prev 2018 Feb 28;19(sup1):S120-S124. [doi: 10.1080/15389588.2017.1401220] [Medline: 29584476]
21. Guttman N, Lotan T. Spying or steering? Views of parents of young novice drivers on the use and ethics of driver-monitoring technologies. Accid Anal Prev 2011 Jan;43(1):412-420. [doi: 10.1016/j.aap.2010.09.011] [Medline: 21094339]
22. Creaser JI, Edwards CJ, Morris NL, Donath M. Are cellular phone blocking applications effective for novice teen drivers? J Safety Res 2015 Sep;54:75-78. [doi: 10.1016/j.jsr.2015.06.014] [Medline: 26403905]
23. Musicant O, Lotan T, Grimberg E. Potential of group incentive schemes to encourage use of driving safety apps. Transport Res Rec 2015 Nov;2516:1-7. [doi: 10.3141/2516-01]
24. Kervick AA, Hogan MJ, O'Hora D, Sarma KM. Testing a structural model of young driver willingness to uptake smartphone driver support systems. Accid Anal Prev 2015 Oct;83:171-181. [doi: 10.1016/j.aap.2015.07.023] [Medline: 26277411]
25. Steinberger F, Schroeter R, Foth M, Johnson D. Designing gamified applications that make safe driving more engaging. New York, NY: ACM; 2017 Presented at: 2017 CHI Conference on Human Factors in Computing Systems; May 6-22, 2017; Denver, CO, USA p. 2826-2839. [doi: 10.1145/3025453.3025511]

26. Steinberger F, Schroeter R, Watling CN. From road distraction to safe driving: evaluating the effects of boredom and gamification on driving behaviour, physiological arousal, and subjective experience. Comput Hum Behav 2017 Oct;75:714-726. [doi: 10.1016/j.chb.2017.06.019]
27. Ginsburg KR, Durbin DR, García-España JF, Kalicka EA, Winston FK. Associations between parenting styles and teen driving, safety-related behaviors and attitudes. Pediatrics 2009 Oct;124(4):1040-1051. [Medline: 19810185]
28. Lee JD. Technology and teen drivers. J Safety Res 2007;38(2):203-213. [doi: 10.1016/j.jsr.2007.02.008] [Medline: 17478191]
29. Mirman JH, Curry AE, Schultheis MT, Brant CL, Wang W, Winston FK, et al. Development of an on-road driving assessment for novice teen drivers. Transport Res Rec 2014;2465(1):64-72. [doi: 10.3141/2465-09]
30. Smith A. Record shares of Americans now own smartphones, have home broadband. Washington, DC: Pew Research Center; 2017 Jan 12. URL: http://www.pewresearch.org/fact-tank/2017/01/12/evolution-of-technology/ [accessed 2018-12-05] [WebCite Cache ID 74RQh65H9]
31. Albert G, Musicant O, Oppenheim I, Lotan T. Which smartphone's apps may contribute to road safety? An AHP model to evaluate experts' opinions. Transport Policy 2016 Aug 19;50(5064):54-62. [doi: 10.1016/j.tranpol.2016.06.004]
32. Botzer A, Musicant O, Perry A. Driver behavior with a smartphone collision warning application – a field study. Saf Sci 2017 Jan;91:361-372. [doi: 10.1016/j.ssci.2016.09.003]
33. Johnson D, Deterding S, Kuhn K, Staneva A, Stoyanov S, Hides L. Gamification for health and wellbeing: a systematic review of the literature. Internet Interv 2016 Nov;6:89-106 [FREE Full text] [doi: 10.1016/j.invent.2016.10.002] [Medline: 30135818]
34. Steinberger F, Schroeter R, Babiac D. Engaged drivers–safe drivers: gathering real-time data from mobile and wearable devices for safe-driving apps. In: Meixner G, Müller C, editors. Automotive User Interfaces: Creating Interactive Experiences in the Car. Cham, Switzerland: Springer Nature Switzerland; 2017:55-77.
35. Hamari J, Koivisto J, Sarsa H. Does gamification work? - A literature review of empirical studies on gamification. 2014 Presented at: 47th Hawaii International Conference on System Sciences (HICSS); Jan 6-9, 2014; Waikoloa, HI, USA p. 3025-3034. [doi: 10.1109/HICSS.2014.377]
36. Smith JG, Ponnuru SK, Patil M. Detection of aggressive driving behavior and fault behavior using pattern matching. 2016 Presented at: 2016 International Conference on Advances in Computing, Communications and Informatics (ICACCI); Sep 21-24, 2016; Jaipur, India p. 207-211. [doi: 10.1109/ICACCI.2016.7732048]
37. Ferreira J, Carvalho E, Ferreira BV, de Souza C, Suhara Y, Pentland A, et al. Driver behavior profiling: an investigation with different smartphone sensors and machine learning. PLoS One 2017;12(4):e0174959 [FREE Full text] [doi: 10.1371/journal.pone.0174959] [Medline: 28394925]
38. Vlahogianni EI, Barmpounakis EN. Driving analytics using smartphones: algorithms, comparisons and challenges. Transport Res Part C Emerg Technol 2017;79:196-206. [doi: 10.1016/j.trc.2017.03.014]
39. Meng R, Mao C, Choudhury RR. Driving analytics: will it be OBDs or smartphones?. 2014. URL: https://s3.amazonaws.com/download.zendrive.com/ResearchPapers/Zendrive+Whitepaper+-+Smartphone+Sensors+vs+OBD.pdf [accessed 2018-12-05] [WebCite Cache ID 74RResdHf]
40. Michie S, Yardley L, West R, Patrick K, Greaves F. Developing and evaluating digital interventions to promote behavior change in health and health care: recommendations resulting from an international workshop. J Med Internet Res 2017 Jun 29;19(6):e232 [FREE Full text] [doi: 10.2196/jmir.7126] [Medline: 28663162]
41. Buckley L, Chapman RL, Sheehan M. Young driver distraction: state of the evidence and directions for behavior change programs. J Adolesc Health 2014 May;54(5 Suppl):S16-S21 [FREE Full text] [doi: 10.1016/j.jadohealth.2013.12.021] [Medline: 24759436]
42. Ouimet MC, Pradhan AK, Brooks-Russell A, Ehsani JP, Berbiche D, Simons-Morton BG. Young drivers and their passengers: a systematic review of epidemiological studies on crash risk. J Adolesc Health 2015 Jul;57(1 Suppl):S24-35.e6 [FREE Full text] [doi: 10.1016/j.jadohealth.2015.03.010] [Medline: 26112735]
43. Stavrinos D, Pope CN, Shen J, Schwebel DC. Distracted walking, bicycling, and driving: systematic review and meta-analysis of mobile technology and youth crash risk. Child Dev 2018 Jan;89(1):118-128. [doi: 10.1111/cdev.12827] [Medline: 28504303]
44. Oviedo-Trespalacios O, Haque MM, King M, Washington S. Understanding the impacts of mobile phone distraction on driving performance: a systematic review. Transport Res Part C Emerg Technol 2016 Nov 01;72:360-380. [doi: 10.1016/j.trc.2016.10.006]
45. Caird JK, Johnston KA, Willness CR, Asbridge M, Steel P. A meta-analysis of the effects of texting on driving. Accid Anal Prev 2014 Oct;71:311-318 [FREE Full text] [doi: 10.1016/j.aap.2014.06.005] [Medline: 24983189]
46. Cazzulino F, Burke RV, Muller V, Arbogast H, Upperman JS. Cell phones and young drivers: a systematic review regarding the association between psychological factors and prevention. Traffic Inj Prev 2014;15(3):234-242. [doi: 10.1080/15389588.2013.822075] [Medline: 24372495]

Abbreviations

ADAS: advanced driver-assistance system

API: application programming interface
GPS: global positioning system

Edited by G Eysenbach; submitted 14.08.18; peer-reviewed by C Friedrich, S Tao, J Mirman; comments to author 13.10.18; revised version received 26.10.18; accepted 17.11.18; published 24.01.19.

Please cite as:
Sezgin E, Lin S
Technology-Based Interventions, Assessments, and Solutions for Safe Driving Training for Adolescents: Rapid Review
JMIR Mhealth Uhealth 2019;7(1):e11942
URL: http://mhealth.jmir.org/2019/1/e11942/
doi:10.2196/11942
PMID:30679149

Review

Mobile Health Interventions for Self-Control of Unhealthy Alcohol Use: Systematic Review

Ting Song[1,2], MIT; Siyu Qian[1,2], PhD; Ping Yu[1,2], PhD

[1]Centre for IT-Enabled Transformation, School of Computing and Information Technology, Faculty of Engineering and Information Sciences, University of Wollongong, Wollongong, Australia

[2]Illawarra Health and Medical Research Institute, University of Wollongong, Wollongong, Australia

Corresponding Author:
Ping Yu, PhD
Centre for IT-Enabled Transformation, School of Computing and Information Technology
Faculty of Engineering and Information Sciences
University of Wollongong
Northfields Avenue
Wollongong, 2522
Australia
Phone: 61 2 4221 5412
Fax: 61 2 4221 4170
Email: ping@uow.edu.au

Abstract

Background: Unhealthy alcohol use (UAU) is one of the major causes of preventable morbidity, mortality, and associated behavioral risks worldwide. Although mobile health (mHealth) interventions can provide consumers with an effective means for self-control of UAU in a timely, ubiquitous, and cost-effective manner, to date, there is a lack of understanding about different health outcomes brought by such interventions. The core components of these interventions are also unclear.

Objective: This study aimed to systematically review and synthesize the research evidence about the efficacy of mHealth interventions on various health outcomes for consumer self-control of UAU and to identify the core components to achieve these outcomes.

Methods: We systematically searched 7 electronic interdisciplinary databases: Scopus, PubMed, PubMed Central, CINAHL Plus with full text, MEDLINE with full text, PsycINFO, and PsycARTICLES. Search terms and Medical Subject Headings "mHealth," "text message," "SMS," "App," "IVR," "self-control," "self-regulation," "alcohol*," and "intervention" were used individually or in combination to identify peer-reviewed publications in English from 2008 to 2017. We screened titles and abstracts and assessed full-text papers as per inclusion and exclusion criteria. Data were extracted from the included papers according to the Consolidated Standards of Reporting Trials-EHEALTH checklist (V 1.6.1) by 2 authors independently. Data quality was assessed by the Mixed Methods Appraisal Tool. Data synthesis and analyses were conducted following the procedures for qualitative content analysis. Statistical testing was also conducted to test differences among groups of studies.

Results: In total, 19 studies were included in the review. Of these 19 studies, 12 (63%) mHealth interventions brought significant positive outcomes in improving participants' health as measured by behavioral (n=11), physiological (n=1), and cognitive indicators (n=1). No significant health outcome was reported in 6 studies (6/19, 32%). Surprisingly, a significant negative outcome was reported for the male participants in the intervention arm in 1 study (1/19, 5%), but no change was found for the female participants. In total, 5 core components reported in the mHealth interventions for consumer self-control of UAU were context, theoretical base, delivery mode, content, and implementation procedure. However, sound evidence is yet to be generated about the role of each component for mHealth success. The health outcomes were similar regardless of types of UAU, deployment setting, with or without nonmobile cointervention, and with or without theory.

Conclusions: Most studies reported mHealth interventions for self-control of UAU appeared to be improving behavior, especially the ones delivered by short message service and interactive voice response systems. Further studies are needed to gather sound evidence about the effects of mHealth interventions on improving physiological and cognitive outcomes as well as the optimal design of these interventions, their implementation, and effects in supporting self-control of UAU.

(JMIR Mhealth Uhealth 2019;7(1):e10899) doi:10.2196/10899

KEYWORDS

systematic review; alcohol drinking; self-control; mobile health; mHealth; treatment outcome

Introduction

Background

Unhealthy alcohol use (UAU) is one of the major causes of preventable morbidity, mortality, and related behavioral risks around the world [1,2]. Approximately 3.3 million deaths, accounting for 5.9% of global deaths, were caused by alcohol-related problems annually [3]. Nearly 81% of adults in Australia and 70% in Europe consume alcohol [3,4]. UAU contributed to around 70,000 Australian emergency department presentations in 2014 and 2015 and 77,000 Canadian hospitalizations in 2015 and 2016 [5,6]. It might cause allergic reactions, hormonal disturbances, and intoxication [7,8]. Over time, it might cause diseases such as alcoholic hepatitis, diabetes, cardiovascular and cerebrovascular diseases [9], or psychological problems such as depression, obsession, mania, and suicide [10,11]. Once the brain and neurons are anesthetized, a person might lose self-control [12], leading to social problems such as conflicts, unprepared sexual activities, drunk driving, and violence [13,14]. Therefore, UAU is not only a profound public health challenge but also a social concern.

As an umbrella term, UAU covers various degrees of negative effects of alcohol use on people's well-being [15]. According to the severity, there are 2 major types of UAU: risky drinking and alcohol use disorder (AUD) [15,16].

Risky drinking is also known as problematic drinking, harmful alcohol use, risky single-occasion drinking (RSOD), or heavy episodic drinking. It refers to alcohol use that leads to the risk of negative health consequences [16]. It can be measured by the number of standard drinks (SDs) consumed. An SD is defined by the amount of pure alcohol contained in a drink, and it varies among countries [14,16,17]. For example, in Australia, an SD contains 10 g of pure alcohol, in the United Kingdom and Iceland, it contains only 8 g, whereas in Austria it is 20 g [17]. It is deemed risky drinking if alcohol consumption is more than 5 SDs for men and 4 for women on a single occasion [18]. If total weekly alcohol consumption is greater than or equal to 15 SDs for men and 13 for women in the United States or over 14 SDs for men and 9 for women in Sweden, it is also considered as risky drinking [19-21]. Risky drinking can also be measured by scales such as fast alcohol screening test (FAST), alcohol use disorders identification test (AUDIT), and AUDIT for consumption (AUDIT-C) by scoring 3 or higher in FAST [22], over 8 for men and 6 for women in AUDIT [23], or 4 for men and 3 for women in AUDIT-C [24].

The other major type of UAU is AUD. It is a chronically recurrent brain impairment in which compulsive and maladaptive alcohol use results in behavior dysregulation and negative mood once alcohol consumption is ceased [16,25]. Alcohol abuse and alcohol dependence are 2 major representatives for moderate and severe degrees of AUD, respectively [16,25,26]. Consumers with either of them can suffer from adverse consequences. Alcohol abuse, that is, unrestrained alcohol use, can make consumers fail to meet their major obligations and cause or exacerbate health and social problems [16,27]. More seriously, alcohol dependence, that is, a constant and strong desire for alcohol use without self-control or consideration of health, might result in physical or mental health problems once a large amount of alcohol is consumed over a long period [28]. To be diagnosed with AUD, a person should meet at least two of the 11 criteria listed in the Diagnostic and Statistical Manual of Mental Disorders 5th Edition in 1 year [29].

Mobile health (mHealth), also known as ecological momentary intervention [30,31], refers to the use of mobile devices, such as mobile phones, personal digital assistants, or other wireless devices, to deliver medical or public health services in a timely manner and in real-world living settings [30,32]. Due to limited human resources available for delivering continuous health care services for community-dwelling consumers suffering from chronic diseases, mHealth interventions are increasingly considered by the decision makers as a potential alternative solution in providing same quality but low-cost services [33]. Similarly, there has been a growing interest in using mobile phones to deliver public health interventions to support consumer self-control of UAU.

mHealth interventions are mainly delivered solely or in combination of 3 channels: short message services (SMS) text messaging, apps, and interactive voice response (IVR). SMS text messaging has been used to guide consumers to change alcohol use behavior, for example, to reduce alcohol intake to enable self-control of UAU [19,34]. Apps have been used to monitor consumers' alcohol use and to provide visual feedback about drinking behavior based on statistical analysis of input data. Raising self-awareness can ignite consumers' self-regulation so as to reduce alcohol use [35,36]. IVR has been used to generate audial interactions and to provide automatic answers to consumer queries on UAU [37,38]. Therefore, these 3 delivery channels can all provide effective and efficient interventions for consumer self-control of UAU.

Objectives

Recent reviews on digital interventions for self-control of UAU focus on the benefit of such interventions on improving health care services. In total, 2 reviews investigated electronic or Web-based interventions and found that despite a small effect, these interventions might improve behavioral outcomes, particularly for the group less likely to access traditional alcohol interventions such as women, youth, and risky drinkers [39,40]. A total of 5 reviews narrowed down the scope on mHealth interventions for self-control of UAU. In total, 2 of them focused on SMS text messaging and found that although the behavioral outcomes were modest, it was still a worthwhile endeavor [41,42]. The other 3 reviews suggested mHealth interventions, especially the ones that can provide personalized feedback, were beneficial for the reduction of UAU with their high fidelity, anonymity, and accessibility [31,43,44]. However, as the mHealth interventions were still nascent in nature, there is still a lack of understanding about how such interventions really

work for changing UAU. Solid evidence about the efficacy of these interventions from empirical field trials is required. Moreover, the other health outcomes, such as physiological and cognitive outcomes, need to be studied. Therefore, this review aimed to synthesize and understand the research evidence about the efficacy of mHealth interventions on various health outcomes for consumer self-control of UAU and to identify the core components to achieve these outcomes.

Methods

Study Design

A mixed-methods systematic review was conducted. Literature search and screening followed the preferred reporting items for systematic reviews and meta-analyses [45]. Data extraction was guided by the Consolidated Standards of Reporting Trials-electronic health checklist (V.1.6.1) [46]. The methodological quality of the studies was assessed by the Mixed Methods Appraisal Tool (MMAT) [47]. Data synthesis and analysis followed the principle of realist synthesis [48] and qualitative content analysis [49].

Literature Search and Screening

The literature search was performed from December 2016 to March 2017 and further refined in August 2018 in 7 electronic interdisciplinary databases: Scopus, PubMed, PubMed Central, CINAHL Plus with full text, MEDLINE with full text, PsycINFO, and PsycARTICLES (see Multimedia Appendix 1). The following terms and medical subject headings were used individually or in combination to identify the relevant publications: “mHealth,” “text message,” “SMS,” “App,” “IVR,” “self-control,” “self-regulation,” “alcohol*,” and “intervention.” To ensure adequate coverage, a manual search was also conducted to identify papers from *Journal of Medical Internet Research* and its sister journals. The search was restricted to peer-reviewed journal papers published in English between 2008 and 2017. In addition, the following criteria were used in the selection of papers.

Inclusion Criteria

The papers were included in which (1) the research focused on supporting consumer self-control of UAU; (2) health intervention was delivered through mobile phone technologies; and (3) the data were collected from empirical randomized controlled trials.

Exclusion Criteria

The papers were excluded that (1) reported clinical therapy such as injection and medication rather than consumer active participation in the daily self-control of UAU; (2) did not report any alcohol-related health outcome; (3) used the intervention not dealing with UAU or containing Web-based components delivered by desktop or Web-based computer applications; or (4) were review papers, study protocols, conceptual papers, editorials, government reports, or guidelines in the topic area.

Data Extraction

Data were extracted using a combination of an Endnote X8 and an Excel spreadsheet by 2 authors independently. These included name(s) of the author(s), year of publication, country of origin, population type, study setting, type of UAU, study type, eligibility, sample size, study arms and grouping, nonmobile cointervention, mHealth intervention theory, delivery mode, mHealth intervention content, implementation procedure, measurement, and outcomes.

Quality Assessment of the Studies

All studies were assessed using the 4 criteria in section 2 of the MMAT, in terms of (1) randomization or sequence generation, checking if there is a clear description about randomization; (2) allocation concealment, verifying if there is a clear description about blinding; (3) outcome data, confirming if more than 80% outcomes were reported; and (4) attrition, assessing if less than 20% of the participants dropped out. Responses to each criterion were “yes,” “no,” or “can’t tell.”

Data Synthesis and Analyses

Data were synthesized and analyzed using an inductive method. We reviewed all data that collected and identified similar notions and tagged them with the same code. Thereafter, we grouped the codes with similar meaning into an overarching concept. Concepts with similar meaning were grouped into a category that addresses our research question. The coding and data management were iteratively developed through constant comparison of the similarities and differences among codes.

To explore the initial outcomes about which components really make the intervention works, chi-square testing was conducted to test the relationship between health outcomes with the following 4 parameters: (1) types of UAU, being risky drinking or AUD; (2) with or without nonmobile cointervention; (3) theory-based or not; (4) deployment setting, being clinical, educational, or community based.

Results

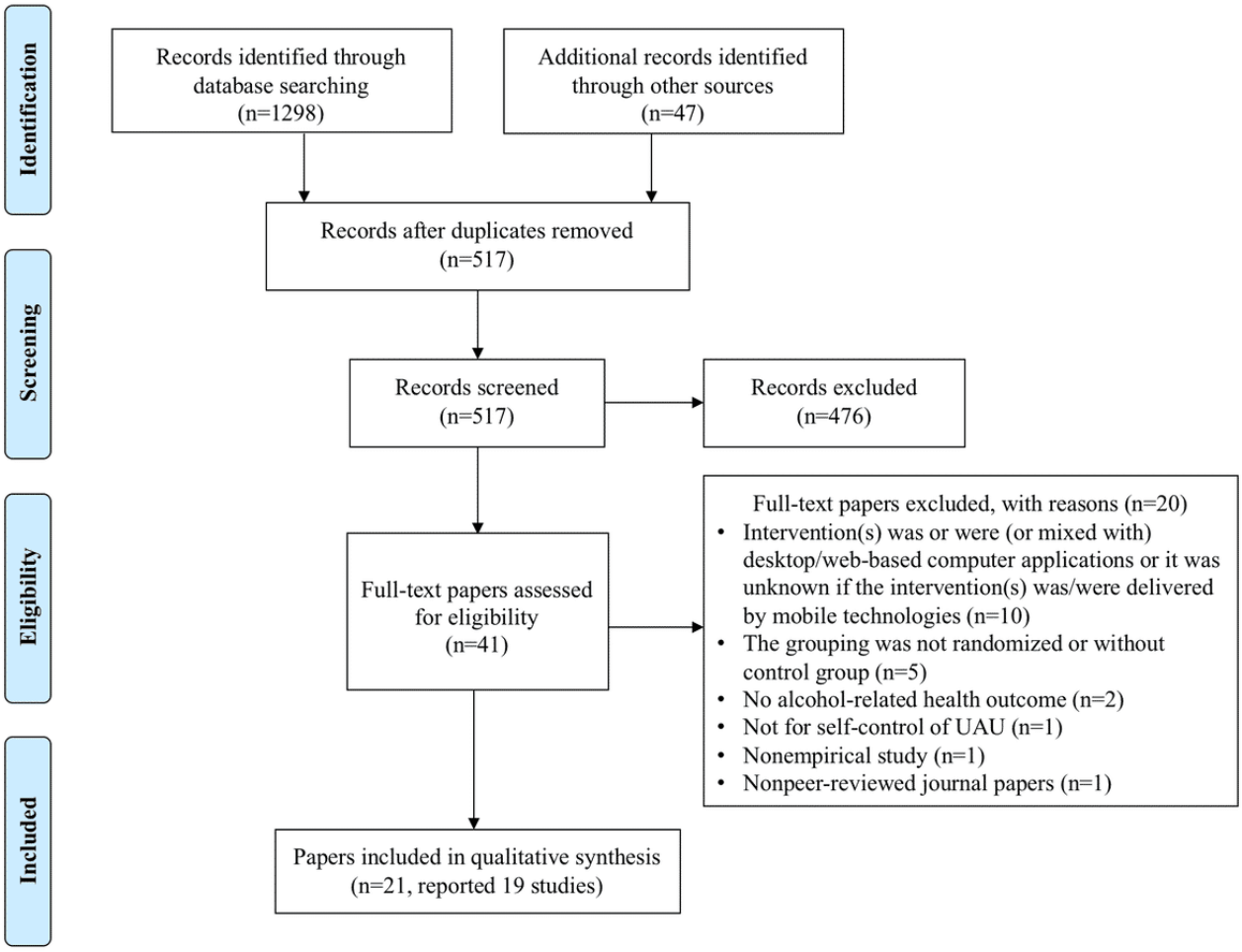
Search Outcome

The primary search yielded 1345 publications. After removing duplicates, 517 papers remained. Their titles and/or abstracts were manually screened against the inclusion and exclusion criteria. This led to 41 candidate papers. Of these, 20 were excluded after further scrutinizing the full paper (see Multimedia Appendix 2). Finally, 21 papers were included. Of these, 4 papers were from 2 studies. Suffoletto et al published 2 papers based on the same study population in 2014 and 2015 [50,51], respectively; so did Agyapong et al in 2012 and 2013 [52,53]. Therefore, a total of 19 studies were eligible for review (see Figure 1 and Multimedia Appendix 3). Among these studies, 58% (11/19) met all 4 MMAT criteria [19,34,37,38,52-59] and 32% (6/19) met 3 criteria [20,21,35,50,51,60,61], indicating high methodological quality in 90% of these studies (see Multimedia Appendix 4).

Characteristics of Studies

Although we searched studies published since 2008, all 19 eligible studies were conducted in 2012 and beyond and were from 7 developed countries (see Multimedia Appendix 5).

Figure 1. Literature search and screening process.



Of these, 9 studies (9/19, 47%) were conducted in the United States [19,34,35,38,50,51,55,57,59,62], 8 in Europe [20,21,37,52-54,56,58,61], and only 2 in New Zealand [60,63].

Study arms ranged from 2 to 6. In total, 12 studies (12/19, 63%) were 2-arm trials with an intervention arm and a control arm [21,34,35,52,53,55,56,58-63]. The control-arm participants received (1) no intervention [35,56,58]; (2) nonmobile intervention with the same content through interview [55,62], email [21], and e-booklet [61]; (3) nonalcohol-related content [34,52,53] or only assessment for monitoring purpose [60,63] through the same mobile devices; or (4) different rewarding mechanisms for their abstinence [59]. A total of 5 studies (5/19, 26%) had 3 arms. Of these, 2 added an assessment-only arm besides the intervention and control arms [50,51,57]. Hasin et al employed an arm in which the participants only received intervention through interview [38]. Gajecki et al used 2 intervention arms delivered by 2 different mobile apps in 1 study [20] and 2 intervention arms that started to use the app at different times in another study [54]. In the last 2 studies (2/19, 11%), Andersson conducted a 5-arm trial in which an mHealth intervention was compared with Web-based intervention and nonintervention. Both the mHealth and Web-based interventions had 2 implementation procedures, single and repeated [37]. Muench et al employed a 6-arm design, including 1 nonintervention arm, 1 assessment-only arm, and 4 intervention arms containing different contents [19].

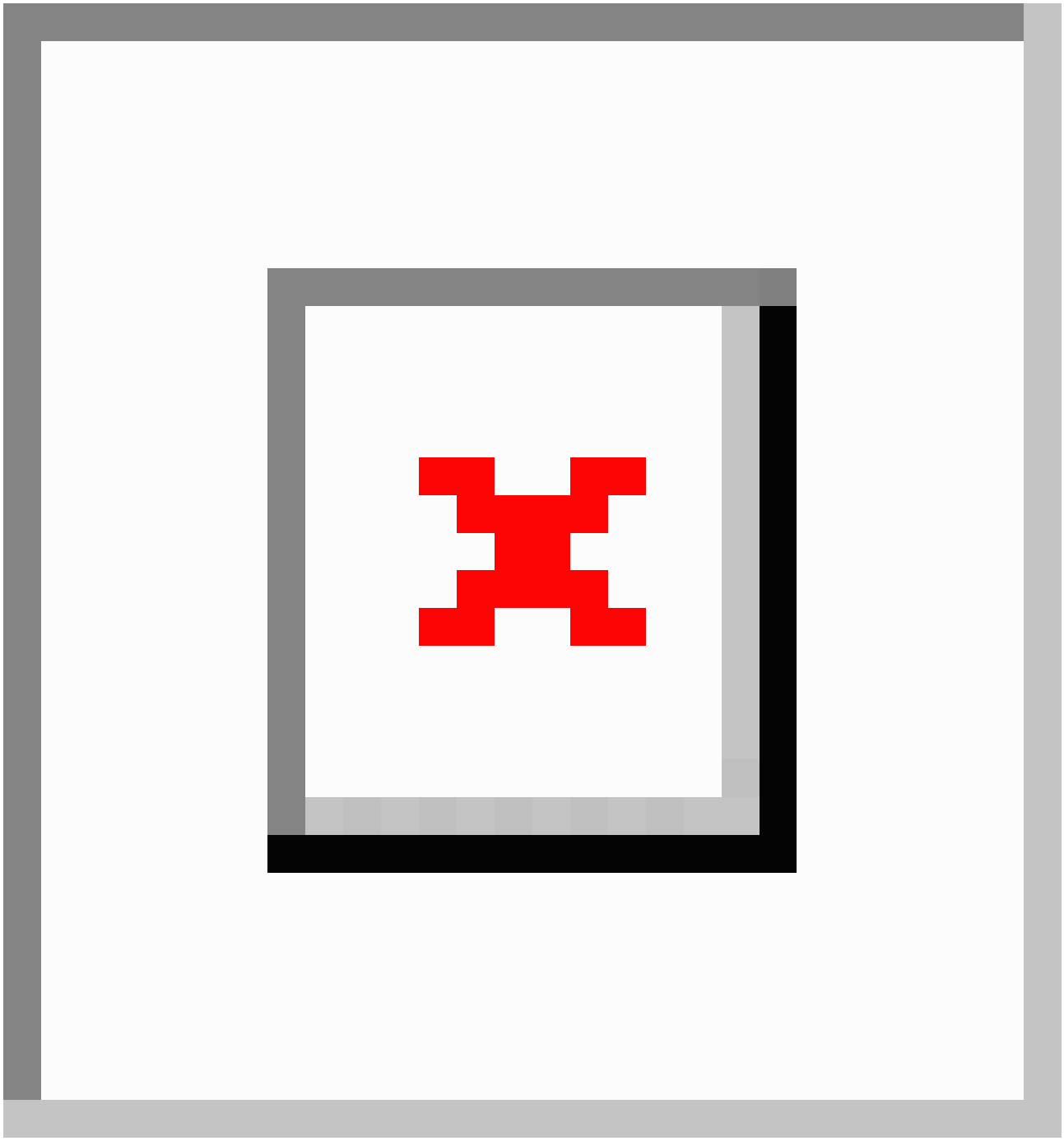
We identified 5 core components of mHealth interventions for UAU: context, theoretical base, content, delivery mode, and implementation procedure and 3 types of potential health outcomes: behavioral, physiological, and cognitive outcome (see Figure 2 and Multimedia Appendix 6).

Five Core Components of Mobile Health Interventions for Self-Control of Unhealthy Alcohol Use

Context

There are 3 types of contexts: participant characteristics, deployment setting, and nonmobile cointervention, which were conducted simultaneously to support the mHealth intervention.

Figure 2. In total, 5 core components of mobile health interventions for self-control of unhealthy alcohol use and 3 types of potential health outcomes.



The participants can be categorized into 2 cohorts according to their age: youth group aged below 35 years [20,21,34,37,50,51,54,57,59,60,62,63] and middle and old aged group, aged above 35 years [19,35,38,52,53,55,56,58,61]. They were either risky drinkers [19-21,34,37,38,50,51,57,59-63] or had AUD [20,35,38,52-56,58]. They suffered from comorbidity of depression [52,53], HIV [38,55], drug dependence [55], or smoking [57]. The interventions were deployed in educational settings [20,21,34,37,54,57,60,62,63], clinical settings [35,38,50-53,56,58], and community-based settings [19,55,59,61]. The nonmobile cointervention included social intervention guided by the theory of motivational interviewing [38,50,51,55,57,62] and paper-based intervention in which participants were provided with guidelines for safe alcohol use about the mHealth intervention [19].

Theoretical Base

In total, 2 types of theories were reported to guide the design and implementation of the mHealth interventions, including behavioral change theories and psychological theories of motivation.

Behavioral change theories included theory of planned behavior [20,21,50,51,56,63], health belief model [19,50,51], social cognitive theory [21,63], theory of reasoned action [50,51], information motivation behavioral model [50,51], cognitive behavioral therapy [57], and social learning theory [19]. Psychological theories of motivation included self-determination

theory [21,35,63], model of action phases [21,63], and contingency management [59]. Notably, although Aharonovich et al did not report any explicit theory applied to their intervention, the design of their app, HealthCall, was theory based [55].

Delivery Mode

A total of 3 delivery modes were identified: SMS text messaging (12/19, 63%) [19,21,34,50-53,56,58-63], app (5/19, 26%) [20,35,54,55,57], and IVR (2/19, 11%) [37,38]. In total, 6 apps tested in the 5 studies were TeleCoach [54], Brief Alcohol and Smoking Intervention for College Students via Mobile (BASICS-Mobile) [57], HealthCall-S [55], Alcohol-Comprehensive Health Enhancement Support System (A-CHESS) [35], PartyPlanner, and Promillekoll [20]. Only 2 studies described the underlying operating system for these apps [20,57]. PartyPlanner and Promillekoll ran on the Android or iOS [20], and BASICS-Mobile ran on Blackberry, Android, or iOS [57].

Content

In total, 3 types of content were designed to support the participants' self-control of UAU. They were information [19-21,34,35,37,38,50-58,60-63], motivation [19,21,34,35,37,50-53,55-57,59,61,62], and reminder [19,35,50,51,55,56,58,59].

Informational content included general and personalized information. The general information facilitated the participants in (1) enriching their knowledge about risks and negative consequences of UAU [19,34,37,50,51,60,61,63], alcohol-related facts [34,37,54,57,61], social drinking norms [50,51,62], and benefits of reducing drinking amount according to safety guidelines [19]; (2) acquiring strategies to control alcohol use [20,34,50,51,54,55,57,58,62], to handle relapse or cravings [52-54,57,61], to manage emotion [54,61], and to reduce intoxication [37]; (3) getting referrals such as alcohol counseling services [55,58], instant library, and weblinks to further alcohol-related information [35]; and (4) conducting recommended actions for self-control of UAU such as tracking and reporting their drinking facts [19,20,35,37,38,50,51,54-56,58,62,63], reasons for drinking or abstinence [38,55], and estimated blood alcohol concentration (eBAC) value [20], mood [38], medication adherence [38], and well-being [38]; introspecting their performance [21]; or simulating a drinking occasion to set personal goal of eBAC and comparing actual eBAC after drinking against this goal [20].

The personalized information helped the participants in (1) providing the tailored feedback according to their responses [19,37,50,51,55,57]; (2) recommending them to set intermittent low-risk drinking goals [61] to replace drinking alcohol by alternative activities [37,57], to celebrate goal attainment [50,51,61], to self-reflect on challenges of UAU [21], to improve the drinking plan, and to reinforce self-control behavior [37,50,51]; and (3) addressing their problems identified at various stages [19].

Motivational content included (1) encouragement messages for reducing alcohol use [21,34,37,50-53,57,61,62], committing to preset drinking goals [19,50,51,55,56] and medical adherence [52,53], and releasing distress [35]; (2) peer support through sharing experiences with others in the anonymous discussion groups [35]; and (3) possible monetary compensation to incent participants to submit their valid on-time video [59].

Reminding content facilitated participants in (1) reminding them to remember and fulfill their promises [50,51,55,56,58,59] and (2) warning them about alcohol risks at their risky drinking times [19] or when they were near high-alcohol places detected by global positioning system [35].

Implementation Procedure

The duration of the interventions varied, ranging from 4 days [62], 1 week [37,60,63], 2 weeks [57], 4 weeks [21,37,59], 6 weeks [34], 7 weeks [20], 2 months [38,55,58], 3 months [19,50-54], 6 months [56,61], to 8 months [35].

With regard to the frequency, SMS text messages were sent once [34], twice [50,51,58], or 4 times [21,60] weekly in 5 studies and once [63], twice [52,53], 1 to 3 times [59], or 4 to 6 times [62] daily in 4 studies. The frequency appeared to reduce when the length of the study increased [56,61]. Haug et al sent 1 SMS text message per week in the first 8 weeks and then 1 per fortnight in the remaining 18 weeks [56]. Brendryen et al sent 1 SMS text message per day for 8 weeks, then 1 per week for 4 weeks, and finally 1 per month in the last 2 months [61]~Brendryen, 2014 #8^. Muench et al sent the SMS text messages with different content at a different frequency, either once daily for educating participants about alcohol use or once weekly for self-monitoring content and feedback [19]. The participants in Alessi and Petry's study were given a breathalyzer and the corresponding accessories to self-measure breath alcohol concentration (BrAC) and submit a valid real-time video containing the whole self-measuring process to the organizer via SMS text message 1 to 3 times per day at the fixed time interval to prove their abstinence [50]. The intervention-arm participants would be rewarded with more vouchers if their BrAC value was normal. In contrast, the control-arm participants were not rewarded although their BrAC value was normal [50].

In 3 studies in which the interventions were delivered by mobile apps, the frequency of data collection was once daily in Aharonovich et al's study [55] and once weekly in Gustafson et al and Gajecki et al's study [35,54]. In most cases, apps were used in real time according to a participant's preference, typically to receive a certain recommendation once a preset condition was met. For example, Promillekoll could send real-time notification and the corresponding strategies to control alcohol use if a participant's eBAC was over 0.06% [20]. A-CHESS would send an alarm when a participant was near a high-risk alcohol place to be detected by the embedded global positioning system [35].

Andersson divided his intervention arm into 2 subgroups, both receiving the same content but through different delivery modes, either delivering single IVR every day for 1 week or delivering repeated IVR for 4 weeks [37]. Hasin et al requested their intervention-arm participants to spend 1 to 3 min per day to send back their answers to a series of questions, asking their compliance with drinking guidelines on the previous day via a

toll-free number [38]. The participants' phone calls were initially answered by the prerecorded IVR in the first 30 days. After evaluating a participant's IVR data, the consultant reset the person's drinking goal for the next 30 days [38].

In terms of timing of delivery, 10 studies reported the fixed time or time slots to deliver the intervention, the rest were flexibly available on demand. The popular days of intervention delivery were Thursday [34,50,51,58,60,63], Saturday [21,34,59,60,63], and Sunday [21,50,51,59,63], mainly at or after 6 pm [19,34,52,53,56,59,60,63].

All studies conducted the baseline assessments. In total, 6 studies conducted an assessment during the intervention period to explore the initial outcome [35,38,54,55,58,61]. The postintervention assessments were conducted in all studies at different time points with different numbers of repetitive measurement. A total of 17 studies conducted 1 assessment immediately after the intervention [19-21,34,35,38,50-61,63]. A total of 8 studies conducted the second assessment 1 month [38,57], 6 weeks [34], 1 academic semester [60,63], 3 months [50-53], and 4 months [35] after the intervention. In total, 2 studies conducted the third assessment 3 [50,51] and 4 months [38] after the intervention. Only 1 study conducted the fourth assessment after 10 months of the intervention [38]. Instead of immediately measuring the outcomes, in 2 studies, the measures were conducted only after 1 month [62]. Of these, 1 study measured the outcome 4 weeks after the intervention for the single IVR intervention arm and 1 week after the intervention for the repeated IVR intervention arm [37].

Health Outcomes

Behavioral Outcome

Behavioral outcome was measured in 18 studies [19,20,34,35,37,38,50-63]. Significant positive outcome was found in 11 of these studies [19,35,37,38,50-54,59-61,63]. These positive outcomes were measured by 1 or more indicators. These included the decreased number of SDs [19,50,51,54,61,63], heavy drinking days [19,35,38,50,51,59], RSOD or binge drinking prevalence [50,51,54], alcohol-related injury prevalence [50,51], and peak eBAC value [37]; increased number of abstinence days [19,52,53,59] or the increased negative affect score in Alcohol Abstinence Self Efficacy Scale [52,53]; and the decreased score in the Alcohol Addiction Severity Index, Drinker Inventory of Consequences [59], or AUDIT [37].

No significant behavioral change was found in 6 studies [34,55-58,62]. In total, 2 studies reported a gender-related behavioral outcome [20,63]. Contrary to the initial objective of reducing UAU, the male participants in the intervention arm significantly increased drinking frequency, whereas no change was found in the female participants and the control arm in 1 study [20]. In the study conducted by Riordan et al, after providing intervention-arm participants with 1-week SMS text messages, the female participants consumed significantly less alcohol 1 week and 1 semester later than their female counterparts in the control arm. However, no intervention effect was found for the male participants [63].

Physiological Outcome

Physiological outcome was measured in only 1 study via BrAC [59]. Alessi and Petry found a significant improvement in the percentage of negative BrAC in the intervention group but no significant change in the control group [59].

Cognitive Outcome

Cognitive outcome was measured in 3 studies [21,34,62] and was significantly positive in only 1 study in which the participants' readiness to change UAU behavior in the intervention arm was significantly improved [62]. No significant cognitive change was found in the other 2 studies in terms of motivation to change and self-confidence to resist alcohol [21,34].

Comparison With the Differences in Health Outcomes Among Different Groups of Studies

Over half of the SMS- and IVR-enabled interventions were effective in reducing alcohol use or increasing readiness to change UAU in 8 out of 12 studies (67%) [19,50-53,59-63] and 2 out of 2 studies (100%) [37,38], respectively. In contrast, app-enabled interventions were only successful in reducing alcohol use in 2 out of 5 studies (40%) [35,54].

Chi-square test did not find any significant differences in health outcomes among groups of studies with different conditions. It suggested that the health outcomes were similar regardless of the types of UAU studied, whether there was nonmobile cointervention, whether the study was theoretical-based, or which setting it was deployed.

Discussion

Principal Findings and Comparison With Previous Work

This study aimed to synthesize and understand the research evidence about efficacy of mHealth interventions on different health outcomes for consumer self-control of UAU and to identify their core components to achieve these outcomes. In total, 19 studies were systematically reviewed and 3 types of health outcomes such as behavioral, physiological, and cognitive outcome and 5 components of these interventions such as context, theoretical base, delivery mode, contents, and implementation procedure were found.

Health Outcomes

As approximately two-thirds (11/18) of the studies that measured the behavioral outcomes identified a significant positive change [19,35,37,38,50-54,59-61,63]: mHealth interventions appear to be more effective in changing UAU behavior in comparison with the traditional methods. The results could be explained by the information-motivation-behavioral skills model, which suggests that a participant's behavior change is attributed to the provided information, motivation, and improved skills [64]. This is also in accordance with the findings of Regmi et al's review in smoking cessation context where the abstinence days of smoking increase after applying mHealth interventions [65].

Despite the significant 100% positive physiological outcome measured by BrAC, only 5% of the included studies assessed

the physiological measurement [59]. There might be two reasons for this. First, people tend to test their biomedical markers in hospital or clinic rather than by themselves as they might lack corresponding skills and it is inconvenient. Second, the corresponding self-testing devices are not cheap and not all research projects can afford them, especially for the projects with a large sample size. Instead, researchers preferred to measure behavioral outcomes because the rough BAC can be calculated simply using Widmark formula once a participant reports his or her alcohol use [66]. In addition, the unique factor to affect the BrAC is alcohol intake. Therefore, the physiological outcome must change when the behavior changes.

For the same reason, as the cognitive outcomes are inconvenient to measure in comparison with behavioral ones, only 16% of the studies [21,34,62] assessed cognitive changes, of which 33% [62] were significantly improved. This might be because cognition can be influenced by various factors, and their measurement can be somewhat subjective and abstract. For example, Mason et al assessed cognitive change using 5 variables: alcohol expectations, readiness to change drinking behavior, importance of change, confidence in ability to change, and intentions to reduce alcohol use. Only the variable of readiness to change drinking behavior was improved [62]. Notably, although all these 3 studies also reported the improvement in behavior changes, it is still not enough to conclude that behavior always changes with cognition. According to the theory of cognitive dissonance proposed by Leon Festinger, a person can be motivated to reduce own psychological inconsistency and discomfort by changing the behavior [67]. No matter whether the interventions were genuinely accepted by the participants or not, most of them modified their behavior in compliance with the information they received from the interventions to reduce their interconflicts [67]. Therefore, when their cognitive changes were assessed, they might provide the real thoughts, which might not be consistent with the behavior that were displayed.

Complementing the traditional interventions such as face-to-face counseling, in which unhealthy alcohol users' access to treatment was provided in a passive manner within a confined time and location, mHealth interventions open new opportunities for engaging consumers in positive self-control with increased flexibility. The effect of control was improved by continuous tracking and monitoring, interactive communication, or personalized feedback from health care providers anytime, anywhere [30,68,69].

Five Components of Mobile Health Interventions for Self-Control of Unhealthy Alcohol Use

Participants in most reviewed studies were risky drinkers without documented pathological conditions [19-21,34,37,38,50,51,57,59-63]. We did not find much difference in the intervention outcome between the types of participants, being risky drinker or AUD. This result is consistent with the finding of Blow et al that health outcomes of an intervention are not influenced by the level of severity of alcohol addiction [70]. However, this is contradictory with the findings in the previous review conducted by White et al that e-interventions can be particularly useful for at-risk users [40]. Kazemi et al also seconded that for this population group, mHealth intervention might be the most cost-effective UAU management strategy [71]. The paradox might be caused by the different conditions such as timing and frequency of the interventions or different population types and settings [42].

The gender difference in intervention outcome found in 2 studies [20,63] might be explained by the observation of Hirschi and Gottfredson that men have lower self-control than women [72]. Notably, these 2 studies were both done on young adult students in university settings. This might suggest that it is much more difficult for males in this setting to change their behavior in terms of UAU. First, there are strong social or peer norms in this cohort, which prevent the change of drinking behavior [73], and second, males seem to be less compliant and agreeable than females, and they lack ability to absorb the meaning of the SMS text messages [74,75]. Riordan et al offered some suggestions for improving SMS text messaging for young men and later demonstrated that using more colloquial tone and sending only messages with the potential social consequences of UAU are better for this population [60]. Similar to the finding of Platt et al [76], we did not find any significant relationship between the health outcome and deployment setting.

Although not having any significant impact on health outcomes, cointervention, such as induction or training to enable a participant to confidently use the apps or IVR, is an integral, vital component for a successful mHealth intervention [77,78]. This might explain why more cases of nonmobile cointervention were reported in interventions delivered via apps (3/5, 60%) and IVR (1/2, 50%). Most likely, the participants were more familiar with SMS text messaging than the other 2 delivery modes; therefore, the cointervention was less reported in the studies delivered by SMS text messaging (3/12, 25%). Notably, the population of the 2 studies in which interventions were delivered via apps without formal reporting of cointervention was university students at a younger age. This might be because of the internet use and mobile phone technologies are popular in this cohort; thus, the app designers did not consider it necessary to provide the students with training to use the app [79].

Behavior change theory provides the foundation for the formation of strategies to incrementally change a consumer's behavior of UAU [80]. Psychological theory of motivation is used to develop motivational strategies to control UAU against psychological craving for alcohol [81]. Although mHealth interventions based on theory can improve instructional design and the effect of self-control of UAU [76], no significant difference in health outcomes was found in this review for the studies based on theory and those otherwise, which is in accordance with the finding of Garnett et al [82]. There might be two reasons to explain this phenomenon. First, from what was described in the Methods, it appears that theory was implicitly applied to the mHealth interventions although a study might not make the claim to be theory based. For example, Bock et al did not report the use of any theory; however, one of the SMS text messages in their intervention "always have an exit plan" indicated the unconscious application of the theory of planned behavior [34]. Second, it takes time to bring in tangible health outcomes for participant's self-control of UAU [52,53].

Almost all SMS- or IVR-enabled interventions were effective in reducing alcohol use or increasing readiness to change except the mobile apps [20,55,57]. This might be because the former 2 types of interventions were delivered proactively, on regular basis, always accessible to the participants regardless of their intention. In contrast, the participants' access to the app-based interventions relied on their self-action of opening the apps, which might not always happen. This is consistent with the findings in Meredith et al's review [83], and it also recommends that the future mHealth apps need push notifications regularly to prompt the active engagement of the users.

Informational content facilitated the participants to develop essential knowledge and skill to build their capacity to change their belief and UAU behavior. It also provided necessary feedback to enable self-awareness of UAU status, which could help execute self-regulation of UAU. Motivational content provided continuous encouragement and peer support through experience sharing to raise the participants' morale in changing UAU behavior. Reminding content provided constant recall to ensure the participants to stay on track in self-control of UAU. Delivery of these 3 types of content is in line with the model of human practical reasoning developed by Michael Bratman [84].

As the length of the reviewed studies was not long enough, ranging from 4 days to 8 months, it is no surprise that there was no obvious improvement in tangible health outcomes in many studies. Longer duration, that is, 6 months or more [35,61], more frequent delivery [52,53,59,62] and certain techniques such as tangible incentives [62], and assessment during the intervention [61] might help achieve positive outcomes. In contrast, a relatively small sample size, less than 100 [34,55,56] and a short follow-up period, less than 2 months [57], might cause a lack of significant health outcomes for the interventions. However, whether the health outcomes can be influenced by these factors still needs to be verified.

With the same content and implementation procedure, Andersson et al found differences in health outcomes measured by peak eBAC and AUDIT scores with different delivery modes in which the efficacy was better delivered by IVR than the Web [37]. Similarly, with the same delivery mode and implementation procedure but different content, Muench et al also found differences in health outcomes measured by numbers of SDs, heavy drinking days, and abstinent days. The content that highlighted the negative consequences of UAU was significantly more likely to bring about positive health outcomes than the content that emphasized the benefits of UAU abstinence [19]. Furthermore, with the same content and delivery mode, Gajeck et al found that the health outcomes measured by SD and drinking frequency were significantly different with different intervals of intervention [54].

Although the first generation of iPhone was released in June 2007, marking the debut of smartphone technology [85], no eligible studies were found before 2012. It appears that using mobile phones to deliver mHealth interventions for UAU was staged in 2012.

Limitations

The first limitation of this study was that the coverage of the studies might not be exhaustive, because of which our search was confined to the 7 databases. However, the comprehensiveness of these databases can ensure the representativeness of the trend suggested by this study. The heterogeneity of participant characteristics, intervention, and health outcome measures makes it difficult to compare rigorously the findings among the studies. A lack of homogenous, quantitative measures in the original studies also deemed it impossible to conduct more rigorous meta-analysis. As only peer-refereed journal papers were included to ensure the rigor of this study, there could be a potential risk of reporting bias toward positive findings.

Conclusions

This systematic review summarized the extant research evidence about the health outcomes of mHealth interventions for consumer self-control of UAU. A total of 3 health outcomes, that is, physiological, behavioral, and cognitive outcomes and 5 core components of these interventions, that is, context, theoretical base, delivery mode, content, and implementation procedure, were synthesized and analyzed. In comparison with the traditional interventions, the evidence to support effectiveness of mHealth interventions for consumer self-control of UAU is modest at best. A majority of studies showed that mHealth interventions brought positive health outcomes in helping unhealthy alcohol users to proactively engage in self-control of their UAU behavior, especially for the ones delivered by SMS text messaging and IVR systems. Sound evidence is yet to be sought about the effects of these interventions in improving the physiological and cognitive outcomes. Further research is needed to gather evidence about the optimal design of mHealth interventions, their implementation, and effects in supporting consumer self-control of UAU.

Acknowledgments

The authors would like to acknowledge and thank Diane Walton for proofreading the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search and screen records.

[PDF File (Adobe PDF File), 260KB - mhealth_v7i1e10899_app1.pdf]

Multimedia Appendix 2

Excluded papers with reasons.

[PDF File (Adobe PDF File), 21KB - mhealth_v7i1e10899_app2.pdf]

Multimedia Appendix 3

List of included studies.

[PDF File (Adobe PDF File), 13KB - mhealth_v7i1e10899_app3.pdf]

Multimedia Appendix 4

Quality appraisal of the included studies.

[PDF File (Adobe PDF File), 254KB - mhealth_v7i1e10899_app4.pdf]

Multimedia Appendix 5

Characteristics of the included studies.

[PDF File (Adobe PDF File), 156KB - mhealth_v7i1e10899_app5.pdf]

Multimedia Appendix 6

Five components and three types of health outcomes.

[PDF File (Adobe PDF File), 312KB - mhealth_v7i1e10899_app6.pdf]

References

1. Johnson NB, Hayes LD, Brown K, Hoo EC, Ethier KA, Centers for Disease Control and Prevention (CDC). CDC National Health Report: leading causes of morbidity and mortality and associated behavioral risk and protective factors--United States, 2005-2013. MMWR Suppl 2014 Oct 31;63(4):3-27. [Medline: 25356673]
2. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2224-2260 [FREE Full text] [doi: 10.1016/S0140-6736(12)61766-8] [Medline: 23245609]
3. World Health Organization. Global Status Report on Alcohol and Health 2014. Geneva, Switzerland: WHO; 2014.
4. Australian Institute of Health and Welfare. Canberra, Australia: AIHW; 2016. Australia's health 2016 URL: https://www.aihw.gov.au/getmedia/9844cefb-7745-4dd8-9ee2-f4d1c3d6a727/19787-AH16.pdf.aspx?inline=true [accessed 2018-04-25] [WebCite Cache ID 6yxA3PG2Q]
5. Australian Bureau of Statistics. 2015. National Health Survey: First Results, 2014-15 URL: http://www.abs.gov.au/AUSSTATS/abs@.nsf/DetailsPage/4364.0.55.0012014-15?OpenDocument [accessed 2018-04-25] [WebCite Cache ID 6yxAESoIa]
6. Canadian Institute for Health Information. Alcohol Harm In Canada: Examining Hospitalizations Entirely Caused by Alcohol and Strategies to Reduce Alcohol Harm. Ottawa, Canada: CIHR; 2017.
7. Rehm J, Gmel G, Sempos CT, Trevisan M. Alcohol-related morbidity and mortality. Alcohol Res Health 2003;27(1):39-51 [FREE Full text] [Medline: 15301399]
8. Emanuele N, Emanuele MA. The endocrine system: alcohol alters critical hormonal balance. Alcohol Health Res World 1997;21(1):53-64. [Medline: 15706763]
9. Shield KD, Parry C, Rehm J. Chronic diseases and conditions related to alcohol use. Alcohol Res 2013;35(2):155-173 [FREE Full text] [Medline: 24881324]
10. Strakowski SM, DelBello MP, Fleck DE, Adler CM, Anthenelli RM, Keck PE, et al. Effects of co-occurring alcohol abuse on the course of bipolar disorder following a first hospitalization for mania. Arch Gen Psychiatry 2005 Aug;62(8):851-858. [doi: 10.1001/archpsyc.62.8.851] [Medline: 16061762]
11. Stickley A, Jukkala T, Norström T. Alcohol and suicide in Russia, 1870-1894 and 1956-2005: evidence for the continuation of a harmful drinking culture across time? J Stud Alcohol Drugs 2011 Mar;72(2):341-347. [doi: 10.15288/jsad.2011.72.341] [Medline: 21388607]
12. Abernathy K, Chandler LJ, Woodward JJ. Alcohol and the prefrontal cortex. Int Rev Neurobiol 2010;91:289-320 [FREE Full text] [doi: 10.1016/S0074-7742(10)91009-X] [Medline: 20813246]

13. Gonzalez-Guarda RM, Ortega J, Vasquez EP, De Santis J. La mancha negra: substance abuse, violence, and sexual risks among Hispanic males. West J Nurs Res 2010 Feb;32(1):128-148 [FREE Full text] [doi: 10.1177/0193945909343594] [Medline: 19915207]
14. Taylor B, Irving HM, Kanteres F, Room R, Borges G, Cherpitel C, et al. The more you drink, the harder you fall: a systematic review and meta-analysis of how acute alcohol consumption and injury or collision risk increase together. Drug Alcohol Depend 2010 Jul 01;110(1-2):108-116 [FREE Full text] [doi: 10.1016/j.drugalcdep.2010.02.011] [Medline: 20236774]
15. Saitz R. Unhealthy alcohol use. N Engl J Med 2005 Feb 10;352(6):596-607. [doi: 10.1056/NEJMcp042262] [Medline: 15703424]
16. Holt S, Tetrault J. Unhealthy alcohol use. Clin Liver Dis 2016 Aug;20(3):429-444. [doi: 10.1016/j.cld.2016.02.003] [Medline: 27373607]
17. International Alliance for Responsible Drinking. Drinking guidelines: General population URL: http://www.iard.org/resources/drinking-guidelines-general-population/ [accessed 2018-04-25] [WebCite Cache ID 6yxAjzep9]
18. Dawson DA. Defining risk drinking. Alcohol Res Health 2011;34(2):144-156 [FREE Full text] [Medline: 22330212]
19. Muench F, van Stolk-Cooke K, Kuerbis A, Stadler G, Baumel A, Shao S, et al. A randomized controlled pilot trial of different mobile messaging interventions for problem drinking compared to weekly drink tracking. PLoS One 2017;12(2):e0167900 [FREE Full text] [doi: 10.1371/journal.pone.0167900] [Medline: 28146560]
20. Gajecki M, Berman AH, Sinadinovic K, Rosendahl I, Andersson C. Mobile phone brief intervention applications for risky alcohol use among university students: a randomized controlled study. Addict Sci Clin Pract 2014;9:11 [FREE Full text] [doi: 10.1186/1940-0640-9-11] [Medline: 24985342]
21. Bendtsen M, Bendtsen P. Feasibility and user perception of a fully automated push-based multiple-session alcohol intervention for university students: randomized controlled trial. JMIR Mhealth Uhealth 2014;2(2):e30 [FREE Full text] [doi: 10.2196/mhealth.3233] [Medline: 25098296]
22. Hodgson R, Alwyn T, John B, Thom B, Smith A. The FAST alcohol screening test. Alcohol Alcohol 2002;37(1):61-66 [FREE Full text] [Medline: 11825859]
23. Saunders JB, Aasland OG, Babor TF, de la Fuente JR, Grant M. Development of the alcohol use disorders identification test (AUDIT): WHO collaborative project on early detection of persons with harmful alcohol consumption--II. Addiction 1993 Jun;88(6):791-804. [doi: 10.1111/j.1360-0443.1993.tb02093.x] [Medline: 8329970]
24. Bush K, Kivlahan DR, McDonell MB, Fihn SD, Bradley KA. The AUDIT alcohol consumption questions (AUDIT-C): an effective brief screening test for problem drinking. Ambulatory Care Quality Improvement Project (ACQUIP). Alcohol use disorders identification test. Arch Intern Med 1998 Sep 14;158(16):1789-1795. [doi: 10.1001/archinte.158.16.1789] [Medline: 9738608]
25. Grant BF, Dawson DA, Stinson FS, Chou PS, Kay W, Pickering R. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. Drug Alcohol Depend 2003 Jul 20;71(1):7-16. [doi: 10.1016/S0376-8716(03)00070-X] [Medline: 12821201]
26. Grant BF, Goldstein RB, Saha TD, Chou SP, Jung J, Zhang H, et al. Epidemiology of DSM-5 alcohol use disorder: results from the national epidemiologic survey on alcohol and related conditions III. J Am Med Assoc Psychiatry 2015 Aug;72(8):757-766 [FREE Full text] [doi: 10.1001/jamapsychiatry.2015.0584] [Medline: 26039070]
27. Renna F. Alcohol abuse, alcoholism, and labor market outcomes: looking for the missing link. Ind Labor Relat Rev 2008 Oct;62(1):92-103. [doi: 10.1177/001979390806200105]
28. Lloyd DA, Turner RJ. Cumulative lifetime adversities and alcohol dependence in adolescence and young adulthood. Drug Alcohol Depend 2008 Mar 01;93(3):217-226 [FREE Full text] [doi: 10.1016/j.drugalcdep.2007.09.012] [Medline: 17980975]
29. National Institute on Alcohol Abuse and Alcoholism. Alcohol use disorder URL: https://www.niaaa.nih.gov/alcohol-health/overview-alcohol-consumption/alcohol-use-disorders [accessed 2018-09-23] [WebCite Cache ID 72eyt3EdY]
30. Heron KE, Smyth JM. Ecological momentary interventions: incorporating mobile technology into psychosocial and health behaviour treatments. Br J Health Psychol 2010 Feb;15(Pt 1):1-39 [FREE Full text] [doi: 10.1348/135910709X466063] [Medline: 19646331]
31. Morgenstern J, Kuerbis A, Muench F. Ecological momentary assessment and alcohol use disorder treatment. Alcohol Res 2014;36(1):101-109 [FREE Full text] [Medline: 26259004]
32. Kay M, Santos J, Takane M. World Health Organization. 2011. Mhealth: New horizons for health through mobile technologies 2011 URL: http://apps.who.int/iris/bitstream/handle/10665/44607/9789241564250_eng.pdf?sequence=1&isAllowed=y [accessed 2018-04-25] [WebCite Cache ID 6yxBHznr3]
33. Yu P, Wu M, Yu H, Xiao G. The challenges for the adoption of m-health. : IEEE; 2006 Presented at: International Conference on Service Operations and Logistics, and Informatics; June 21-23, 2006; Shanghai, China p. 181-186. [doi: 10.1109/SOLI.2006.329059]
34. Bock BC, Barnett NP, Thind H, Rosen R, Walaska K, Traficante R, et al. A text message intervention for alcohol risk reduction among community college students: TMAP. Addict Behav 2016 Dec;63:107-113. [doi: 10.1016/j.addbeh.2016.07.012] [Medline: 27450909]

35. Gustafson DH, McTavish FM, Chih MY, Atwood AK, Johnson RA, Boyle MG, et al. A smartphone application to support recovery from alcoholism: a randomized clinical trial. J Am Med Assoc Psychiatry 2014 May;71(5):566-572 [FREE Full text] [doi: 10.1001/jamapsychiatry.2013.4642] [Medline: 24671165]
36. Gonzalez VM, Dulin PL. Comparison of a smartphone app for alcohol use disorders with an internet-based intervention plus bibliotherapy: a pilot study. J Consult Clin Psychol 2015 Apr;83(2):335-345. [doi: 10.1037/a0038620] [Medline: 25622202]
37. Andersson C. Comparison of web and interactive voice response (IVR) methods for delivering brief alcohol interventions to hazardous-drinking university students: a randomized controlled trial. Eur Addict Res 2015;21(5):240-252. [doi: 10.1159/000381017] [Medline: 25967070]
38. Hasin DS, Aharonovich E, O'Leary A, Greenstein E, Pavlicova M, Arunajadai S, et al. Reducing heavy drinking in HIV primary care: a randomized trial of brief intervention, with and without technological enhancement. Addiction 2013 Jul;108(7):1230-1240 [FREE Full text] [doi: 10.1111/add.12127] [Medline: 23432593]
39. Dedert EA, McDuffie JR, Stein R, McNiel JM, Kosinski AS, Freiermuth CE, et al. Electronic interventions for alcohol misuse and alcohol use disorders: a systematic review. Ann Intern Med 2015 Aug 4;163(3):205-214. [doi: 10.7326/M15-0285] [Medline: 26237752]
40. White A, Kavanagh D, Stallman H, Klein B, Kay-Lambkin F, Proudfoot J, et al. Online alcohol interventions: a systematic review. J Med Intern Res 2010;12(5):e62 [FREE Full text] [doi: 10.2196/jmir.1479] [Medline: 21169175]
41. Keoleian V, Polcin D, Galloway GP. Text messaging for addiction: a review. J Psychoactive Drugs 2015;47(2):158-176 [FREE Full text] [doi: 10.1080/02791072.2015.1009200] [Medline: 25950596]
42. Riordan B, Flett J, Conner T, Scarf D. Text message interventions for alcohol use: Current research and future directions. In: Gutierrez W, editor. Alcohol and Drug Abuse: Alcohol Consumption: Patterns, Influences and Health Effects. Hauppauge, New York: Nova Science Publishers; 2016:187-191.
43. Fowler LA, Holt SL, Joshi D. Mobile technology-based interventions for adult users of alcohol: a systematic review of the literature. Addict Behav 2016 Nov;62:25-34. [doi: 10.1016/j.addbeh.2016.06.008] [Medline: 27310031]
44. Kazemi DM, Borsari B, Levine MJ, Li S, Lamberson KA, Matta LA. A systematic review of the mHealth interventions to prevent alcohol and substance abuse. J Health Commun 2017 May;22(5):413-432. [doi: 10.1080/10810730.2017.1303556] [Medline: 28394729]
45. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. 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]
46. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of web-based and mobile health interventions. J Med Intern Res 2011;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
47. Pluye P, Robert E, Cargo M, Bartlett G, O'Cathain A, Griffiths F, et al. Department of Family Medicine, McGill University. 2011. Proposal: A mixed methods appraisal tool for systematic mixed studies URL: http://mixedmethodsappraisaltoolpublic.pbworks.com/w/file/fetch/84371689/MMAT%202011%20criteria%20and%20tutorial%202011-06-29updated2014.08.21.pdf [accessed 2018-04-25] [WebCite Cache ID 6yxBYcfF4]
48. Rycroft-Malone J, McCormack B, Hutchinson AM, DeCorby K, Bucknall TK, Kent B, et al. Realist synthesis: illustrating the method for implementation research. Implement Sci 2012 Apr 19;7:33 [FREE Full text] [doi: 10.1186/1748-5908-7-33] [Medline: 22515663]
49. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
50. Suffoletto B, Kristan J, Chung T, Jeong K, Fabio A, Monti P, et al. An interactive text message intervention to reduce binge drinking in young adults: a randomized controlled trial with 9-month outcomes. PLoS One 2015;10(11):e0142877 [FREE Full text] [doi: 10.1371/journal.pone.0142877] [Medline: 26580802]
51. Suffoletto B, Kristan J, Callaway C, Kim KH, Chung T, Monti PM, et al. A text message alcohol intervention for young adult emergency department patients: a randomized clinical trial. Ann Emerg Med 2014 Dec;64(6):664-72.e4 [FREE Full text] [doi: 10.1016/j.annemergmed.2014.06.010] [Medline: 25017822]
52. Agyapong VI, McLoughlin DM, Farren CK. Six-months outcomes of a randomised trial of supportive text messaging for depression and comorbid alcohol use disorder. J Affect Disord 2013 Oct;151(1):100-104. [doi: 10.1016/j.jad.2013.05.058] [Medline: 23800443]
53. Agyapong VI, Ahern S, McLoughlin DM, Farren CK. Supportive text messaging for depression and comorbid alcohol use disorder: single-blind randomised trial. J Affect Disord 2012 Dec 10;141(2-3):168-176. [doi: 10.1016/j.jad.2012.02.040] [Medline: 22464008]
54. Gajecki M, Andersson C, Rosendahl I, Sinadinovic K, Fredriksson M, Berman AH. Skills training via smartphone app for university students with excessive alcohol consumption: a randomized controlled trial. Int J Behav Med 2017 Dec;24(5):778-788 [FREE Full text] [doi: 10.1007/s12529-016-9629-9] [Medline: 28224445]

55. Aharonovich E, Stohl M, Cannizzaro D, Hasin D. HealthCall delivered via smartphone to reduce co-occurring drug and alcohol use in HIV-infected adults: a randomized pilot trial. J Subst Abuse Treat 2017 Dec;83:15-26 [FREE Full text] [doi: 10.1016/j.jsat.2017.09.013] [Medline: 29129192]
56. Haug S, Lucht MJ, John U, Meyer C, Schaub MP. A pilot study on the feasibility and acceptability of a text message-based aftercare treatment programme among alcohol outpatients. Alcohol Alcohol 2015 Mar;50(2):188-194. [doi: 10.1093/alcalc/agu107] [Medline: 25600249]
57. Witkiewitz K, Desai SA, Bowen S, Leigh BC, Kirouac M, Larimer ME. Development and evaluation of a mobile intervention for heavy drinking and smoking among college students. Psychol Addict Behav 2014 Sep;28(3):639-650. [doi: 10.1037/a0034747] [Medline: 25000269]
58. Lucht MJ, Hoffman L, Haug S, Meyer C, Pussehl D, Quellmalz A, et al. A surveillance tool using mobile phone short message service to reduce alcohol consumption among alcohol-dependent patients. Alcohol Clin Exp Res 2014 Jun;38(6):1728-1736. [doi: 10.1111/acer.12403] [Medline: 24730528]
59. Alessi SM, Petry NM. A randomized study of cellphone technology to reinforce alcohol abstinence in the natural environment. Addiction 2013 May;108(5):900-909 [FREE Full text] [doi: 10.1111/add.12093] [Medline: 23279560]
60. Riordan BC, Conner TS, Flett JA, Scarf D. A text message intervention to reduce first year university students' alcohol use: a pilot experimental study. Digit Health 2017;3:2055207617707627 [FREE Full text] [doi: 10.1177/2055207617707627] [Medline: 29942597]
61. Brendryen H, Lund IO, Johansen AB, Riksheim M, Nesvåg S, Duckert F. Balance--a pragmatic randomized controlled trial of an online intensive self-help alcohol intervention. Addiction 2014 Feb;109(2):218-226. [doi: 10.1111/add.12383] [Medline: 24134709]
62. Mason M, Benotsch EG, Way T, Kim H, Snipes D. Text messaging to increase readiness to change alcohol use in college students. J Prim Prev 2014 Feb;35(1):47-52. [doi: 10.1007/s10935-013-0329-9] [Medline: 24114551]
63. Riordan BC, Conner TS, Flett JA, Scarf D. A brief orientation week ecological momentary intervention to reduce university student alcohol consumption. J Stud Alcohol Drugs 2015 Jul;76(4):525-529. [doi: 10.15288/jsad.2015.76.525] [Medline: 26098027]
64. Fisher W, Fisher J, Harman J. The information-motivation-behavioral skills model: A general social psychological approach to understanding and promoting health behavior. In: Social Psychological Foundations of Health and Illness. Hoboken, New Jersey: Wiley-Blackwell; 2003:82-106.
65. Regmi K, Kassim N, Ahmad N, Tuah N. Effectiveness of mobile apps for smoking cessation: a review. Tob Prev Cessation 2017 Apr 12;3(April):12 [FREE Full text] [doi: 10.18332/tpc/70088]
66. Brouwer IG. The Widmark formula for alcohol quantification. S Afr Dent J 2004 Nov;59(10):427-428. [Medline: 15726763]
67. Festinger L. A Theory of Cognitive Dissonance. Redwood, CA: Stanford University Press; 1962.
68. Gurman TA, Rubin SE, Roess AA. Effectiveness of mHealth behavior change communication interventions in developing countries: a systematic review of the literature. J Health Commun 2012;17(Suppl 1):82-104. [doi: 10.1080/10810730.2011.649160] [Medline: 22548603]
69. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. J Med Intern Res 2014;16(6):e146 [FREE Full text] [doi: 10.2196/jmir.3077] [Medline: 24905070]
70. Blow FC, Ilgen MA, Walton MA, Czyz EK, McCammon R, Chermack ST, et al. Severity of baseline alcohol use as a moderator of brief interventions in the emergency department. Alcohol Alcohol 2009;44(5):486-490 [FREE Full text] [doi: 10.1093/alcalc/agp031] [Medline: 19692345]
71. Kazemi DM, Borsari B, Levine MJ, Lamberson KA, Dooley B. REMIT: Development of a mHealth theory-based intervention to decrease heavy episodic drinking among college students. Addict Res Theory 2017 Dec 26;26(5):377-385 [FREE Full text] [doi: 10.1080/16066359.2017.1420783]
72. Hirschi T, Gottfredson M. Commentary: testing the general theory of crime. J Res Crime Delinq 2016 Aug 18;30(1):47-54. [doi: 10.1177/0022427893030001004]
73. Perkins HW. Social norms and the prevention of alcohol misuse in collegiate contexts. J Stud Alcohol Suppl 2002 Mar(14):164-172. [doi: 10.15288/jsas.2002.s14.164] [Medline: 12022722]
74. Costa PT, Terracciano A, McCrae RR. Gender differences in personality traits across cultures: robust and surprising findings. J Pers Soc Psychol 2001 Aug;81(2):322-331. [doi: 10.1037/0022-3514.81.2.322] [Medline: 11519935]
75. Feingold A. Gender differences in personality: a meta-analysis. Psychol Bull 1994 Nov;116(3):429-456. [Medline: 7809307]
76. Platt L, Melendez-Torres GJ, O'Donnell A, Bradley J, Newbury-Birch D, Kaner E, et al. How effective are brief interventions in reducing alcohol consumption: do the setting, practitioner group and content matter? Findings from a systematic review and metaregression analysis. BMJ Open 2016 Dec 11;6(8):e011473 [FREE Full text] [doi: 10.1136/bmjopen-2016-011473] [Medline: 27515753]
77. Chan SR, Torous J, Hinton L, Yellowlees P. Mobile tele-mental health: increasing applications and a move to hybrid models of care. Healthcare (Basel) 2014 May 06;2(2):220-233 [FREE Full text] [doi: 10.3390/healthcare2020220] [Medline: 27429272]

78. Kirk GD, Himelhoch SS, Westergaard RP, Beckwith CG. Using mobile health technology to improve hiv care for persons living with HIV and substance abuse. AIDS Res Treat 2013;2013:194613 [FREE Full text] [doi: 10.1155/2013/194613] [Medline: 24381751]
79. Haug S, Paz Castro R, Kowatsch T, Filler A, Dey M, Schaub MP. Efficacy of a web- and text messaging-based intervention to reduce problem drinking in adolescents: results of a cluster-randomized controlled trial. J Consult Clin Psychol 2017 Dec;85(2):147-159. [doi: 10.1037/ccp0000138] [Medline: 27606700]
80. Paige SR, Alber JM, Stellefson ML, Krieger JL. Missing the mark for patient engagement: mHealth literacy strategies and behavior change processes in smoking cessation apps. Patient Educ Couns 2018 May;101(5):951-955. [doi: 10.1016/j.pec.2017.11.006] [Medline: 29153592]
81. Goodwin N, Nelson JA, Ackerman F, Weisskopf T. Global Development And Environment Institute, Tufts University. 2008. Consumption and the Consumer Society URL: http://www.ase.tufts.edu/gdae/education_materials/modules/Consumption_and_the_Consumer_Society.pdf [accessed 2018-04-25] [WebCite Cache ID 6yxCU96D8]
82. Garnett C, Crane D, Brown J, Kaner E, Beyer F, Muirhead C, et al. Reported theory use by digital interventions for hazardous and harmful alcohol consumption, and association with effectiveness: meta-regression. J Med Intern Res 2018 Feb 28;20(2):e69 [FREE Full text] [doi: 10.2196/jmir.8807] [Medline: 29490895]
83. Meredith SE, Alessi SM, Petry NM. Smartphone applications to reduce alcohol consumption and help patients with alcohol use disorder: a state-of-the-art review. Adv Health Care Technol 2015;1:47-54 [FREE Full text] [doi: 10.2147/AHCT.S65791] [Medline: 27478863]
84. Bratman M. Intention, Plans, and Practical Reason. Stanford, CA: Center for the Study of Language and Information; 1999.
85. West J, Mace M. Browsing as the killer app: explaining the rapid success of Apple's iPhone. Telecomm Policy 2010 Jun;34(5-6):270-286. [doi: 10.1016/j.telpol.2009.12.002]

Abbreviations

A-CHESS: Alcohol-Comprehensive Health Enhancement Support System
AUD: alcohol use disorder
AUDIT: alcohol use disorders identification test
AUDIT-C: alcohol use disorders identification test for consumption
BASICS-Mobile: Brief Alcohol and Smoking Intervention for College Students via Mobile
BrAC: breath alcohol concentration
eBAC: estimated blood alcohol concentration
FAST: fast alcohol screening test
IVR: interactive voice response
mHealth: mobile health
MMAT: Mixed Methods Appraisal Tool
RSOD: risky single-occasion drinking
SD: standard drink
SMS: short message services
UAU: unhealthy alcohol use

Edited by G Eysenbach; submitted 29.04.18; peer-reviewed by T Conner, M Bendtsen; comments to author 07.08.18; revised version received 02.10.18; accepted 30.10.18; published 29.01.19.

Please cite as:
Song T, Qian S, Yu P
Mobile Health Interventions for Self-Control of Unhealthy Alcohol Use: Systematic Review
JMIR Mhealth Uhealth 2019;7(1):e10899
URL: http://mhealth.jmir.org/2019/1/e10899/
doi:10.2196/10899
PMID:30694200

Review

Effects of Social Media and Mobile Health Apps on Pregnancy Care: Meta-Analysis

Ko Ling Chan[1], PhD; Mengtong Chen[1], PhD

Department of Applied Social Sciences, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

Corresponding Author:
Mengtong Chen, PhD
Department of Applied Social Sciences
The Hong Kong Polytechnic University
Hung Hom
Kowloon
Hong Kong,
China (Hong Kong)
Phone: 852 5425 9462
Email: jenna.mt.chen@polyu.edu.hk

Abstract

Background: The use of social media and mobile health (mHealth) apps has been increasing in pregnancy care. However, the effectiveness of these interventions is still unclear.

Objectives: We conducted a meta-analysis to examine the effectiveness of these interventions with regard to different health outcomes in pregnant and postpartum women and investigate the characteristics and components of interventions that may affect program effectiveness.

Method: We performed a comprehensive literature search of major electronic databases and reference sections of related reviews and eligible studies. A random effects model was used to calculate the effect size.

Results: Fifteen randomized controlled trial studies published in and before June 2018 that met the inclusion criteria were included in the meta-analysis. The interventions were effective in promoting maternal physical health including weight management, gestational diabetes mellitus control, and asthma control with a moderate to large effect size (d=0.72). Large effect sizes were also found for improving maternal mental health (d=0.84) and knowledge about pregnancy (d=0.80). Weight control interventions using wearable devices were more effective.

Conclusion: Social media and mHealth apps have the potential to be widely used in improving maternal well-being. More large-scale clinical trials focusing on different health outcomes are suggested for future studies.

(JMIR Mhealth Uhealth 2019;7(1):e11836) doi:10.2196/11836

KEYWORDS

mHealth; social media; pregnancy; postpartum; maternal health

Introduction

Every pregnancy is unique but carries risks of a number of physical and psychological problems. Low maternal well-being during pregnancy can negatively impact women's health outcomes and child development [1]. For example, overweight and obesity have become a common health problem associated with pregnancy in both developed and developing countries with dramatically increased prevalence over the past two decades [2,3]. Overweight and obesity before, during, and after pregnancy increase the risk of diseases such as metabolic syndrome, cardiovascular disease, and diabetes, as well as a number of child developmental problems such as preterm birth, low birth weight, neurodevelopmental delay, and immune and infectious disease, further increasing medical costs and negatively influencing family well-being [3-5]. Approximately 7.5% of pregnant women suffer from gestational diabetes mellitus (GDM), and the prevalence is significantly higher among Asian and Pacific Islanders [6]. Pregnant women with GDM, in particular those having obesity and overweight problems, are at significantly higher risk of adverse pregnancy outcomes [7]. Depression among women during and after pregnancy can also have negative effects on maternal health and interpersonal functioning, which is a common and persistent mental health problem [8,9].

Effective interventions that can help reduce risks during pregnancy and improve maternal well-being therefore play an important role. Research has shown that in addition to regular check-ups, several other forms of pregnancy care provided by medical professionals, therapists, and social workers are useful means to improve maternal well-being during pregnancy, such as yoga and physical activity, lifestyle, mindfulness, and psychotherapeutic interventions [10-14]. However, these traditional health services are often restricted by time and place, as working parents may not be able to attend during the daytime. Women from disadvantaged groups often have limited resources, which prevent their access to health services [15,16]. In addition, these women were found to have poor treatment adherence and high attrition, which resulted in nonsignificant changes after the services [17]. From a service provider point of view, traditional services for pregnancy care often involve a number of health professionals providing face-to-face treatment, which is quite expensive and cannot reach different populations [18].

In recent years, mobile technologies have been widely used in the provision of pregnancy care services, benefiting from the rapid development of information communication technology (ICT) and universal access to these technologies [16]. More social media and mobile health (mHealth) apps are being used today, taking the place of traditional text message or email services. Social media websites provide women with a platform for obtaining health information and interactions with health professionals and peers [19]. Because of the increasing ownership rate of mobile phones, a large number of mHealth apps on health topics have been developed and are installed by consumers [18]. In addition to quick and easy access to health information, mHealth apps can improve interactions with the health care system—for example, consumers can monitor their health conditions by recording or uploading health status data using the apps [20]. Many apps can also promote health behaviors such as maintaining sufficient physical activity and having a healthy diet [21].

Pregnant and postpartum women are increasingly relying on social media and mHealth apps as sources of health information and services for self-care and infant care [22,23]. Systematic reviews show that the use of mHealth apps and social media is feasible and acceptable to support pregnancy care, including promoting a healthy lifestyle and providing health information in high-income countries [16,24]. However, the effectiveness of the interventions using mHealth apps and social media is still unclear, and the ways that diverse intervention components contribute to program effectiveness is also unclear.

We conducted a meta-analysis to examine the effectiveness of mHealth apps and social media interventions for pregnant and postpartum women by calculating the effect size and examining the characteristics of these interventions that may be related to program effectiveness.

Methods

Search Method

Study procedures followed Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. We conducted a comprehensive literature search in online databases including PsycInfo, PsycARTICLES, Sociological Abstracts, Social Services Abstracts, Medline, ERIC, CINAHL Complete, and PubMed. We searched relevant studies published in and before June 2018. Advanced searches in titles, keywords, and abstracts were performed using the combinations of three groups of terms: (1) mobile technology and social media, including smartphone, mobile phone, social networking, Facebook, Twitter, WhatsApp, WeChat, and virtual reality; (2) pregnancy status, including pregnancy, pregnant, gestation, postnatal, and postpartum; (3) pregnancy care, including intervention, program, treatment, prevention, education, and therapy. In addition to the electronic database search, we hand-searched the reference lists of retrieved studies and relevant reviews, as well as grey literature including conference abstracts and dissertations.

Inclusion and Exclusion Criteria

The literature search aimed at identifying original evaluation studies on mHealth apps and social media interventions for pregnancy care. Eligible studies should (1) focus on interventions providing pregnancy care, including prenatal and postpartum health care for expectant mothers; (2) use advanced technology, such as mobile apps for social media or health care; (3) aim at health outcomes such healthy pregnancy and maternal well-being; (4) use experimental or quasi-experimental design; (5) report enough data to calculate the effect size; and (6) be published in English or Chinese.

Studies were not included if they (1) examined the use of mobile technology among health workers, (2) used a qualitative evaluation method only, or (3) used a traditional method such as providing short message services (texts) or sending emails.

Data Extraction

In the first step, we designed a standardized form to code study characteristics. The study publication information (author, contacts, publication year, and country), methodological characteristics (study design, sample size, and the use of clinical sample or community sample), intervention details (aim, content, device, mHealth or social media apps, duration, attrition rate, and service provider), and participant profiles (age, pregnancy status, health status, and socioeconomic status) were recorded using this form. In the second step, we coded study outcomes and extracted data (eg, mean, standard deviation, *P* values, sample sizes) for effect size calculation. The outcomes include health outcomes of pregnant or postpartum women such as pregnancy weight control, asthma control, health knowledge, and stress and depression management. Two authors performed data extraction separately, and disagreements were resolved by consensus.

Quality Assessment

To obtain a valid estimate of intervention effectiveness and reduce the risk of bias in the meta-analysis, we used a checklist to assess the methodological quality of the included studies. The checklist (see Multimedia Appendix 1) is composed of eight items, measuring study design, participant eligibility criteria, sample size calculation, randomization process, intervention details, participant profiles, primary outcomes, and

statistical methods. Each item is allocated 1 point; therefore, the highest score is 8 if all criteria are met, and scores of 5 and above are regarded as satisfactory. Two researchers in the research team evaluated the studies independently. To measure rater agreement, the Cohen kappa coefficient was used. The level of agreement was high between the two raters. Disagreements were resolved by discussion with the first author through a consensus-building process.

Statistical Analysis

First, we provided a complete summary of the studies by tabulating publication information, methodology, intervention, and participant characteristics. Second, we calculated the effect size of each study using the Cohen *d* statistic. The Cohen *d* was calculated using the formula seen in Figure 1, in which the difference between two means is divided by a standard deviation for the two groups [25]. If a study reported multiple outcomes, the mean effect size of these outcomes was used. If there were studies based on the same intervention program, they were merged into one study and we calculated the mean effect size for these studies. The overall pooled effect size of the included studies was calculated based on a random effects model because of the different features of the interventions. The *Q* statistic and I^2 were used to measure the variation in study outcomes between different studies. In addition, we used the *Q* statistic to test the effect of moderator variables, which may be related to program effectiveness. To examine whether publication bias occurred in our meta-analysis, we constructed a funnel plot. A symmetric inverted funnel plot indicates an absence of publication bias with high probability. Statistical analyses were performed using the Comprehensive Meta-Analysis (CMA) 3.0 program (Biostat Inc).

Figure 1. Formula for calculating the Cohen *d* statistic.

$$d = \frac{\overline{X1} - \overline{X2}}{Swithin}$$

Results

Study Characteristics

Study Selection Process

Figure 2 shows the results of the literature search and the study selection process. The literature search yielded 577 citations after removing duplicate records. A total of 149 articles were excluded because they were not published in English or Chinese or were focused on irrelevant topics. Then, on the basis of title, abstract, and full-text screening, 412 research articles were excluded based on the inclusion and exclusion criteria. Finally, 16 articles were found to be suitable for inclusion. Because two research articles [26,27] were based on one intervention program, a total of 15 studies were synthesized in the meta-analysis.

Methodological Characteristics

Multimedia Appendix 2 summarizes the publication information and methodological characteristics of each included study. All of the included studies were published in or after 2014, in line with the rapid development and spread of ICT in recent years. The studies were conducted in diverse countries and regions, including the United States, Australia, United Kingdom, Ireland, Israel, Indonesia, China, and Taiwan.

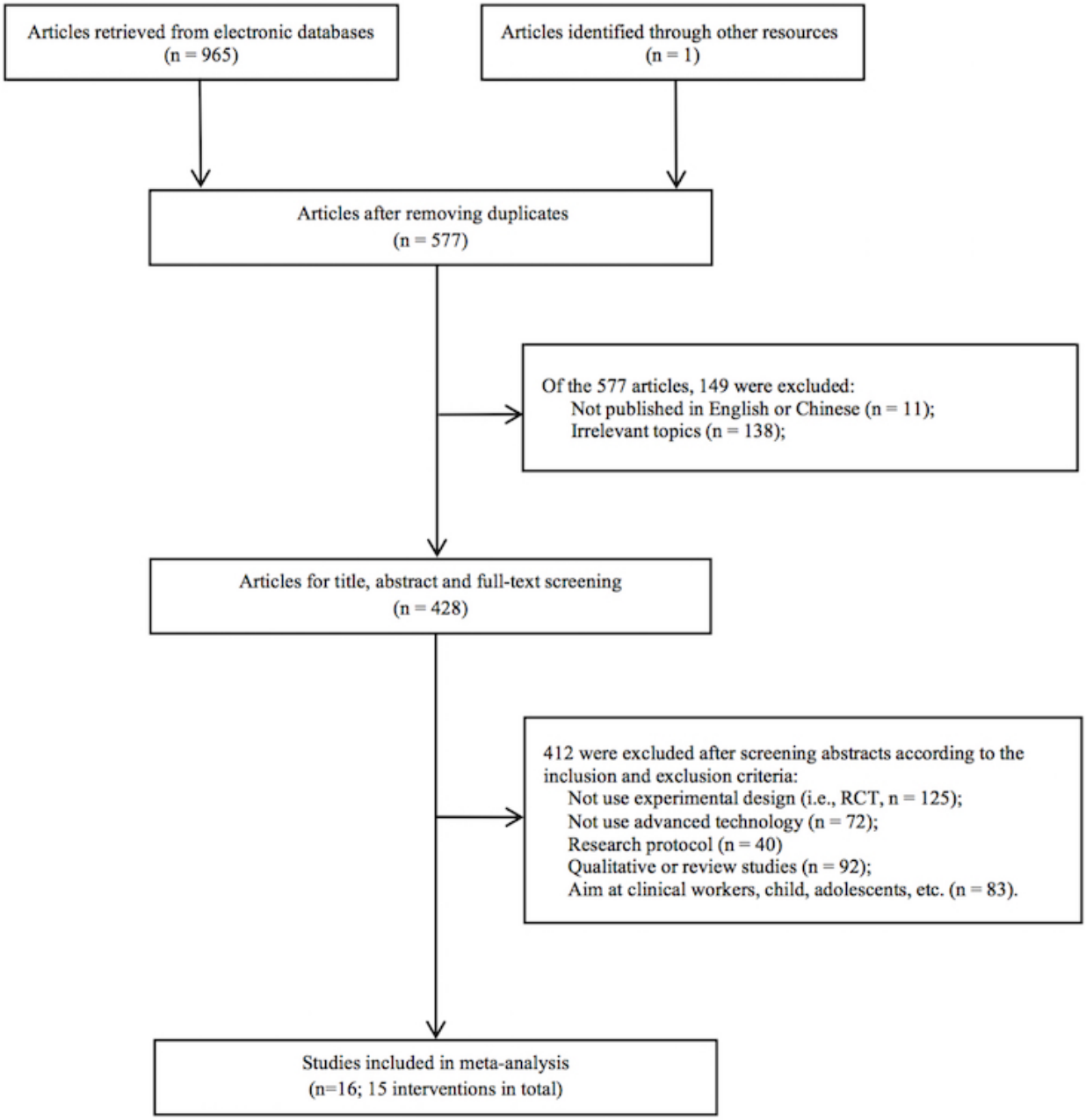
All 15 studies used randomized controlled trials (RCTs) to evaluate the effectiveness of these technology-based interventions. The sample size of the included studies ranged from 16 to 1689 participants, with a mean sample size of 225 and a median sample size of 87. All of the studies investigated a clinical sample of pregnant or postpartum women. With satisfactory methodological quality scores, all of the studies were graded at low risk of bias.

Intervention Components

Multimedia Appendix 3 summarizes the intervention characteristics and outcomes of each included study. Although all of the selected interventions aimed at improving maternal well-being, they involved different contents, and the approaches providing services differed to some extent.

Intervention in lifestyle was the major content in 14 selected studies. One approach was through psychoeducation. For example, pregnant and postpartum women obtained health information, identified risk behaviors and situations, learned to set achievable goals, and used behavior skills [28,29]. Thus, the participants were expected to manage their weight or control glucose by increasing their physical activity and changing dietary intake [27,30-32]. In addition, with knowledge about maternal and infant health, the participants increased their birth preparedness, complication readiness, and feeding behaviors [33,34]. Participants often accompanied the psychoeducation approach with self-monitoring to promote their lifestyle change, in which parents were required to maintain regular physical activities and pay attention to dietary intake. Patient monitoring devices such as handheld respiratory devices [35] and wearable devices such as Fitbit (Fitbit Inc) [30,31,36] were often used as support tools to track participants' physical activities and record health status.

Figure 2. Flow diagram of study selection.



In addition to lifestyle intervention, there was one social support program developed by Cheng et al [37] delivered via mobile phone to reduce postpartum perceived stress and depression. Participants received emotional support as well as information about maternal and infant care from professionals via the social media instant communication app Line.

Ten interventions included a socially interactive component. Social media platforms (eg, Facebook, Line, and WeChat) not only provided a forum for knowledge sharing and behavior skills training, but also enabled participants easy access to support from clinical professionals and peer groups. Interactive components also increased peer support and promoted participant engagement [17,27]. Although clinical professionals were not necessarily involved in service providing because of the nature of mHealth, professional consultations by health coaches, psychologists, dietitians, physicians and obstetric doctors were still provided in interventions with interactive components.

Pregnancy care was either provided via social media platforms or via mobile phone platforms such as mHealth apps. A reminder function was used in apps to encourage participants to use the service and change their health behaviors. The interventions were generally long term, which covered most of the long gestation period and/or postpartum period.

Participant Profiles

With the exception of the study by Santoso et al [34], which included pregnant couples, all of the interventions were designed exclusively for women. The female participants were aged between 24 and 34 years with diverse socioeconomic characteristics and ethnicities. Participants who were African American or Hispanic and received Medicaid were particularly

selected in the study by Herring et al [17,26,27]. Ten studies focused on overweight or obese women with a body mass index (BMI) above 25 kg/m^2. Multimedia Appendix 4 summarizes the demographic characteristics of the sample.

Intervention Effectiveness

Table 1 shows the effect size of each study pooled by different outcomes and time points. The results are displayed in a forest plot as shown in Figure 3. With the exception of the study by Olson et al [38], all of the interventions reported positive effects with an overall random effect size of 0.74 (P<.001, Q=146.45, I^2=90.44). The overall effect size is considered medium to large, according to the Cohen criteria for effect size interpretation [39]. However, the large Q statistic indicated the wide variance in the effect sizes of different interventions, and it was estimated that 90.44% of the variance was due to heterogeneity. One reason for the large amount of heterogeneity is that the included studies aimed at different health outcomes. When we examined the effect sizes of different outcomes, the Q value decreased, showing that the variation is smaller for different specific outcomes.

As shown in Table 2, twelve studies aimed to improve the physical health outcomes of pregnant or postpartum women, including weight management (d=0.45, P=.003), GDM control (d=0.41, P=.03), and asthma control (d=3.43, P<.001), with an overall random effect size of 0.72 (P<.001, Q=127.3, I^2=91.36). One study aimed to improve maternal mental health (eg, reducing postpartum stress and depression), and the effect size was 0.84 (P<.001). Two studies aimed to improve knowledge about birth preparedness and infant feeding, and the effect size was 0.8 (P=.04, Q=3.55, I^2=71.82).

Table 1. Effect size for each study pooled by outcomes and time points.

Study name	Effect size	Standard error	Lower limit	Upper limit	*P* value
Herring, SJ (2014)	0.80	0.49	–0.17	1.75	.11
Cheng, HY (2016)	0.84	0.19	0.48	1.21	<.001
Choi, J (2016)	0.48	0.37	–0.25	1.20	.20
Herring, SJ (2016)	0.45	0.28	–0.09	1.00	.10
Zairina, E (2016)	3.43	0.38	2.69	4.17	<.001
Fiks, AG (2017)	0.45	0.23	0.003	0.90	.048
Gilmore, LA (2017)	2.05	0.43	1.21	2.89	<.001
Redman, LM (2017)	0.63	0.34	–0.05	1.30	.07
Santoso, HY (2017)	1.25	0.36	0.55	1.94	<.001
Dodd, JM (2018)	0.10	0.20	–0.29	0.49	.62
Olson, CM (2018)	–0.002	0.06	–0.12	0.11	.97
Kennelly, MA (2018)	0.26	0.10	0.06	0.46	.01
Mackillop, L (2018)	0.04	0.14	–0.24	0.31	.78
Miremberg, H (2018)	0.94	0.19	0.56	1.32	<.001
Yang, P (2018)	0.57	0.20	0.18	0.96	.004
Total	0.74	0.16	0.43	1.04	<.001

Figure 3. Effect size for each study.

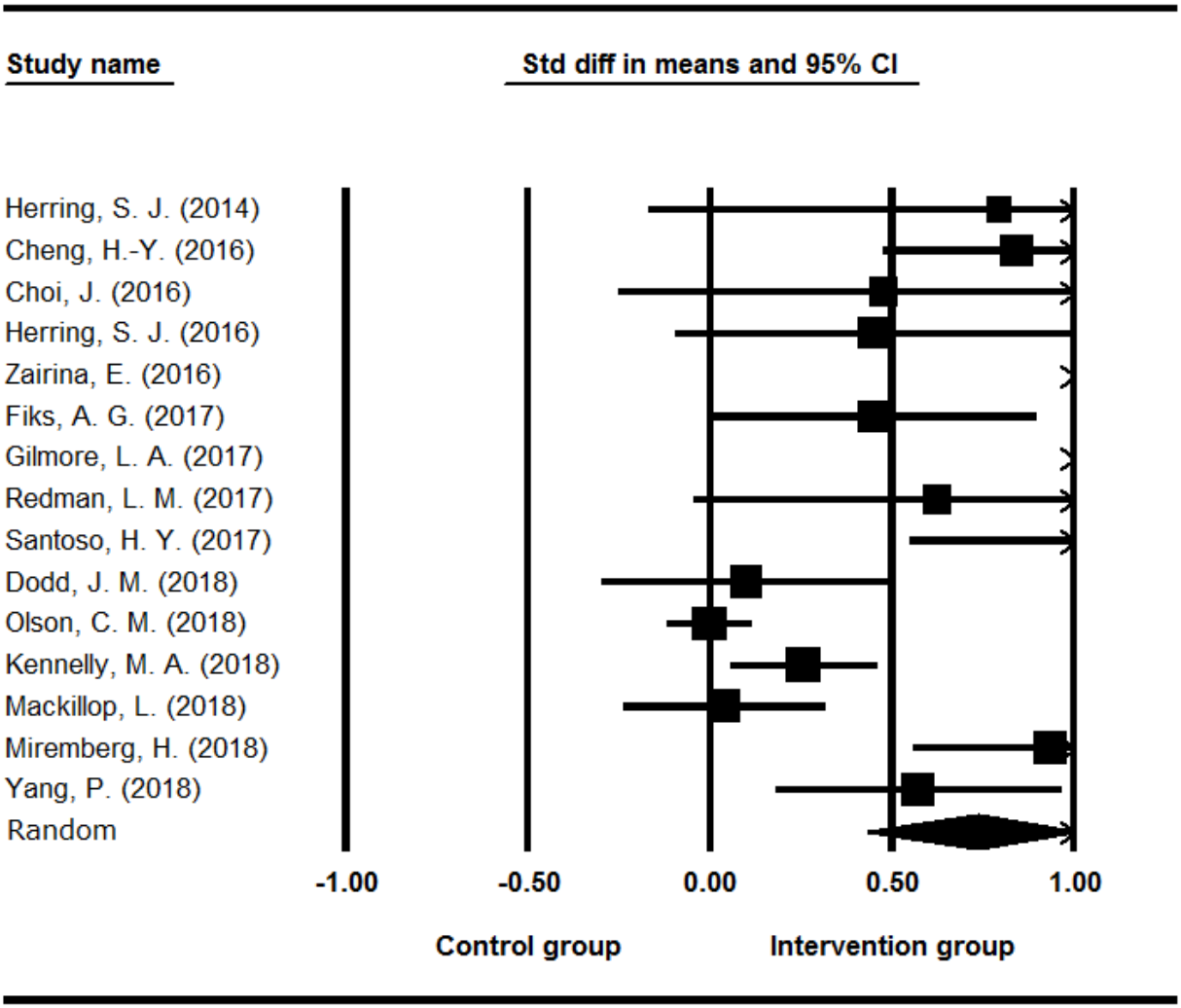


Table 2. Effect sizes of social media and mHealth apps for different health outcomes.

Outcome	*k* [b]	ES[a] and 95% CI				Test of null		Test of heterogeneity		
		d [c]	*SE*	*LL* [d]	*UL* [e]	*Z*	*P* value	*Q*	*P* value	I^2
Physical health	12	0.72	0.18	0.37	1.07	4.04	<.001	127.30	<.001	91.36
Weight management	8	0.45	0.15	0.16	0.74	3.01	.003	36.85	<.001	81.00
Gestational diabetes mellitus control	4	0.41	0.19	0.04	0.78	2.16	.03	17.38	<.001	82.74
Asthma control	1	3.43	0.38	2.69	4.17	9.06	<.001	0	>.99	0
Stress and postnatal depression	1	0.84	0.19	0.47	1.21	4.53	<.001	0	>.99	0
Birth preparedness knowledge	2	0.80	0.40	0.03	1.57	2.03	.04	3.55	.06	71.82
Total	15	0.74	0.16	0.43	1.04	4.72	<.001	146.45	<.001	90.44

[a]ES: effect size

[b]*k*: number of studies.

[c]*d*: effect size.

[d]LL: lower limit.

[e]UL: upper limit.

Subgroup Analyses

We investigated factors that may moderate the program effectiveness, including whether the intervention included interactive treatment content, use of professional consultation or not, the type of technology used, use of a wearable device or not, and whether the participants were overweight or obese. As shown in Table 3, the use of a wearable device to track physical activity in interventions aiming at weight management was a significant moderator (Q_b=5.91, P=.02). Using such a device resulted in a larger pooled effect size (d=0.97). Sample size was also a significant moderator of the effect size (Q_b=7.38, P=.007). Studies with smaller sample sizes resulted in a larger pooled effect size (d=1.13).

Moderator analyses in Table 3 show that interactive content and professional consultation were not significant moderators (Q_b=1.5, P=.22). The effects of interventions providing interactive treatment content and involving professional consultations were not significantly better compared with the interventions without these components. The interventions using social networking (d=0.67) and health and fitness mobile phone apps (d=0.77) were also similarly effective as the result of the Q test was insignificant.

Table 3. Moderator variable analyses.

Moderator group	k [a]	d [b]	LL [c]	UL [d]	Q_b [e]	P value
Interactive content						
Yes	10	0.60	0.18	1.03	1.51	.22
No	5	1.08	0.46	1.70		
Professional consultation						
Yes	10	0.60	0.18	1.03	1.51	.22
No	5	1.08	0.46	1.70		
Technology						
Social networking	6	0.67	0.21	1.14	0.10	.75
Health and fitness mobile phone app	9	0.77	0.39	1.15		
Wearable device[f]						
Yes	3	0.97	0.45	1.49	5.91	.02
No	5	0.24	–0.05	0.52		
Sample size						
Above 100	7	0.38	0.02	0.74	7.38	.007
Below 100	8	1.13	0.73	1.54		

[a]k: number of studies.

[b]d: effect size.

[c]LL: lower limit.

[d]UL: upper limit.

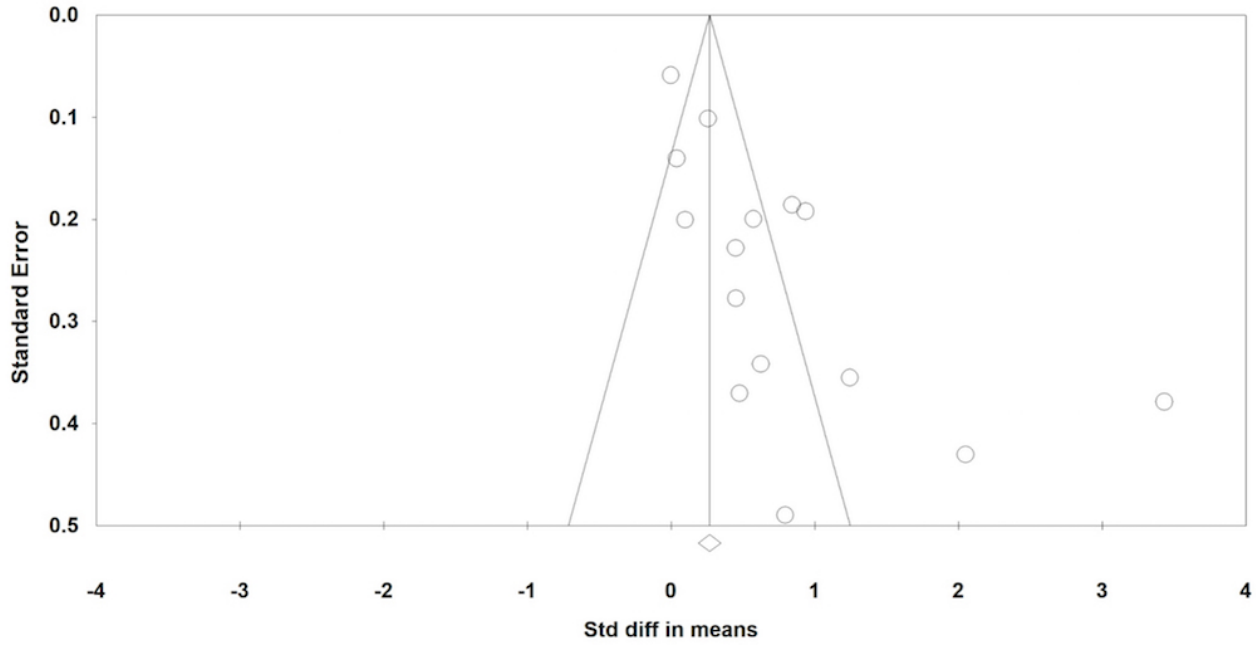
[e]Q_b: between-group heterogeneity

[f]Moderator effect in weight management.

Publication Bias Analysis

A funnel plot was used to examine publication bias of our meta-analysis, as shown in Figure 4. Almost half of the studies used a larger sample (n>100), and they concentrated around the top of the funnel plot. However, there were more studies on the right side of the mean effect size, especially studies with smaller samples, which made the funnel plot asymmetric. This means that positive results of the interventions based on small sample sizes were more likely to be published. Therefore, there is evidence that publication bias exists.

Figure 4. Funnel plot for publication bias.



Discussion

Principal Findings

The use of social media and mHealth apps has been increasing in pregnancy care because of the low cost and their easy access regardless of time and geographic location [40]. This meta-analysis synthesized the findings of 15 RCT studies conducted in different countries and regions and provides evidence on the effectiveness of these technology-based interventions in providing health care services to pregnant and postpartum women. Moderate to large effect sizes were found in regard to different health outcomes including maternal physical health, mental health, and knowledge about pregnancy. In addition, we investigated the characteristics and components of the interventions that may affect program effectiveness.

Effects on Maternal Physical Health

Pregnancy is a life-changing experience with risks of excessive weight gain and obesity [36]. Postpartum weight retention is a prevalent problem among US women, especially within racial and ethnic minorities [17]. However, it can be difficult for pregnant or postpartum women to manage their weight by increasing physical activity or changing dietary intake because of limited resources in health care or a low level of engagement in health management [16,41]. This meta-analysis finds that there was a moderate effect in maternal weight control and maintaining optimal body composition by promoting lifestyle change and self-monitoring via mHealth apps and social media.

Similarly to pregnancy weight control, GDM control also relies on self-monitoring the change of unhealthy lifestyle and listening to clinical decisions, and patient compliance is particularly important [42,43]. The results of the meta-analysis show that mHealth apps and social media were also effective for pregnant women with GDM, with a small to moderate effect size. As participants in the intervention group only had half of the clinic visits compared with participants in the standard care group in the research of Mackillop et al [42], we estimate that the intervention can be more effective with the same number of clinical visits.

An mHealth app was found very effective in asthma control during pregnancy [35]. However, this study is just an initial step toward understanding the effect of the social media and mHealth apps in physical health outcomes other than weight or GDM control. The result of a case-control study showed that the mHealth app can be effective in urinary incontinence management during pregnancy [44]. We can estimate that mHealth apps can be applied in other health services. To conclude, the positive and significant effects demonstrate that lifestyle intervention using advanced technology can be effective in improving maternal health.

Effects on Maternal Mental Health and Birth Preparedness

Psychological interventions delivered via mobile phones have been found effective in reducing depression and anxiety in existing meta-analyses [45,46]. However, whether the interventions are effective among pregnant or postpartum women is unclear. This study provided additional evidence that mHealth apps and social media can be useful in reducing pregnancy-related stress and depression. However, there was only one RCT study examining the effectiveness in maternal mental health outcomes [37].

Social media and mobile phone apps are becoming increasingly popular among pregnant women and their partners to access health knowledge and learn to identify risk behaviors and danger signs during pregnancy. The findings from two RCT studies [33,34] demonstrate the usefulness of the intervention programs to prepare the participants to become mothers. A pretest-posttest study found that providing information about maternal and infant care via the mHealth app can reduce maternal stress during pregnancy [47]. Therefore, with the improvement in health knowledge, maternal mental health may also improve.

Factors Related to Intervention Effectiveness

Finally, we were interested in investigating whether intervention characteristics or components can affect their effectiveness. Interventions using social media and interventions using mobile phone apps resulted in similar effect sizes. In addition, it seems to make little difference whether interactive treatment content or professional consultation was provided in the intervention. One explanation may be that all apps have a reminder function and provide health information, which is similar to some functions of the interactive treatment and professional consultation. Another possible explanation may be that the usefulness of the interventions is more likely to rely on women's self-monitoring. Therefore, different forms of mHealth apps and social media providing pregnancy care may have similar benefits. Interventions without interactive content or professional consultation can be more cost effective.

The results of moderator analyses also showed that using wearable devices to track participants' physical activities has the potential to enhance program effectiveness in weight control during the prenatal or postnatal periods. The use of wearable devices may be a good way to improve self-monitoring. Another moderator variable that significantly contributed to the variance in the effect size was sample size. Interventions with smaller sample sizes seem to be more effective, whereas interventions with larger samples were less effective. Olson et al [38] and Dodd et al [28] argued that in their study, the similar contents of intervention and control groups or the low use of self-monitoring tools in the intervention may explain the low program effectiveness. However, the absence of small studies with small effect sizes also indicates publication bias among the studies in this meta-analysis.

Strengths and Limitations

Because there was a lack of quantitative integration of the evidence on effectiveness of the social media and mHealth apps, further investigation was recommended before the implementation of the intervention [24]. This study includes rigorous studies that offer high-quality evidence. Our review is the first meta-analysis evaluating program effectiveness through credible statistical analyses.

There are several limitations to this study. First, the presence of publication bias indicates that more studies need to be included, as the studies included in the meta-analysis were more likely to report larger effect sizes. Second, as most studies included provided only limited details about participant profiles, several factors cannot be examined through moderator analysis such as participant socioeconomic status and health status.

Implications for Research

Although the use of advanced technology in pregnancy care has been increasing in recent years and there are promising results in improving maternal health outcomes, research in this area is still in its early stages. First, more large-scale clinical trials are suggested in future studies. This is because interventions with smaller sample sizes are more likely to report larger effect sizes, which can result in skewed distribution of effect sizes in the meta-analysis. Also, more studies are suggested to be included in future reviews. Second, because interventions included in this review were used predominantly for managing health problems, the effectiveness in improving mental health of pregnant and postpartum women needs to be examined. Third, cost effectiveness could be an important feature of the use of mHealth apps and social media in pregnancy care [42]; however, it was not examined in most studies. Therefore, cost analysis is necessary in future studies.

Implications for Practice

This meta-analysis of the effectiveness of social media and mHealth apps has several implications for future practice. First, interventions with the use of social media and mHealth apps can be effective in promoting maternal well-being. The positive effects in developing countries such as Indonesia and China imply that the use of mobile technologies in pregnancy care can be less restricted by social and economic development. Social media and mHealth apps can be widely adopted in different areas and have greater public health impact.

Second, the study of Santoso et al [34] demonstrated that fathers can also be positively involved in pregnancy care and birth preparedness by using social media and health apps. The inclusion of fathers could improve health outcomes for the whole family [1]. Future practice should consider attracting fathers to use the related services.

Third, the use of mHealth apps was poor among participants in some interventions, which may lead to low effectiveness [28,38]. Therefore, it is important for researchers, service providers, and app developers to consider how to increase the use of interventions and customer stickiness. It is also necessary to find useful ways to improve participant self-monitoring.

Fourth, even though this review included a number of mHealth apps, most commonly used mHealth apps are commercial and the credibility of their information is unknown [23]. Therefore, it is necessary to examine the quality and effectiveness of their services. Evidence-based mHealth apps and social media interventions for pregnant women are recommended in the practice.

Conclusion

Social media and mHealth apps are increasingly used in pregnancy care with emerging promising findings. In this meta-analysis, we found the interventions were useful with moderate to large effect sizes in regard to maternal health, mental health, and knowledge about pregnancy. We conclude that social media and mHealth apps have the potential to be widely used in improving maternal well-being during the prenatal and postnatal periods. More large-scale clinical trials with comprehensive aims are suggested for future studies.

Acknowledgments

The work described in this paper was supported by the Health Medical Research Fund, the Government of the Hong Kong Special Administrative Region (project number 15161101).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Methodological quality assessment checklist.

[PDF File (Adobe PDF File), 21KB - mhealth_v7i1e11836_app1.pdf]

Multimedia Appendix 2

Publication information and methodological characteristics.

[PDF File (Adobe PDF File), 47KB - mhealth_v7i1e11836_app2.pdf]

Multimedia Appendix 3

Intervention characteristics and study outcomes.

[PDF File (Adobe PDF File), 49KB - mhealth_v7i1e11836_app3.pdf]

Multimedia Appendix 4

Characteristics of female participants.

[PDF File (Adobe PDF File), 51KB - mhealth_v7i1e11836_app4.pdf]

References

1. Stein A, Pearson RM, Goodman SH, Rapa E, Rahman A, McCallum M, et al. Effects of perinatal mental disorders on the fetus and child. Lancet 2014 Nov 15;384(9956):1800-1819. [doi: 10.1016/S0140-6736(14)61277-0] [Medline: 25455250]
2. Amugsi DA, Dimbuene ZT, Mberu B, Muthuri S, Ezeh AC. Prevalence and time trends in overweight and obesity among urban women: an analysis of demographic and health surveys data from 24 African countries, 19912014. BMJ Open 2017 Oct 27;7(10):e017344 [FREE Full text] [doi: 10.1136/bmjopen-2017-017344] [Medline: 29079606]
3. McDonald SD, Han Z, Mulla S, Beyene J, Knowledge Synthesis Group. Overweight and obesity in mothers and risk of preterm birth and low birth weight infants: systematic review and meta-analyses. BMJ 2010 Jul 20;341:c3428 [FREE Full text] [Medline: 20647282]
4. Arabin B, Stupin JH. Overweight and obesity before, during and after pregnancy: Part 2: evidence-based risk factors and interventions. Geburtshilfe Frauenheilkd 2014 Jul;74(7):646-655 [FREE Full text] [doi: 10.1055/s-0034-1368462] [Medline: 25100879]
5. Godfrey KM, Reynolds RM, Prescott SL, Nyirenda M, Jaddoe VWV, Eriksson JG, et al. Influence of maternal obesity on the long-term health of offspring. Lancet Diabetes Endocrinol 2017 Jan;5(1):53-64 [FREE Full text] [doi: 10.1016/S2213-8587(16)30107-3] [Medline: 27743978]
6. Lawrence JM, Contreras R, Chen W, Sacks DA. Trends in the prevalence of preexisting diabetes and gestational diabetes mellitus among a racially/ethnically diverse population of pregnant women, 1999-2005. Diabetes Care 2008 May;31(5):899-904. [doi: 10.2337/dc07-2345] [Medline: 18223030]
7. Langer O, Yogev Y, Xenakis EMJ, Brustman L. Overweight and obese in gestational diabetes: the impact on pregnancy outcome. Am J Obstet Gynecol 2005 Jun;192(6):1768-1776. [doi: 10.1016/j.ajog.2004.12.049] [Medline: 15970805]
8. Gentile S. Untreated depression during pregnancy: short- and long-term effects in offspring. A systematic review. Neuroscience 2017 Feb 07;342:154-166. [doi: 10.1016/j.neuroscience.2015.09.001] [Medline: 26343292]
9. Underwood L, Waldie K, D'Souza S, Peterson ER, Morton S. A review of longitudinal studies on antenatal and postnatal depression. Arch Womens Ment Health 2016 Dec;19(5):711-720. [doi: 10.1007/s00737-016-0629-1] [Medline: 27085795]
10. Barth J, Munder T, Gerger H, Nüesch E, Trelle S, Znoj H, et al. Comparative efficacy of seven psychotherapeutic interventions for patients with depression: a network meta-analysis. PLoS Med 2013;10(5):e1001454 [FREE Full text] [doi: 10.1371/journal.pmed.1001454] [Medline: 23723742]
11. Davis K, Goodman SH, Leiferman J, Taylor M, Dimidjian S. A randomized controlled trial of yoga for pregnant women with symptoms of depression and anxiety. Complement Ther Clin Pract 2015 Aug;21(3):166-172. [doi: 10.1016/j.ctcp.2015.06.005] [Medline: 26256135]

12. Matvienko-Sikar K, Lee L, Murphy G, Murphy L. The effects of mindfulness interventions on prenatal well-being: a systematic review. Psychol Health 2016 Dec;31(12):1415-1434. [doi: 10.1080/08870446.2016.1220557] [Medline: 27539908]
13. Poyatos-León R, García-Hermoso A, Sanabria-Martínez G, Álvarez-Bueno C, Cavero-Redondo I, Martínez-Vizcaíno V. Effects of exercise-based interventions on postpartum depression: a meta-analysis of randomized controlled trials. Birth 2017 Sep;44(3):200-208. [doi: 10.1111/birt.12294] [Medline: 28589648]
14. Watson ED, Macaulay S, Lamont K, Gradidge PJ, Pretorius S, Crowther NJ, et al. The effect of lifestyle interventions on maternal body composition during pregnancy in developing countries: a systematic review. Cardiovasc J Afr 2017;28(6):397-403 [FREE Full text] [doi: 10.5830/CVJA-2017-003] [Medline: 28906530]
15. Bushar J, Pirretti A. Reaching at-risk families through mobile phones: successes from the Text4baby program. Zero to Three 2017 Mar;37(4):25-32.
16. Chen H, Chai Y, Dong L, Niu W, Zhang P. Effectiveness and appropriateness of mhealth interventions for maternal and child health: systematic review. JMIR Mhealth Uhealth 2018 Jan 09;6(1):e7 [FREE Full text] [doi: 10.2196/mhealth.8998] [Medline: 29317380]
17. Herring SJ, Cruice JF, Bennett GG, Davey A, Foster GD. Using technology to promote postpartum weight loss in urban, low-income mothers: a pilot randomized controlled trial. J Nutr Educ Behav 2014;46(6):610-615. [doi: 10.1016/j.jneb.2014.06.002] [Medline: 25069621]
18. Scott KM, Gome GA, Richards D, Caldwell PHY. How trustworthy are apps for maternal and child health? Health Technol 2015 Mar 12;4(4):329-336. [doi: 10.1007/s12553-015-0099-x]
19. Huesch MD, Galstyan A, Ong MK, Doctor JN. Using social media, online social networks, and internet search as platforms for public health interventions: a pilot study. Health Serv Res 2016 Jun;51 Suppl 2:1273-1290. [doi: 10.1111/1475-6773.12496] [Medline: 27161093]
20. Krebs P, Duncan DT. Health app use among US mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
21. Knight E, Stuckey MI, Prapavessis H, Petrella RJ. Public health guidelines for physical activity: is there an app for that? A review of android and apple app stores. JMIR Mhealth Uhealth 2015;3(2):e43 [FREE Full text] [doi: 10.2196/mhealth.4003] [Medline: 25998158]
22. Guerra-Reyes L, Christie VM, Prabhakar A, Harris AL, Siek KA. Postpartum health information seeking using mobile phones: experiences of low-income mothers. Matern Child Health J 2016 Dec;20(Suppl 1):13-21 [FREE Full text] [doi: 10.1007/s10995-016-2185-8] [Medline: 27639571]
23. Lee Y, Moon M. Utilization and content evaluation of mobile applications for pregnancy, birth, and child care. Healthc Inform Res 2016 Apr;22(2):73-80 [FREE Full text] [doi: 10.4258/hir.2016.22.2.73] [Medline: 27200216]
24. Overdijkink SB, Velu AV, Rosman AN, van Beukering MD, Kok M, Steegers-Theunissen RP. The usability and effectiveness of mobile health technology-based lifestyle and medical intervention apps supporting health care during pregnancy: systematic review. JMIR Mhealth Uhealth 2018 Apr 24;6(4):e109 [FREE Full text] [doi: 10.2196/mhealth.8834] [Medline: 29691216]
25. Borenstein M, Hedges L, Higgins J, Rothstein H. Introduction to Meta-Analysis. Chichester: John Wiley; 2009.
26. Herring SJ, Cruice JF, Bennett GG, Darden N, Wallen JJ, Rose MZ, et al. Intervening during and after pregnancy to prevent weight retention among African American women. Prev Med Rep 2017 Sep;7:119-123 [FREE Full text] [doi: 10.1016/j.pmedr.2017.05.015] [Medline: 28660118]
27. Herring SJ, Cruice JF, Bennett GG, Rose MZ, Davey A, Foster GD. Preventing excessive gestational weight gain among African American women: a randomized clinical trial. Obesity (Silver Spring) 2016 Jan;24(1):30-36 [FREE Full text] [doi: 10.1002/oby.21240] [Medline: 26592857]
28. Dodd JM, Louise J, Cramp C, Grivell RM, Moran LJ, Deussen AR. Evaluation of a smartphone nutrition and physical activity application to provide lifestyle advice to pregnant women: the SNAPP randomised trial. Matern Child Nutr 2018 Jan;14(1). [doi: 10.1111/mcn.12502] [Medline: 28836373]
29. Wilcox S, Liu J, Addy CL, Turner-McGrievy G, Burgis JT, Wingard E, et al. A randomized controlled trial to prevent excessive gestational weight gain and promote postpartum weight loss in overweight and obese women: Health In Pregnancy and Postpartum (HIPP). Contemp Clin Trials 2018 Mar;66:51-63. [doi: 10.1016/j.cct.2018.01.008] [Medline: 29371061]
30. Choi J, Lee JH, Vittinghoff E, Fukuoka Y. mHealth physical activity intervention: a randomized pilot study in physically inactive pregnant women. Matern Child Health J 2016 May;20(5):1091-1101. [doi: 10.1007/s10995-015-1895-7] [Medline: 26649879]
31. Gilmore LA, Klempel MC, Martin CK, Myers CA, Burton JH, Sutton EF, et al. Personalized mobile health intervention for health and weight loss in postpartum women receiving women, infants, and children benefit: a randomized controlled pilot study. J Womens Health (Larchmt) 2017 Jul;26(7):719-727. [doi: 10.1089/jwh.2016.5947] [Medline: 28338403]
32. Yang P, Lo W, He Z, Xiao X. Medical nutrition treatment of women with gestational diabetes mellitus by a telemedicine system based on smartphones. J Obstet Gynaecol Res 2018 Jul;44(7):1228-1234. [doi: 10.1111/jog.13669] [Medline: 29797375]

33. Fiks AG, Gruver RS, Bishop-Gilyard CT, Shults J, Virudachalam S, Suh AW, et al. A social media peer group for mothers to prevent obesity from infancy: the grow2gether randomized trial. Child Obes 2017 Oct;13(5):356-368. [doi: 10.1089/chi.2017.0042] [Medline: 28557558]
34. Santoso HYD, Supriyana S, Bahiyatun B, Widyawati MN, Fatmasari D, Sudiyono S, et al. Android application model of "suami siaga plus" as an innovation in birth preparedness and complication readiness (BP/CR) intervention. J Family Reprod Health 2017 Mar;11(1):30-36 [FREE Full text] [Medline: 29114266]
35. Zairina E, Abramson MJ, McDonald CF, Li J, Dharmasiri T, Stewart K, et al. Telehealth to improve asthma control in pregnancy: a randomized controlled trial. Respirology 2016 Jul;21(5):867-874. [doi: 10.1111/resp.12773] [Medline: 27037722]
36. Redman LM, Gilmore LA, Breaux J, Thomas DM, Elkind-Hirsch K, Stewart T, et al. Effectiveness of SmartMoms, a novel ehealth intervention for management of gestational weight gain: randomized controlled pilot trial. JMIR Mhealth Uhealth 2017 Sep 13;5(9):e133 [FREE Full text] [doi: 10.2196/mhealth.8228] [Medline: 28903892]
37. Cheng H, Huang T, Chien L, Cheng Y, Chen F. [The effects of a mobile application social support program on postpartum perceived stress and depression]. Hu Li Za Zhi 2016 Dec;63(6):52-60. [Medline: 27900745]
38. Olson CM, Groth SW, Graham ML, Reschke JE, Strawderman MS, Fernandez ID. The effectiveness of an online intervention in preventing excessive gestational weight gain: the e-moms roc randomized controlled trial. BMC Pregnancy Childbirth 2018 May 09;18(1):148 [FREE Full text] [doi: 10.1186/s12884-018-1767-4] [Medline: 29743026]
39. Cohen J. A power primer. Psychol Bull 1992 Jul;112(1):155-159. [Medline: 19565683]
40. Hall CS, Fottrell E, Wilkinson S, Byass P. Assessing the impact of mHealth interventions in low- and middle-income countries—what has been shown to work? Glob Health Action 2014;7:25606 [FREE Full text] [Medline: 25361730]
41. Kennelly MA, Ainscough K, Lindsay KL, O Sullivan E, Gibney ER, McCarthy M, et al. Pregnancy exercise and nutrition with smartphone application support: a randomized controlled trial. Obstet Gynecol 2018 May;131(5):818-826. [doi: 10.1097/AOG.0000000000002582] [Medline: 29630009]
42. Mackillop L, Hirst JE, Bartlett KJ, Birks JS, Clifton L, Farmer AJ, et al. Comparing the efficacy of a mobile phone-based blood glucose management system with standard clinic care in women with gestational diabetes: randomized controlled trial. JMIR Mhealth Uhealth 2018 Mar 20;6(3):e71 [FREE Full text] [doi: 10.2196/mhealth.9512] [Medline: 29559428]
43. Miremberg H, Ben-Ari T, Betzer T, Raphaeli H, Gasnier R, Barda G, et al. The impact of a daily smartphone-based feedback system among women with gestational diabetes on compliance, glycemic control, satisfaction, and pregnancy outcome: a randomized controlled trial. Am J Obstet Gynecol 2018 Apr;218(4):453.e1-453.e7. [doi: 10.1016/j.ajog.2018.01.044] [Medline: 29425836]
44. Kinouchi K, Ohashi K. Smartphone-based reminder system to promote pelvic floor muscle training for the management of postnatal urinary incontinence: historical control study with propensity score-matched analysis. PeerJ 2018;6:e4372 [FREE Full text] [doi: 10.7717/peerj.4372] [Medline: 29441244]
45. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]
46. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
47. Krishnamurti T, Davis AL, Wong-Parodi G, Fischhoff B, Sadovsky Y, Simhan HN. Development and testing of the MyHealthyPregnancy app: a behavioral decision research-based tool for assessing and communicating pregnancy risk. JMIR Mhealth Uhealth 2017 Apr 10;5(4):e42 [FREE Full text] [doi: 10.2196/mhealth.7036] [Medline: 28396302]

Abbreviations

BMI: body mass index
GDM: gestational diabetes mellitus
ICT: information communication technology
mHealth: mobile health
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 04.08.18; peer-reviewed by C Myers, E Zairina, D Mohan; comments to author 24.10.18; revised version received 04.11.18; accepted 05.11.18; published 30.01.19.

Please cite as:
Chan KL, Chen M
Effects of Social Media and Mobile Health Apps on Pregnancy Care: Meta-Analysis
JMIR Mhealth Uhealth 2019;7(1):e11836
URL: https://mhealth.jmir.org/2019/1/e11836/
doi:10.2196/11836
PMID:30698533

Original Paper

Physical Activity Surveillance Through Smartphone Apps and Wearable Trackers: Examining the UK Potential for Nationally Representative Sampling

Tessa Strain[1], PhD; Katrien Wijndaele[1], PhD; Søren Brage[1], PhD

MRC Epidemiology Unit, Institute of Metabolic Science, School of Clinical Medicine, University of Cambridge, Cambridge, United Kingdom

Corresponding Author:
Tessa Strain, PhD
MRC Epidemiology Unit
Institute of Metabolic Science, School of Clinical Medicine
University of Cambridge
Addenbrookes Biomedical Campus
Cambridge, CB2 0QQ
United Kingdom
Phone: 44 1223 330315
Email: tessa.strain@mrc-epid.cam.ac.uk

Abstract

Background: Smartphones and wearable activity trackers present opportunities for large-scale physical activity (PA) surveillance that overcome some limitations of questionnaires or researcher-administered devices. However, it remains unknown whether current users of such technologies are representative of the UK population.

Objective: The objective of this study was to investigate potential sociodemographic biases in individuals using, or with the potential to use, smartphone apps or wearable activity trackers for PA surveillance in the United Kingdom.

Methods: We used data of adults (aged ≥16 years) from two nationally representative surveys. Using the UK-wide 2018 Ofcom Technology Tracker (unweighted N=3688), we derived mutually adjusted odds ratios (ORs; 95% CI) of personal use or household ownership of a smartwatch or fitness tracker and personal use of a smartphone by age, sex, social grade, activity- or work-limiting disability, urban or rural, and home nation. Using the 2016 Health Survey for England (unweighted N=4539), we derived mutually adjusted ORs of the use of wearable trackers or websites or smartphone apps for weight management. The explanatory variables were age, sex, PA, deprivation, and body mass index (BMI). Furthermore, we stratified these analyses by BMI, as these questions were asked in the context of weight management.

Results: Smartphone use was the most prevalent of all technology outcomes, with 79.01% (weighted 2085/2639) of the Technology Tracker sample responding affirmatively. All other outcomes were <30% prevalent. Age ≥65 years was the strongest inverse correlate of all outcomes (eg, OR 0.03, 95% CI 0.02-0.05 for smartphone use compared with those aged 16-44 years). In addition, lower social grade and activity- or work-limiting disability were inversely associated with all Technology Tracker outcomes. Physical inactivity and male sex were inversely associated with both outcomes assessed in the Health Survey for England; higher levels of deprivation were only inversely associated with websites or phone apps used for weight management. The conclusions did not differ meaningfully in the BMI-stratified analyses, except for deprivation that showed stronger inverse associations with website or phone app use in the obese.

Conclusions: The sole use of PA data from wearable trackers or smartphone apps for UK national surveillance is premature, as those using these technologies are more active, younger, and more affluent than those who do not.

(JMIR Mhealth Uhealth 2019;7(1):e11898) doi:10.2196/11898

KEYWORDS

adult; exercise; fitness trackers; health surveys; smartphone; surveys and questionnaires; United Kingdom; mobile phone

Introduction

National-level physical activity (PA) surveillance usually involves data collection through questionnaires, although some countries also use devices such as accelerometers [1]. The United Kingdom generates prevalence figures from a number of different survey questionnaires; research-grade devices have only been used in small subsamples of some surveys and not yet on a regular basis. Both methods require randomly sampling a proportion of the population to infer representative prevalence and trends; samples are typically small because the resources required are substantial [2]. The representativeness of this sample, however, may be compromised by lower response rates, although a sufficiently large sample size allows advanced statistical modeling to be used to minimize selection bias. It is, therefore, worth considering all surveillance methods that decrease researcher and participant burden, while still achieving large sample sizes. Two such potential options for PA surveillance are smartphone apps and personal wearable activity trackers.

A recent study has demonstrated the potential scale of PA data collection through smartphone apps, describing step count data from 717,527 iPhone users from 111 countries [3]. The combined size and geographical coverage of this dataset make this a potentially useful resource for PA epidemiology. However, the sample was restricted to iPhone users, who may not be representative of the general population. Unsurprisingly, most of the data originated from people living in richer countries. Among the 46 countries for which demographic data were presented, the median age was under 40 years, and there was a strong tendency toward overrepresentation of men. Such demographic selection biases would be problematic for global and national surveillance unless they were taken into account in the analyses.

The aim of this study was to investigate potential sociodemographic biases in individuals using, or with the potential to use, smartphone apps or wearable activity trackers for PA surveillance in the United Kingdom.

Methods

Data Sources

We used two nationally representative surveys that collected data relating to the use of smartphone apps or wearable activity trackers: the 2018 Ofcom Technology Tracker (TT) survey and the 2016 Health Survey for England (HSE); the former covered all 4 home nations in the United Kingdom, while the latter covered England only.

The 2018 TT data were obtained on May 18, 2018, through contact with Ofcom but have since been made publicly available on their website [4]. The 2016 HSE data were downloaded from the UK Data Archive on April 17, 2018 [5].

Ofcom Technology Tracker

The Ofcom TT survey measures awareness, access, use of, and attitudes toward fixed and mobile telecoms, internet, multichannel television, and radio of adults (aged ≥16 years) in the United Kingdom [6]. The 2018 survey was run by Saville-Rossiter Base on behalf of Ofcom, the UK communications regulator [7]. Data were collected between January 3 and February 28, 2018, by interviewer-led, tablet computer-assisted interviews carried out at respondents' homes. A quota sample of 3730 adults was selected to match the 2011 Census data on age, sex, and social grade [8]. Weighting matched the sample to the geographical and demographic population profile of the United Kingdom [6].

Device Ownership and Use

Two main outcomes were derived from the responses to questions on device use:

1. *Personal use of a smartphone*. Respondents were provided with the following description: "a smartphone is a phone on which you can easily access emails, download files and applications, as well as view websites and generally surf the internet. Popular brands of smartphone include BlackBerry, iPhone, and Android phones such as the Samsung Galaxy S6."
2. *Personal use of "a smartwatch or wearable tech such as fitness trackers*." The following description was provided: "a wearable computer that may be compatible with a smartphone. Brands include Apple Watch, Pebble, Fitbit, and Garmin."

In addition, we derived "household ownership of a smartwatch or fitness tracker" as a supplementary outcome to identify any differences between ownership and use.

Explanatory Variables

Respondents reported their age in years and a 3-category variable was derived: 16-44, 45-64, and ≥65 years. Sex was coded by the interviewer but not asked directly of the respondent. Social grade was prederived on the dataset according to the National Readership Survey categories [8]. This was based on the self-reported occupational details of the main earner in the household: position or rank, industry, qualifications, and the number of staff members responsible for. The commonly used 2-category variable was derived—ABC1: higher; intermediate, supervisory, or junior managerial, administrative, or professional occupations and C2DE: skilled manual, semiskilled or unskilled manual workers, state pensioners, casual or lowest-grade workers, or unemployed with state benefits only. Respondents who self-reported any of the following conditions were deemed to have an activity or work-limiting disability: breathlessness or chest pains, visual, hearing, mobility, speaking or communicating difficulties, limited ability to reach, mental health problems, dyslexia, or any other self-reported health problems that limit daily activities or work. Postcodes were not included on the dataset, but 2 geographical variables were prederived from them: urban or rural location and UK home nation. Rural was defined as a postcode in villages with a population <2000 that are at least 10 miles from a town or city with a population >15,000. All other locations were defined as urban.

Statistical Analysis

The analysis sample consisted of 3688 individuals who provided complete data for all relevant variables. Logistic regressions were used to calculate the crude and mutually adjusted (for age, sex, social grade, disability status, urban or rural, and UK home nation) odds ratios (ORs) for the likelihood of reporting (1) personal use of a smartphone; (2) personal use of a wearable tracker; and (3) household ownership of a wearable tracker. All analyses were weighted using the sampling weights provided.

The 2016 Health Survey for England

The HSE is an annual survey commissioned by the Health and Social Care Information Centre, undertaken by the NatCen Social Research and University College London [9]. It aims to provide nationally representative data on the prevalence and trends of health conditions and behaviors for the population living in private households in England.

The majority of information, including the demographic data, was collected through a computer-assisted, interviewer-led interview carried out at respondents' homes, spread throughout the year [10]. In addition, respondents' height and weight were measured at the main interview. A follow-up visit by a nurse was offered to all participants. This consisted of a further questionnaire, including items on the use of technology for weight management, more anthropometric measurements, and a blood sample. The questions about technology use were relevant to the present analysis (see below). A total of 5049 adults (aged ≥16 years) participated in both the main interview and nurse visit. Sampling weights were provided for this subsample that accounted for selection probability and nonresponse bias, calibrating to mid-year population estimates for sex and age groups by region. Further details are available elsewhere [9].

Use of Technology for Weight Management

As part of the nurse visit, respondents were asked whether they had used any devices or services to help manage or change their weight (multiple responses allowed). The 2 responses of interest were (1) activity trackers or fitness monitors such as a Fitbit, FuelBand, or Jawbone Up and (2) websites or mobile phone apps. For the activity tracker question, nurses were given the prompt "explain if necessary, activity trackers or fitness monitors are often a band worn on the wrist like a watch. They keep track of the number of steps people take and track activity over time" [10].

Explanatory Variables

Age, sex, and PA in the 28 days prior to the interview were reported. The following 3-category variable for age was derived: 16-44, 45-64, and ≥65 years. We used the prederived variable on compliance to the UK Chief Medical Officers' PA recommendation of inactive (0-<150 minutes/week) and active (≥150 minutes/week) [11]. This was derived from questions on the duration and frequency of different domains of PA according to the protocol used to derive the national prevalence estimates. All heavy housework, heavy manual nonoccupational activity, gardening, and do-it-yourself home maintenance were counted as a moderate-intensity activity; examples of activities were provided to assist participants identify whether an activity was intense enough. Time spent climbing stairs or ladders, lifting, carrying or moving heavy loads, and walking at work was reported but only counted as moderate intensity if the respondents' Standard Occupational Classification 2000 code was in a predetermined list [12,13]. Sport and exercise activities were counted as a moderate or vigorous activity dependent on a predetermined list, which, for some activities, factored in response to a question as to whether it made them out of breath or sweaty. For those aged <65 years, walking counted as moderate intensity if the self-reported pace was "fairly brisk" or "fast pace—at least 4 miles per hour." All walking counted as moderate-intensity activity in those aged ≥65 years. The total weekly duration of vigorous intensity activity was counted as double that of moderate intensity and summed to give a total that was used to determine compliance with the PA recommendation.

The body mass index (BMI; weight, kg/height, m^2) was calculated using the measurements obtained at the main interview. A 3-category variable was derived: normal or underweight (<25 kg/m^2), overweight (25-<30 kg/m^2), and obese (≥30 kg/m^2). A score on the 2015 Index of Multiple Deprivation (a multidomain measure of area deprivation [14]) was prederived from respondents' postcodes. Quintiles of this score (based on the main interview sample) were provided on the downloaded dataset. We derived a binary variable to identify the most deprived 20%.

Statistical Analysis

The analysis sample consisted of 4539 individuals who provided complete data for all relevant variables. Logistic regressions were used to calculate the crude and mutually adjusted (for age, sex, activity status, deprivation status, and BMI) ORs for the likelihood of reporting the use of (1) an activity tracker or fitness monitor and (2) a website or mobile phone app, for weight management. All analyses were weighted using the sampling weights provided. As these questions were asked in the context of weight management, and our interest here is more generic activity tracking, we also ran the analyses stratified by BMI category.

Results

Sample Characteristics and Prevalence of Activity Tracking Technology

Tables 1 and 2 show the sociodemographic characteristics of weighted TT and HSE samples, respectively (see Multimedia Appendix 1 for the BMI-stratified HSE sample data). Figure 1 and Multimedia Appendix 2 show that smartphone use was the most prevalent of all the investigated TT outcomes (weighted 2085/2639, 79.01%). Prevalence of personal use of a smartwatch or fitness tracker was 13.86% (weighted 366/2639). Those aged ≥65 years, those who had an activity- or work-limiting disability, or those with a lower social grade reported the lowest prevalence figures. Prevalence of household ownership of a smartwatch or fitness tracker was slightly higher than that for personal use but followed a similar pattern among subgroups (see Multimedia Appendix 3). Data from the HSE (Figure 2 and Multimedia Appendix 4) showed that 6.53% (weighted 286/4380) of the

sample reported using a wearable tracker for weight management and 8.86% (weighted 388/4380) of the sample reported using websites or phone apps for weight management.

Sociodemographic Correlates of Activity Tracking Technology Use in the 2018 Technology Tracker

Figure 3 shows that age ≥65 years is the characteristic associated with the lowest odds of personal use of a smartwatch or fitness tracker, as well as of the personal use of a smartphone in the TT survey. The mutually adjusted ORs for this group compared with those aged 16-44 years were 0.14 (95% CI 0.09-0.24) and 0.03 (95% CI 0.02-0.05), respectively. In addition, age between 45 and 64 years was associated with a lower likelihood of reporting smartphone use (mutually adjusted OR 0.27, 95% CI 0.20-0.36) but the respective OR CI for personal use of a smartwatch or fitness tracker just crossed one.

Lower social grade (C2DE compared with ABC1) was inversely associated with the use of tracking technology, with mutually adjusted ORs ranging between 0.31 and 0.42 (see Multimedia Appendix 2). Reporting an activity- or work-limiting disability inversely correlated with the personal use of a smartwatch or fitness tracker and smartphone use (mutually adjusted ORs 0.55, 95% CI 0.35-0.86 and 0.45, 95% CI 0.35-0.57, respectively). There were mixed results regarding the geographical explanatory variables of urban-rural and home nation; those in urban areas were less likely to own a smartwatch or fitness tracker in the household (mutually adjusted OR 0.69, 95% CI 0.53-0.90), but the CIs crossed one for the other outcomes. Those living in Northern Ireland were less likely to report personal use of a smartwatch or fitness tracker than those living in England (mutually adjusted OR 0.56, 95% CI 0.38-0.81). Other comparisons between nations and for other outcomes did not present clear patterns. We observed no differences by sex for any outcome. Furthermore, there were no substantial differences between the results for personal use and household ownership of a smart watch or activity tracker (see Multimedia Appendix 5).

Table 1. Sociodemographic profile of the 2018 Ofcom Technology Tracker sample (unweighted N=3688, weighted N=2639).

Characteristics	Sample, weighted n (%)	SE
Age group (years)		
16-44	1259 (47.72)	1.0
45-64	869 (32.93)	0.9
>65	511 (19.35)	0.7
Sex		
Women	1350 (51.17)	1.0
Men	1289 (48.83)	1.0
Social grade[a]		
ABC1	1417 (53.69)	1.0
C2DE	1222 (46.31)	1.0
Disability status		
No activity or work-limiting disability	2192 (83.05)	0.7
Activity or work-limiting disability	447 (16.95)	0.7
Location		
Rural	351 (13.29)	0.6
Urban	2288 (86.71)	0.6
UK home nation		
England	2201 (83.42)	0.6
Northern Ireland	73 (2.78)	0.1
Scotland	232 (8.81)	0.5
Wales	132 (4.99)	0.3

[a]ABC1 includes those where the main household earner is in a higher, intermediate, supervisory, or junior managerial, administrative, or professional occupation and C2DE includes those where the main household earner is a skilled manual, semiskilled or unskilled manual worker, state pensioner, casual or lowest-grade worker, or unemployed with state benefits only.

Table 2. Sociodemographic profile of the 2016 Health Survey for England sample (unweighted N=4539, weighted N=4380).

Characteristic	Sample, weighted n (%)	SE
Age group (years)		
16-44	2015 (45.99)	0.9
45-64	1408 (32.14)	0.7
>65	958 (21.86)	0.6
Sex		
Women	2198 (50.19)	0.8
Men	2182 (49.81)	0.8
Physical activity		
Active	3292 (75.15)	0.7
Inactive	1088 (24.85)	0.7
Deprivation		
Top 80%	3539 (80.79)	0.7
Most deprived 20%	841 (19.21)	0.7
Body mass index		
Under or normal weight	2013 (46.96)	0.8
Overweight	1506 (34.39)	0.8
Obese	861 (19.65)	0.6

Figure 1. Percentage reporting the use of activity tracking-related technology in the 2018 Ofcom Technology Tracker survey (unweighted N=3688, weighted N=2639).

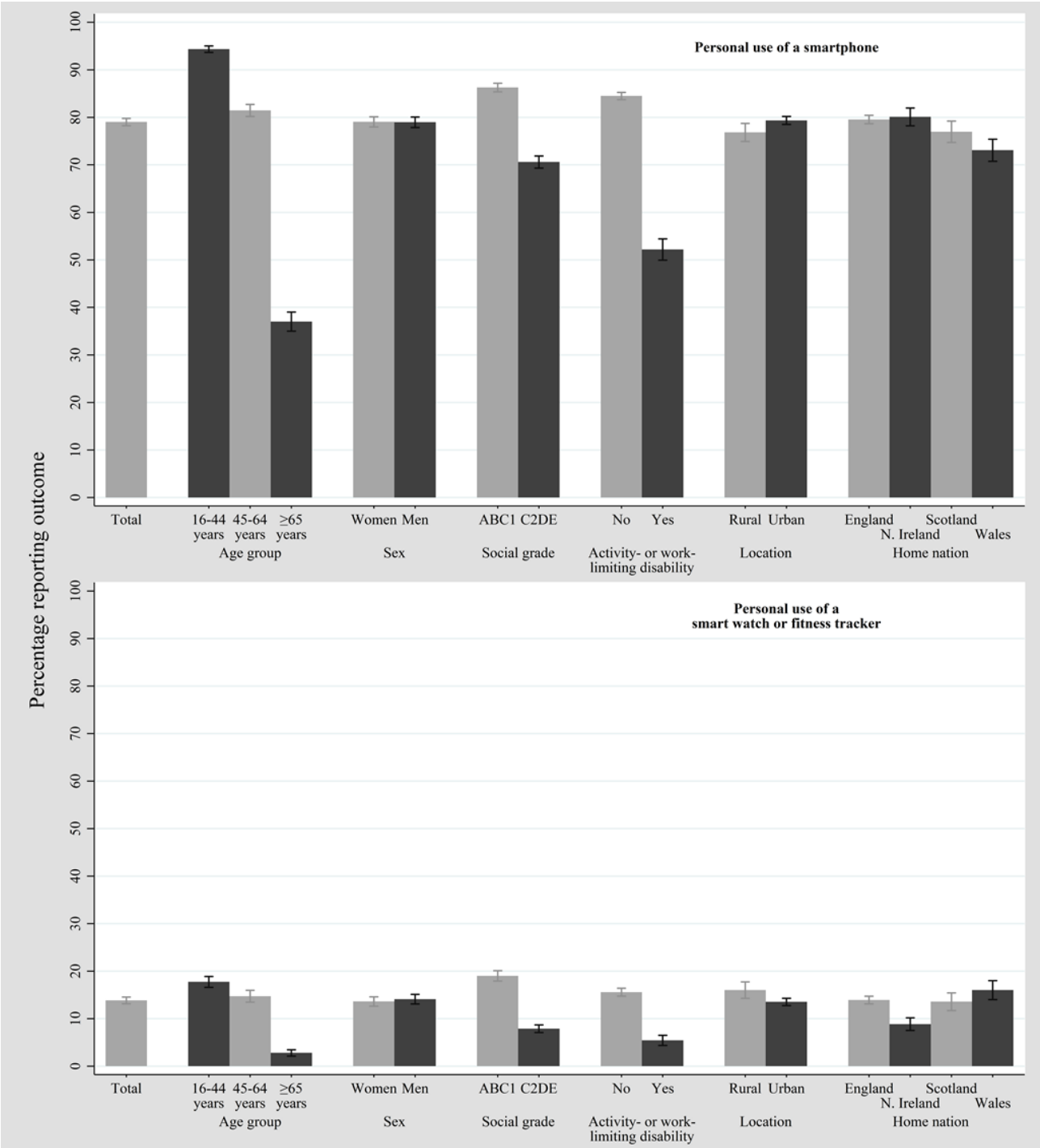


Figure 2. Percentage reporting the use of activity tracking-related technology in the 2016 Health Survey for England (unweighted N=4539, weighted N=4380).

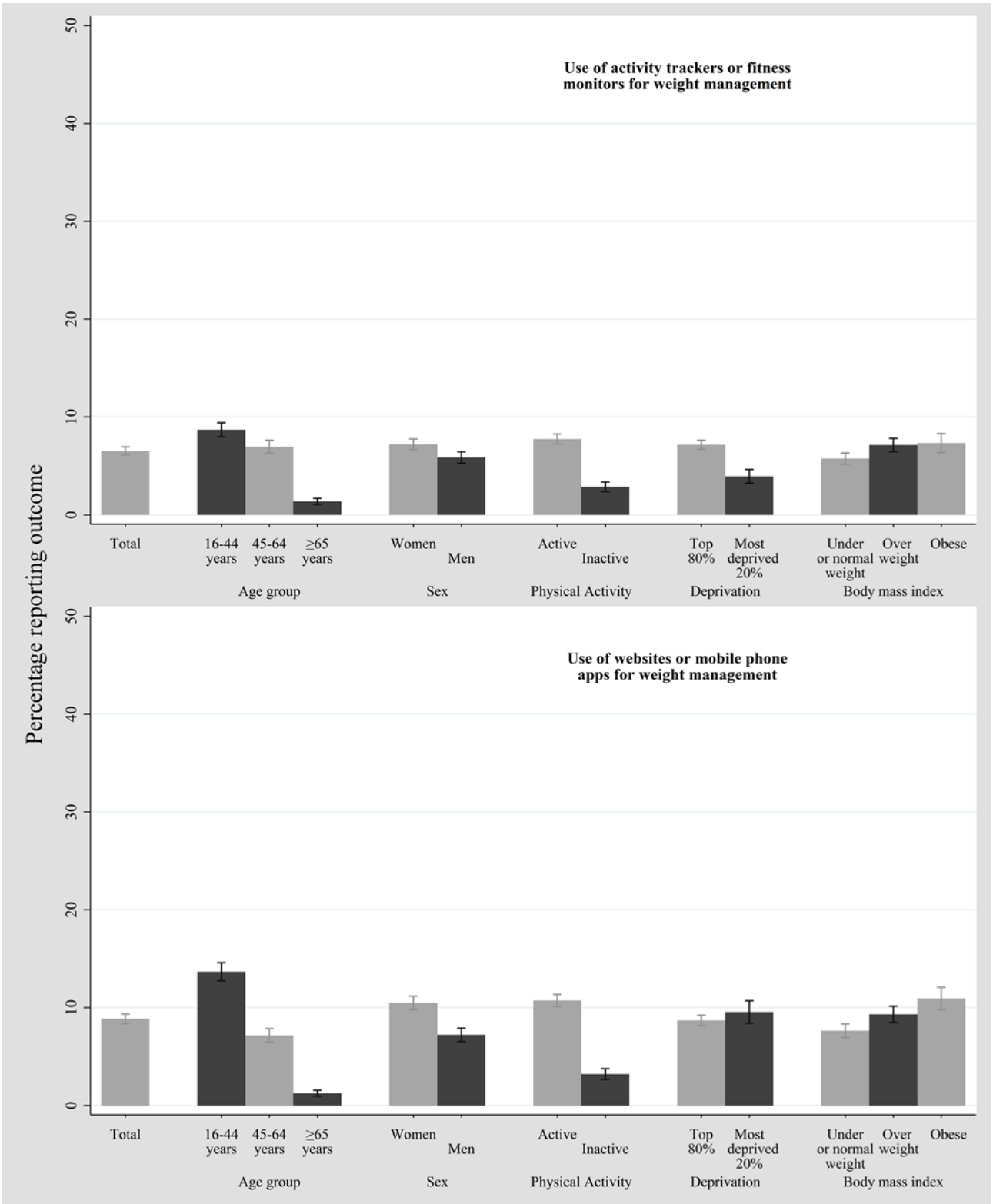
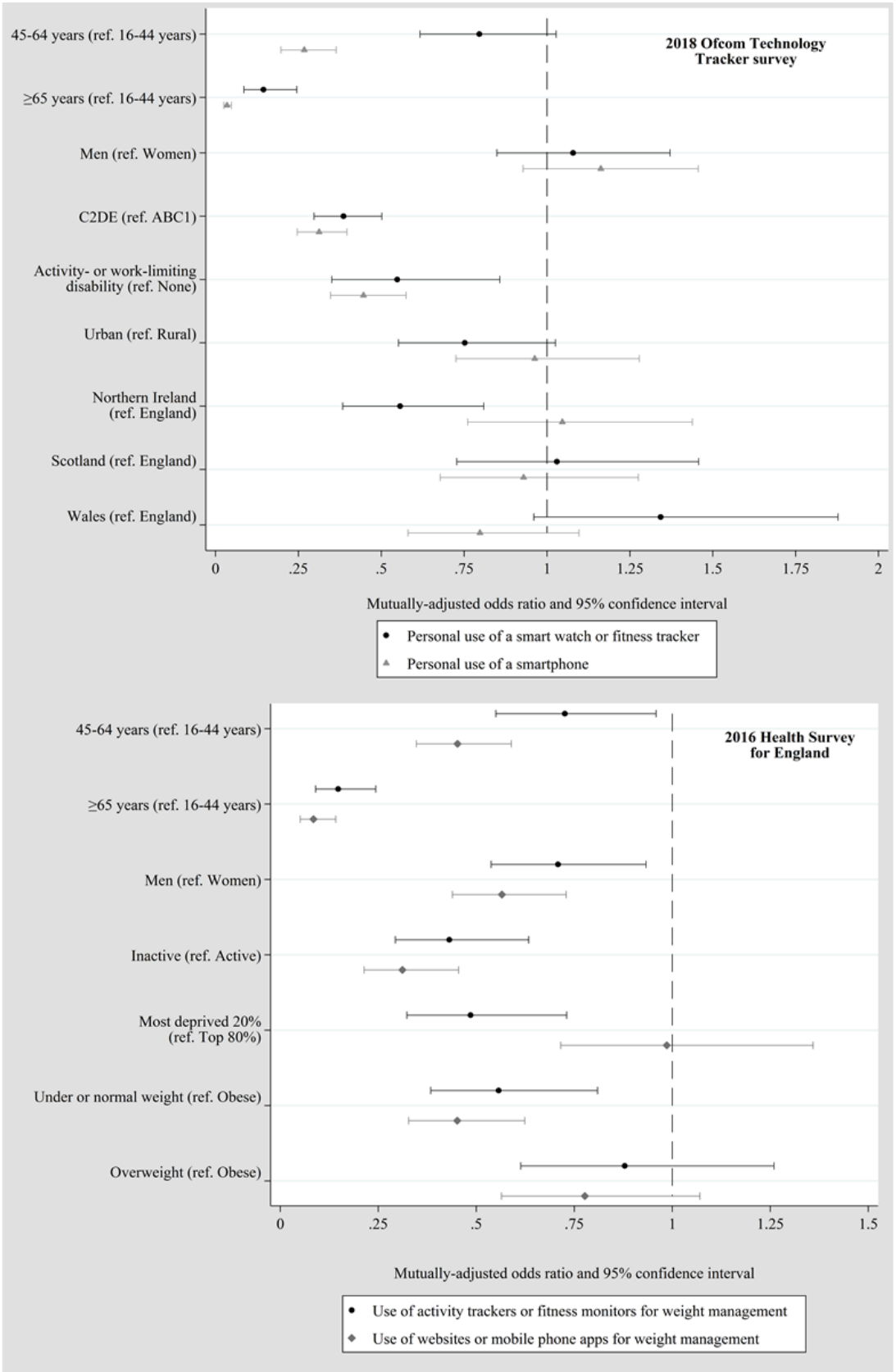


Figure 3. Mutually adjusted odds ratios of reporting the use or ownership of activity tracking-related technology by sociodemographic characteristics in the 2018 Ofcom Technology Tracker survey (unweighted N=3688, weighted N=2639) and the 2016 Health Survey for England (unweighted N=4539, weighted N=4380).



Sociodemographic Correlates of Activity Tracking Technology Use in the 2016 Health Survey for England

In the 2016 HSE, age ≥65 years showed the strongest inverse relationships with the use of tracking technology, with mutually adjusted ORs between 0.08 and 0.15 (Figure 3; Multimedia Appendix 4). In addition, not meeting the PA guidelines (compared with meeting them) and male sex (compared with female) were inversely associated with both uses of technology for weight management (mutually adjusted ORs 0.31-0.43 and 0.57-0.71, respectively). Those in the 20% most deprived areas were less likely to report using websites or phone apps for weight management compared with those in the top 80%; however, there was no evidence of a difference in the use of wearable trackers for weight management. Conversely, those aged 45-64 years were less likely to use a wearable tracker for weight management compared with those aged 16-44 years; however, there were no differences in the website or phone app use. A majority of the conclusions did not differ meaningfully when the analyses were stratified by BMI. One notable exception was area deprivation, which showed stronger inverse associations with the smartphone use in the obese individuals (see Multimedia Appendices 6 and 7).

Discussion

Principal Findings

This is the first study to consider the issue of representativeness of users of tracking technology in UK data in relation to PA surveillance. This is timely as the Expert Group reviewing the UK PA guidelines has recommended that all technological advances in the field of PA measurement are considered for the long-term future of surveillance (report due to be published in 2019).

Our results show that users or owners of smartphones and wearable activity trackers are not representative of the general population in the United Kingdom; this was also true for the use of wearable activity trackers or websites or apps in the context of weight management. Statistical weighting, that is, attempting to make the sample results more reflective of the population distribution of key sociodemographic variables, is unlikely to be able to resolve these issues for 2 reasons. First, our results indicate that PA levels themselves may be correlated with the use of such technologies, albeit it is important to note that these data from the HSE are asked within the context of weight management. If users are more active than nonusers, adjusting for other population demographic characteristics will still lead to an overestimate. Second, some of the biases are strong (eg, age >65 years), meaning that certain sample substrata would be weighted heavily and be highly influential in the estimates. When such a minority of a population use the technology required for measurement, such as would be the case for some subgroups, it is unlikely that the assumption that users and nonusers are similar with respect to the relevant characteristics would hold. Further discussion on the issues of statistical weighting in population surveys is provided elsewhere [15].

Comparison With Prior Work

Despite smartphone usage being almost ubiquitous among people aged 16-44 years, it remains much less common in those aged >65 years, at around one-third. Age ≥65 years was the strongest inverse correlate for all outcomes. This is comparable to similar studies looking at smartphone use undertaken in Canadian [16], Swiss [17], German [18], and American [19] samples. These studies also found differences by activity levels [17-19], some indications of health status [16], and measures of socioeconomic position [16,18,19].

This is a fast-moving field, and trends indicate that activity trackers will become more prevalent in the coming years. The Ofcom TT data indicate that the percentage reporting using a smartwatch has increased from 2% in 2015 to 14% in 2018 [20,21]. The percentage of people aged ≥55 years (no older age group breakdown available) using smartphones has increased from 32% to 51% over that period [20,21]. Although it is hard to reach conclusions with such small starting prevalence figures, it does appear that it is the more affluent driving the increase, but that it is relatively uniform across the age groups [20,21]. As more data are collected, this will be an important trend to monitor.

Strengths and Limitations

A major strength of this analysis is that it uses the most up-to-date nationally representative data. Both datasets were only released in April 2018; the Ofcom TT data were even collected this year. Although the HSE questions were asked in a weight management context, they are the only source of data to provide paired information on PA and device use. We examined this potential bias by performing BMI-stratified sensitivity analyses (see Multimedia Appendices 6 and 7). The results were similar between weight groups, except for deprivation where the inverse association was stronger in the obese.

The PA levels in the HSE “nurse interview sample” were higher than reported for the “main interview” sample in 2016 (58% women and 66% of men [13]), even after weighting by age, sex, and geographical location. This bias is likely to affect (overestimate) the prevalence estimates for those using wearable trackers or websites or mobile phone apps for weight management. For our specific purpose, it would have been advantageous for the TT survey to also have included a measure of PA, as that is the potential bias that most limits the use of this technology for PA surveillance.

A limitation of this study was that we were only able to investigate differences in the use or ownership due to the questions asked in the surveys. Ownership of a smartphone will not necessarily mean that an individual is willing to download and use an activity tracking app and then share the data for national surveillance purposes. Even among willing individuals, there may be further biases concerning what activities are recorded: for example, a smartphone app is unlikely to be used to record swimming, and wrist-worn devices may not be able to adequately quantify activity when cycling. This may also be influenced by how they are worn (eg, trouser or breast pocket, handbag). As both the types of activities that adults participate

in and the method of wearing a smartphone have been shown to differ systematically by characteristics such as sex, age, disabilities, and cultural norms [22,23], this is another layer of representativeness that should be considered. In addition, we were unable to examine why people were using devices in the TT survey, whereas in the HSE survey, the questions were only asked for weight management. Reportedly, many who use these devices do so to improve their health [24]. This may mean that individuals' behavior while using a device is not representative of their habitual levels. More detailed data will be needed to understand whether this introduces further or different biases. Ideally, we should also have been able to identify users of different smartphone operating systems (asked in the TT survey but not on the publicly available dataset), as this can have a bearing on what apps are available for download and the practicalities of obtaining the data for researchers. Furthermore, data on the use of specific PA tracking apps would have added useful information. This investigation does not allow us to make any conclusions regarding the validity of these technologies for measuring the metrics of PA. This issue is equally important when considering their potential use for PA surveillance, particularly as some evidence suggests that there may be systematic biases for some estimates. For example, walking metrics, such as step count and distance, appear to be underestimated at slower speeds, higher BMI, female sex, and among certain ethnic groups [25,26]. Finally, the scope of this study was to consider these data sources for surveillance purposes. Other study designs, most notably those using smartphones and activity trackers as an intervention aids for changing PA, may well conclude that these methods have utility [27]. In addition, the study of within-person patterns across the week or year using these data sources may well generalize better to the general population, but no data are currently allowing us to examine that.

Conclusions

We conclude that the sole use of PA data from personal trackers or smartphone apps for national surveillance in the United Kingdom is premature as those using these devices are more active, younger, and more affluent than those who do not.

Acknowledgments

This work was supported by the Medical Research Council (grant number MC_UU_12015/3). The authors would like to thank participants of the Ofcom TT survey and the HSE. We would also like to thank those who funded the surveys, undertook the fieldwork, and processed the data.

Conflicts of Interest

TS is a member of the UK Expert Group reviewing PA surveillance methods. She has no vested interests regarding any specific measurement methods. KW and SB declare no competing interests.

Multimedia Appendix 1

Sociodemographic profile of the 2016 Health Survey for England sample (unweighted N=4539, weighted N=4380), stratified by body mass index.

[PDF File (Adobe PDF File), 28KB - mhealth_v7i1e11898_app1.pdf]

Multimedia Appendix 2

Crude and mutually-adjusted odds ratios of reporting personal use of a smart watch or fitness tracker, household ownership of a smart watch or fitness tracker, or personal use of a smartphone, by socio-demographic characteristic in the 2018 Ofcom Technology Tracker survey (unweighted N=3688, weighted N=2639).

[PDF File (Adobe PDF File), 67KB - mhealth_v7i1e11898_app2.pdf]

Multimedia Appendix 3

Percentage reporting household ownership of a smart watch or activity tracker in the 2018 Ofcom Technology Tracker survey (unweighted N=3688, weighted N=2639).

[PNG File, 934KB - mhealth_v7i1e11898_app3.png]

Multimedia Appendix 4

Crude and mutually-adjusted odds ratios of reporting use of activity trackers or fitness monitors or websites or mobile phone applications for weight management, by socio-demographic characteristic in the 2016 Health Survey for England (unweighted N=4539, weighted N=4380).

[PDF File (Adobe PDF File), 51KB - mhealth_v7i1e11898_app4.pdf]

Multimedia Appendix 5

Mutually-adjusted odds ratios of reporting household ownership of smart watch or activity tracker by socio-demographic characteristic, in the 2018 Ofcom Technology Tracker survey (unweighted N=3688, weighted N=2639).

[PNG File, 771KB - mhealth_v7i1e11898_app5.png]

Multimedia Appendix 6

Crude and mutually-adjusted odds ratios of reporting use of activity trackers or fitness monitors or websites or mobile phone applications for weight management, by sociodemographic characteristic, stratified by body mass index, in the 2016 Health Survey for England (unweighted N=4539, weighted N=4380).

[PDF File (Adobe PDF File), 80KB - mhealth_v7i1e11898_app6.pdf]

Multimedia Appendix 7

Mutually-adjusted odds ratios of reporting use of activity trackers or fitness monitors or websites or mobile phone applications for weight management, by sociodemographic characteristic, stratified by body mass index, in the 2016 Health Survey for England (unweighted N=4539, weighted N=4380).

[PNG File, 157KB - mhealth_v7i1e11898_app7.png]

References

1. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U, Lancet Physical Activity Series Working Group. Global physical activity levels: surveillance progress, pitfalls, and prospects. Lancet 2012 Jul 21;380(9838):247-257. [doi: 10.1016/S0140-6736(12)60646-1] [Medline: 22818937]
2. Mindell JS, Moody A, Vecino-Ortiz AI, Alfaro T, Frenz P, Scholes S, et al. Comparison of Health Examination Survey Methods in Brazil, Chile, Colombia, Mexico, England, Scotland, and the United States. Am J Epidemiol 2017 Sep 15;186(6):648-658 [FREE Full text] [doi: 10.1093/aje/kwx045] [Medline: 28486584]
3. Althoff T, Sosič R, Hicks JL, King AC, Delp SL, Leskovec J. Large-scale physical activity data reveal worldwide activity inequality. Nature 2017 Dec 20;547(7663):336-339 [FREE Full text] [doi: 10.1038/nature23018] [Medline: 28693034]
4. Ofcom. 2018 May 21. Technology Tracker URL: https://data.gov.uk/dataset/eb673e35-1a59-47d3-b5f1-914a67d85baf/technology-tracker [accessed 2018-08-06] [WebCite Cache ID 71TLuhxUi]
5. NatCen Social Research, University College London. Department of Epidemiology and Public Health. UK Data Service. 2018. Health Survey for England, 2016 URL: https://beta.ukdataservice.ac.uk/datacatalogue/doi/?id=8334 [accessed 2018-11-27] [WebCite Cache ID 74Evhjzet]
6. Ofcom. 2018. Technology Tracker H1 2018 technical report URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0015/113172/Technology-Tracker-H1-2018-technical-report.pdf [accessed 2018-08-06] [WebCite Cache ID 71TMDDQg2]
7. Ofcom. About Ofcom URL: https://www.ofcom.org.uk/about-ofcom [accessed 2018-08-07] [WebCite Cache ID 71UQFN6Ju]
8. National Readership Survey. Social Grade URL: http://www.nrs.co.uk/nrs-print/lifestyle-and-classification-data/social-grade/ [accessed 2018-08-07] [WebCite Cache ID 71UQLNbqp]
9. NatCen Social Research, University College London. NHS Digital. Leeds, UK; 2017 Dec 13. Health Survey for England 2016: Methods URL: https://files.digital.nhs.uk/publication/m/3/hse2016-methods-text.pdf [accessed 2018-08-07] [WebCite Cache ID 71UQVRc7W]
10. NatCen Social Research, University College London. NHS Digital. Leeds, UK; 2016. Health Survey for England 2016: Field documents and measurement protocols URL: https://files.digital.nhs.uk/pdf/q/j/hse2016-documentation.pdf [accessed 2018-08-07] [WebCite Cache ID 71UQbJ6rd]
11. Department of Health, Physical Activity, Health Improvement and Protection. gov.uk. London, UK; 2011 Jul 11. Start Active, Stay Active: A report on physical activity for health from the four home countries' Chief Medical Officers URL: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/216370/dh_128210.pdf [accessed 2018-08-07] [WebCite Cache ID 71UQeaZNx]
12. NatCen Social Research, University College London Department of Epidemiology and Public Health. UK Data Service. Colchester, UK; 2016. Health Survey for England 2016: Dataset documentation URL: http://doc.ukdataservice.ac.uk/doc/8334/mrdoc/pdf/8334_hse_2016_dataset_documentation.pdf [accessed 2018-08-07] [WebCite Cache ID 71UQjpevN]
13. Scholes S, University College London Department of Epidemiology and Public Health. NHS Digital. Leeds, UK; 2017 Dec 13. Health Survey for England 2016 Physical activity in adults URL: http://healthsurvey.hscic.gov.uk/media/63730/HSE16-Adult-phy-act.pdf [accessed 2018-06-21] [WebCite Cache ID 71UQo6i8V]
14. Department for Communities and Local Government. Ministry of Housing, Communities & Local Government. London, UK; 2015. The English Index of Multiple Deprivation (IMD) 2015 – Guidance URL: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/464430/English_Index_of_Multiple_Deprivation_2015_-_Guidance.pdf [accessed 2018-07-11] [WebCite Cache ID 71UQs7obj]

15. Parsons VL, Moriarity C, Jonas K, Moore TF, Davis KE, Tompkins L. Design and estimation for the national health interview survey, 2006-2015. Vital Health Stat 2 2014 Apr(165):1-53 [FREE Full text] [Medline: 24775908]
16. 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]
17. Seifert A, Schlomann A, Rietz C, Schelling HR. The use of mobile devices for physical activity tracking in older adults' everyday life. Digit Health 2017;3:2055207617740088 [FREE Full text] [doi: 10.1177/2055207617740088] [Medline: 29942617]
18. Ernsting C, Dombrowski SU, Oedekoven M, O Sullivan JL, Kanzler M, Kuhlmey A, et al. Using Smartphones and Health Apps to Change and Manage Health Behaviors: A Population-Based Survey. J Med Internet Res 2017 Dec 05;19(4):e101 [FREE Full text] [doi: 10.2196/jmir.6838] [Medline: 28381394]
19. Omura JD, Carlson SA, Paul P, Watson KB, Fulton JE. National physical activity surveillance: Users of wearable activity monitors as a potential data source. Prev Med Rep 2017 Mar;5:124-126 [FREE Full text] [doi: 10.1016/j.pmedr.2016.10.014] [Medline: 28101443]
20. Ofcom. London, UK; 2018. Half 1 Technology Tracker data tables URL: https://www.ofcom.org.uk/__data/assets/file/0020/113168/Technology-Tracker-H1-2018-data-tables.csv [accessed 2018-08-07] [WebCite Cache ID 71URBD7S0]
21. Ofcom. 2015. Ofcom Nations & Regions Tracker - Main Set - Quarter 1 2015 URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0028/82819/ofcom_technology_tracker_w1_2015_-_questions_qb1-qb5_main.pdf [accessed 2018-08-07] [WebCite Cache ID 71UR7dOw2]
22. Antos SA, Albert MV, Kording KP. Hand, belt, pocket or bag: Practical activity tracking with mobile phones. J Neurosci Methods 2014 Jul 15;231:22-30 [FREE Full text] [doi: 10.1016/j.jneumeth.2013.09.015] [Medline: 24091138]
23. Strain T, Fitzsimons C, Foster C, Mutrie N, Townsend N, Kelly P. Age-related comparisons by sex in the domains of aerobic physical activity for adults in Scotland. Prev Med Rep 2016 Jun;3:90-97 [FREE Full text] [doi: 10.1016/j.pmedr.2015.12.013] [Medline: 26844194]
24. Maher C, Ryan J, Ambrosi C, Edney S. Users' experiences of wearable activity trackers: a cross-sectional study. BMC Public Health 2017 Dec 15;17(1):880 [FREE Full text] [doi: 10.1186/s12889-017-4888-1] [Medline: 29141607]
25. Brodie MA, Pliner EM, Ho A, Li K, Chen Z, Gandevia SC, et al. Big data vs accurate data in health research: Large-scale physical activity monitoring, smartphones, wearable devices and risk of unconscious bias. Med Hypotheses 2018 Oct;119:32-36. [doi: 10.1016/j.mehy.2018.07.015] [Medline: 30122488]
26. Evenson KR, Goto MM, Furberg RD. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act 2015 Dec 18;12:159 [FREE Full text] [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]
27. Stuckey MI, Carter SW, Knight E. The role of smartphones in encouraging physical activity in adults. Int J Gen Med 2017;10:293-303 [FREE Full text] [doi: 10.2147/IJGM.S134095] [Medline: 28979157]

Abbreviations

BMI: body mass index
HSE: Health Survey for England
OR: odds ratio
PA: physical activity
TT: Technology Tracker

Edited by G Eysenbach; submitted 10.08.18; peer-reviewed by J Omura, S Kozey-Keadle; comments to author 05.09.18; revised version received 19.09.18; accepted 22.09.18; published 29.01.19.

Please cite as:
Strain T, Wijndaele K, Brage S
Physical Activity Surveillance Through Smartphone Apps and Wearable Trackers: Examining the UK Potential for Nationally Representative Sampling
JMIR Mhealth Uhealth 2019;7(1):e11898
URL: http://mhealth.jmir.org/2019/1/e11898/
doi:10.2196/11898
PMID:30694198

Original Paper

Digital Pain Drawings Can Improve Doctors' Understanding of Acute Pain Patients: Survey and Pain Drawing Analysis

Nour Shaballout[1], MD; Anas Aloumar[1], MD; Till-Ansgar Neubert[1]; Martin Dusch[2*], Dr Med; Florian Beissner[1*], Dr Phil Nat

[1]Somatosensory and Autonomic Therapy Research, Institute for Diagnostic and Interventional Neuroradiology, Hannover Medical School, Hannover, Germany

[2]Section Pain Medicine, Clinic of Anaesthesiology and Intensive Care Medicine, Hannover Medical School, Hannover, Germany

[*]these authors contributed equally

Corresponding Author:
Florian Beissner, Dr Phil Nat
Somatosensory and Autonomic Therapy Research
Institute for Diagnostic and Interventional Neuroradiology
Hannover Medical School
Carl-Neuberg-Strasse 1
Hannover, 30625
Germany
Phone: 49 511 53508413
Email: Beissner.Florian@mh-hannover.de

Related Article:

This is a corrected version. See correction statement: https://mhealth.jmir.org/2019/9/e16017/

Abstract

Background: Pain drawings (PDs) are an important tool to evaluate, communicate, and objectify pain. In the past few years, there has been a shift toward tablet-based acquisition of PDs, and several studies have been conducted to test the usefulness, reliability, and repeatability of electronic PDs. However, to our knowledge, no study has investigated the potential role of electronic PDs in the clinical assessment and treatment of inpatients in acute pain situations.

Objective: The aim of this study was to evaluate whether knowledge of the patients' electronic PD has the potential to improve the doctors' understanding of their patients and to influence their clinical decision making. Furthermore, we sought to identify differences between electronic PDs of patients and their treating pain specialists in an acute pain situation and to find those specific characteristics derived from the PDs that had the largest impact on doctors' understanding.

Methods: We obtained electronic PDs from 47 inpatients in acute pain situations before their consultation with a pain specialist on a tablet personal computer with a stylus. Before looking at their patients' drawings, these specialists drew their own conception of the patients' pain after anamnesis and physical examination. Patients' drawings were then revealed to the doctors, and they were asked to evaluate how much the additional information improved their understanding of the case and how much it influenced their clinical decision on an 11-point Likert scale (0="not at all" and 10="very much"). Similarities and differences of patients' and doctors' PDs were assessed by visual inspection and by calculating Jaccard index and intraclass correlation coefficient (ICC) of the pain area and the number of pain clusters. Exploratory analyses were conducted by means of correlation tables to identify specific factors that influenced doctors' understanding.

Results: Patients' PDs significantly improved the doctors' understanding (mean score 4.81, SD 2.60, $P<.001$) and to a lesser extent their clinical decision (mean 2.68, SD 1.18, $P<.001$). Electronic PDs of patients and doctors showed fair to good similarity for pain extent ($r=.454$, $P=.001$) and widespreadness ($P=.447$, $r=.002$) were important factors helping doctors to understand their patients.

Conclusions: In a clinical setting, electronic PDs can improve doctors' understanding of patients in acute pain situations. The ability of electronic PDs to visualize differences between doctors' and patients' conception of pain has the potential to improve doctor-patient communication.

(JMIR Mhealth Uhealth 2019;7(1):e11412) doi:10.2196/11412

KEYWORDS

pain drawing; symptom drawing; manikins; tablet computers; eHealth; app; acute pain

Introduction

Background

Pain is a very complex and subjective phenomenon. It is regarded as a symptom of an underlying condition or as a condition of its own. For adequate medical treatment, however, it is compulsory to classify the reported pain. Hints to the correct pain diagnosis are given by a pain assessment looking at the intensity of the pain, its distribution and duration, as well as the quality of the pain. Despite many new technological advances, however, objectification of pain is still an unsolved problem [1]. Asking patients to draw their pain has been used for half a century to overcome the complexity of communicating a subjective sensation from patient to physician. This method has different names in the literature, the most common being pain drawing (PD). Different instruments have been used to obtain PDs, starting from pen-on-paper drawings [2] and recently developing toward electronic PDs collected on tablet personal computers (PCs) [3-7]. Several studies have been conducted to test the usability, reliability, and repeatability of PDs in chronic pain situations such as shoulder pain [8], knee pain [9], back pain [10], and neck pain [11] as well as in acute low back pain [12-14], whiplash disorder [15], or experimentally triggered pain [16]. Regardless of the method used, it has been proven that using PDs together with anamnesis and physical examination can aid the differential diagnosis in many pain situations [9,10].

Objectives

In this study, we examined the potential role of electronic PDs collected on a tablet PC in the clinical assessment and treatment of inpatients in acute pain situations. Therefore, we evaluated if knowing the patients' PDs improved the pain specialists' understanding of their patients and influenced their clinical decision making. Furthermore, we sought to identify differences between PDs of patients and doctors and to find those specific characteristics of the drawings that had the largest impact on doctors' understanding.

To this end, we collected electronic PDs from a sample of inpatients that received a consultation by a pain specialist from acute pain service (APS) because of severe pain (average score of 7.3, SD 2.0 on an 11-point numeric rating scale [NRS]). We then asked the pain specialists to draw their own impression of the patients' pain after anamnesis and physical examination but without having seen the patients' drawings. PDs of the patients were then revealed to the doctors, and they were asked to evaluate how much the additional information improved their understanding of the case and how much it influenced their clinical decision. Similarities and differences of patients' and doctors' PDs were assessed by statistical image analysis and visual inspection. Finally, exploratory analyses were conducted to identify specific factors that had the greatest impact on doctors' understanding of the patients.

Methods

Study Population

Our study population were inpatients from different departments of Hannover Medical School. All patients were in acute pain situations that required a consultation by a pain specialist, which is provided by members of the APS. APS members visit the patients and adjust their pain management individually according to the requirements of each patient's situation. Eligible patients were adult (age ≥18 years in Germany) inpatients of Hannover Medical School with acute pain and the ability to give written informed consent. Furthermore, they had to be physically able to complete an electronic PD on a tablet PC. We recruited 69 patients (37 females) for participation in our study, and all of them prepared a PD. Due to a technical failure of the tablet PC, 1 drawing was lost. The treating pain specialist prepared a PD for 61 of the remaining patients, all of which were included in the analysis. Of these 61 patients with complete data, however, only 47 (24 females) were discussed and rated because of the absence of some of the treating pain specialists during the meeting of the APS in the afternoon (see below). Characteristics of the final study population can be found in Table 1.

Procedures

The study was approved by the Ethics Committee of Hannover Medical School (#2987-2015) and was conducted in accordance with the Declaration of Helsinki. All patients gave written informed consent after information of the purpose of the study.

Data collection was divided into 3 steps: (1) the patient draws a PD shortly before the consultation, (2) the pain specialist draws a PD of the same patient directly after the consultation based on anamnesis and bodily examination results, and (3) the pain specialist rates the patient's PD's influence on his or her understanding of the patient as well as on clinical decision making.

Patients' Pain Drawings

Two of the authors (NS and AA) screened all patients on schedule for consultation by the APS. All eligible patients were informed about the study purpose and asked for participation after checking inclusion criteria. Written informed consent was obtained, and patients were asked to rate their pain intensity on an 11-point NRS ranging from 0 for "no symptom" to 10 for "maximum imaginable intensity." Next, the use of the tablet PC and the SymptomMapper app [17] was explained to the participant, and an electronic drawing of the acute pain and related sensations was acquired using the following instructions: "Please draw the location of your sensations as accurately as possible. Make sure to draw all sensations that you perceive as unpleasant or unnormal." These instructions were complemented by a graphical depiction in the app's drawing instructions module (see below), and patients were supervised by NS or AA during the drawing process.

Table 1. Demographics of our study population.

Characteristics	Statistics
Age in years, mean (SD)	59.2 (15.9)
Age range in years, n (%)	
18-39	7 (15)
40-59	15 (32)
60-79	21 (45)
80+	4 (9)
Women, n (%)	24 (51)
Numeric rating scale pain intensity, mean (SD)	7.3 (2.0)
Origin of pain, n (%)	
Cancer	18 (38)
Infection	8 (17)
Postsurgical	5 (11)
Neurological	3 (6)
Other	13 (28)

Doctors' Pain Drawings

Consultation by the APS consisted of anamnesis and clinical examination by a member of the service. All members participating in the study were anesthesiologists with at least 5 years of clinical experience and either pain specialists or in training for pain specialization. On the basis of the anamnesis and findings of the physical examination, doctors drew their patients' pain and related sensations during or immediately after the consultation but before seeing the patients' drawing. They used the same kind of tablet PC and app like the patients. However, in addition to pain and related sensations, they were also free to draw pain-related symptoms such as allodynia, hyperalgesia, erythema, swelling, hyperhidrosis, and muscular defense. After doctors had finished their PD, they continued with their usual clinical procedures, such as starting the pain treatment or modification of an ongoing treatment.

Rating of Knowledge Gained From Pain Drawings

Patients' PDs were shown to the doctors during the meeting of the APS in the afternoon, where each new patient is reviewed and treatment options are discussed. When discussing a study participant, we revealed his or her PD to all doctors, and they were free to discuss it. After the meeting, the doctors were asked for their anonymous rating of the following 2 questions: (1) "How much did the electronic PD improve your understanding of the patient?" and (2) "How much did the electronic PD influence your clinical decision?" Both questions were followed by an 11-point Likert scale from 0 for "not at all" to 10 for "very much." Doctors were only allowed to rate PDs from patients that they themselves were treating.

Data Acquisition

Tablet Computer

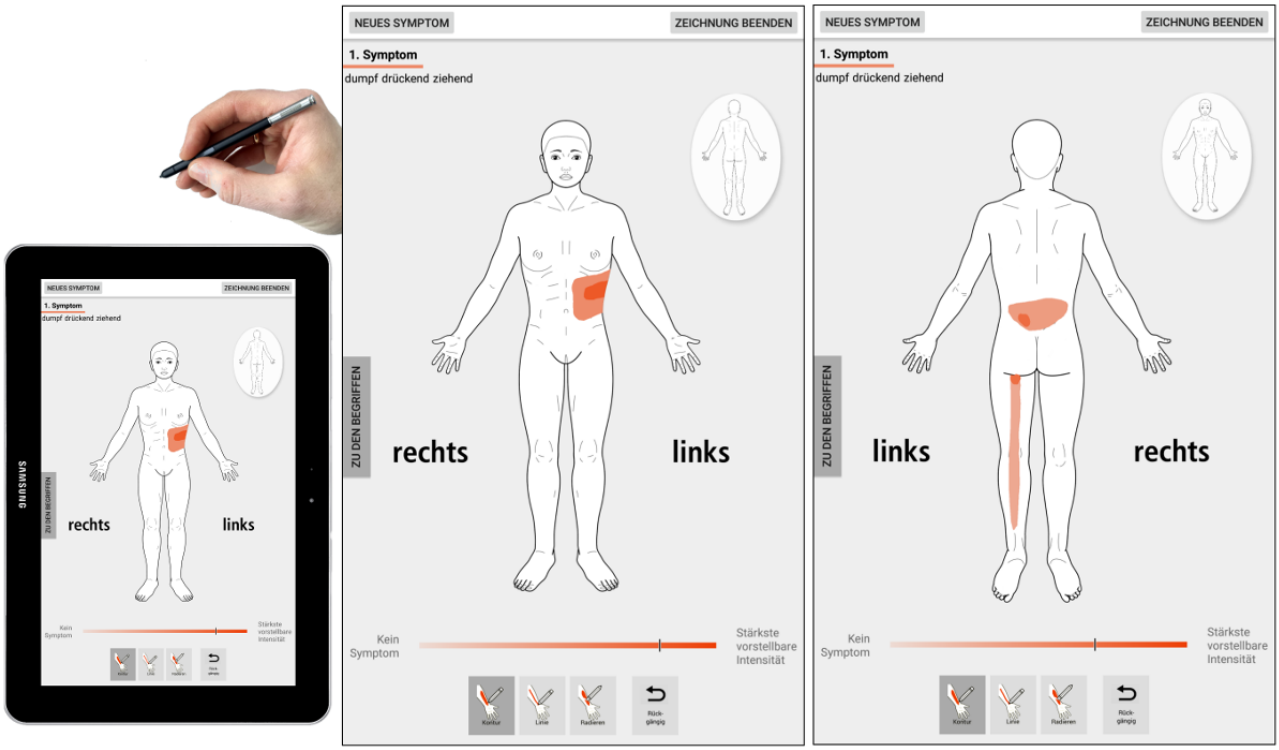
All electronic PDs were acquired on tablet PCs type Samsung Galaxy Note 2014 edition 10.1 (SM-P600) running Android 5.1.1 with an electronic pen (stylus) based on inductive digitizing technology. The tablet had a 10.1-inch touch screen with a resolution of 800×1280 pixels, and its stylus was used for all data entry. In contrast to entering data by finger on the capacitive touchscreen, the tablet records stylus interactions with a separate inductive digitizer, which allows for a higher resolution while eliminating unwanted activation of the screen, for example, by the palm.

Software App

We used SymptomMapper [17], a software app developed by our group to acquire the PDs from both patients and doctors (see Figure 1). The app consisted of 3 different modules: drawing instructions, symptom specification, and drawing. App versions for patients and doctors used the same modules but with slightly different content. The average time to complete a drawing ranges from 1 to 10 min depending on the level of details and number of symptoms drawn. We have previously shown that SymptomMapper has a good usability for patients and doctors and that test-retest reliability of symptom drawings by chronic pain patients have fair reproducibility for the exact symptom pattern but excellent reproducibility for symptom extent [17].

In the symptom specification module, users were asked to specify any pain-related symptom in an iterative process. They first chose the type of sensation from the following list of descriptors (in German): burning, cold, cramping, dull, electric, heavy, hot, numb, pressing, pricking, shooting, stabbing, tender, throbbing, tingling, and tugging. Next, they rated the intensity of the sensation on a visual analog scale, ranging from "no symptom" to "strongest imaginable intensity." Finally, they entered the perceived depth of the sensations by choosing among different depth descriptors.

Figure 1. Graphical user interface of the SymptomMapper app that was used in our study. Its drawing module allows for quick and easy data entry without previous training, a crucial prerequisite when studying patients in acute pain situations. Sides are emphasized by the words left ("links") and right ("rechts"). Doctors and patients used the same app for their pain drawings.



In the drawing instructions module, the user was asked to color every point of the body outline where the specified pain sensation or related symptom was present and to use all available body views. Other ways to mark a body region, such as hatching, ticking, or marking by arrows or other symbols, were explicitly prohibited.

In the final module, users were shown a body outline of matched sex to draw the location of the symptom or finding specified in the previous module. Drawings could only be made using the tablet's stylus, and drawing was restricted by the app to within the borders of the body outline. Adherence to the other drawing instructions was checked by the author supervising the drawing process (NS or AA). After finishing the drawing, users could either choose to end data entry or to add another symptom, which would bring them back to symptom specification.

The doctors' version of the app differed only slightly from the patients', in that its symptom specification module provided the user with a list of common pain-related diagnostic findings in addition to the list of pain descriptors.

Data Analysis

Statistical Analysis

All statistical calculations were done in R version 3.4.3 (R Foundation for Statistical Computing, Vienna, Austria) [18] and Microsoft Excel (Microsoft Corporation, Redmont, WA) using the Real Statistics Resource Pack release 5.4.1 [19]. PDs were converted from Portable Network Graphics format to Neuroimaging Informatics Technology Initiative format [20], with a custom-written Python script (Python 2.7, Python Software Foundation [21]), and analyzed using FMRIB Software Library (FSL) version 5.0 (FMRIB Analysis Group, Oxford University, UK) [22]. Final figures were prepared using VINCI 4.86.0 (Max Planck Institute for Metabolism Research, Cologne, Germany) [23] and GNU Image Manipulation Program version 2.8.16 (GIMP, The GIMP Team) [24].

Impact on Understanding of Pain and Clinical Decision

The average of doctors' ratings of improvement in understanding and influence on clinical decision were individually tested for their difference from zero using a 1-sided 1-sample *t* test. Here and in all further statistical tests, a *P* value of .05 or less was considered significant.

Pain Drawing Analysis

We extracted the following data from each PD for statistical analysis: (1) number of drawn pixels (pain extent), (2) number of clusters, (3) number of body views used in PD, (4) number of symptom descriptors, (5) average intensity per symptom, and (6) widespread pain index [25]. The first 2 quantities were also calculated for images thresholded at pain intensity larger or equal to 6 to focus on the most severe symptoms. The pain extent (thresholded and unthresholded) was normalized to percent template surface for each view by dividing the pixel count by the total number of pixels of the respective view of the body outline. Widespread pain index (WPI) was calculated by masking the PD with a custom-made template of the 19 body regions used in the WPI and counting the number of nonempty body regions. Please note that signs and symptoms other than pain or paresthesia as recorded by the doctors' version of the app were not included in the PD analysis.

Comparison of Patients' and Doctors' Pain Drawings

To identify potential systematic differences between doctors' and patients' PDs, we calculated 2-sided paired *t* tests for all

quantities mentioned above. Furthermore, we assessed the similarity of doctors' and patients' PDs by means of the Jaccard index and by calculating intraclass correlation coefficients (ICCs (3,1); Shrout and Fleiss classification) for the pain area and the number of symptom clusters. Both indices were calculated separately for each body view and averaged over all body views excluding those that were empty in both the doctors' and the patients' PD. When the drawing contained multiple pain symptoms, they were merged, and the maximum intensity was used for each pixel.

We further calculated the average pain distribution of all patients for visual comparison by pixel-wise averaging of all these individual drawings using FSL Maths. Doctors' and patients' PDs were analyzed independently. Due to the large diversity of pain states and syndromes encountered in our study sample, we did not attempt a direct statistical comparison of the data. The final images were thresholded to show only those body regions where at least 10% (5/47) of all users had drawn a symptom. This was an arbitrary threshold used to reduce the impact of single drawings with very large pain areas. As doctors' and patients' drawings were thresholded the same way, it does not obscure any differences between the 2 but allows the reader to focus on relevant areas.

Exploratory Analysis of Relevant Factors

To identify factors of relevance that improved doctors' understanding of their patients' pain, we calculated cross-correlation coefficients for the ratings of improvement in understanding and the quantities derived from the patients' PDs (see above).

Results

Impact on Understanding of Pain and Clinical Decision

Knowing patients' PDs significantly improved the doctors' understanding of their patients (average rating: 4.81, SD 2.60, $P<.001$) and to a lesser extent their clinical decision (average rating: 2.68, SD 1.18, $P<.001$). Results are shown in Figure 2.

Comparison of Patients' and Doctors' Drawings

Patients drew on average 1.25 (SD 0.53) pain symptoms, a number closely matched by the doctors' who drew 1.34 (SD 0.64) symptoms. With 3.34 (SD 2.82) different pain descriptors, patients described their pain more detailed ($P=.03$) than the doctors, who used 2.43 (SD 1.30) descriptors. The average pain distribution drawn by patients and doctors and the frequencies of pain descriptors are shown in Figure 3 and Table 2, respectively. Visual comparison of the averaged PDs suggested a high similarity of doctors' and patients' drawings. Similarity analysis of the individual PDs revealed fair reproducibility for pain extent with an ICC of 0.565 (95% CI 0.459-0.655) but poor reproducibility for the number of pain clusters with an ICC of 0.368 (95% CI 0.238-0.485). The Jaccard index was 0.217 (SD 0.171). Detailed results for each body view are listed in Table 3. The poor reproducibility of the number of pain clusters also showed when we compared PD characteristics directly between the 2 groups (Table 4). Here, we found that patients drew significantly more pain clusters when comparing unthresholded ($P<.001$) and thresholded clusters ($P=.01$). Pain extent, average pain intensity, and the number of nonempty body views on the other hand showed no significant differences between patients and doctors.

Figure 2. Impact of knowing patients' pain drawings (PDs) on understanding of the pain and clinical decision making as rated by the doctors. Patients' PDs significantly improved the doctors' understanding of the pain and to a lesser but still significant extent influenced their clinical decision.

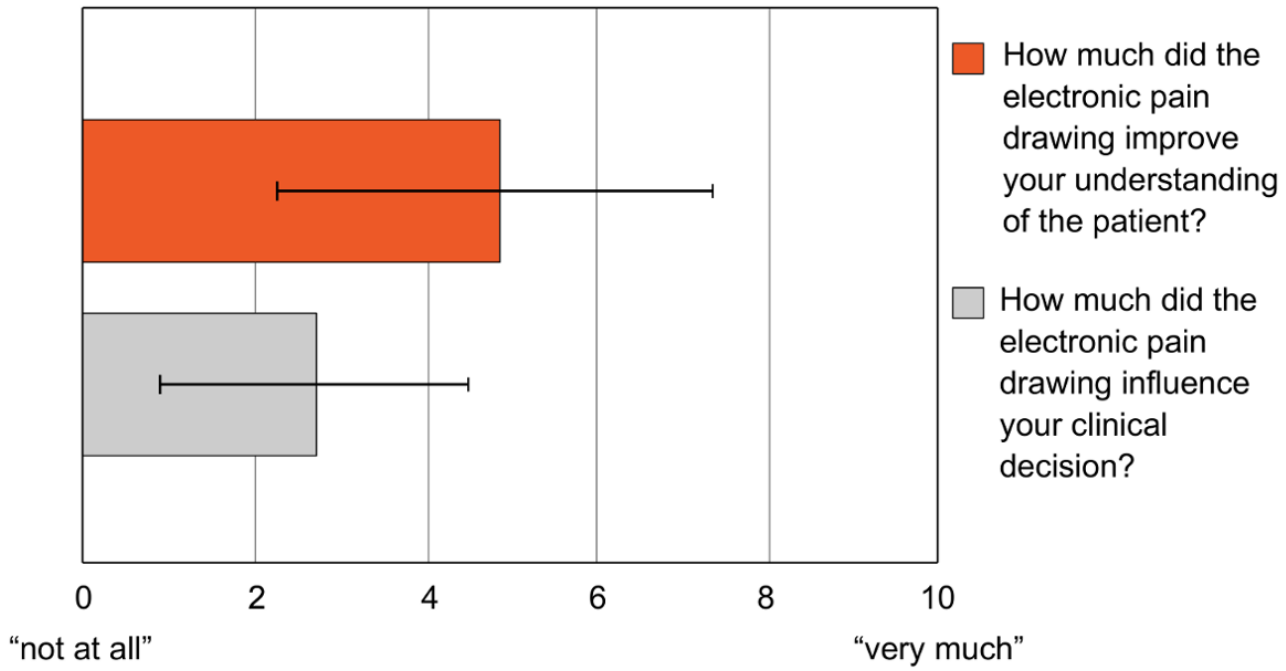


Figure 3. Descriptive comparison of patients' (top line) and doctors' (lower line) perception of pain in our final sample of 47 acute pain patients. Average pain distribution thresholded at 10% overlap between patients.

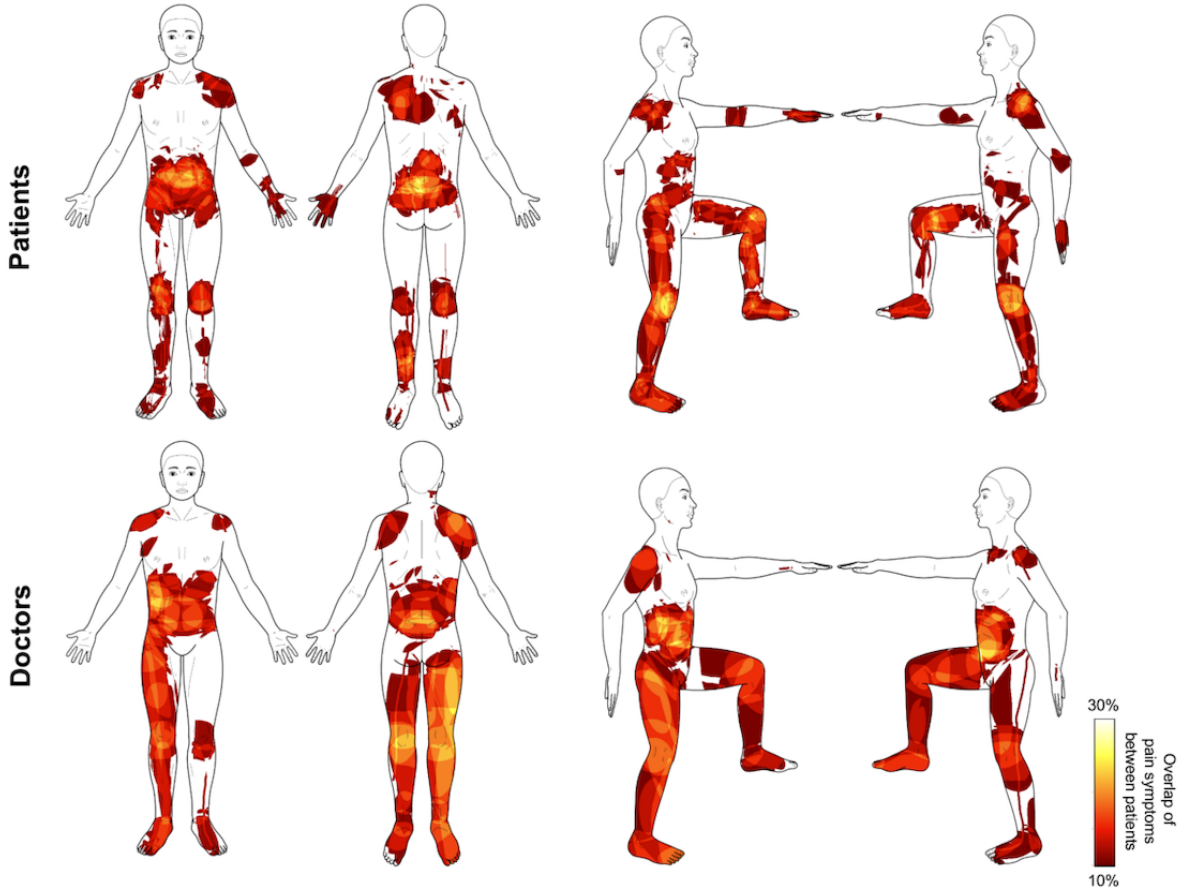


Table 2. Frequency of symptom descriptors.

Symptom descriptor	Patients, n	Doctors, n
Stinging	22	28
Burning	18	16
Pressing	16	15
Tugging	15	11
Radiating	13	6
Dull	10	7
Cramping	11	5
Tingling	10	4
Shooting	5	8
Electric	4	7
Heavy	7	1
Tender	6	2
Throbbing	7	1
Pricking	4	2
Numb	4	1
Hot	4	0
Cold	1	0
Total	157	114

Table 3. Similarity of doctors' and patients' pain drawings.

Analysis	Result
Jaccard index of symptom pattern, mean (SD)	0.22 (0.17)
ICC[a] of symptom extent (95% CI)	
Whole drawing (all body views)	0.57 (0.46-0.66)
Single views	
Front	0.51 (0.26-0.69)
Back	0.52 (0.28-0.70)
Left	0.56 (0.32-0.73)
Right	0.70 (0.51- 0.82)
ICC of number of symptom clusters (95% CI)	
Whole drawing (all body views)	0.37 (0.24-0.49)
Single views	
Front	0.32 (0.04-0.55)
Back	0.33 (0.05-0.56)
Left	0.42 (0.15-0.63)
Right	0.43 (0.17-0.64)

[a]ICC: intraclass correlation coefficient.

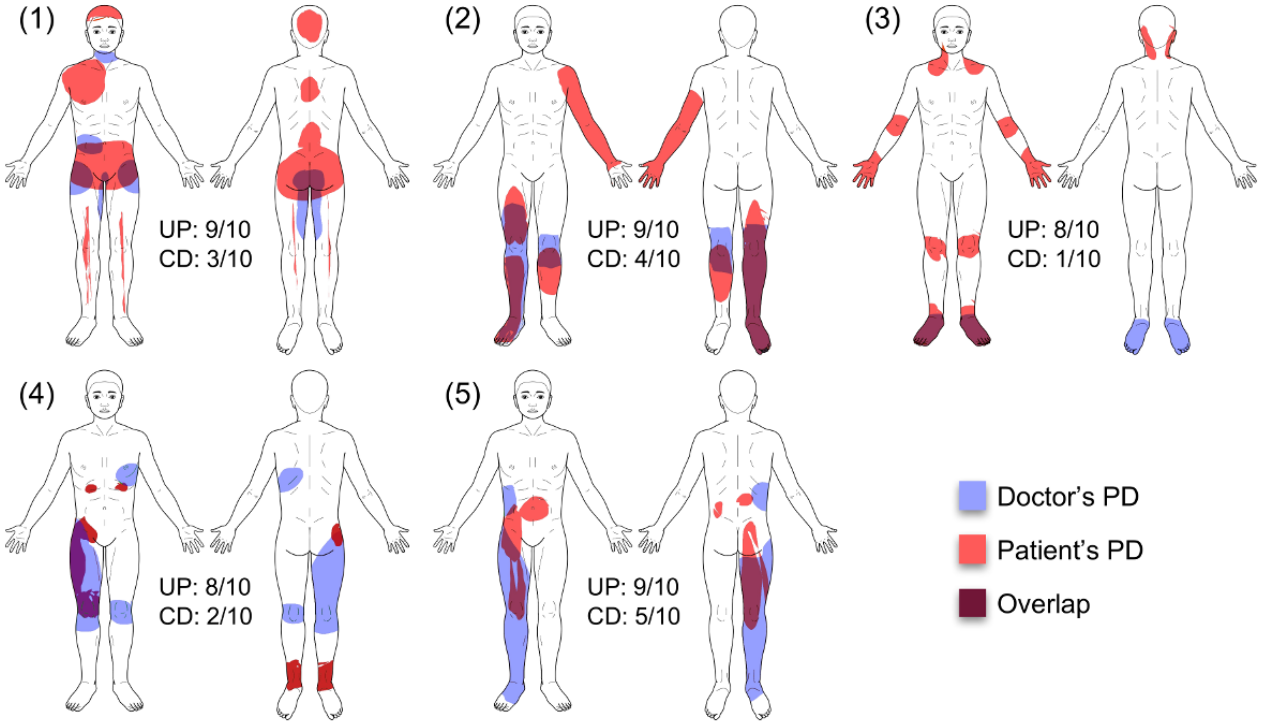
Table 4. Comparison of doctors' and patients' pain drawing characteristics.

Pain drawings characteristics	Patients, mean (SD)	Doctors, mean (SD)	*P* value[a]
Pain extent[b]	7.08 (9.66)	8.12 (14.13)	.55
Pain extent (Visual Analog Scale >6)	5.69 (9.51)	7.15 (14.03)	.39
Number of pain clusters	3.63 (3.23)	1.81 (1.33)	<.001
Number of pain clusters (Visual Analog Scale >6)	2.59 (3.18)	1.48 (1.33)	.01
Number of nonempty body views	3.40 (0.74)	3.30 (0.95)	.40
Total number of symptom descriptors	3.34 (2.82)	2.43 (1.30)	.03
Average pain intensity	7.19 (2.17)	7.46 (1.82)	.33

[a]Paired 2-tailed *t* test.

[b]In percent template surface.

Figure 4. A comparison of patients' and doctors' pain drawings (PDs) for individual patients, in which knowledge of the PD led to strong improvement of the doctor's understanding of the patient. CD: impact on clinical decision; UP: understanding of the patient.



Patient Characteristics of Representative Cases

We identified 5 patients who received either very low or very high ratings from the doctors, that is, patients in which seeing the PD was of very little or very high value for the doctors understanding. The group with high ratings is shown in Figure 4, the group with the low ratings in Multimedia Appendix 1. The individual clinical cases are discussed in Table 5.

Exploratory Analysis of Relevant Factors

Exploratory cross-correlation analysis revealed that the understanding of pain was influenced most strongly by 2 factors: the area of the pain as drawn by the patient (r=.454, P=.001) and the WPI (r=.447, P=.001) as calculated from the PD (see Figure 5). In both cases, the correlation was positive, which means that a larger pain area and higher WPI were associated with greater improvement in understanding of pain. When testing the same factors but looking at their absolute differences in doctors' and patients' drawings, area of pain showed the only significant correlation with understanding of pain (r=.313, P=.03), whereas WPI showed a tendency (r=.255, P=.08).

Table 5. Discussion of the patients in which knowledge of the PD led to strong improvement of the doctor's understanding of them.

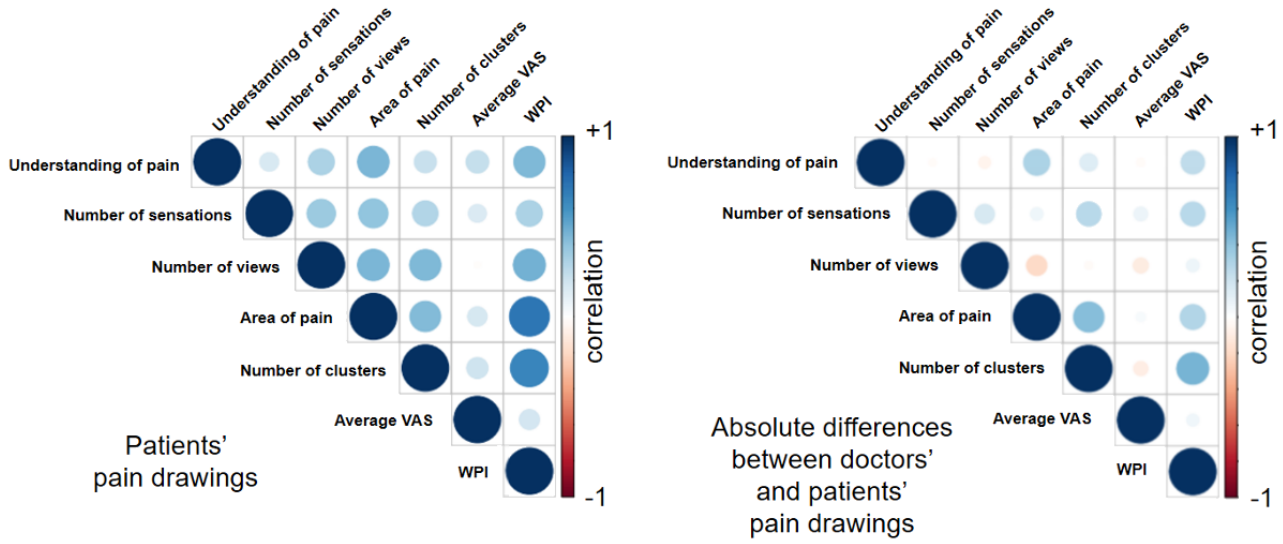
Patient	Description
Patient 1 (female, 45 years)	
Indication for hospital admission	Unexplained abdominal pain
Indication for presentation to APS[a]	Severe abdominal pain
Diagnosis	Somatization disorder
History	Diagnostic laparoscopy (10 weeks before admission) and hysterectomy (3 years ago)
Notes	Pain cluster in the neck appeared after laparoscopy and can be explained by irritation upon endotracheal intubation
Knowledge gained from patient's PD[b]	Additional pain clusters in the patient's PD supported the clinical diagnosis of somatization disorder
Implications for treatment	Referral to further psychiatric and psychosomatic treatment; discontinuation of antinociceptive therapy
Patient 2 (male, 63 years)	
Indication for hospital admission	Surgery: transcatheter aortic valve implantation for aortic stenosis
Indication for presentation to APS	Acute pain in the right leg
Diagnosis	Exacerbation of pain in the right leg with mixed nociceptive, ischemic, and neuropathic pain states in the course of peripheral arterial occlusive disease
History	Transtibial amputation of the left leg; pain syndrome of the cervical spine
Notes	No phantom limb pain in the left leg; pain cluster in the left arm and hand can be explained by pre-existing pain syndrome of the cervical spine
Knowledge gained from patient's PD	Comprehensive overview of pain clusters originating from different causes
Implications for treatment	None
Patient 3 (female, 36 years)	
Indication for hospital admission	Surgery: cyclophotocoagulation status post chronic open-angle glaucoma
Indication for presentation to APS	Acute pain in both feet
Diagnosis	Exacerbation of pre-existing pain in both feet from polyneuropathy in the course of Wegener granulomatosis
History	Wegener granulomatosis with joint involvement; polyneuropathy
Knowledge gained from patient's PD	Additional pain clusters in the patients' PD supported the clinical understanding of the widespread manifestations of the underlying disease
Implications for treatment	Referral to specialized outpatient pain treatment
Patient 4 (male, 80 years)	
Indication for hospital admission	Acute pain exacerbation with suspicion of cancer
Indication for presentation to APS	Acute pain in the right upper limb, right knee, and costal arch
Diagnosis	Exacerbation of pre-existing pain due to because of multiple cancerous osteolytic lesions from unknown primary
History	Pre-existing pain in the abovementioned regions starting 3 to 1 weeks before admission
Knowledge gained from patient's PD	Comprehensive overview of all pain sites
Patient 5 (male, 82 years)	
Indication for hospital admission	Urinary tract infection and deterioration of the patient's general condition
Indication for presentation to APS	Acute pain in the right leg and flank
Diagnosis	Exacerbation of 2 different pre-existing pain states; neuropathic pain in the right leg; visceral pain in the area of the right kidney
History	Urothelial carcinoma (UICC-Classification (Union for International Cancer Control-Classification) pTx, pNx, G3, L1, V1) and recurrent urinary tract infections under treatment with a double-J catheter; pre-existing pain in the abovementioned regions starting 3 to 1 months before admission
Knowledge gained from patient's PD	Comprehensive overview of pain clusters originating from different causes; pain pattern confirmed the neuropathic origin of the pain in the leg

Patient	Description
Implications for treatment	Start of an antineuropathic treatment

[a]APS: acute pain service.

[b]PD: pain drawing.

Figure 5. Factors with the potential to influence doctors' understanding of the patients. The left image shows correlations of pain drawing characteristics extracted from the patients' drawings, whereas the right image is based on absolute differences of those characteristics between patients' and doctors' drawings. Correlation strength is encoded in color brightness and circle size. Blue color indicates positive values and red color indicates negative values. Both the areas of pain (in percent body area) and the widespread pain index (WPI) showed significant correlations with the doctors' understanding of pain. VAS: visual analog scale.



Discussion

Overview

In this study, we have tested whether knowledge of the electronic PD of a patient in acute pain can improve the doctors' understanding of the patient's pain and potentially influence clinical decision making. We have furthermore sought to identify similarities and differences of electronic PDs made by patients and their treating pain specialists. Finally, we wanted to find specific characteristics of the drawings that have a large impact on doctors' understanding.

Impact on Understanding of Pain and Clinical Decision

Our results show that PDs significantly improved the doctors' understanding of their patients' pain based on their own judgment. On average, doctors rated this improvement with 4.81 out of 10 points. The impact on clinical decision was also significant but of smaller size (2.68 out of 10 points). There are 2 possible explanations for the fact that a relatively large improvement in understanding resulted in rather modest changes in clinical decision. First, in the majority of cases, additional pain clusters drawn by the patients do not lead the physicians to new diagnoses that would require additional medical investigation or intervention. Instead, most of these clusters reveal previously diagnosed chronic pain sources that are unrelated to the acute problem. To take one of our examples (patient number 2 from Figure 4 and Table 5), a patient requiring consultation by the APS for acute leg pain from peripheral arterial occlusive disease draws pain clusters in the leg region but also adds a large cluster in the arm. The latter stems from a previously diagnosed pain syndrome of the cervical spine. Although this additional knowledge gives a more complete image of the patient and, therefore, improves the doctor's understanding, it has little impact on the clinical decision.

Comparison of Patients' and Doctors' Drawings

The similarity of PDs from patients and their treating doctors observed in our study was considerably lower compared with studies looking at test-retest reliability of PDs in chronic pain patients. We found only fair reproducibility for pain extent (ICC: 0.57) and poor reproducibility for the number of pain clusters (ICC: 0.37). In contrast, Barbero et al and Neubert et al, using PDs from different chronic pain populations, found ICCs from 0.92 to 0.97 for pain extent [3,17]. Results regarding the reproducibility of the number of pain clusters differed even more. Here, Neubert et al report an ICC of 0.70, which is almost twice as high as the value found in our study. The Jaccard index of 0.22 (as compared with 0.46-0.49 in the above-mentioned studies) indicates an average overlap of only 22% between the PD of a patient and the associated drawing of the doctor in our study. Of course, we are comparing apples with pears here, as repeated PDs by the same person will be much more similar than PDs based on information derived from verbal and nonverbal communication. A comparison like this, however, allows us to get an estimate for how much information is lost or changed in patient-doctor communication.

Which Drawing Is "Correct"?

The above-mentioned issue raises the rather philosophical question, which of the drawings contains the correct information.

On the one hand, only the patient is able to perceive the symptoms that are expressed in the PD. As several studies have shown, generating a PD is a highly reproducible and reliable process [15,16]. Thus, there are good reasons to argue that patients' PDs constitute the ground truth in this comparison. However, other studies have based their analysis on the assumption that the doctors' drawings contain the correct information. For example, Cummings et al compared patients' PDs with drawings by their doctors and found that the usage of patients' PDs to measure pain extent, number of pain clusters, and related symptoms may lead to inaccurate diagnoses [26]. It should also be noted that anamnesis and physical examination themselves can lead to more detailed PDs by the doctors as they may reveal additional symptoms omitted by the patient.

In our opinion, both patients' and doctors' PDs are valid drawings in their own right. When analyzing their contents, however, it must be acknowledged that they do not represent the same thing and that there may be systematic differences that will influence the results of the analysis. These differences will be discussed in the next section.

Potential Sources of Systematic Differences

In our study, we did not find significant differences between patients' and doctors' PDs regarding pain extent, average pain intensity, and the number of body views used in for the PD. Together with our findings regarding the low similarity of doctors' and patients' drawings, this means that both groups drew about the same number of pixels with the same intensity but in different places on the body. We further found that doctors used significantly fewer pain clusters and symptom descriptors than patients when drawing their pain symptoms. There are several possible explanations for this. On the one hand, it is likely that doctors' drawings exhibit less specificity of pain location than the patients', simply because some level of information loss is to be expected for verbal communication of bodily symptoms. This can be seen in Multimedia Appendix 1, where clusters drawn by doctors are clearly overlapping with those of the patients but are generally larger and, thus, less specific. On the other hand, doctors may have focused more on single symptoms with larger extent. This could be explained by the principles of the APS [27]. In contrast to outpatient care settings, inpatients are well diagnosed and known to their treating medical team. Therefore, these patients are presented with a specific question to the specialists of the APS. Pre-existing and already treated pain diagnoses are not as much in the focus of interest as they would be in an outpatient setting.

Pain Drawing Characteristics That Can Improve the Understanding

Our innovative methodology of electronic PD analysis allowed us to extract a variety of information from the acquired PDs. This included characteristics such as the area of pain, average intensity, and WPI [25] as well as the number of clusters, pain sensations, and body views used in the drawing. Availability of this information enabled us to perform an exploratory analysis to identify those characteristics that significantly improved doctors' understanding of their patients' pain. We found that both pain area and WPI had the largest impact on doctors' understanding. Thus, not only drawings with more pain area received higher ratings but also those where the pain was more widespread. Both effects can be observed when comparing patients with the highest ratings (Figure 4) with those with the lowest (Multimedia Appendix 1). It is evident that the latter show much smaller pain areas and less widespreadness than the former. However, pain area and WPI also showed a high level of correlation with each other, indicating similar information content.

When looking at differences between patients' and doctors' PDs, however, only the absolute differences in pain area improved the doctors' understanding. Thus, an over- or underestimation of pain area by the doctor made seeing the patient's PD valuable for the doctor. Among other things, this finding indicates that doctors do consider information from the patients' PD as being "correct."

Advantages of Electronic Pain Drawings

Although our study was not aimed at comparing electronic PDs with their pen-on-paper counterparts, we would nevertheless like to emphasize the advantages of using the electronic version. In the last 10 years, several research groups have developed PD apps to be used on tablet computers [3-7,28]. A central advantage of the electronic drawings acquired this way is the possibility to analyze results right after completion of the drawings and without the need for time-consuming digitization. Although measuring pain area is possible for pain-on-paper drawings by using grid-based methods [13,29], such analyses usually take several minutes and require the doctor to sit down at a table with adequate lighting. Furthermore, the calculation of more complicated but relevant variables, such as average pain intensity or pain overlap as used in our paper, would take even longer to extract from pen-on-paper drawings. Finally, well-designed PD apps allow the user to zoom in and, thus, can be used by people with severe visual impairments for whom conventional drawings would be challenging or even impossible.

Limitations

Although our study has reached the planned aims, there were some limitations that we could not avoid. First, our sample consisted of inpatients in acute and often severe pain. Thus, the accuracy of completing PDs may have been lower compared with, for example, chronic pain patients that have had some time to adapt to their pain. Of course, such lower accuracy will also influence all further analyses, for example, regarding similarity of patients' and doctors' PDs. Second, our procedure of rating the improvement in understanding and influence on clinical decision was suboptimal as only the treating doctor was allowed to give a rating and this rating was anonymous. This made it impossible to assess potential bias, for example, by certain doctors, giving only good or bad ratings. Furthermore, we did not ask for the explicit reasons why a PD was considered helpful or why its knowledge did or did not influence clinical decision. Third, our sample size of 47 patients was rather small and may have led to false-negative results in all analyses directly comparing patients with doctors. Finally, the fact that all doctors that rated the impact of the patients' drawings on their understanding and clinical decision also prepared drawings themselves may be seen as a confound. Although we believe that the instructions for rating were unmistakably aimed at the

impact of seeing the patients' drawing (and not, eg, of comparing the patients with their own drawing), we cannot rule out the possibility that the act of drawing may have confounded the rating in some individuals.

Future studies should assess the importance of patients' PDs independently from that of PDs made by the doctors, that is, compare doctors using the app with those not using it. Furthermore, it would be desirable to assess improved understanding of the patient by more objective means than self-report.

Conclusions

We have shown that in a clinical setting, electronic PD can improve doctors' understanding of patients in acute pain situations based on their own judgment. The ability of electronic PDs to visualize differences between doctors' and patients' conception of pain has the potential to improve doctor-patient communication.

Acknowledgments

The authors would like to thank the German Academic Exchange Service and Horst Görtz Foundation for their financial support.

Conflicts of Interest

None declared.

Multimedia Appendix 1

A comparison of patients' and doctors' pain drawings (PDs) for individual patients, in which knowledge of the PDs led to small or no improvement of the doctor's understanding of the patient.

[PNG File, 1MB - mhealth_v7i1e11412_app1.png]

References

1. Giordano J, Abramson K, Boswell MV. Pain assessment: subjectivity, objectivity, and the use of neurotechnology. Pain Physician 2010;13(4):305-315 [FREE Full text] [Medline: 20648198]
2. Palmer H. Pain charts; a description of a technique whereby functional pain may be diagnosed from organic pain. N Z Med J 1949 Apr;48(264):187-213. [Medline: 18138692]
3. Barbero M, Moresi F, Leoni D, Gatti R, Egloff M, Falla D. Test-retest reliability of pain extent and pain location using a novel method for pain drawing analysis. Eur J Pain 2015 Sep;19(8):1129-1138. [doi: 10.1002/ejp.636] [Medline: 25565607]
4. Serif T, Ghinea G. Recording of time-varying back-pain data: a wireless solution. IEEE Trans Inf Technol Biomed 2005 Sep;9(3):447-458. [Medline: 16167699]
5. Ghinea G, Spyridonis F, Serif T, Frank AO. 3-D pain drawings-mobile data collection using a PDA. IEEE Trans Inf Technol Biomed 2008 Jan;12(1):27-33. [doi: 10.1109/TITB.2007.903266] [Medline: 18270034]
6. Jaatun EA, Haugen DF, Dahl Y, Kofod-Petersen A. Proceed with caution: transition from paper to computerized pain body maps. In: Procedia Computer Science. 2013 Apr Presented at: The 3rd International Conference on Current and Future Trends of Information and Communication Technologies in Healthcare; 21-24, 2013; Ontario, Canada p. 398-406. [doi: 10.1016/j.procs.2013.09.052]
7. Boudreau SA, Badsberg S, Christensen SW, Egsgaard LL. Digital pain drawings: assessing touch-screen technology and 3D body schemas. Clin J Pain 2016 Feb;32(2):139-145. [doi: 10.1097/AJP.0000000000000230] [Medline: 25756558]
8. Bayam L, Ahmad MA, Naqui SZ, Chouhan A, Funk L. Pain mapping for common shoulder disorders. Am J Orthop (Belle Mead NJ) 2011 Jul;40(7):353-358. [Medline: 22013572]
9. Elson DW, Jones S, Caplan N, Stewart S, St Clair Gibson A, Kader DF. The photographic knee pain map: locating knee pain with an instrument developed for diagnostic, communication and research purposes. Knee 2011 Dec;18(6):417-423. [doi: 10.1016/j.knee.2010.08.012] [Medline: 20850976]
10. Tucker KJ, Fels M, Walker SR, Hodges PW. Comparison of location, depth, quality, and intensity of experimentally induced pain in 6 low back muscles. Clin J Pain 2014 Sep;30(9):800-808. [doi: 10.1097/AJP.0000000000000026] [Medline: 25098553]
11. Gioia F, Gorga D, Nagler W. The value of pain drawings in the care of neck and back pain. J Back Musculoskelet Rehabil 1997 Jan 01;8(3):209-214. [doi: 10.3233/BMR-1997-8305] [Medline: 24572821]
12. Sivik TM. Personality traits in patients with acute low-back pain. A comparison with chronic low-back pain patients. Psychother Psychosom 1991;56(3):135-140. [doi: 10.1159/000288546] [Medline: 1836880]
13. Fordyce WE, Brockway JA, Bergman JA, Spengler D. Acute back pain: a control-group comparison of behavioral vs traditional management methods. J Behav Med 1986 Apr;9(2):127-140. [Medline: 2940370]
14. Hüllemann P, Keller T, Kabelitz M, Freynhagen R, Tölle T, Baron R. Pain drawings improve subgrouping of low back pain patients. Pain Pract 2017 Dec;17(3):293-304. [doi: 10.1111/papr.12470] [Medline: 27334429]

15. Southerst D, Stupar M, Côté P, Mior S, Stern P. The reliability of measuring pain distribution and location using body pain diagrams in patients with acute whiplash-associated disorders. J Manipulative Physiol Ther 2013 Sep;36(7):395-402. [doi: 10.1016/j.jmpt.2013.05.023] [Medline: 23891480]
16. Leoni D, Falla D, Heitz C, Capra G, Clijsen R, Egloff M, et al. Test-retest reliability in reporting the pain induced by a pain provocation test: further validation of a novel approach for pain drawing acquisition and analysis. Pain Pract 2017 Dec;17(2):176-184. [doi: 10.1111/papr.12429] [Medline: 26992099]
17. Neubert T, Dusch M, Karst M, Beissner F. Designing a tablet-based software app for mapping bodily symptoms: usability evaluation and reproducibility analysis. JMIR Mhealth Uhealth 2018 May 30;6(5):e127 [FREE Full text] [doi: 10.2196/mhealth.8409] [Medline: 29848470]
18. R Project. The R Project for Statistical Computing URL: https://www.r-project.org/
19. Real Statistics. 2018. Real statistics using Excel URL: http://www.real-statistics.com [accessed 2018-01-18] [WebCite Cache ID 6wYxPcqQO]
20. Neuroimaging Informatics Technology Initiative. URL: https://nifti.nimh.nih.gov/ [accessed 2017-07-11] [WebCite Cache ID 6rsV42huH]
21. Python Software Foundation. URL: https://www.python.org [accessed 2018-03-15] [WebCite Cache ID 6xvuWs8iZ]
22. FMRIB Software Library v5.0. Oxford, UK URL: http://fsl.fmrib.ox.ac.uk/fsl/ [WebCite Cache ID 6rsVGH76M]
23. Max Planck Institute for Metabolism Research Cologne. VINCI 4. URL: http://vinci.sf.mpg.de/doc/vinci-about.html?1.2 [accessed 2018-06-27] [WebCite Cache ID 70TOexyCc]
24. The GNU Image Manipulation Program Team. GNU Image Manipulation Program (GIMP) version 2.8.16. URL: https://www.gimp.org/ [accessed 2018-06-27] [WebCite Cache ID 70TNm1ORs]
25. Wolfe F, Clauw DJ, Fitzcharles M, Goldenberg DL, Katz RS, Mease P, et al. The American College of Rheumatology preliminary diagnostic criteria for fibromyalgia and measurement of symptom severity. Arthritis Care Res (Hoboken) 2010 May;62(5):600-610 [FREE Full text] [doi: 10.1002/acr.20140] [Medline: 20461783]
26. Cummings GS, Routan JL. Accuracy of the unassisted pain drawings by patients with chronic pain. J Orthop Sports Phys Ther 1987;8(8):391-396. [Medline: 18797035]
27. Kishore K, Agarwal A, Gaur A. Acute pain service. Saudi J Anaesth 2011 Apr;5(2):123-124 [FREE Full text] [doi: 10.4103/1658-354X.82777] [Medline: 21804789]
28. Cruder C, Falla D, Mangili F, Azzimonti L, Araújo LS, Williamon A, et al. Profiling the location and extent of musicians' pain using digital pain drawings. Pain Pract 2018 Jan;18(1):53-66. [doi: 10.1111/papr.12581] [Medline: 28466572]
29. Gatchel RJ, Mayer TG, Capra P, Diamond P, Barnett J. Quantification of lumbar function. Part 6: the use of psychological measures in guiding physical functional restoration. Spine (Phila Pa 1976) 1986;11(1):36-42. [Medline: 2939568]

Abbreviations

APS: acute pain service
FSL: FMRIB Software Library
ICC: intraclass correlation coefficient
NRS: numeric rating scale
PC: personal computer
PD: pain drawing
WPI: widespread pain index

Edited by G Eysenbach; submitted 27.06.18; peer-reviewed by S Boudreau, J Kahn; comments to author 05.09.18; revised version received 05.10.18; accepted 08.10.18; published 10.01.19.

Please cite as:
Shaballout N, Aloumar A, Neubert TA, Dusch M, Beissner F
Digital Pain Drawings Can Improve Doctors' Understanding of Acute Pain Patients: Survey and Pain Drawing Analysis
JMIR Mhealth Uhealth 2019;7(1):e11412
URL: https://mhealth.jmir.org/2019/1/e11412/
doi:10.2196/11412
PMID:30632970

Original Paper

An mHealth App for Users with Dexterity Impairments: Accessibility Study

Daihua Yu[1*], PhD; Bambang Parmanto[1*], PhD; Brad Dicianno[2*], MD, MS

[1]Health & Rehab Informatics, Department of Health Information Management, University of Pittsburgh, Pittsburgh, PA, United States

[2]Department of Physical Medicine and Rehabilitation, University of Pittsburgh, Pittsburgh, PA, United States

[*]all authors contributed equally

Corresponding Author:
Bambang Parmanto, PhD
Health & Rehab Informatics
Department of Health Information Management
University of Pittsburgh
6026 Forbes Tower
Pittsburgh, PA, 15260
United States
Phone: 1 412 383 6649
Fax: 1 412 383 6655
Email: parmanto@pitt.edu

Abstract

Background: A mobile health (mHealth) system called iMHere (interactive mobile health and rehabilitation) was developed to support individuals with chronic conditions and disability in their self-management regimens. The initial design of iMHere, however, lacked sufficient accessibility for users with a myriad of dexterity impairments. The accessibility of self-management apps is essential in ensuring usability.

Objective: This study aims to increase the usability of the iMHere system for users with dexterity impairments by increasing the app's accessibility.

Methods: We targeted the accessibility redesign by focusing on the physical presentation and the navigability of the iMHere apps. Six participants presenting with dexterity impairments were included in the usability study of the original and redesigned apps.

Results: We observed a lower number of touches needed to complete tasks (P=.09) and time to complete individual tasks (P=.06) with the redesigned app than with the original app; a significantly lower time for users to complete all tasks (P=.006); and a significantly lower error rate (P=.01) with the redesigned app than with the original app. In fact, no errors occurred with use of the redesigned app. Participant-reported overall average usability of the redesigned app (P=.007) and usability of individual modules (P<.001) were significantly higher than that of the original app due mostly to better ease of use and learnability, interface quality, and reliability.

Conclusions: Improved usability was achieved using a redesigned app. This study offers insight into the importance of personalization in enhancing the accessibility and also identifies strategies for improving usability in app development.

(JMIR Mhealth Uhealth 2019;7(1):e202) doi:10.2196/mhealth.9931

KEYWORDS

accessibility; dexterity impairments; disability; mHealth; self-management; smartphone apps; spina bifida; spinal cord injury; wellness; mobile phone

Introduction

Mobile health (mHealth) technologies, an emergent form of treatment support, offer a variety of health services and information through mobile devices such as phones and tablets [1,2]. Using mobile devices to wirelessly link remote and highly mobile populations, mHealth links users directly with health care providers and systems. Mobile apps have become a popular mode for delivering reminders to conduct self-management activities, collect data, and provide treatment support [3], all

with the goal of encouraging behavioral changes and improving health care delivery [4-6]. Some specific self-management techniques include frequent communication between patients and clinicians, as well as continuous adherence to, and adjustment of, complex treatment regimens [7].

Improving users' self-management skills is of critical importance for improving health outcomes and fostering independent living in persons with disabilities (PwDs) [8-10]. This is especially true for individuals with conditions such as spina bifida (SB) and spinal cord injuries (SCI)—a population of 442,000 in the United States—because these individuals are susceptible to secondary complications such as urinary tract infections, constipation, skin breakdown (due to paralysis and loss of sensation), and sepsis [11-13]. These secondary complications are, in part, preventable, but this requires active involvement on the part of patients, caregivers, and clinicians in adherence to self-management regimens. Therefore, developing technologies that promote self-management skills in this population could have a profound impact on health outcomes.

Investigators at the University of Pittsburgh have developed a novel mHealth system aimed at empowering persons with chronic conditions like SB and SCI and clinicians to be engaged in improving patient health [14]. This mHealth system, iMHere (interactive mobile health and rehabilitation; Figure 1), is a platform consisting of a smartphone app with a suite of modules aimed at managing various medical conditions, a Web-based clinician portal, and a communication system connecting patients with clinicians and caregivers. Some specific modules within the iMHere self-management app target medication management (MyMeds), skin integrity (SkinCare), bowel management (BMQ), bladder self-catheterization (TeleCath), and mental health (Mood).

The first version of iMHere (v1.0) did not offer sufficient accessibility—and, thus, usability—to persons with intellectual disabilities or dexterity impairments. Our prior work [15] revealed that the personalized user interface (UI) design may improve accessibility. In addition, this work generated a list of design requirements for the next iteration of the software. These design requirements were as follows:

1. Using simple and common words to ensure the readability and understandability of the text to help users better understand the app by simplifying the cognitive processes needed for completing tasks.
2. Using shortcuts in navigation to make a given task easier to complete.
3. Reducing the number of touches to reduce the burden of navigation and text entry.
4. Implementing contrasting colors between the text and background, as well as adding text-shadows, to enhance the contrast and improve readability.
5. Providing a short, one-sentence reminder offering directional guidance to prevent mistakes related to task procedures.
6. Using large icons and buttons to improve accessibility, especially for users with dexterity impairments.
7. Implementing colors to indicate the status of medications to let users know whether or not a medication is scheduled.
8. Separating the modules by color to easily signal which module is in use.
9. Using color-coded body parts on a map of the body to help users correctly specify the location of a skin problem.
10. Hiding the unused modules from the iMHere dashboard, selecting text display size, and changing contrast and display theme to make the system more personalized.

In general, users expressed a desire to have a simpler app that is easy to understand and physically use [15,16]. Because an mHealth app is a user's data point of input, such accessibility is essential for users in performing their self-management-related activities and reporting or communicating with their clinicians.

Identifying patient needs and preferences with respect to using an iMHere app delineates only one step in the process of creating greater levels of accessibility. We believe that the accessibility of mHealth can be enhanced with user-centered design and implementation. Better accessibility of smartphone apps may benefit some of the 4.04 million adults in the United States with dexterity impairments [17] whose medical problems can be addressed with iMHere.

Figure 1. Architecture of the iMHere (interactive mobile health and rehabilitation) system.

This study aims to design accessible features in the iMHere self-management app for persons with intellectual disabilities and dexterity impairments. We hypothesized that use of the redesigned app would result in significantly improved usability measures compared with the use of the original app. Results from this study will be used to develop a new version of the software.

Methods

Development Method

An earlier evaluation study [15] suggested that possible accessibility issues could be mitigated with better app design and development. We believe that the approach to designing an accessible interface involves working with two primary UI components: physical presentation and navigation (Figure 2). The physical presentation includes the following:

- *Presentation of widgets*: Focuses on the size and contrast of text and the use of buttons. The size of the widgets (icons) and text and the contrast can be adjusted to users' preferences.
- *Visual impact*: Focuses on the use of charts, images, and visual cues.

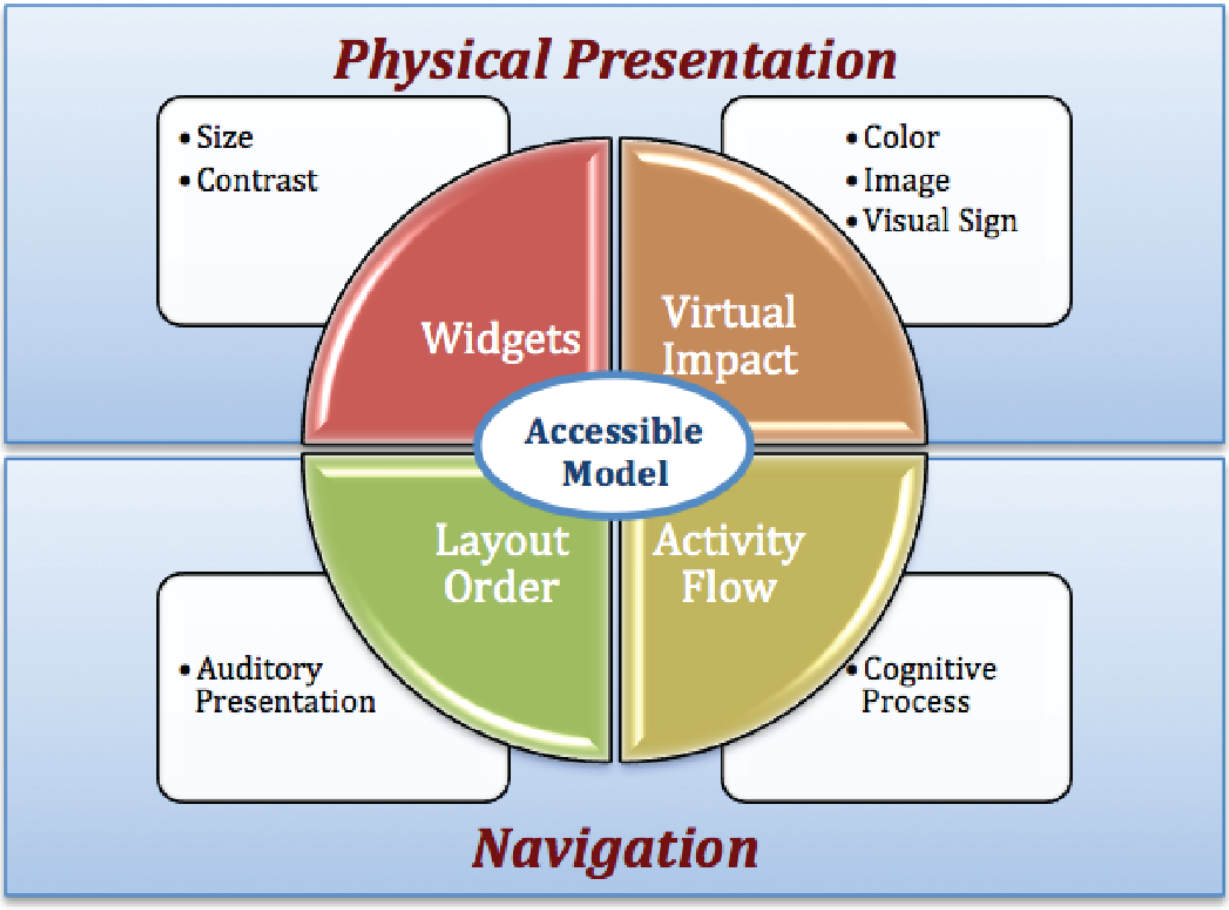
Navigation refers to activity flow and layout order in terms of effectiveness. Simple navigation is important for all users, but especially important for people with dexterity or cognitive impairments. The proposed design approaches the app's accessibility in terms of navigation from the following aspects:

- *Activity flow*: Focuses on the cognitive process, on providing straight-line experiences for a user to complete a task. Good activity flow means the user is able to effectively and efficiently locate the needed information in the smartphone app.
- *Layout order*: Focuses on the presentation of individual screens. Placing related information in close proximity makes it easier for a user to understand the presented information. In addition, having consistent layouts across the modules within the app provides a smooth learning curve for users.

Usability Study Method

After the development of new accessibility features, a usability study was conducted. Inclusion criteria were as follows: users must have participated in the prior usability study [15], be aged 18-55 years, have dexterity issues in the fingers or hands, have an active condition or past history of skin breakdown from using a wheelchair or having insensate areas of skin, and be taking at least one prescription or nonprescription medication. Exclusion criteria were as follows: users having any problem in vision, hearing, or conversation that completely precluded the use of a mobile phone.

Figure 2. Four elements of the user interface.



Manufacturers have moved to replace the physical keyboard with virtual or soft keys to reduce the size and weight of smartphone devices. To not leave PwDs behind in the area of smartphone touch screen technologies, this research examined the use of apps on a smartphone with virtual or soft keys (touch screen). Specifically, this research utilized Samsung Galaxy, a lightweight, touch screen-enabled, slate format Android smartphone with no physical keyboard (dimensions: 4.82 in×2.53 in×0.55 in; weight=5.5 oz); this screen size is smaller than the current standard screen size, which is ≥5.5 in.

The Institutional Review Board of the University of Pittsburgh approved this study. All participants were asked to provide informed consent. We enrolled all nine participants (9/9, 100%) from the prior evaluation study [15] were enrolled. All of these individuals could be classified as experienced participants but had abstained from using the iMHere app for >4 months before participating in this study. This abstinence is aimed at minimizing the potential learning effects that could ostensibly carry over from the previous experiences.

The Purdue Pegboard Assessment, a popularly utilized diagnostic tool for measuring the movements of a person's fingers, hands, and arms, was used to measure the baseline for participant dexterity levels [18-22]. We used 4 tests from the Purdue Pegboard Assessment in this study. The assessment comprised 4 tests with 30-second intervals using the right hand, left hand, and both hands, yielding a composite score of "right+left+both hands." During these tests, participants were asked to pick up pins, collars, or washers from the top of the board and drop them into the peg holes. The score for each test was based on the total number of pins, collars, or washers dropped into the holes correctly. The "right+left+both" hand score was used as the basis for evaluating a participant's dexterity, with lower "right+left+both" hand scores indicating a higher degree of dexterity impairment.

This study focused on two specific modules within the iMHere self-management app: MyMeds for medication management and Skincare for skin monitoring and reporting of skin breakdown. These two modules were selected not only on the basis of their critical importance to self-management for individuals with chronic conditions like SB and SCI but also for their relative complexity.

A 1-week field trial was completed, in which participants were asked to use the two modules in their daily lives. Afterwards, a laboratory-setting evaluation and in-depth interview were conducted. A "think-aloud" protocol [23] requires participants to verbalize their thoughts as they attempt to complete the tasks, thereby allowing investigators to identify further usability or accessibility issues that need to be addressed. The "think-aloud" method required participants to describe, in words, what they see, think, do, and feel while performing the tasks needed to navigate through the two modules. The following tasks were included in the laboratory test:

- *Task 1*: Scheduling a new medication alert; this includes searching for and finding the correct medication as well as setting up a medication schedule.
- *Task 2*: Modifying a medication reminder, which includes changing the alert time for a medication.
- *Task 3*: Responding to a medication alert, which includes indicating whether the participant took a particular medication.
- *Task 4*: Scheduling an alert to remind oneself to check the skin for any issues or problems.
- *Task 5*: Responding to a skincare reminder, which involves taking a picture and describing any dermatological issues through a series of survey questions.
- *Task 6*: Setting personalized configurations for UI presentations, including choosing a preferred list of modules, modifying the reading size of text, and choosing the size of onscreen buttons.

The researcher first explained the tasks to the participant until he or she understood the details of each activity (approximately 15 min). Once the participant was well informed of his or her expectations in performing the tasks, a quantitative evaluation was performed, and the following usability measures were collected:

- *Importance ranking*: Participants were asked to rate the new accessibility features on a scale from 1 to 10 (1=most important feature; 10=the least important feature).
- *User effort*: The minimum number of times the participant needed to touch the screen to complete all tasks.
- *Individual task time*: Average time to complete a specific task.
- *Overall task time*: Average time to complete all tasks.
- *Error rate*: The number of errors or mistakes committed during all tasks.
- *Usability*: Participants were asked to complete a modified version of the Telehealth Usability Questionnaire (TUQ) [24,25]. The TUQ is a qualitative survey covering the following factors—usefulness, ease of use and learnability, interface quality, interaction quality, reliability, and satisfaction and future use [24,25]. In assessing these factors, the TUQ utilizes a 7-point Likert scale (with a value of 1 as least usable and 7 as most usable). An overall average score and individual factor scores were calculated.

An in-depth interview was subsequently conducted to gather participant feedback and impressions regarding the iMHere app.

Statistical Analysis

All the data collected from this study were uploaded to SPSS (IBM Corp. Released 2016, IBM Statistics for Windows, Version 24.0, Armonk, NY: IBM Corp) for statistical analysis. The sum and average task completion times were utilized to measure participants' performance levels. Error rate was calculated as the number of errors or mistakes divided by the total of steps taken to complete tasks. SDs were calculated to reveal any possible dispersion patterns. The results from the previous evaluation study of the originally designed iMHere app [15] were used here for comparison.

Because our sample size was smaller than 50, Shapiro-Wilk test was used to determine whether the data were normally distributed. As all data were normally distributed, paired *t* tests were utilized to evaluate differences between the original and new app with regard to usability measures. Statistical significance was set at $P<.05$.

Results

Backgrounds of Participants

Of the 9 participants from the earlier evaluation study [15], 3 were lost due to follow-up issues (ie, changed phone number or had relocated). Overall, 6 participants completed this study. Of all participants, 5 had SB and 1 had SCI. All participants with SB had some degree of cognitive impairment related to shunted hydrocephalus.

All 6 participants were right-hand dominant and all met the inclusion and exclusion criteria. All individuals with SB had spinal lesion levels at the low thoracic or lumbosacral levels. The participant with SCI had a cervical lesion level.

As shown in Table 1, all participants' "right+left+both" hand scores were below −2 SD from the mean score of general factory workers (46.76−2 SD=38.68) [26]. Participants 1, 5, 6, and 7 tried picking up pins using both hands and dropping the pins in the holes at the same time to speed up their performance. This led to scores for the "both-hand test" that were around the mean of general factory workers at 16.01. Participant 8 had experienced a traumatic SCI (C5) resulting in minimal movement of the arms, a slight movement of the thumb and index figure, and an inability to hold or pick up objects. In addition, participant 8 was unable to perform the pegboard assessment test, but could access a smartphone either using the side of the fifth digit or a stylus mounted to a custom orthosis.

Development Results

Table 2 shows the number of individuals assigning high (1-3, very important), medium (4-7, important but not essential), and low (8-10, less important) ranks for each newly developed accessibility feature.

Table 1. Background of participants (P).

Question	P01	P03	P04	P05	P07	P08
Age (in years)	36	27	25	20	33	22
Highest education	Graduate	High school	High school	High school	Undergraduate	Graduate
Gender	Female	Male	Male	Male	Female	Male
Regular phone versus smartphone	Regular	Regular	Smartphone	Smartphone	Smartphone	Smartphone
Physical keypad versus touch screen	Physical	Physical	Touch	Touch	Touch	Touch
Mobile phone experience (in years)	0-2	>5	>5	>5	>5	>5
Daily use (in minutes)	>60	>60	>60	>60	>60	>60
Pegboard score right+left+both	33.00	27.00	23.67	36.33	37.00	0.00

Table 2. Importance ranking.

#	Features	Number of individuals assigning ranks		
		Ranks 1-3	Ranks 4-7	Ranks 8-10
1	Customized app list	2	2	2
2	Customized text display size	2	1	3
3	Customized theme	0	2	4
4	Customized button size	2	3	1
5	Customized keyboard	2	2	2
6	Ability to take a picture of a pill or med bottle	2	4	0
7	Color-coding	1	2	3
8	Text guide	2	4	0
9	Voice guide	4	1	1
10	Short cut for navigation	2	2	2

Figure 3. Screenshots of the use of color-coding at the app level. (Source: Created by the authors).

1. *Customized app list*: This feature provides the ability for a user to hide or show a selected module from the home screen. Overall, 67% (4/6) participants thought that the first feature was important to hide the TeleCath and BMQs apps because they did not need to catheterize the bladder (TeleCath) or perform bowel management (BMQs).
2. *Size of display text*: A user can specify his or her minimal and comfortable reading size. This display size is then used as the foundation for all other configuration parameters for text display in iMHere modules. Overall, 50% (3/6) participants thought that using customized text size was important; participants 1, 3, and 8 ranked this feature 2, 3, and 4, respectively.
3. *Customized theme*: The feature allows the user to select his or her preferred background and text color. Although all participants reported liking this feature, 67% (4/6) participants, that is, participants 1, 4, 7, and 8, thought it to be unnecessary for improving the accessibility of the modules. These participants ranked this feature as 10, 10, 9, and 8, respectively.
4. *Customizedbutton size*: The system asks the user to press his or her index finger on the screen to record his or her fingertip size. This touch size was used as the minimum target size for buttons or icons in the accessible design. Overall, 83% (5/6) participants thought this feature was important. Participants 4 and 8—notably individuals who presented with a higher degree of dexterity impairments—ranked it as the second most important accessibility feature.
5. *Customized keyboard*: A customized keypad with softer keys, larger key sizes, and preconfigured characters was designed to reduce the number of required touches on the smartphone screen. When using the customized keypad to enter “2 tablets,” of a medication, for instance, the users would touch “2” and “tablet.” This 2-touch entry can be contrasted with the 8-touch entry necessitated by using a traditional keypad for text entry. Overall, 67% (4/6) participants identified this feature as important for them. In particular, participant 8 (with severe dexterity

impairments) ranked the customizable keyboard as the most important feature.

6. *Ability to take a picture of a pill or bottle*: This feature provides the ability for a user to take a photo of a pill or medication bottle and upload it into his or her medication schedule. With this feature, a user can "double verify" the medication is correct by comparing it with a picture before taking his or her prescribed dose. Overall, 33% (2/6) participants ranked this feature as one of the most important.
7. *Color-coding*: As suggested by participants in the earlier evaluation study [15], color-coding was utilized in the new design to help a user navigate within the modules. For example, the title for the SkinCare module has been highlighted in red, and all screens under the SkinCare module now have a red bar to remind the user which module is being used (Figure 3). Participant 5 indicated that this feature was very important for him, as it provided a way to remember which module he was using. This participant ranked the color-coding feature as 3. Participants 3 and 6 thought this feature was important but might not be essential. Participant 7 thought this feature might be beneficial to users with intellectual disabilities.
8. *Text guidance*: Text containing self-training instructional notes is displayed on the screen and highlighted in a particular color (such as orange in Figure 2). Participants 3 and 4 ranked the text guidance as a very important feature to them, ranking this feature as 2 and 3, respectively. The remainder thought the text guidance was important but not essential, providing respective rankings of 4 and 6.
9. *Voice guidance*: Using text-to-speech technology, users can listen to text guidance as audio output. Participants 4, 5, 7, and 8 (ie, 4/6, 67%, participants) thought this voice guidance ability was important, ranking it as 3, 1, 1, and 3, respectively.
10. *Navigational short cut*: The newly designed app allows for personalization on the level of navigation. For example, the system checks the database for personalized settings first (Figure 4). If no personalized settings are found, the system will then lead the new user to set his or her preferences before going to the home screen (a list of modules). Overall, 33% (2/6) participants indicated that the ability to create shortcuts in navigation was very important to them. Participants 1 and 5 ranked this feature as 1 and 3, respectively, while participants 5 and 8 thought this feature was important but not essential, ranking it as 4 and 7, respectively.

Usability Study Results

Table 3 displays user effort results. Overall, user effort to complete all tasks was reduced by an average of about 25% in the redesigned modules. A lower average number of touches was needed for completing tasks with the redesigned modules (mean 7.20, SD 4.82) than with the original modules (mean 10.80, SD 8.04), but this difference was not statistically significant (t_4=2.25; P=.09).

Table 4 shows individual task time results. The average time to complete individual tasks was reduced by just over 50% in the redesigned modules. Participants spent the most time on tasks that required scheduling a medication or reporting a new skin problem. Particularly, task 3, responding to a medication alert, showed only a small improvement in completion time (7.7%).

Figure 4. Navigation for personalized configuration.

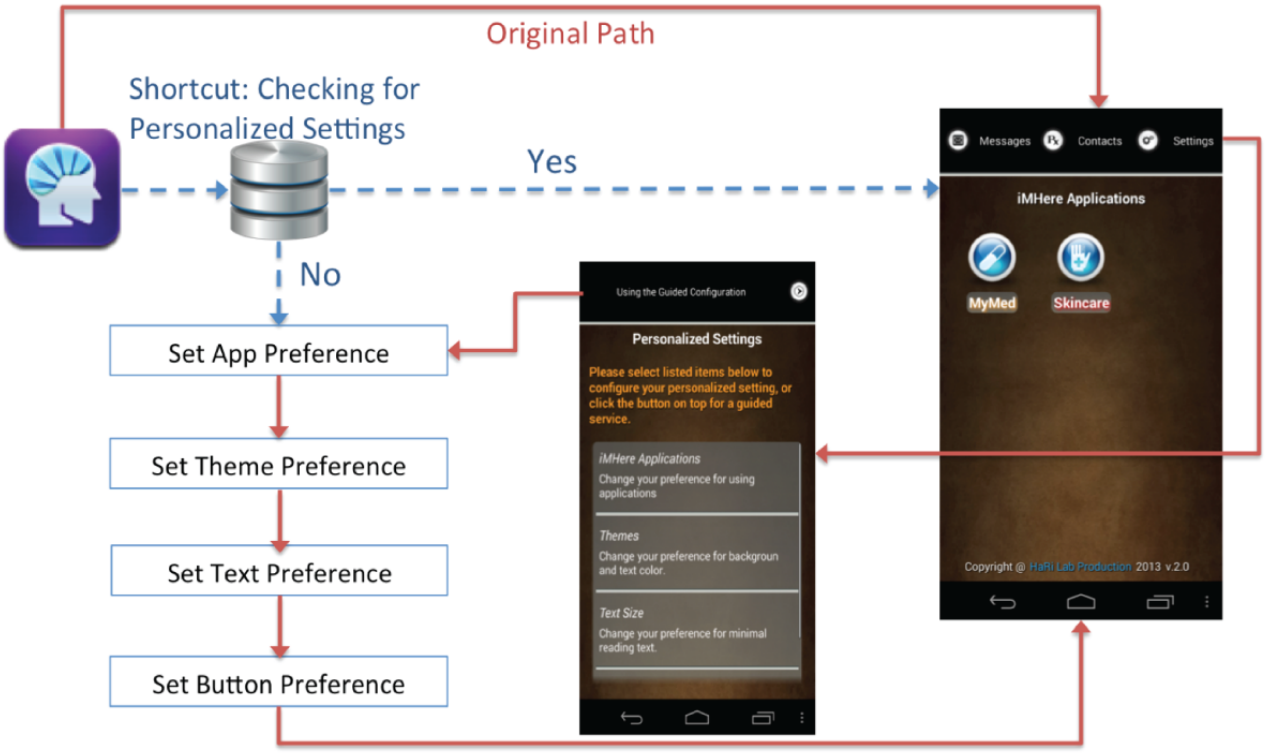


Table 3. User effort: minimum number of screen touches to complete a task.

Tasks	Original modules (n)	Redesigned modules (n)	Difference (n)	Change in effort (%)
Schedule med alert	20	11	−9	−45
Modify med alert	9	6	−3	−33
Respond to med alert	1	1	0	0
Schedule a skin check	6	5	−1	−17
Report a new skin problem	18	13	−5	−28
Average effort	10.80	7.20	−3.60	−24.60

Table 4. Individual task time: average time needed to complete individual tasks.

Task #	Tasks	Original modules (seconds), Mean (SD)	Redesigned modules (seconds), Mean (SD)	Time difference	
				Seconds	Percentage
1	Schedule medication alert	203.2 (122.8)	89.2 (49.5)	−114.1	−56.1
2	Modify medication	61.8 (43.6)	18.8 (5.6)	−43.0	−69.5
3	Respond to medication alert	2.9 (1.4)	2.7 (1.0)	−0.2	−7.7
4	Schedule skin check	42.8 (31.3)	15.7 (4.5)	−27.1	−63.4
5	Report new skin problem	147.9 (87.1)	61.3 (22.5)	−86.6	−58.5
	Average task time (seconds)	91.72 (57.24)	37.54 (16.62)	−54.20	−51.04

Table 5. Telehealth Usability Questionnaire (TUQ) scores and overall task time for each participant.

Parameter	P01	P03	P04	P05	P07	P08
TUQ score, mean (SD)						
Original modules	6.55 (0.7)	6.35 (0.9)	5.55 (0.9)	6.10 (0.7)	6.35 (0.5)	5.60 (1.1)
Redesigned modules	6.90 (0.3)	7.00 (0.0)	6.89 (0.3)	7.00 (0.0)	6.60 (0.5)	6.70 (0.5)
Overall task time in seconds, mean (SD)						
Original modules	127.20 (108.7)	68.40 (59.4)	79.40 (73.4)	104.00 (114.5)	44.60 (40.9)	66.60 (58.6)
Redesigned modules	33.50 (24.6)	38.00 (22.8)	38.83 (29.9)	53.00 (69.7)	22.17 (21.7)	26.50 (27.3)

This small increase may be attributed to the fact that this task involved only a single click on the alert screen for both the original and redesigned modules. The average time to complete individual tasks was higher using the original modules (mean 91.72, SD 81.79, seconds) than using the redesigned modules (mean 37.54, SD 36.32, seconds), but this difference was not statistically significant (t_4=2.64; P=.06).

Table 5 shows the average time in seconds for each participant to complete all 5 tasks and TUQ scores. A significantly lower average time for users to complete all 5 tasks was observed with the use of the redesigned modules (mean 35.33, SD 10.83, seconds) than with the use of the original modules (mean 81.70, SD 29.51, seconds; t_5=−4.52; P=.006). Significantly higher overall average TUQ scores were observed with the use of the redesigned modules (mean 6.85, SD 0.16) than with the use of the original modules (mean 6.08, SD 0.42; t_5=4.39; P=.007).

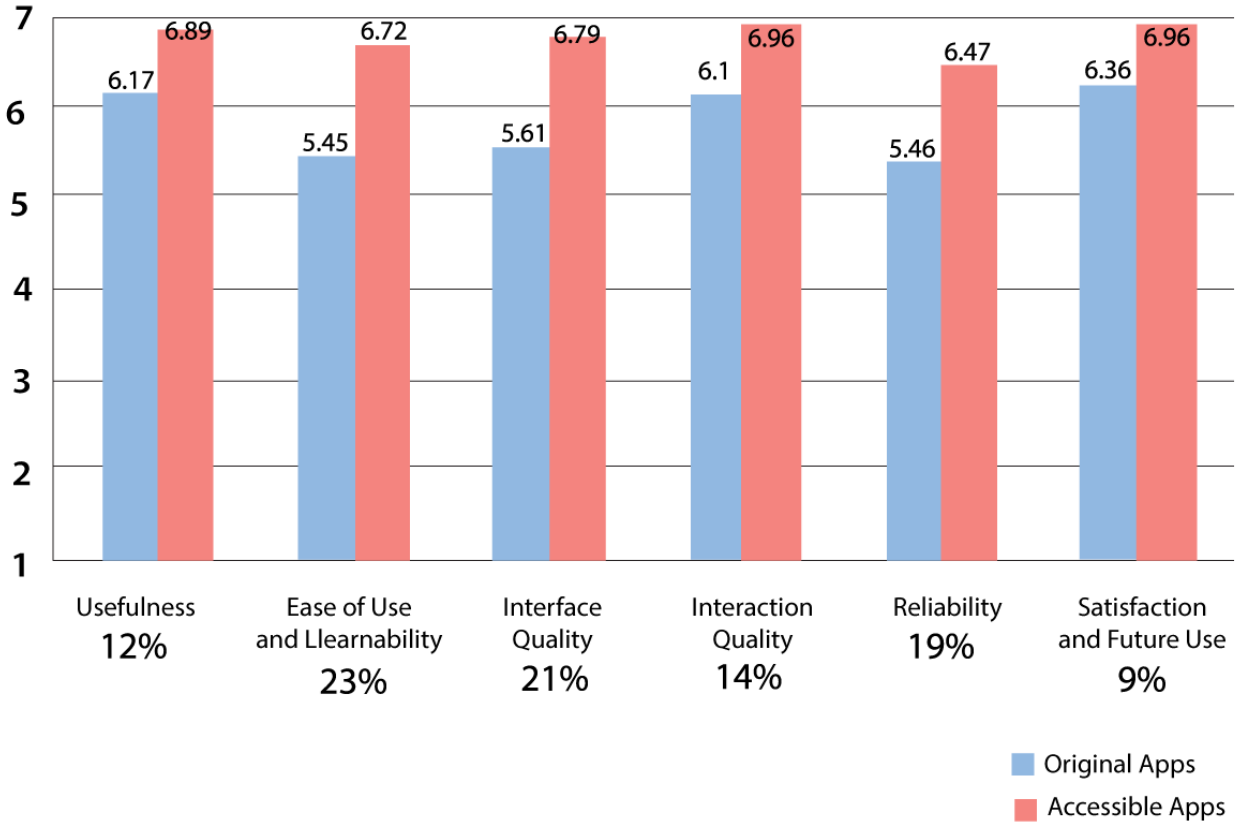
As shown in Table 6, the error rate using the redesigned modules (mean 0, SD 0) was significantly lower than that using the original modules (mean 8.51, SD 5.55; t_5=3.76; P=.01). In fact, no participants made errors using the redesigned modules.

When comparing the average subscale scores for the 6 individual domains of the TUQ with the subscale scores in the earlier evaluation study [15], usability improved significantly from the original app (mean 5.86, SD 0.40) to the redesigned app (mean 6.80, SD 0.19; t_5=−8.81; P<.001). As shown in Figure 5, pronounced improvements were noted for the factors “ease of use and learnability,” “interface quality,” and “reliability” (>15% improvements).

Table 6. Comparison of the error rates.

Participant	Original modules (%)	Redesigned modules (%)
P01	7.17	0.00
P03	0.00	0.00
P04	16.08	0.00
P05	5.75	0.00
P07	10.00	0.00
P08	12.08	0.00
Average	8.51	0.00

Figure 5. Telehealth Usability Questionnaire factors, scores, and percent increase.



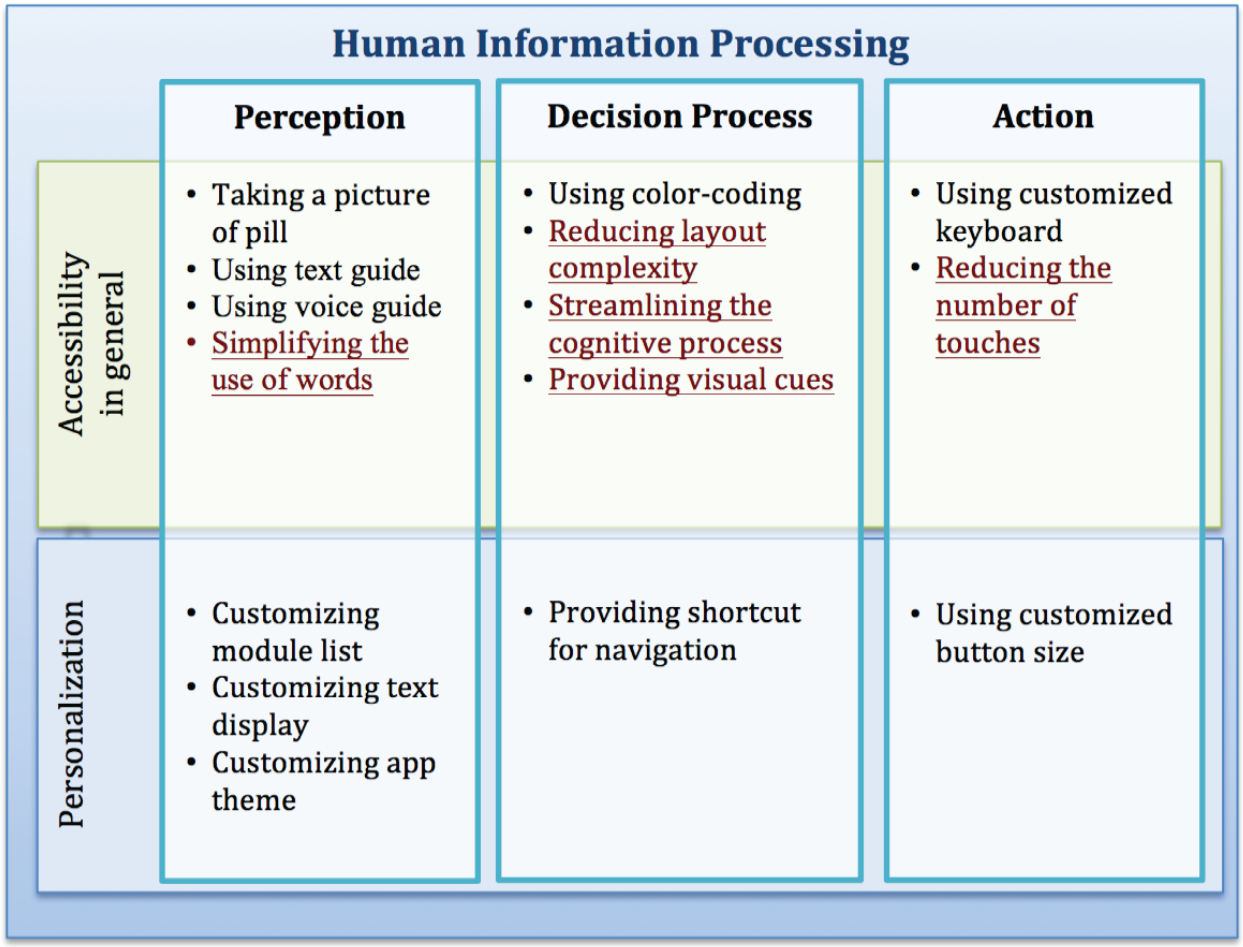
Discussion

Principal Findings

A smartphone is an ideal tool for implementing self-management programs for PwDs [27], but it does pose accessibility challenges. The size of the screen and the mobile device itself is the main obstacle to accessibility [28-30]. The small screen becomes easily cluttered when a designer wishes to fill the space with attractive text, images, and widgets [30]. This small size of the screen leads to an issue with usability [31] because it is difficult for users to read [32]. The small target or touch size, low contrast, and inappropriate text size presented on a small screen might be problematic for users with visual or dexterity problems to access [33-35]. In addition, unnecessary options and functions create difficulties for users with intellectual disabilities to understand the process, as well as to recall procedures [32].

Some of the abovementioned accessibility issues can be mitigated with design and development of a better UI. The results of this study and our prior studies [15,16] reveal strategies important to improving accessibility of smartphone apps. These strategies are presented in Figure 6, organized according to the different stages of human information processing. The text underlined in Figure 6 indicates the accessibility strategies that are important for general users; the other text indicates important features for persons with dexterity impairments.

Figure 6. Accessibility strategies.



Some features such as the customized app list, reading size, theme, and button size made the system simpler and more conducive to personal use. Particularly, the small target size of icons or buttons presented a problem for users with dexterity impairments due to the decreased strength and sensation in their fingers.

The redesign implemented in this study was based on the findings that the size of buttons has a significant impact on usability. Chen et al found that users without disabilities plateaued with a minimal button size of 20 mm and users with disabilities plateaued at 30 mm [36]. Colle and Hiszem found that 20 mm^2 buttons resulted in optimal user performance for younger participants [37], while Jin et al suggested a button size of 19.05 mm for elderly users [38]. Monterey Technologies Inc recommends the button size to be at least 19.05 mm [39]. In addition, Apple recommends a minimum target size of 44 pixels wide and 44 pixels long (by 11.64 mm) [40]. Notably, all these prior studies assumed a fixed button size.

We introduced the ability to measure the finger or touch size of a user via the smartphone as well as the ability to leverage that measurement toward creating an optimum target button or icon size. This feature is especially beneficial for users with a higher degree of dexterity impairment.

In addition to the abovementioned features, participants also found the following strategies implemented in the redesigned apps to be helpful:

- *Multiple-choice questions in place of text entry*: All participants found that making a selection was easier than entering long lines of text. Text entry, however, should always be an option in the list; if a user selects “other” he or she can then operate the text function and answer the given prompt in more detail.
- *The volume button has been appropriated as the camera button*: Except for participant 8, who was unable to hold a smartphone, all other participants liked being able to use the volume control button to take a picture, especially when taking photographs of a skin wound located in a difficult to reach area.
- *A self-directed questionnaire has been utilized to simplify the cognitive procedures of tasks*: Compared with the regular format, the redesigned modules show only one question at a time. The system automatically proceeds to the next question after a user makes a selection. In this study, 4 of 6 participants indicated that the process flow in the self-directed questionnaire was easier to understand and follow as a result of offering more guidance and fewer functions per screen.

Most notably, the average time to complete tasks in this study was reduced by about 60% in the redesigned modules. Usability

of the redesigned apps as measured by TUQ showed a significant increase. Pronounced improvements were particularly noted for the factors "ease of use and learnability," "interface quality," and "reliability." Finally, the redesigned modules were able to eliminate all errors that occurred during use of the original modules.

A surprising finding of this study was the degree of dexterity impairment identified in participants with SB. All participants had spinal lesion levels in the low thoracic or lumbosacral areas, which means that there was no paralysis of the arms or hands. Impairments in fine motor control in SB are thought to be due to the abnormal organization of the cerebral cortex [41,42]. All participants with SB, however, had pronounced impairments in dexterity, measured as >2 SDs below normative values. Little is known about the extent of fine motor control problems in SB and how it affects the use of mHealth technologies.

Limitations and Future Studies

Only a limited number of participants were involved in this study of the redesigned iMHere modules. The development and usability study follow the iterative design [43], which consists of a cycle process of prototyping, testing, analyzing, and refining a system. This study is at the later stage of the iterative cycle that follows previous studies [15,16]. By limiting the evaluation to participants from the earlier studies, we were able to probe deeper into the usability of the fundamental structure of the mHealth apps and to find majority of the usability problems [43].

The results of this study should be viewed with the nature of participants' impairments in mind. The next study should include more participants with varying levels of dexterity impairments—as well as a wider range in the diagnoses underlying these impairments—to better assess the overall acceptance and preference of the redesigned modules. In addition, more studies into the various degrees of dexterity impairments in individuals with SB, and the effect(s) of these impairments on the use of mHealth technologies, are warranted. Furthermore, future studies are warranted on the usability of the iMHere clinician portal and caregiver app—work that is conducted in parallel with studies performed on the patient app.

Conclusion

The accessibility standards and guidelines such as the Web Content Accessibility Guidelines 1.0 [44] and 2.0 [45] are mainly aimed at improving the general accessibility of the Web, not specifically of smartphone apps. The cross-platform technology for developing smartphone apps, which is based on Web technology, is increasingly popular. We plan to implement the strategies and accessibility principles in this study to the cross-platform app development environments that is based on Web technology in our future studies.

This study proposes a design and developmental model to approach accessibility through two primary elements of UI: physical presentation and navigation. A usability study showed that the effectiveness and efficiency of, and user satisfaction with, the redesigned modules significantly improved after implementing accessibility strategies into the UI design. As the results suggested, the meaningful presentation and navigation flow also helped us achieve a smoother activity flow during task completion. By extending the concept of personalization to navigation and task flow, the efficiency of users' performance could be significantly improved.

The aforementioned accessibility strategies and features could be used for other developers to design and develop smartphone apps. This paper focuses on the general principles of accessible mHealth design. Most of the UI elements can be implemented as an accessibility personalization setting of an mHealth app. We plan to implement accessibility personalization feature in our future mHealth developments.

Acknowledgments

This research was supported in part by the National Institute on Disability, Independent Living, and Rehabilitation Research grant #90RE5018 (Rehabilitation Engineering Research Center From Cloud to Smartphone: Empowering and Accessible Information and Communication Technologies and grant #90DP0064 Disability and Rehabilitation Research Project Promoting Independence & Self-management using mHealth), by the National Institutes of Health grant #1R21HD071810, and by a grant from the Craig H. Neilsen Foundation.

Conflicts of Interest

DY, BP, and BD are inventors of the iMHere mHealth system.

References

1. Cipresso P, Serino S, Villani D, Repetto C, Albani G, Mauro A, et al. Is your phone so smart to affect your state? An exploratory study based on psychophysiological measures. Neurocomputing 2012 May 01;84:23-30 [FREE Full text] [Medline: 25904163]
2. UN Foundation-Vodafone Foundation. 2009. mHealth for Development: The Opportunity of Mobile Technology for Healthcare in the Developing World URL: http://www.globalproblems-globalsolutions-files.org/unf_website/assets/publications/technology/mhealth/mHealth_for_Development_full.pdf [accessed 2018-11-15] [WebCite Cache ID 73xljThEz]
3. Kosaraju A, Barrigan CR, Poropatich RK, Casscells SW. Use of mobile phones as a tool for United States health diplomacy abroad. Telemed J E Health 2010 Mar;16(2):218-222. [doi: 10.1089/tmj.2009.0095] [Medline: 20156128]

4. Asangansi I, Braa K. The emergence of mobile-supported national health information systems in developing countries. Stud Health Technol Inform 2010;160(Pt 1):540-544. [Medline: 20841745]
5. Boyer EW, Smelson D, Fletcher R, Ziedonis D, Picard RW. Wireless Technologies, Ubiquitous Computing and Mobile Health: Application to Drug Abuse Treatment and Compliance with HIV Therapies. J Med Toxicol 2010 Jun;6(2):212-216 [FREE Full text] [doi: 10.1007/s13181-010-0080-z] [Medline: 20623215]
6. Han D, Lee M, Park S. THE-MUSS: Mobile u-health service system. Comput Methods Programs Biomed 2010 Feb;97(2):178-188. [doi: 10.1016/j.cmpb.2009.08.004] [Medline: 19762108]
7. Hays R, Kraft G, Stolov W. Chronic Disease and Disability : A Contemporary Rehabilitation Approach to the Practice of Medicine. New York: Demos Medical Publishing; 1994.
8. Lorig KR, Holman H. Self-management education: history, definition, outcomes, and mechanisms. Ann Behav Med 2003 Aug;26(1):1-7. [doi: 10.1207/S15324796ABM2601_01] [Medline: 12867348]
9. Clark NM. Management of chronic disease by patients. Annual Review Public Health 2003;24:289-313. [doi: 10.1146/annurev.publhealth.24.100901.141021] [Medline: 12415147]
10. Fairman AD, Dicianno BE, Datt N, Garver A, Parmanto B, McCue M. Outcomes of Clinicians, Caregivers, Family Members and Adults with Spina Bifida Regarding Receptivity to use of the iMHere mHealth Solution to Promote Wellness. Int J Telerehabil 2013;5(1):3-16 [FREE Full text] [doi: 10.5195/ijt.2013.6116] [Medline: 25945209]
11. National Institute of Health. 2018 Jul 06. Spina Bifida Fact Sheet URL: https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Fact-Sheets/Spina-Bifida-Fact-Sheet [WebCite Cache ID 6fWLfUlKb]
12. National Spinal Cord Injury Statistical Center (NSCISC). 2015. Spinal Cord Injury (SCI) Facts and Figures at a Glance URL: https://www.nscisc.uab.edu/Public/Facts%202015%20Aug.pdf [accessed 2016-02-23] [WebCite Cache ID 6fWMQja9n]
13. Dicianno BE, Kurowski BG, Yang JMJ, Chancellor MB, Bejjani GK, Fairman AD, et al. Rehabilitation and medical management of the adult with spina bifida. Am J Phys Med Rehabil 2008 Dec;87(12):1027-1050. [doi: 10.1097/PHM.0b013e31818de070] [Medline: 18923330]
14. Parmanto B, Pramana G, Yu DX, Fairman AD, Dicianno BE, McCue MP. iMHere: A Novel mHealth System for Supporting Self-Care in Management of Complex and Chronic Conditions. JMIR Mhealth Uhealth 2013 Jul 11;1(2):e10 [FREE Full text] [doi: 10.2196/mhealth.2391] [Medline: 25100682]
15. Yu DX, Parmanto B, Dicianno BE, Watzlaf VJ, Seelman KD. Accessibility needs and challenges of a mHealth system for patients with dexterity impairments. Disability and Rehabilitation Assistive Technology 2017 Dec;12(1):56-64. [doi: 10.3109/17483107.2015.1063171] [Medline: 26153097]
16. Yu DX, Parmanto B, Dicianno BE, Pramana G. Accessibility of mHealth Self-Care Apps for Individuals with Spina Bifida. Perspectives in Health Information Management 2015;12:1-19 [FREE Full text] [Medline: 26755902]
17. Adams PF, Martinez ME, Vickerie JL. Summary health statistics for the U.S. population: National Health Interview Survey, 2009. Vital Health Stat 10 2010 Dec(248):1-115 [FREE Full text] [Medline: 21696068]
18. Desrosiers J, Hébert R, Bravo G, Dutil E. The Purdue Pegboard Test: normative data for people aged 60 and over. Disabil Rehabil 1995 Jul;17(5):217-224. [Medline: 7626768]
19. Ozçelik IB, Purisa H, Sezer I, Mersa B, Kabakaş F, Tuncer S, et al. [Evaluation of long-term results in mutilating hand injuries]. Ulus Travma Acil Cerrahi Derg 2009 Mar;15(2):164-170 [FREE Full text] [Medline: 19353320]
20. Delp HL, Newton RA. Effects of brief cold exposure on finger dexterity and sensibility in subjects with Raynaud's phenomenon. Phys Ther 1986 Apr;66(4):503-507. [Medline: 3960975]
21. Wilson BC, Iacoviello JM, Wilson JJ, Risucci D. Purdue Pegboard performance of normal preschool children. J Clin Neuropsychol 1982 May;4(1):19-26. [Medline: 7096583]
22. Smoot B, Wong J, Cooper B, Wanek L, Topp K, Byl N, et al. Upper extremity impairments in women with or without lymphedema following breast cancer treatment. J Cancer Surviv 2010 Jun;4(2):167-178 [FREE Full text] [doi: 10.1007/s11764-010-0118-x] [Medline: 20373044]
23. Lewis C. Using the“ thinking-aloud” method in Cognitive Interface Design (Technical Report). IBM TJ Watson Research Center 1982.
24. Parmanto B, Saptono A, Pramana G, Pulantara W, Schein RM, Schmeler MR, et al. VISYTER: versatile and integrated system for telerehabilitation. Telemed J E Health 2010 Nov;16(9):939-944. [doi: 10.1089/tmj.2010.0033] [Medline: 21034239]
25. Schutte J, Gales S, Filippone A, Saptono A, Parmanto B, McCue M. Evaluation of a telerehabilitation system for community-based rehabilitation. Int J Telerehabil 2012;4(1):15-24 [FREE Full text] [doi: 10.5195/ijt.2012.6092] [Medline: 25945193]
26. Lafayette Instrument Evaluation. 2002. Purdue Pegboard Test URL: http://lafayetteevaluation.com/products/purdue-pegboard [accessed 2018-01-25] [WebCite Cache ID 6wk8LNNPu]
27. Holman H. Chronic disease--the need for a new clinical education. JAMA 2004 Sep 01;292(9):1057-1059. [doi: 10.1001/jama.292.9.1057] [Medline: 15339897]
28. Brewster S, Cryer PG. Maximising screen-space on mobile computing devices. In: CHI'99 extended abstracts on Human factors in computing systems. New York: ACM; 1999:224-225.

29. Brewster S. The design of sonically-enhanced widgets. Interacting with Computers 1998 Dec 1;11(2):211-235. [doi: 10.1016/S0953-5438(98)00028-9]
30. Brewster S. Overcoming the Lack of Screen Space on Mobile Computers. Personal and Ubiquitous Computing 2002 May 1;6(3):188-205. [doi: 10.1007/s007790200019] [Medline: 26811110]
31. O'Neill E, Kaenampornpan M, Kostakos V, Warr A, Woodgate D. Can we do without GUIs? Gesture and speech interaction with a patient information system. Pers Ubiquit Comput 2005 Nov 10;10(5):269-283. [doi: 10.1007/s00779-005-0048-1]
32. Kurniawan S, Mahmud M, Nugroho Y. A study of the use of mobile phones by older persons. In: CHI '06 Extended Abstracts on Human Factors in Computing Systems. Montreal, Quebec, Canada: ACM; Apr 22, 2006.
33. Abascal J, Civit A. A Mobile Communication for Older People: New Opportunities for Autonomous Life. The 6th ERCIM Workshop 2000;27:57-66.
34. Kane SK, Bigham JP, Wobbrock JO. Slide Rule: Making mobile touch screens accessible to blind people using multi-touch interaction techniques. New York: ACM; 2008 Presented at: Proceedings of the 10th International ACM SIGACCESS Conference on Computers and Accessibility; October 13-15, 2008; Halifax, Nova Scotia, Canada p. 73-80. [doi: 10.1145/1414471.1414487]
35. Kane SK, Jayant C, Wobbrock JO, Ladner RE. Freedom to roam: a study of mobile device adoption and accessibility for people with visual and motor disabikities. 2009 Oct 25 Presented at: Proceedings of the 11th international ACM SIGACCESS conference on Computers and accessibility; October 25-28, 2009; Pittsburgh, PA. [doi: 10.1145/1639642.1639663]
36. Chen KB, Savage AB, Chourasia AO, Wiegmann DA, Sesto ME. Touch screen performance by individuals with and without motor control disabilities. Appl Ergon 2013 Mar;44(2):297-302 [FREE Full text] [doi: 10.1016/j.apergo.2012.08.004] [Medline: 23021630]
37. Colle H, Hiszem K. Standing at a kiosk: effects of key size and spacing on touch screen numeric keypad performance and user preference. Ergonomics 2004 Oct 22;47(13):1406-1423. [doi: 10.1080/00140130410001724228] [Medline: 15513716]
38. Jin ZX, Plocher T, Kiff L. Touuch screen user interfaces for older adults: button size and spacing. In: Universal Access in Human Computer Interaction Coping with Diversity. New York: Springer; 2007:933-941.
39. Monterey TINC. Resource Guide for Accessible Design of Consumer Electronics. Washington, DC 1996:1996.
40. Apple Developer. 2014. Human Interface Guidelines URL: https://developer.apple.com/ios/human-interface-guidelines/overview/themes/ [accessed 2018-01-29] [WebCite Cache ID 6wpH6oU4h]
41. Treble A, Juranek J, Stuebing KK, Dennis M, Fletcher JM. Functional significance of atypical cortical organization in spina bifida myelomeningocele: relations of cortical thickness and gyrification with IQ and fine motor dexterity. Cereb Cortex 2013 Oct;23(10):2357-2369 [FREE Full text] [doi: 10.1093/cercor/bhs226] [Medline: 22875857]
42. Ware AL, Kulesz PA, Williams VJ, Juranek J, Cirino PT, Fletcher JM. Gray matter integrity within regions of the dorsolateral prefrontal cortical-subcortical network predicts executive function and fine motor dexterity in spina bifida. Neuropsychology 2016 Dec;30(4):492-501 [FREE Full text] [doi: 10.1037/neu0000266] [Medline: 26752120]
43. Nielsen J. Nielsen Norman Group. 2018 Jun 19. Why You Only Need to Test with 5 Users URL: https://www.nngroup.com/articles/why-you-only-need-to-test-with-5-users/ [accessed 2018-06-19] [WebCite Cache ID 70IT9rNx5]
44. Chisholm W, Vanderheiden G, Jacobs I. World Wide Web Consortium (W3C). 1999 May 05. Web Content Accessibility Guidelines 1.0 URL: https://www.w3.org/TR/1999/WAI-WEBCONTENT-19990505/ [accessed 2018-06-19] [WebCite Cache ID 70ITS4tgg]
45. Caldwell B, Cooper M, Reid LG, Vanderheiden G. World Wide Web Consortium (W3C). 2008 Dec 11. Web Content Accessibility Guidelines (WCAG) 2.0 URL: https://www.w3.org/TR/WCAG20/ [accessed 2018-06-19] [WebCite Cache ID 70ITgsLeA]

Abbreviations

iMHere: interactive mobile health and rehabilitation
mHealth: mobile health
PwDs: persons with disabilities
SB: spina bifida
SCI: spinal cord injuries
TUQ: Telehealth Usability Questionnaire
UI: user interface

Edited by C Dias; submitted 29.01.18; peer-reviewed by B Smaradottir, J Moon; comments to author 18.02.18; revised version received 21.06.18; accepted 27.08.18; published 08.01.19.

Please cite as:
Yu D, Parmanto B, Dicianno B
An mHealth App for Users with Dexterity Impairments: Accessibility Study
JMIR Mhealth Uhealth 2019;7(1):e202
URL: https://mhealth.jmir.org/2019/1/e202/
doi:10.2196/mhealth.9931
PMID:30622096

Original Paper

Barriers to and Facilitators of Engagement With mHealth Technology for Remote Measurement and Management of Depression: Qualitative Analysis

Sara Simblett[1], PhD, DClinPsy; Faith Matcham[1,2], PhD, CPsychol; Sara Siddi[3,4], PhD; Viola Bulgari[5], PhD; Chiara Barattieri di San Pietro[5,6], MA; Jorge Hortas López[7], MA; José Ferrão[8], PhD; Ashley Polhemus[8], MSE; Josep Maria Haro[3,4], PhD; Giovanni de Girolamo[5], MD; Peter Gamble[8], MBA; Hans Eriksson[9], MBA, MD, PhD; Matthew Hotopf[1,2], PhD, FRCPsych; Til Wykes[1,2], PhD; RADAR-CNS Consortium

[1]Institute of Psychology, Psychiatry and Neuroscience, King's College London, London, United Kingdom

[2]National Institute for Health Research Biomedical Research Centre for Mental Health, South London and Maudsley National Health Service Foundation Trust, King's College London, London, United Kingdom

[3]Parc Sanitari Sant Joan de Déu, Sant Boi de Llobregat, Centro de Investigacion Biomedica en Red CIBERSAM, Madrid, Spain

[4]Department of Psychiatry and Clinical Psychobiology, University of Barcelona, Barcelona, Spain

[5]IRCCS Istituto Centro San Giovanni di Dio Fatebenefratelli, Brescia, Italy

[6]Department of Psychology, University of Milano-Bicocca, Milan, Italy

[7]Research Department, QITERIA Investigación Social Aplicada, Madrid, Spain

[8]Information Technology Department, MSD Czech Republic, Prague, Czech Republic

[9]Clinical Development, Depression and Paediatrics, H Lundbeck A/S, Copenhagen, Denmark

Corresponding Author:
Sara Simblett, PhD, DClinPsy
Institute of Psychology, Psychiatry and Neuroscience
King's College London
De Crespigny Park
London, SE5 8AF
United Kingdom
Phone: 44 207 848 0762
Email: sara.simblett@kcl.ac.uk

Abstract

Background: Mobile technology has the potential to provide accurate, impactful data on the symptoms of depression, which could improve health management or assist in early detection of relapse. However, for this potential to be achieved, it is essential that patients engage with the technology. Although many barriers to and facilitators of the use of this technology are common across therapeutic areas and technology types, many may be specific to cultural and health contexts.

Objective: This study aimed to determine the potential barriers to and facilitators of engagement with mobile health (mHealth) technology for remote measurement and management of depression across three Western European countries.

Methods: Participants (N=25; 4:1 ratio of women to men; age range, 25-73 years) who experienced depression participated in five focus groups held in three countries (two in the United Kingdom, two in Spain, and one in Italy). The focus groups investigated the potential barriers to and facilitators of the use of mHealth technology. A systematic thematic analysis was used to extract themes and subthemes.

Results: Facilitators and barriers were categorized as health-related factors, user-related factors, and technology-related factors. A total of 58 subthemes of specific barriers and facilitators or moderators emerged. A core group of themes including motivation, potential impact on mood and anxiety, aspects of inconvenience, and ease of use was noted across all countries.

Conclusions: Similarities in the barriers to and facilitators of the use of mHealth technology have been observed across Spain, Italy, and the United Kingdom. These themes provide guidance on ways to promote the design of feasible and acceptable cross-cultural mHealth tools.

(JMIR Mhealth Uhealth 2019;7(1):e11325) doi:10.2196/11325

KEYWORDS

acceptability; barriers; depression; facilitators; feasibility; mHealth; qualitative

Introduction

Depression is a major cause of disability in Europe and worldwide. It is associated with a range of negative outcomes including premature mortality [1], reduced quality of life [2], loss of occupational function [3], poor social integration and loneliness [4], and increased risk of other psychiatric problems such as comorbid anxiety disorders [5] and alcohol dependence [6]. Experiences of depression are commonly episodic, and the risk of recurrence following an initial episode is high [7].

With the global increase in availability of mobile phones and wearable devices [8,9], there is potential for more frequent health assessment that might help identify signals indicative of relapse, such as changes in behaviors, circadian rhythms, stresses, or symptoms [10]. Identification of such indicators might lead to fast treatment, possibly preventing relapse through early interventions [10]. However, a critical challenge is user acceptance of these technologies, particularly, the extent to which people are willing to engage with the technologies, considering the level of intrusiveness and possible discomfort. By engagement, we refer to the extent and manner in which people actively use resources. The level of engagement should specifically be important for people who are experiencing depression, as symptoms such as lack of motivation and interest to carry out activities (anhedonia) have been shown to influence the pursuit of potential rewards [11]. Clinical trials of mobile technologies for individuals with depression have highlighted engagement as a specific challenge [12].

In order to build on the potential of mobile technologies, we need to determine the views of people living with or having a history of depression, so that these views can be embedded at the start of the mobile health (mHealth) technology-designing process to ensure maximum applicability, acceptability, and adoption. This study builds on a recent systematic review of barriers to and facilitators of engagement with remote measurement technology [13]. This review used data from single-country studies, but engagement with mHealth technology may also be influenced by cultural context [14] in addition to individual differences. These differences would affect building of platforms that span across Europe and would need to be taken into account in the design of mHealth systems to maximize the value of interventions. This study aimed to identify these differences through focus groups from three European countries (Italy, United Kingdom, and Spain), providing an opportunity to identify a broader range of potential barriers to and facilitators of engagement and problems with adherence early in order to support the design of mHealth systems.

Methods

Design

A qualitative approach following a thematic analysis was employed to identify different experiences and potential barriers to and facilitators of engagement with mHealth technology among people with a history of or living with depression. The topic guide and coding frame were built on a recent systematic review on barriers to and facilitators of engagement with remote measurement technology [13]. Within the coding frame, several pre-established major and minor codes and subthemes emerged through the use of grounded-theory methods.

Context

Researcher Characteristics

Native speakers in all countries managed the focus groups. Coordination among the three groups was agreed upon via telephone and email contact prior to commencing the study, and a facilitator with training in clinical psychology led each group. None of the facilitators were directly involved in the clinical care of the participants. All facilitators were female, apart from those in Spain, where the facilitators were a man and a woman. Notably, these characteristics may have influenced the collection and interpretation of data. To reduce some of this bias, the coding was replicated by a qualitative researcher who was not present in the focus group and did not have a background of clinical psychology. Disagreements in coding were resolved as a pair, and a joint decision was made about the allocation of a code to each quotation.

Participant Characteristics

Participants were eligible if they were above the age of 18 years and were currently experiencing clinically significant symptoms of major depressive disorder

or had experienced such symptoms in the past 2 years. Individuals with a history of a psychotic disorder, including bipolar disorder and schizoaffective disorder, and substance misuse in the last 6 months were excluded. Participants were recruited via different sources in the three countries. In the United Kingdom, potential participants were screened telephonically by using a self-report measure of depression (World Health Organization's Composite International Diagnostic Interview - Short Form [15]). In Spain and Italy, clinicians selected patients diagnosed with major depressive disorder, who attended psychiatric services. Participants were identified by convenience sampling and their eligibility to participate. All participants provided written informed consent to participate in this study.

Procedure

The local research ethics committees for each country approved the procedures (Ethics codes: United Kingdom, 16/LO/1513; Italy, Parere 5/2017; Spain, PIC-149-16). All participants were screened for their eligibility to participate over the phone or in person. Subsequently, they were invited to participate in a face-to-face focus group session. In this session, they first completed a consent form and a demographics questionnaire before participating in a focus group, as detailed below. All travel expenses were covered.

Focus Group

The discussion was semistructured using a prespecified topic guide (available on request) that was designed to elicit discussions about barriers to and facilitators of engagement with mHealth technology in the context of living with a long-term mental health condition. The open-discussion format allowed people to share a range of examples. Each group discussion lasted for 60-120 minutes. This format was developed and tested in the United Kingdom, where a second focus group with the same participants was conducted to validate the emerging findings.

Data Analysis

Focus group discussions were audio recorded and transcribed verbatim. Both the Italian and Spanish transcripts were translated into English, allowing combined analyses by two researchers working independently with the use of the software package NVivo (version 10; QSR International, Melbourne, Australia). Subthemes emerging from the data were identified in the final analysis.

Results

Participant Characteristics

Focus groups were conducted with 25 participants across three countries (United Kingdom, n=8; Spain, group 1: n=3, group 2: n=5; Italy, n=9). Participants in Spain and Italy were living with depression for longer than those in the United Kingdom, and all participants were Caucasian. In Spain, all participants were female, but the age of the participants was similar across all three countries (Table 1).

Validation

Textbox 1 displays the subthemes emerging from the data, which were categorized into prespecified major and minor themes of the coding frame. Subthemes emerged in all major and minor codes of the coding frame, except physical ability. This evidence was taken as validation of the coding frame. Table 2 displays all the subthemes that emerged for the five different focus groups separately. Only a small number of additional subthemes emerged from the Spanish and Italian groups (10/58) after the focus group in the United Kingdom had taken place.

Table 1. Participant characteristics in each country.

Characteristics	United Kingdom (n=8)	Spain (n=8)	Italy (n=9)
Female, n (%)	5 (63)	8 (100)	7 (78)
Age (years), mean (SD)	51.9 (9.4)	47.1 (11.4)	52.8 (11.6)
Time since diagnosis (years), mean (SD)	2.9 (1.6)	13.2 (12.5)	11.5 (4.3)
Ethnicity, n (%)			
White	5 (63)	8 (100)	9 (100)
Black	2 (25)	—[a]	—
Asian	1 (13)	—	—

[a]Not applicable.

Textbox 1. Final major and minor codes and subthemes emerging from the discussions

Health-related barriers and facilitators

1. Symptom intensity or severity
   - Times of crisis
   - Accommodating fluctuations in symptoms
2. Emotional resources
   - Lack of motivation
   - Doubt
3. Awareness
   - Insight
4. Cognition
   - Poor memory (forgetfulness)
   - Difficulty reading
   - Difficulty with spoken expression
5. Physical ability

User-related barriers and facilitators

1. Technology acceptance
   - Attitude toward technology
   - Nonstigmatizing or familiar
   - Digital literacy (self and others)
   - Not ready to change
   - Codes of practice (eg, dress codes)
2. Perceived utility
   - Motivating action
   - Raising awareness or understanding
   - Sense of control
   - Opportunities for connection
   - Sense of achievement
   - Novelty or enjoyment
   - Measuring treatment response
   - Thinking more positively
   - Improving health and safety
   - Sharing data improves care
   - Reassuring (others)
   - Reassuring (others)
   - Contributing to research (others)
3. Perceived costs
   - Fears about privacy
   - Fears about security
   - Negative impact on mood or anxiety
   - Time and effort

- Increased dependency
- Fear of discrimination and stigma
- Unavailable or burden on resources (others)

4. Overall value
   - Inaccurate, ineffective, or meaningless
   - Balancing utility and costs
   - Value of human contact
   - Managing expectations
   - Inability to sustain resources
   - Curiosity
   - Trust in experts

Technology-related barriers and facilitators

1. Convenience
   - Fitting with routine or lifestyle
   - Inconvenience of charging
   - Inconvenience of notifications
   - Automatic and simplifies life
   - Loss of connection
2. Accessibility
   - Tailored or personalized
   - Expense
   - Lacking equipment
3. Convenience
   - Ease of use
   - Wearable
   - Data visualization
   - Short assessments
   - Poorly designed systems
4. Intrusiveness
   - Passive data collection
   - Obtrusiveness or discomfort
   - Invasion of body

Table 2. Summary of themes across major and minor codes for all countries.

Theme	Group				
	UK (Group 1a)[a], (n)	UK (Group 1b)[b], (n)	Spain (Group 1), (n)	Spain (Group 2), (n)	Italy, (n)
Health-related theme					
Symptom intensity or severity	• Times of crisis (3)	—[c]	—	—	• Times of crisis (1) • Accommodating fluctuations (1)
Emotional resources	• Lack of motivation (2)	• Lack of motivation (2)	• Doubting benefits (1)	• Lack of motivation (1)	• Motivation as a moderator (3)
Awareness	• Insight as a moderator (2)	—	• Poor insight (2)	• Poor insight (3)	—
Cognition	• Forgetfulness (1)	—	—	• Poor memory (2) • Difficulty reading (3) • Difficulty with spoken expression (1)	• Forgetfulness (1)
User-related theme					
Technology acceptance: self	• Skepticism towards technology (4) • Nonstigmatizing or familiar (1)	• Skepticism towards technology (4) • Nonstigmatizing or familiar (10) • Digital literacy as a moderator (2) • Not ready to change (1) • Dress codes (1)	—	• Familiar (1) • Poor digital literacy (1)	• Liking technology as a moderator (3)
Technology acceptance: others	• Lack of digital skills (2)	—	—	—	—
Perceived utility: self	• Motivating action (4) • Raising awareness or understanding (4) • Sense of control (4) • Opportunities for connection (3) • Sharing data improves care (2) • Sense of achievement (1) • Novelty or enjoyment (1)	• Motivating action (6) • Raising awareness or understanding (2) • Sharing data improves care (1) • Novelty or enjoyment (3) • Measure treatment response (1) • Thinking more positively (1)	• Motivating action (1) • Opportunities for connection (2) • Sharing data improves care (1) • Improves health and safety (1)	• Sense of control (2) • Sharing data improves care (2) • Improves health and safety (2)	• Reassuring (2)
Perceived utility: others	• Aiding decisions/communication (1)	• Aiding decisions/communication (1) • Contributing to research (8)	—	• Aiding decisions/communication (1)	—

Theme	Group				
	UK (Group 1a)[a], (n)	UK (Group 1b)[b], (n)	Spain (Group 1), (n)	Spain (Group 2), (n)	Italy, (n)
Perceived costs: self	• Fears about privacy (5) • Negative impact on mood/anxiety (5) • Fears about security and safety (3) • Time and effort (1) • Increased dependency (1)	• Fears about privacy (4) • Negative impact on mood/anxiety (3) • Fears about security and safety (1) • Time and effort (1) • Fear of discrimination and stigma (7)	—	• Negative impact on mood/anxiety (4)	• Fears about privacy (5) • Negative impact on mood/anxiety (1) • Time and effort (2) • Fear of discrimination and stigma (7)
Perceived costs: others	• Unavailable or burden on resources (4)	• Unavailable or burden on resources (1)	—	—	• Unavailable or burden on resources (1)
Overall value	• Inaccurate, ineffective, or meaningless (7) • Balancing utility and costs (4) • Value of human contact (2) • Managing expectations (1) • Inability to sustain resources (1)	• Inaccurate, ineffective, or meaningless (6) • Balancing utility and costs (3) • Managing expectations (1) • Sustainability of resources (2)	• Inaccurate, ineffective, or meaningless (2) • Balancing utility and costs (2) • Curiosity (2)	• Inaccurate, ineffective, or meaningless (1) • Curiosity (2) • Trust in experts (2)	• Value of human contact (3)
Technology-related theme					
Convenience	• Fitting with routine/lifestyle (2)	• Fitting with routine/lifestyle (9) • Inconvenience of charging (1) • Inconvenient notifications (1)	• Inconvenience of charging (1) • Inconvenient notifications (1) • Automatic and simplifies life (2) • Loss of connection (1)	—	• Inconvenient notifications (3) • Simplifies life (3) • Loss of connection (3)
Accessibility	• Tailored or personalized (9)	• Tailored or personalized (14) • Expense as a moderator (2)	• Expense as a moderator (3)	—	• Tailored or personalized (2) • Lacking equipment (1)
Usability	• Ease of use (3) • Wearable (1) • Data visualization (1) • Short assessments (1) • Poorly designed systems (1)	• Data visualization (1)	• Ease of use (1)	—	• Ease of use (3)
Intrusiveness	• Passive data collection (4) • Obtrusiveness (1) • Live sharing (1)	• Passive data collection (1) • Obtrusiveness or discomfort (3) • Invasion of body (1)	• Obtrusiveness (1)	• Discomfort (2)	—

[a]This group discussed prespecified points on the topic guide.

[b]This group reviewed topics raised in the first focus group to validate the findings.

[c]Not applicable.

Barriers and Facilitators of Engagement

We present our results in 3 main categories: health-related barriers and facilitators, which included the impact of the health status of the individual on engagement with technology; user-related barriers and facilitators, which summarized the impact of user attitudes, preferences, and beliefs about engagement with technology; and technology-related barriers and facilitators, which focus on direct interaction with the technology.

Participants in the United Kingdom and Italy discussed the impact of depression on their ability to engage with mHealth technology. *Times of crisis* was the most difficult period to adhere to treatment; one participant mentioned, "at that stage you just don't want to do anything. You're just living in a self-imposed prison" (UK23). There may be a window of opportunity for clinical prediction:

> *Once you get over the edge, there is no going back. Until the wave passes, and then you get back to normality, but when we get to that stage, no advice, no nothing can help us, except ourselves.* [UK23]

Adjusting technology to *accommodate fluctuations in symptoms* may be important.

Emotional Resources

Lack of *motivation* was noted in all countries. Some participants spoke about reduced motivation during depression as "when I get in a downer, part of the issue is that I just cannot get on with anything" (UK24). In relation to remote measurement, one participant said, "I had to fill it in, in the morning, afternoon and evening. I did it for the first two days, then the third I did just at morning and afternoon and then stop, I didn't do it anymore" (IT6).

Awareness

A subtheme of poor *insight* into the health status emerged across the United Kingdom and Spain; a participant stated, "I don't always realise that I've suffered a dip or a rise" (SP8).

Cognition

The impact of difficulties with cognition that caused *problems with memory, reading, and expression* was only mentioned in Spain. Single participants in the United Kingdom and Italy mentioned that they might be forgetful, but they did not attribute this to cognitive difficulties.

Target Users

Participants' general attitude towards mHealth technology emerged as a potential moderator of engagement. Participants in the United Kingdom demonstrated a *skeptical attitude:*

> *I don't think for me personally technology would work, to be honest, because I'm a person more about feeling and touching, rather than kind of connecting with something cold things, and um, electronical* [UK19]

Acceptance may be influenced by *digital literacy*. One person was willing to accept technology with extra support, saying "you'd have to download the application for me because I don't really know how those things work" (SP8). Others felt they had the required skills. Alternatively, some people *may not be ready to change* the way they manage their condition. This would be particularly important for individuals who do not own mobile technologies.

Acceptance of wearable devices that were *nonstigmatizing or familiar* was endorsed. One participant said, "I heard on TV that almost everyone nowadays has some sort of wearable device" (SP1); another participant added, "It wouldn't be stigmatising. In fact, they're quite trendy" (UK24). However, some participants raised concerns about employers not allowing people to wear devices due to *dress codes:*

> *I was just thinking about doctors and nurses and they're not allowed to wear anything below their elbow.* [UK22]

UK participants discussed the impact of *digital literacy* amongst healthcare professionals, who may also find the use of new technologies difficult; one participant stated, "I worked with older GPs and they struggled with the new technology coming in" (UK18).

Perceived Utility

Target Users

Participants discussed aspects that would provide a utility and facilitate use. A function was deemed useful if the technology could *motivate action*, for example, "go for a walk...do some meditation" (UK2), or "call your doctor" (UK5):

> *I sometimes go out for a run and my phone tells me that it has detected physical activity. Of course, and when it picks up on that, it also tells me: you still have time to achieve your goal today* [SP1]

One participant in the United Kingdom said that this type of feedback might help to *think more positively*; another suggested that it could lead to a *sense of achievement*. Some thought that mHealth technology was *novel and enjoyable* besides useful.

Raising awareness and understanding of one's health emerged as themes from the UK group. One person said "by measuring, you might discover things that people are not aware of already" (UK24). Feeling a *sense of control* and providing *opportunities for connection* with others may have further utility, as would using prompts or alerts to *improve health and safety* as, for example, a way to respond to symptoms early:

> *I have periods when I take medication and periods when I don't well...until now, I've been the one to notice that oh, I'm not doing very well, or I'm a bit, I don't know. And then after 3 or 4 weeks I've touched rock bottom. Well, maybe if I had some monitoring before that, then I could take the meds sooner and not get to that point, so, in my case, maybe it would be good for me* [SP1]

Sharing data with healthcare professionals was considered a way of *improving care* by this individual and others in the Spanish and UK groups. In the Italian group, health monitoring was considered *reassuring*.

Participants felt that there was scope for mHealth technology to *support clinical decision making and communication*; one

participant said, "I could see that if um the tracking information would be useful for my doctor, to help with trying to find the right medication" (UK22). This view was shared across the United Kingdom and Spain. In addition, UK participants noted benefits of *contributing to research* and the potential wider impact on others with depression; one participant stated, "I know there's a potentially bigger benefit—that's worthwhile" (UK20).

Perceived Costs

Target Users

In addition to opportunities for utility, costs were identified. Participants feared about their *privacy and security:*

> *I don't care if it knows I've been to Tesco's this morning, don't give a monkeys. But, I don't particularly want people to know I'm in Tesco's now.* [UK20]

> *I was given one of those new fashion ones, but I wanted one that looked cheap, otherwise I'm just going to get mugged* [UK22]

Although these issues were deemed important by some participants in Italy, one person mentioned, "I really don't care about privacy." [IT9]

Further costs were associated with feeling *increased anxiety* about one's health:

> *The technology which could remind you, not remind you, but tell you that you're going down or something. That would increase the anxiety, to be honest* [UK23]

> *I wouldn't recommend it to a hypochondriac. Because they'd spend all day obsessed, keeping an eye on what's happening to them.* [SP5]

> *I'm scared of relapses* [SP6]

Concerns about spending *time and effort* were mentioned by participants in Italy and the United Kingdom. In addition, there were concerns about *increasing dependency* and *fear of discrimination and stigma*. Participants suggested that data gathered might have "implications for travel insurance" (UK24) or prevent them from being promoted at work; one person stated, "I wouldn't want to declare. I wouldn't want to have a little badge on me saying I'm depressed." (UK20).

The main area of concern was the *increased burden on resources* for healthcare professionals and its potential negative impact on care; one participant said, "The more that they're bombarded with technology, the less energy there is for normal, human interaction" (UK18). Healthcare professionals and carers may not be *available to help* process information, and signs of deterioration may not be acted upon even if discovered.

Overall Value

People expressed *curiosity* about trying new technology. Hope for the future may provide motivation for engagement:

> *I'd quite happily do something that was two years, as long as I thought that if it was successful, there would be a hope for something afterwards* [UK22]

However, others in the UK group questioned the *sustainability of resources*, and the importance of *balancing utility and costs* was apparent. Investing money and time or making some sort of sacrifice to benefit from the rewards of the system was mentioned. One person felt that, overall, the perceived costs might outweigh the perceived utility. Due to current levels of information security, they said, "I think it is better not to collect this kind of personal data in the first place" (UK23). There were concerns across countries that the data gathered by the technology might be *inaccurate, ineffective, or meaningless*. Nonetheless, participants in Spain expressed *trust in researchers as experts* and were willing to be led by their guidance, but *managing participants' expectations* of the achievements through remote measurement was highlighted as an important role of researchers in the United Kingdom.

UK and Italian participants emphasized on the *value of human contact*. In the Italian group, some participants raised general concerns about technology limiting the relationship with their clinician and preferred face-to-face contact rather than telehealth.

Convenience

Participants felt that technology played a role in *simplifying activities* and serving a purpose; one participant noted, "if there is a purpose, if it simplifies my life, I am glad to use it" (IT6). There were discussions about the pros and cons of wearing devices that doubled up as watches. The participants believed that technology should easily *fit within a daily routine*. Practical challenges were noted, such as losing opportunities to log data due to the appearance of *notifications at inconvenient timings*, the need for *charging*, and the *loss of connection*.

Accessibility

The *financial expense* associated with the devices was a potential moderator of accessibility, and practical issues including *lack of equipment* were considered a barrier. Even if technology was available, for it to be accessible, resources need to be *tailored or personalized* to meet the specific requirements of individuals. When a person feels more unwell than usual, this issue may affect usage. Comments such as "It would depend also on the severity of symptoms, it must be adjusted" (IT4) and "it must be tailored to the person's mood and feelings" (IT8) highlighted this point.

Usability

mHealth resources should be *easy to use* and not "fiddly" (UK24). To reduce the effort needed to engage in surveys, one participant said, "I'd prefer something that is very short that I can complete within a minute" (UK20). Simplicity and low effort appear to be key facilitators, whereas complicated features or poor design were barriers:

> *I've got a watch my brother gave to me and it measures your heart rate. But it's so sophisticated, you've got to stick a cable down here, it's a bit much and I say: I'm thinking that I'm not going to wear this* [SP3]

Wearable monitors were endorsed, and the ability to *visualize data* was declared important for usability.

Unobtrusive and comfortable devices were important for acceptability. Similarly, discreet devices and *passive* collection of data were preferred. Only the UK group enquired about how *invasive devices* would be implanted under the skin.

A separate issue related to the theme of intrusiveness was the level of comfort participants felt with *live sharing* of data with others:

> *If its location, I'd rather it didn't know, that data wasn't live imported, instead when I'm not where the watch is telling, because you can get into live imports and just, everyone knows where you are all the time. And some of my cousins are quite happy to know where each other are 24/7, I find that scary… horrible, I don't want that.* [UK20]

Knowing who the data would be shared with was deemed important, and some participants suggested that sharing data with clinicians may be more acceptable than sharing them with profit-driven organizations.

Cross-Country Comparisons

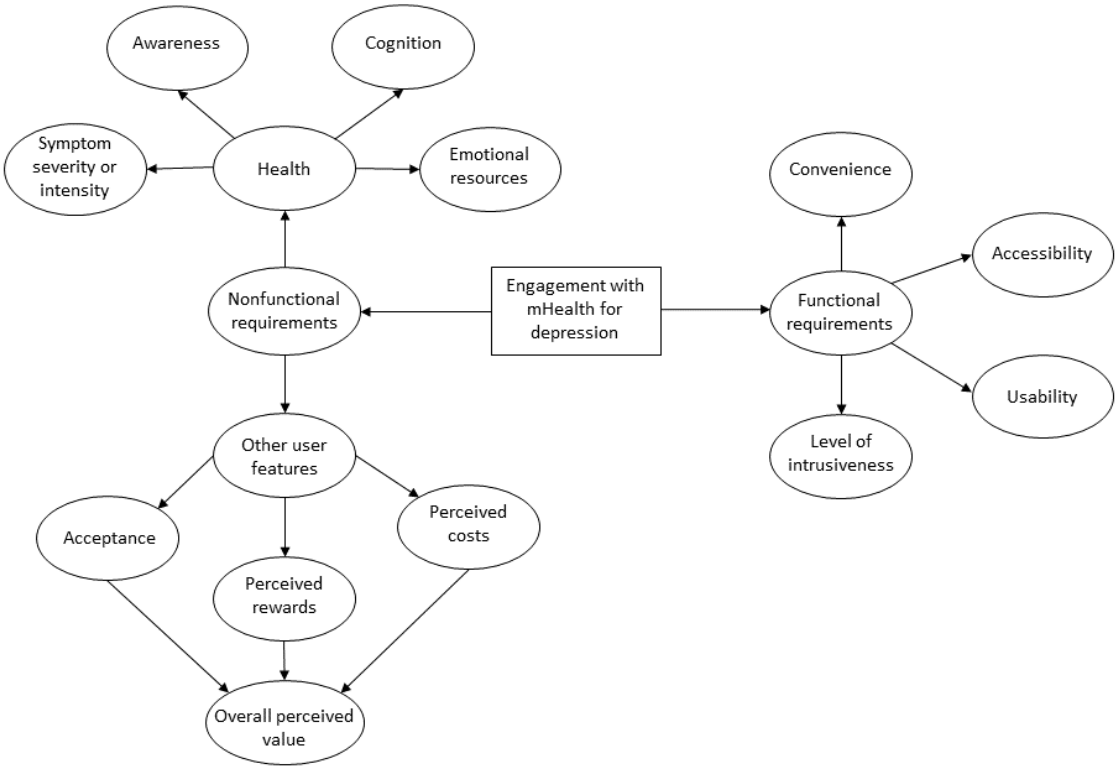
Almost half the themes were similar across at least two countries, suggesting replication and an acceptable level of data saturation [16]. A core group of themes was repeated across all countries: the need for *motivation***,** the potential *negative impact on anxiety and mood*, the inconvenience of *too-frequent notifications*, and the importance of *ease of use.* A number of key differences regarding additional subthemes emerged between regions. First, although the UK group provided an extensive list of utility examples, they were skeptical about the use of mHealth technology. Similarly, the Spanish participants had many issues with perceived utility. In contrast, the Italian group focused more on perceived costs. Participants in Spain were the only group to trust the experts. The UK group was uncertain about the digital skills and availability of resources in clinical practice. Acceptance of technology from the perspective of other people in their health systems, such as clinicians, was not raised as a concern in Spain or Italy. One older participant textboxin Italy expressed the inability to access equipment. Second, issues regarding usability were discussed in greater depth in the UK group. Although the ease of use was the only subtheme in the Italian and Spanish groups, some specific suggestions about data visualization, length of assessments, and the ease of wearables emerged in the UK group. Few technology-related barriers and facilitators emerged in Spain, where participants focused more on health-related and user-related themes than technology-related themes.

Discussion

Principal Findings

In this study, 3 major, 14 minor, and 58 subthemes emerged from the data; some were related to functionality of technology and others were about users' abilities, perceptions, and attitudes toward technology. These nonfunctional requirements have been reported previously [17]. Our nonfunctional requirements were categorized as health-related and user-related barriers and facilitators (Figure 1).

Figure 1. Requirements for engagement with mobile health (mHealth) technology for depression.



Nonfunctional Requirements

In terms of health-related barriers and facilitators, the severity of symptoms may moderate engagement with mHealth technology. Low motivation and intermittent poor insight and memory are known to be symptoms of depression and may be specific to this population [18], which may affect mHealth systems that require direct interaction with an app as well as users' decisions to wear devices. However, the absence of data could be as informative as its presence in algorithms created to identify the risk of relapse.

In terms of user-related barriers and facilitators, participants' attitudes toward mHealth technology affected their engagement. Similarly, digital literacy moderates the use of technology [19], and as the results of this study suggest, affects all users including the healthcare providers who support patients. Familiarity with technology and employment regulations may either further facilitate technology adoption or pose a barrier for use.

Our participants emphasized the importance of weighing costs against utility in order to make a decision about the overall value of mHealth technology. Utility included factors such as opportunities to connect with others; prompts; raising awareness and understanding; a sense of control; and responding to early warning signs by, for example, supporting clinical decision making. Perceived costs included reduced privacy and security; lack of availability and limited resources to support use; increased health anxiety or dependency; and expending time and effort, especially if there were inaccuracies in measurement.

Functional Requirements

Previous research demonstrated a relationship between perceived convenience and usability, and the acceptance of technology [20]. Similarly, mHealth technologies were thought to be easier to accept if they reduced effort, served a clear purpose, fit into one's daily routine, were comfortable, and promoted choice or control. Barriers included the receipt of notifications at inconvenient times and the need to charge devices or fix technical malfunctions. In addition to convenience and usability, previous literature has advocated the development of resources that are accessible or equally available to all users (eg, "universal design") [21], and this work has reiterated the need for such development with respect to depression-specific symptoms.

Geographical Requirements

The subthemes that emerged from multiple countries demonstrate some of the most important considerations for developing mHealth resources across Europe. *Motivation* is a key moderator of engagement. Two barriers across countries were the *potential negative effect on anxiety and mood* and the *inconvenience of too-frequent notifications*, which may be related. Focus on an *easy-to-use design* was clear. Some differences between the countries may relate to diversity in health care experience and the availability of or familiarity with mHealth technology. There is variation in the percentage of adults using mobile phones and internet-based technologies across Europe; Italy has lower access to these technologies than Spain and the United Kingdom [22]. Varied familiarity with mHealth technology may account for fewer examples of utility and greater concerns about potential costs in Italy, where some people may not be fully aware of the benefits and may have raised potential concerns about the loss of human interaction.

Results from this study are similar to those of a systematic review on barriers to and facilitators of engagement with remote measurement technology [13]. However, our study focused on the attitudes of individuals with depression toward technology and the nonfunctional, rather than functional, factors. Motivation was clearly an important category, but was incorporated into health-related barriers and facilitators in this study, due to the inextricable link between mood and emotional resources for people with depression; their physical abilities were never discussed. Although a few previous studies reported the acceptability and feasibility of mHealth resources for people with mental health conditions (eg, [23-26]), none of them explored barriers and facilitators across several countries.

This study uniquely provides views from participants living in different countries and revealed both similar and potentially different issues that were considered by the different groups. Although mHealth resources should take into account the similarities of views, it is essential to continue monitoring engagement across different countries, as these differences may affect their efficient implementation.

Strengths and Limitations

Strengths of this study were the inclusion of a varied sample that represented three European countries that place similar emphasis on community-based or "remote" treatment interventions for mental health. The qualitative approach enabled a rich, in-depth discussion of possible barriers to and facilitators of engagement with mHealth technology. It was not constrained to responses to specific questions, which allowed the discovery of themes that may not have emerged otherwise.

Although another strength of the study was the cross-national approach to understand factors influencing engagement, it is important to note that the translation may have influenced the findings. A further key limitation is the dependence of our results on hypothetical scenarios rather than actual experience. We have identified several themes that can guide research design and technological development, but we should be cautious about the anticipated risks or benefits that may not be sustained when people are exposed to technology. Further user testing with specific prototypes is required to maximize acceptability and usability. Such user testing will include a wider sample of the population with a history or current symptoms of depression, which will involve purposive sampling.

Future Research

Future research should consider other stakeholders. Craven et al [17] advocate the involvement of all possible end users including carers and clinicians, which may result in systems that are easily implemented in practice. A few studies that implemented this holistic perspective and involved several users [27] found commonalities in terms of universal support for technology innovation and potential barriers to the use of mHealth technology, similar to those identified in this study.

Conclusion

This qualitative study investigated the potential barriers to and facilitators of engagement with mHealth technology. A number of functional and nonfunctional categories emerged with both similarities and differences across European countries. The themes form a platform for future research on engagement with mHealth technology as a part of healthcare. A number of hypotheses have been generated: Increased familiarity and perceived utility, improved choice and control, greater convenience and accessibility, and lower intrusiveness may influence decisions about the use and engagement of mHealth technology and should be encouraged and evaluated in future studies, as the data might provide useful to improve existing models.

Acknowledgments

This paper was written as a part of the development of useful mobile health (mHealth) and remote measurement technology systems in the RADAR-CNS project. The RADAR-CNS project has received funding from the Innovative Medicines Initiative (IMI) 2 Joint Undertaking under grant agreement no. 115902. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation program and the European Federation of Pharmaceutical Industries and Associations (EFPIA); two pharmaceutical companies—H Lundbeck A/S and MSD Czech Republic—supported the preparation of this manuscript as part of this precompetitive public and private partnership. This communication reflects the views of the RADAR-CNS consortium, but neither IMI nor the European Union or EFPIA are liable for any use of the information contained herein. This paper also represents independent research, partly funded by the National Institute for Health Research (NIHR) Biomedical Research Centre in South London, the Maudsley National Health Service (NHS) Foundation Trust, and King's College in London. The views expressed are those of the author(s) and not necessarily those of the NHS, NIHR, or Department of Health and Social Care. TW and MH would specifically like to acknowledge support from the NIHR for their Senior Investigator Awards. Dr Giovanni Battista Tura provided further support with recruitment in Italy, and Nicholas Meyer provided insight into the initial structuring of the results based on observations from unpublished results of a study exploring engagement with mHealth technology among people with psychosis.

Conflicts of Interest

None declared.

References

1. Cuijpers P, Schoevers RA. Increased mortality in depressive disorders: a review. Curr Psychiatry Rep 2004 Dec;6(6):430-437. [Medline: 15538991]
2. Lenox-Smith A, Macdonald MTB, Reed C, Tylee A, Peveler R, Quail D, et al. Quality of Life in Depressed Patients in UK Primary Care: The FINDER Study. Neurol Ther 2013 Dec;2(1-2):25-42 [FREE Full text] [doi: 10.1007/s40120-013-0006-1] [Medline: 26000214]
3. Lerner D, Adler DA, Chang H, Lapitsky L, Hood MY, Perissinotto C, et al. Unemployment, job retention, and productivity loss among employees with depression. Psychiatr Serv 2004 Dec;55(12):1371-1378 [FREE Full text] [doi: 10.1176/appi.ps.55.12.1371] [Medline: 15572564]
4. Barger SD, Messerli-Bürgy N, Barth J. Social relationship correlates of major depressive disorder and depressive symptoms in Switzerland: nationally representative cross sectional study. BMC Public Health 2014 Mar 24;14:273 [FREE Full text] [doi: 10.1186/1471-2458-14-273] [Medline: 24656048]
5. Hofmeijer-Sevink M, Batelaan NM, van Megen HJ, Penninx BW, Cath DC, van den Hout MA, et al. Clinical relevance of comorbidity in anxiety disorders: a report from the Netherlands Study of Depression and Anxiety (NESDA). J Affect Disord 2012 Mar;137(1-3):106-112 [FREE Full text] [doi: 10.1016/j.jad.2011.12.008] [Medline: 22240085]
6. Boschloo L, Vogelzangs N, Smit JH, van den Brink W, Veltman DJ, Beekman AT, et al. Comorbidity and risk indicators for alcohol use disorders among persons with anxiety and/or depressive disorders: findings from the Netherlands Study of Depression and Anxiety (NESDA). J Affect Disord 2011 Jun;131(1-3):233-242 [FREE Full text] [doi: 10.1016/j.jad.2010.12.014] [Medline: 21247636]
7. Hardeveld F, Spijker J, De Graaf R, Hendriks SM, Licht CM, Nolen WA, et al. Recurrence of major depressive disorder across different treatment settings: results from the NESDA study. J Affect Disord 2013 May;147(1-3):225-231 [FREE Full text] [doi: 10.1016/j.jad.2012.11.008] [Medline: 23218899]
8. Statista. 2016. Number of smartphone users worldwide from 2014 to 2020 (in billions) URL: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ [accessed 2018-11-29]
9. Statista. 2017. Number of connected wearable devices worldwide from 2016 to 2021 (in millions) URL: https://www.statista.com/statistics/487291/global-connected-wearable-devices/ [accessed 2018-11-29]
10. Wicks P, Hotopf M, Narayan VA, Basch E, Weatherall J, Gray M. It's a long shot, but it just might work! Perspectives on the future of medicine. BMC Med 2016 Nov 07;14(1):176 [FREE Full text] [doi: 10.1186/s12916-016-0727-y] [Medline: 27817747]

11. Sherdell L, Waugh CE, Gotlib IH. Anticipatory pleasure predicts motivation for reward in major depression. J Abnorm Psychol 2012 Feb;121(1):51-60 [FREE Full text] [doi: 10.1037/a0024945] [Medline: 21842963]
12. Anguera J, Jordan JT, Castaneda D, Gazzaley A, Areán PA. Conducting a fully mobile and randomised clinical trial for depression: access, engagement and expense. BMJ Innov 2016 Jan;2(1):14-21 [FREE Full text] [doi: 10.1136/bmjinnov-2015-000098] [Medline: 27019745]
13. Simblett S, Greer B, Matcham F, Curtis H, Polhemus A, Ferrão J, et al. Barriers to and Facilitators of Engagement With Remote Measurement Technology for Managing Health: Systematic Review and Content Analysis of Findings. J Med Internet Res 2018 Jul 12;20(7):e10480 [FREE Full text] [doi: 10.2196/10480] [Medline: 30001997]
14. Rydén L, Stokoe G, Breithardt G, Lindemans F, Potgieter A, Task Force 2 of the Cardiovascular Round Table of the European Society of Cardiology. Patient access to medical technology across Europe. Eur Heart J 2004 Apr;25(7):611-616. [doi: 10.1016/j.ehj.2004.02.015] [Medline: 15120057]
15. Kessler R, Andrews G, Mroczek D, Ustun B, Wittchen H. The World Health Organization Composite International Diagnostic Interview short-form (CIDI-SF). Int. J. Method. Psychiat. Res 2006 Nov;7(4):171-185. [doi: 10.1002/mpr.47]
16. Fusch PI, Ness LR. Are We There Yet? Data Saturation in Qualitative Research. The Qualitative Report 2015;21(7):1408-1416 [FREE Full text] [doi: 10.1177/1049732311401424] [Medline: 21368018]
17. Craven M, Lang AR, Martin JL. Developing mHealth Apps with Researchers: Multi-Stakeholder Design Considerations. In: Design, User Experience, and Usability. User Experience Design for Everyday Life Applications and Services. Cham: Springer; 2014:15-24.
18. Beck AT, Alford BA. Depression: Causes and Treatment. Pennsylvania, PA: University of Pennsylvania Press, Inc; 2009.
19. Mohammadyari S, Singh H. Understanding the effect of e-learning on individual performance: The role of digital literacy. Computers & Education 2015 Mar;82(C):11-25 [FREE Full text] [doi: 10.1016/j.compedu.2014.10.025]
20. Chang CC, Yan C, Tseng J. Perceived convenience in an extended technology acceptance model: Mobile technology and English learning for college students. AJET 2012 Jul 26;28(5). [doi: 10.14742/ajet.818]
21. Iwarsson S, Ståhl A. Accessibility, usability and universal design--positioning and definition of concepts describing person-environment relationships. Disabil Rehabil 2003 Jan 21;25(2):57-66. [Medline: 12554380]
22. Poushter J. Pew Research Center. 2016. Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies URL: http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/ [accessed 2018-11-29]
23. Ben-Zeev D, Brenner CJ, Begale M, Duffecy J, Mohr DC, Mueser KT. Feasibility, acceptability, and preliminary efficacy of a smartphone intervention for schizophrenia. Schizophr Bull 2014 Nov;40(6):1244-1253 [FREE Full text] [doi: 10.1093/schbul/sbu033] [Medline: 24609454]
24. Ben-Zeev D, Scherer EA, Gottlieb JD, Rotondi AJ, Brunette MF, Achtyes ED, et al. mHealth for Schizophrenia: Patient Engagement With a Mobile Phone Intervention Following Hospital Discharge. JMIR Ment Health 2016 Jul 27;3(3):e34 [FREE Full text] [doi: 10.2196/mental.6348] [Medline: 27465803]
25. Naslund JA, Aschbrenner KA, Barre LK, Bartels SJ. Feasibility of popular m-health technologies for activity tracking among individuals with serious mental illness. Telemed J E Health 2015 Mar;21(3):213-216 [FREE Full text] [doi: 10.1089/tmj.2014.0105] [Medline: 25536190]
26. Price M, Ruggiero KJ, Ferguson PL, Patel SK, Treiber F, Couillard D, et al. A feasibility pilot study on the use of text messages to track PTSD symptoms after a traumatic injury. Gen Hosp Psychiatry 2014;36(3):249-254 [FREE Full text] [doi: 10.1016/j.genhosppsych.2014.02.004] [Medline: 24636721]
27. Simons L, Valentine AZ, Falconer CJ, Groom M, Daley D, Craven MP, et al. Developing mHealth Remote Monitoring Technology for Attention Deficit Hyperactivity Disorder: A Qualitative Study Eliciting User Priorities and Needs. JMIR Mhealth Uhealth 2016 Mar 23;4(1):e31 [FREE Full text] [doi: 10.2196/mhealth.5009] [Medline: 27009498]

Abbreviations

mHealth: mobile health

Edited by G Eysenbach; submitted 20.06.18; peer-reviewed by A Cerga Pashoja, J Apolinário-Hagen; comments to author 13.09.18; revised version received 01.10.18; accepted 03.10.18; published 30.01.19.

Please cite as:
Simblett S, Matcham F, Siddi S, Bulgari V, Barattieri di San Pietro C, Hortas López J, Ferrão J, Polhemus A, Haro JM, de Girolamo G, Gamble P, Eriksson H, Hotopf M, Wykes T, RADAR-CNS Consortium
Barriers to and Facilitators of Engagement With mHealth Technology for Remote Measurement and Management of Depression: Qualitative Analysis
JMIR Mhealth Uhealth 2019;7(1):e11325
URL: http://mhealth.jmir.org/2019/1/e11325/
doi:10.2196/11325
PMID:30698535

Original Paper

Investigating the Adoption of Mobile Health Services by Elderly Users: Trust Transfer Model and Survey Study

Fanbo Meng[1,2*], MSc, PhD; Xitong Guo[1*], PhD; Zeyu Peng[3*], PhD; Kee-Hung Lai[2*], PhD; Xinli Zhao[1*], PhD

[1]eHealth Research Institute, School of Management, Harbin Institute of Technology, Harbin, China

[2]Logistics and Maritime Studies, Faculty of Business, Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[3]Department of Information Systems, School of Business, East China University of Science and Technology, Shanghai, China

[*]all authors contributed equally

Corresponding Author:
Xitong Guo, PhD
eHealth Research Institute
School of Management
Harbin Institute of Technology
92 Xidazhi Street
Nangang District
Harbin,
China
Phone: 86 451 86414022
Email: xitongguo@hit.edu.cn

Abstract

Background: Although elderly users comprise a major user group in the field of mobile health (mHealth) services, their adoption rate of such services is relatively low compared with their use of traditional health services. Increasing the adoption rate of mHealth services among elderly users is beneficial to the aging process.

Objective: This study aimed to examine the determinants of mHealth service use intentions using a trust transfer model among elderly users facing declining physiological conditions and lacking support from hospitals.

Methods: A survey comprising 395 users aged 60 years and above was conducted in China to validate our research model and hypotheses.

Results: The results reveal that (1) trust in mHealth services positively influences use intentions, (2) trust in offline health services positively influences trust in mHealth services, (3) declining physiological conditions strengthen the effects of trust in offline health services regarding trust in mHealth services, (4) support from hospitals weakens the effects of trust in mHealth services on use intentions, and (5) the relationship between trust in offline health services and intention to use mHealth services is partially mediated by trust in mHealth services. The independent variables and moderators collectively explain a 48.3% variance in the use intention of mHealth services.

Conclusions: We conclude that the trust transfer theory is useful in explaining the development of initial trust in mHealth services. In addition, declining physiological conditions and support from hospitals are important factors for investigating the adoption of mHealth services among elderly users.

(JMIR Mhealth Uhealth 2019;7(1):e12269) doi:10.2196/12269

KEYWORDS

mobile health; trust; health services for the elderly; adoption; health behavior

Introduction

Background

With advances in health care assisting more people to live longer, the number of people aged over 60 years is projected to reach nearly 2.1 billion, representing 22% of the world's total population [1]. In China, the proportion of the population aged 60 years and above will increase from 12.4% (168 million) in 2010 to 28% (402 million) in 2040 [2]. This demographic change further intensifies the conflict between the medical care demands of elders and the limited medical care resources. Indeed, mobile health (mHealth) has the potential to enable the

elderly to experience longer and healthier lives by transforming health care services and clinical interventions for elderly users [3]. mHealth services are believed to be beneficial for elderly people because they can bring about multiple positive outcomes, such as health care costs savings, individually tailored health information and services, and a more effective health service process [4,5]. Therefore, mHealth is regarded as an important approach of extending the traditional (ie, offline) health services and satisfying the medical demands of an increasingly aging population.

However, older people are often found to be less innovative toward mHealth services [6]. There is limited usage of mHealth services to manage chronic diseases among the elderly [7]. In China, the development of mHealth services continues to remain at the infancy stage, and the adoption rate of such services among older people remains low [8]. The low adoption rate of mHealth services by the elderly has prompted the Chinese government to reconsider their developing strategies, to encompass the rapidly growing number of elderly users to alleviate the pressures of the aging population. In addition, although the elderly could obtain great potential benefits by using health information technology (HIT), previous research has not examined the relationship between HIT use among elderly users from their use of health services [9,10].

Due to their unique physical and psychological characteristics, elderly users may need to expend much more effort and time familiarizing themselves with information technology (IT) than the younger user groups [11]. In addition, declining physiological conditions [12] and limited social resources [13] are hurdles for elderly individuals in making health-related decisions independently. To serve the elderly population through mHealth services, service providers need to understand the behavior of this special group and the antecedents that influence their acceptance and usage of mHealth.

Building user trust is the key to promoting the adoption of mHealth services by elderly users. Trust could alleviate uncertainties and risks when users encounter new IT or services (eg, purchasing books on the Web and mobile payment services) [14,15]. In this context, trust transfer is a means of building customer trust in an unknown target through a trusted party [16]. Customer trust can be developed through trust transference from offline health care services, which are trusted by the elderly, to mHealth services with which they are not familiar. Although trust plays an important role in mHealth adoption, very few previous studies have investigated the development of trust in mHealth services [17,18]. Furthermore, as mHealth services are developed based on the traditional service in a new context, that is, a mobile context, users' perceptions on mHealth may be influenced by the offline traditional services, such as trust. However, research on the trust transfer from an offline context to a mobile channel is also underexplored in both the information system (IS) and health care fields. In addition, although considerable research on trust transfer has been conducted in the IS domain, inadequate research has been undertaken on trust transfer in mHealth research.

Our study aims to develop a trust transfer model to investigate the adoption of mHealth services by the elderly and to address the following research question: To what extent is trust transfer a means of establishing the initial trust of elderly users in their adoption of mHealth services? In this study, based on the trust transfer theory, we incorporate both declining physiological conditions and support from hospitals into our research model to explore their moderating effects on the trust transfer process. To validate the research model and proposed hypotheses, a survey comprising 395 elderly users in China was conducted to analyze the research model.

Literature Review and Hypotheses

Trust Transfer

Trust can develop through a transference process, suggesting that trust can be transferred from a trusted object to an unfamiliar object [19]. Channels of trust transfer include the intrachannel and interchannel transfers [20]. An intrachannel trust transfer is one where trust can be transferred within the same channel that is deemed as trustworthy. For example, trust transfer can occur either from offline to offline contexts or from Web to Web contexts. Stewart illustrated that trust can be transferred from a trusted hypertext link to an unfamiliar one on a website [16]. In fact, Stewart also suggested that trust transfer may work based on a cognitive process that is based on the mere knowledge of the relationship between the trusted target and trusted source.

On the other hand, interchannel trust transfer refers to trust transfers from one context to another, mainly from the offline to the Web channels or from the Web to mobile channels. Turel et al indicated that in a transfer from an offline to a Web channel, the e-customer service provider can improve user trust and use intention by associating itself with a known human service representative [21]. Belanche et al purported that trust in offline public administration recommendations can be transferred to trust in the online public e-service [22]. From the perspective of transfer from a Web to a mobile channel, the level of trust in internet payment services is positively associated with the initial trust in mobile payment services offered by the same company [23]. Similarly, in a study on brokerage services, trust in an online environment is positively related to initial trust in a mobile environment [20]. On their part, Wang et al studied the mobile e-word of mouth (eWOM) services and found that trust in Web services has a positive effect on trust in mobile services, thus influencing the intention to use mobile eWOM services [24].

Offline health services are the pivotal source of health services for the elderly. These users' prior experiences and familiarity with offline health services have resulted in the development of their deeply ingrained belief that offline health services are trustworthy. Accordingly, in the interchannel trust transfer context, we anticipate that trust in well-established offline health services would affect trust in corresponding mHealth services.

Thus, we hypothesize the following:

> Hypothesis 1: Trust in offline health services is positively associated with trust in mHealth services.

Trust in Mobile Health Services

Prior research demonstrates that the conceptualizations of trust are diverse, incomplete, and inconsistent [25]. Integrating with the shared characteristics of trust across different disciplines, Mayer defined trust as "the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party" [26]. According to prior research, beliefs in competence, integrity, and benevolence lead to a general belief in trust and behavioral intentions [14,27,28]. All these 3 main trusting beliefs are the trustee's internal trust-related characteristics observed by the trustor, and these trusting beliefs are formatted in the cognitive process of a trustee's attributes.

As trust can reduce risks and uncertainties, trust plays a significant role in the adoption of a new IT [29]. Prior studies underscore that trust is a significant prerequisite of social behavior and positively associated with users' use intentions of an internet store [27], purchasing books on the internet [14], e-government services [30,31], e-commerce [25,32,33], and mobile payment [34].

From the perspective of mHealth services, Guo et al found that trust in mHealth service providers enables the reduction of individuals' privacy concerns and an increase in adoption intentions [18]. Zhao et al indicated that trust is positively associated with the behavioral intention to use mHealth services [35]. Deng et al demonstrated that patients' trust positively affects the adoption intention of mHealth services [36]. On the basis of their findings of the post adoption stage, Akter et al theoretically conceptualized trustworthiness (trusting belief) in mHealth service research and indicated that trustworthiness positively influences consumer trust (trusting intentions), which directly affects consumers' continuance intentions [17]. In a later research, Akter et al demonstrated that perceived trust positively affects satisfaction with mHealth services and continuance intention [37].

As the mHealth service is a credence product and personalized service, trust plays a significant role in predicting individuals' adoption intentions. Elderly users, in particular, have less experience with the use of mHealth services and encounter more difficulties in using this emerging technology. As health is a sensitive subject, these users may pay more attention to health services assessed through mobile channels.

Thus, we hypothesize the following:

> Hypothesis 2: Trust in mHealth services is positively associated with the intention to use mHealth services.

Declining Physiological Conditions

Individuals' physiological conditions (such as hearing, vision, speech, locomotion, and memory capabilities) are known to decline in the aging process, thus influencing their physical and cognitive capabilities [38]. Timmermann demonstrated that the declining physical and cognitive capabilities may cause elders to experience considerably greater difficulty in the use of computers, and these declining physiological conditions can serve as internal controls or inhibiting conditions that increase the effort expectancy associated with IT use [39]. Mathur indicated that age-related decline in physiological conditions will increase the elderly's need for family assistance in coping with these declines and revealed that family assistance had a positive effect on the adoption of IT [13].

Phang et al introduced new constructs that are used in gerontology and IS literature, comprising preference for human contact, self-actualization, and resource savings as antecedents of perceived usefulness, whereas computer anxiety, computing support, and declining physiological conditions reflect perceived ease of use in the research of e-government service use intentions [12]. Heart and Kalderon demonstrated that cognitive and physical impairments that increase with age are negatively associated with health-related IT use [40].

Xue et al integrated aging-specific constructs including perceived use resources, technology anxiety, and bio-physical age (perceived physical conditions) with the technology acceptance model to reveal that perceived use resources and technology anxiety were antecedents for perceived usefulness, whereas perceived physical conditions significantly influenced perceived ease of use [41]. Deng et al found that elderly users' aging characteristics, including declining physiological conditions, technology anxiety, and self-actualization needs, positively influence their use intentions of an mHealth service [8].

Therefore, declining physiological conditions may cause elderly users to expend greater energy on evaluating the competence, reliability, and dependability of mHealth services. Hence, elderly users are more likely to develop trust in mHealth services based on trust in offline health services.

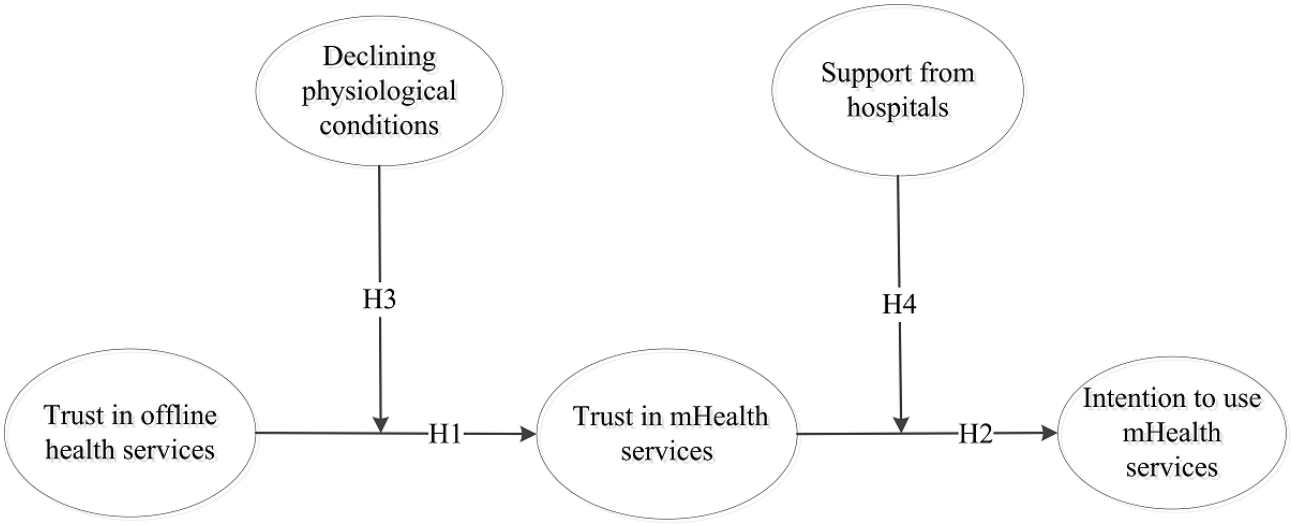
Thus, we hypothesize the following:

> Hypothesis 3: The relationship between trust in offline health services and trust in mHealth services will be stronger in the case of declining physiological conditions.

Support From Hospitals

On the basis of the previous literature on IT acceptance and usage, the source of support is mainly an organization [42] or an expert user [43]. An organization can offer managerial interventions and support resources to encourage users to accept IS and improve performance outcomes [44]. In their study of personal computer usage, Compeau and Higgins observed that an organization's support for computer users influences individuals' judgments of self-efficacy because support and assistance from an organization assisted users in increasing their ability [45]. Hoque and Sorwar found that elderly users believe an organizational support or technical infrastructure support positively affect their behavioral intentions toward mHealth services [46]. In the context of the adoption of mHealth services, with support from hospitals, elderly users are able to efficiently evaluate the competence of mHealth services and develop trust in mHealth services. This leads to an intention to use mHealth services. When elderly users receive support from hospitals, they are more likely to develop a high level of trust in mHealth services.

Figure 1. Research model. mHealth: mobile health.



Thus, we hypothesize the following:

> *Hypothesis 4: The relationship between trust in mHealth services and the intention to use mHealth services will be stronger when hospital support is provided.*

In summary, this study proposed a trust transfer model to predict elderly users' intentions to use mHealth services, as depicted in Figure 1. According to the trust transfer theory [16,47], trust in offline health care services positively influences trust in mHealth services. A declining physiological condition, as a salient variable of gerontology, influences the trust transfer process from an offline channel to a mobile channel. In addition, support from hospitals has a moderating effect on the relationship between trust in mHealth services and use intentions.

Methods

Development of the Study Questionnaire

All the measures in the survey were adapted or adopted from existing prevalidated instruments. The measures of trust in offline health services and trust in mHealth services were adapted from the study by Gefen et al [25]. The intention to use the mHealth service was measured using the items from the study by Deng et al [8]. The measures of declining physiological conditions were adapted from the study by Phang et al [12]. The measures of support from hospitals were adapted from the measures of vendor support in the study by Thong et al [48]. In fact, Thong et al indicated that vendor support results in a higher level of IS effectiveness. In the context of our study, the hospital of our study launched an mHealth service to extend the traditional offline health services. The role of the hospital is very similar to the role of the vendor. Therefore, we modified the measurement items of vendor support according to our research context.

To improve the content validity of the questionnaire, we revised the questions and deleted similarly phrased items based on feedback from 2 mHealth researchers and a pretest with 30 elderly users. All items included in the survey were measured on a 7-point Likert scale ranging from 1 to 7, with "1" representing strongly disagree and "7" representing strongly agree. The details of the constructs and measurements are presented in Multimedia Appendix 1.

Data Collection

We conducted our survey at Aerospace Center hospital in Beijing, the capital of China. At the end of 2016, this hospital had launched a health management platform aiming to provide reliable daily health care services for patients through mobile services. Therefore, this target hospital is an appropriate site for data collection. The mHealth services provided by this platform included a routine appointment out-patient clinic, follow-up visits, medicine reminders, medical records, and real-time positioning. Permission was obtained, and proper arrangements were made by the management board of the hospital for successful data collection. We followed 3 screening criteria to select the target participants: (1) patients coming to a physical examination center for medical examination, from which we could assume that they cared about their health; (2) patients who did not have a serious disease; and (3) patients aged 60 years or above. As these elderly participants could find it difficult to complete questionnaires unaided because of literacy problems, we accordingly employed 6 postgraduate students to assist the participants in completing the questionnaires. Furthermore, considering that most of the participants might be unfamiliar with mHealth services, a copy of an instruction manual of mHealth services was shown to the participants before they completed the questionnaire.

In fact, 500 copies of the survey questionnaire were distributed, and 395 usable copies were obtained, accounting for a response rate of 79.0% (395/500). We excluded 105 participants because, owing to several reasons, they were not our target subjects. First, these participants were illiterate. Second, their declining abilities regarding reading, writing, and speaking hindered their use of mobile services. Third, these participants did not own mobile phones, which could support the use of mHealth services. Accordingly, we selected the remaining respondents as our target subjects to use in the data analysis. Among these participants, approximately 49.1% (194/395) were males and 50.9% (201/395) were females. In fact, 316 out of 395 participants (80.0%) were aged from 60 years to 70 years. In

addition, 75.9% (300/395) of the participants had attended high school. All respondents were given ¥10 (US $1.5) supermarket coupons for their participation.

Data Analysis

Partial least squares (PLS) was used to test the research model because of the several advantages of this technique. First, PLS can predict all loadings and weights of indicators and causal relationships among constructs in multistage models [49,50]. Second, compared with covariance-based (CB) structural equation modeling (SEM), PLS is the most suitable technique for models with formative constructs and is appropriate for relatively small samples [51], which is the case in our study. In addition, PLS provides a good approximation of CB-SEM in terms of final estimates [50,51]. Given these considerations, we adopted the PLS to analyze our research model.

The data analysis was conducted in 2 stages. In the first stage, the measurement model (ie, based on the reliability, validity, common method bias, and multicollinearity of the constructs) was assessed to ensure its appropriateness. In the second stage, the structural model was examined and the stated hypotheses were tested [52].

Results

Results of the Measurement Model Testing

The internal reliability, convergent validity, and discriminant validity were examined to assess the measurement model [53]. The reliability of the constructs was tested using Cronbach alpha, composite reliability (CR), and average variance extracted (AVE) [54]. In our study, the threshold values of the CR and AVE were .70 and .50, respectively, which is consistent with the work by Chin [55]. The value of Cronbach alpha, which was greater than the threshold .70, indicated adequate construct reliability [56]. CRs for these constructs ranged from .906 to .957, and the AVE varied from .765 to .881. All the values of Cronbach alpha and CR were above the threshold values, indicating good construct reliability [54]. All the item loadings of each construct were significant and above the suggested cut-off value (.700), indicating convergent validity [55]. As shown in Table 1, the loadings of all items were much greater than the cross-loadings on other constructs, and the correlations of any 2 constructs were significantly smaller than the square root of the AVE of each construct, further indicating acceptable discriminant validity. In addition, we tested the potential issue of multicollinearity through the use of the variance inflation factor (VIF). As a rule of thumb, the threshold value of the VIF is less than or equal to 10, indicating no presence of multicollinearity. The results indicate that all VIFs are less than 7, thereby suggesting that there is no multicollinearity among or between the independent variables.

Common method bias could be a potential concern as the data were collected from a self-reported survey [57]. We first tested for common method bias according to Harman single-factor test [58]. According to the results, we found that the factors accounted for 81.5% of the variance and the first factor only explained 25.1% of the variance, thus indicating that common method bias was not likely to have been an issue.

Results of the Structural Model Testing

The results of the structural model are recorded in Figure 2. They indicate that trust in mHealth services (beta=.556; t_{394}=11.174; P<.001) had significant effects on the intention to use mHealth services. Hence, hypothesis 1 is supported. The results demonstrate that trust in offline health services has a significant effect on trust in mHealth services (beta=.583; t_{394}=14.528; P<.001), thus supporting hypothesis 2.

The moderating effects of declining physiological conditions and support from physicians were further tested. Declining physiological conditions were perceived to have a positive moderating effect on the relationship between trust in offline health services and trust in mHealth services (beta=.140; t_{394}=2.723; P=.003), thus lending support to hypothesis 3.

Table 1. Correlations and discriminant validity.

Construct	Cronbach alpha	Composite reliability	AVE[a]	Use intention	Trust in offline health services	Trust in mHealth[b] services	Declining physiological conditions	Support from hospitals
Use intention	.888	.930	.816	.903[c]	—[d]	—	—	—
Trust in offline health services	.932	.957	.881	.531	.938[c]	—	—	—
Trust in mHealth services	.914	.945	.853	.683	.582	.923[c]	—	—
Declining physiological conditions	.915	.906	.765	.033	−.007	.052	.874[c]	—
Support from hospitals	.929	.949	.824	.535	.581	.682	.093	.907[c]

[a]AVE: average variance extracted.

[b]mHealth: mobile health.

[c]Square root of average variance extracted.

[d]Not applicable.

Figure 2. Results of the research model. mHealth: mobile health. Asterisk indicates *P*<.10; two asterisks *P*<.01; three asterisks *P*<.001.

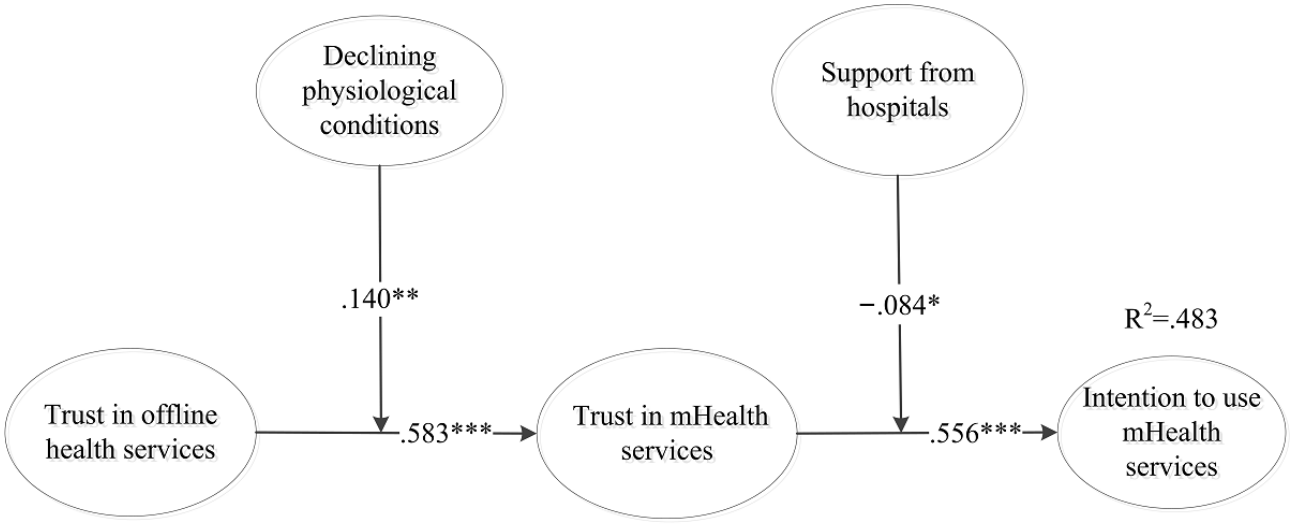


Table 2. Summary of the hypotheses testing.

Hypothesis	Description	Path coefficient	*P* value	Supported
1	Trust in offline health services is positively associated with trust in mHealth[a] services.	.583	<.001	Yes
2	Trust in mHealth services is positively associated with intention to use mHealth services.	.510	<.001	Yes
3	The relationship between trust in offline health services and trust in mHealth services will be stronger in the case of declining physiological conditions.	.140	.003	Yes
4	The relationship between trust in mHealth services and intention to use mHealth services will be stronger in the case of support from hospitals.	−.084	.04	No

[a]mHealth: mobile health.

Support from hospitals was seen to negatively moderate the relationship between trust in mHealth services and intention to use mHealth services (beta=−.084; t_{394}=1.725; *P*=.04). Hence, hypothesis 4 is not supported.

In addition, the mediating effect of trust in mHealth services in the proposed model was tested according to the procedures provided by Baron and Kenny [59]. First, the relationship between trust in offline health services and intention to use mHealth services was tested (beta=.532; t_{394}=10.252; *P*<.001). Then, we tested the relationship between trust in mHealth services and trust in offline health services (beta=.582; t_{394}=13.191; *P*<.001). Finally, We tested the relationship between trust in offline health services and intention to use mHealth services (beta=.202; t_{394}=3.420; *P*<.001), and the relationship between trust in mHealth services and intention to use mHealth services (beta=.566; t_{394}=11.988; *P*<.001). Therefore, the relationship between trust in offline health services and intention to use mHealth services is partially mediated by trust in mHealth services. The results of each hypothesis are summarized in Table 2.

Discussion

Principal Findings

This study yields several important findings. First, trust in mHealth services positively affects elderly users' use intentions. This highlights the proposition that trust in mHealth services is of importance in predicting the adoption of mHealth services. Second, trust in offline health services has significant effects on trust in mHealth services, which indicates that trust in offline health services can be transferred to the mobile environment. This suggests that mHealth service providers may be able to swiftly build elderly users' trust in mHealth services through leveraging their existing trust in offline health services, thus leading to a higher adoption rate of mHealth services. Third, declining physiological conditions are seen to strengthen the relationship between trust in offline health services and trust in mHealth services. As physiological capabilities decline with age, elderly users are less capable of using new IT apps. As elderly users with declining physiological conditions need to exert more effort to evaluate the competence and ability of mHealth services, they rather tend to rely more on their prior experience and knowledge of offline health services to cultivate trust in mHealth services. Fourth, support from hospitals is seen to weaken the association between trust in mHealth services and the intention to use mHealth services. Contrary to our hypotheses, support from hospitals decreases patients' trust in mHealth services because they depend more on physicians' support instead of expending extra effort on evaluating the competence and ability of mHealth services. There is possibly a mismatch between their perceptions and their physicians' advice.

Theoretical and Practical Implications

This study can enrich and advance our theoretical understanding in several ways. First, it extended the trust transfer theory to

the offline context in transition to mHealth services. Previous studies of the trust transfer theory have mainly focused on the perspectives of offline to online and online to mobile channels [16,20,22,60]. On our part, we explored the trust transfer process in the context of the new mHealth services from a cross-environment perspective. As the offline to mobile service transition has become a trend for most offline services, our research into this phenomenon provides a valuable reference. Our study can be seen as an attempt to fill this research gap by providing a cornerstone for further theoretical development.

Second, to comprehensively understand the IT acceptance behavior of the elderly, this research introduced a variable from gerontology—declining physiological conditions. Previous IS research on the IT acceptance behavior of a specific group is rare, except for the studies by Phang et al [12] and Deng et al [8]. On our part, the moderating effect of declining physiological conditions was tested in the trust transfer from offline to mobile channels. Indeed, exploring the moderating role of declining physiological conditions on trust transfer will facilitate the understanding of which conditions are more effective and thus further extend our knowledge of the trust transfer theory.

Third, support from hospitals was introduced in the context of research on mHealth services. Social support is positively associated with health behavior in the health care literature [61-63]. However, studies on the role of social support in moderating the associations between these 2 trust elements and intention to use have been limited. We suggest that future studies should take social support into account in an attempt to better understand how trust elements influence elderly users' attitude changes and health behaviors.

This study also reflects several practical implications. First, the important role of the relationship between offline health services and mHealth services suggests that offline trust can be used as an enabler that encourages a provider of mHealth services to expand from the offline to mobile channels. Accordingly, mHealth service providers are encouraged to increase their cooperation with hospitals that are trusted by patients when marketing their mHealth services to elderly users. This may lead to higher adoption rates than simply promoting mHealth services in isolation.

Second, hospitals should undertake responsibility for ensuring that the elderly use mHealth services because they possess limited health literacy and limited experience with mHealth services. However, hospitals need to provide support to elderly users in a proper way to avoid them being overly dependent on hospitals. This is because elderly users who rely too much on support from hospitals are less likely to develop trust in mHealth services, which can weaken their intention to use mHealth services.

Third, mHealth service providers need to understand the different behaviors among different groups of users of mHealth services. Service providers are urged to employ tailored strategies to promote their mHealth services to elderly users when they develop marketing campaigns targeting the building of trust in a mobile environment.

Limitations

As with all empirical research, this study has its limitations. First, the study did not include users of all age groups. Elderly users were taken as the sample in our study because this specific group accounts for a large portion of all users of mHealth services. Our results need to be interpreted with caution for applications in other population and age groups. Second, this study was conducted in China, and the results may be applicable only in cultural contexts similar to those of the Chinese mainland. We suggest a similar study in a Western context for comparing the results across different cultures. Third, although the explanatory power of the model is acceptable (48.3% for intention to use), we still advocate the potential to enhance our explanatory power through taking additional factors into consideration, in future research.

Conclusions

In conclusion, we reiterate that mHealth services are regarded as an essential means to alleviate the conflicts between the medical demands of an increasingly aging population and limited medical care resources. However, the adoption rate of mHealth services among elderly users still remains low. Our research draws upon the trust transfer theory and builds a comprehensive framework by integrating declining physiological conditions and support from hospitals to investigate the initial trust-building mechanism of mHealth services. The results indicate that trust in offline health services has a significant effect on trust in mHealth services, thereby leading to an intention to use mHealth services. Declining physiological conditions is seen to positively moderate the association between trust in offline health services and trust in mHealth services; however, contrary to our hypotheses, support from hospitals weakens the association between trust in mHealth services and the intention to use mHealth services. The results provide a good explanatory power to predict the mHealth use intentions of elderly users. These findings have advanced the trust transfer theory and enriched the literature on mHealth services. Accordingly, mHealth services practitioners can better understand how to leverage the benefits of trust transfer and the characteristics of elderly users to promote their mHealth services.

Acknowledgments

This study was partially funded by the National Natural Science Foundation of China (numbers 71531007, 71622002, 71471048, 71871074, 71871073).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaire: measurement of the major constructs.

[PDF File (Adobe PDF File), 24KB - mhealth_v7i1e12269_app1.pdf]

References

1. World Health Organization. Geneva: WHO; 2015. World report on ageing and health URL: http://www.who.int/ageing/publications/world-report-2015/en/ [accessed 2018-11-02] [WebCite Cache ID 73dGRbRb8]
2. World Health Organization. Geneva: WHO; 2015. China country assessment report on ageing and health URL: http://www.who.int/ageing/publications/china-country-assessment/en/ [accessed 2018-11-02] [WebCite Cache ID 73dGigHrw]
3. Chiarini G, Ray P, Akter S, Masella C, Ganz A. mHealth technologies for chronic diseases and elders: a systematic review. IEEE J Sel Areas Commun 2013 Sep;31(9):6-18. [doi: 10.1109/JSAC.2013.SUP.0513001]
4. Angst CM, Agarwal R. Adoption of electronic health records in the presence of privacy concerns: the elaboration likelihood model and individual persuasion. MIS Quarterly 2009 Jun;33(2):339-370. [doi: 10.2307/20650295]
5. Mishra AN, Anderson C, Angst CM, Agarwal R. Electronic health records assimilation and physician identity evolution: an identity theory perspective. Info Syst Res 2012 Sep;23(3-part-1):738-760. [doi: 10.1287/isre.1110.0407]
6. 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 2;15(8):e149. [doi: 10.2196/jmir.2635] [Medline: 23912839]
7. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: a scoping review. JMIR mHealth uHealth 2016 Jun 9;4(2):e29. [doi: 10.2196/mhealth.5127] [Medline: 27282195]
8. Deng Z, Mo X, Liu S. Comparison of the middle-aged and older users' adoption of mobile health services in China. Int J Med Inform 2014 Mar;83(3):210-224. [doi: 10.1016/j.ijmedinf.2013.12.002] [Medline: 24388129]
9. Choi N. Relationship between health service use and health information technology use among older adults: analysis of the US National Health Interview Survey. J Med Internet Res 2011 Apr 20;13(2):e33 [FREE Full text] [doi: 10.2196/jmir.1753] [Medline: 21752784]
10. Gracia E, Herrero J. Internet use and self-rated health among older people: a national survey. J Med Internet Res 2009 Dec 2;11(4):e49 [FREE Full text] [doi: 10.2196/jmir.1311] [Medline: 19955041]
11. Morris MG, Venkatesh V. Age differences in technology adoption decisions: implications for a changing work force. Pers Psychol 2000 Jun;53(2):375-403. [doi: 10.1111/j.1744-6570.2000.tb00206.x]
12. Phang CW, Sutanto J, Kankanhalli A, Li Y, Tan BC, Teo HH. Senior citizens' acceptance of information systems: a study in the context of e-government services. IEEE Trans Eng Manag 2006 Nov;53(4):555-569. [doi: 10.1109/TEM.2006.883710]
13. Mathur A. Adoption of technological innovations by the elderly: a consumer socialization perspective. J Mark Manag 1999;9(3):21-35.
14. Gefen D. E-commerce: the role of familiarity and trust. Omega 2000 Dec;28(6):725-737. [doi: 10.1016/S0305-0483(00)00021-9]
15. Kim E, Tadisina S. A model of customers' trust in e-businesses: micro-level inter-party trust formation. J Comput Inform Syst 2007;48(1):88-104. [doi: 10.1080/08874417.2007.11645998]
16. Stewart KJ. Trust transfer on the World Wide Web. Organ Sci 2003 Jan;14(1):5-17. [doi: 10.1287/orsc.14.1.5.12810]
17. Akter S, D'Ambra J, Ray P. Trustworthiness in mHealth information services: an assessment of a hierarchical model with mediating and moderating effects using partial least squares (PLS). J Am Soc Inf Sci Technol 2011 Jan;62(1):100-116. [doi: 10.1002/asi.21442]
18. Guo X, Zhang X, Sun Y. The privacy–personalization paradox in mHealth services acceptance of different age groups. Electron Commer Res Appl 2016 Mar;16:55-65. [doi: 10.1016/j.elerap.2015.11.001]
19. Doney PM, Cannon JP. An examination of the nature of trust in buyer-seller relationships. J Mark 1997 Apr;61(2):35-51. [doi: 10.2307/1251829]
20. Lin J, Lu Y, Wang B, Wei KK. The role of inter-channel trust transfer in establishing mobile commerce trust. Electron Commer Res Appl 2011 Nov;10(6):615-625. [doi: 10.1016/j.elerap.2011.07.008]
21. Turel O, Yuan Y, Connelly CE. In justice we trust: predicting user acceptance of e-customer services. J Manag Inform Syst 2008 Dec;24(4):123-151. [doi: 10.2753/MIS0742-1222240405]
22. Belanche D, Casaló LV, Flavián C, Schepers J. Trust transfer in the continued usage of public e-services. Inform Manag 2014 Sep;51(6):627-640. [doi: 10.1016/j.im.2014.05.016]
23. Lu Y, Yang S, Chau PY, Cao Y. Dynamics between the trust transfer process and intention to use mobile payment services: a cross-environment perspective. Inform manage 2011 Dec;48(8):393-403. [doi: 10.1016/j.im.2011.09.006]

24. Wang N, Shen XL, Sun Y. Transition of electronic word-of-mouth services from web to mobile context: a trust transfer perspective. Decis Support Syst 2013 Feb;54(3):1394-1403. [doi: 10.1016/j.dss.2012.12.015]
25. Gefen D, Karahanna E, Straub DW. Trust and TAM in online shopping: an integrated model. MIS Quarterly 2003 Mar;27(1):51-90. [doi: 10.2307/30036519]
26. Mayer RC, Davis JH, Schoorman FD. An integrative model of organizational trust. Acad Manag Rev 1995 Jul;20(3):709-734. [doi: 10.2307/258792]
27. Jarvenpaa SL, Tractinsky N, Saarinen L. Consumer trust in an internet store: a cross cultural validation. J Comput Mediat Commun 1999 Dec;5(2). [doi: 10.1111/j.1083-6101.1999.tb00337.x]
28. McKnight DH, Choudhury V, Kacmar C. Developing and validating trust measures for e-Commerce: an integrative typology. Inf Syst Res 2002 Sep;13(3):334-359. [doi: 10.1287/isre.13.3.334.81]
29. Kim KK, Prabhakar B. Initial trust and the adoption of B2C e-commerce: the case of internet banking. ACM sigmis database 2004;35(2):50-64. [doi: 10.1145/1007965.1007970]
30. Carter L, Bélanger F. The utilization of e-government services: citizen trust, innovation and acceptance factors. Inf Syst J 2005 Jan;15(1):5-25. [doi: 10.1111/j.1365-2575.2005.00183.x]
31. Lim ET, Tan CW, Cyr D, Pan SL, Xiao B. Advancing public trust relationships in electronic government: the Singapore e-filing journey. Inf Syst Res 2012 Dec;23(4):1110-1130. [doi: 10.1287/isre.1110.0386]
32. Pavlou PA. Consumer acceptance of electronic commerce: integrating trust and risk with the technology acceptance model. Int J Electron Commer Re 2003;7(3):101-134. [doi: 10.1080/10864415.2003.11044275]
33. Palvia P. The role of trust in e-commerce relational exchange: a unified model. Inform Manage 2009 May;46(4):213-220. [doi: 10.1016/j.im.2009.02.003]
34. Gao L, Waechter KA. Examining the role of initial trust in user adoption of mobile payment services: an empirical investigation. Inform Syst Front 2017 Jun;19(3):525-548. [doi: 10.1007/s10796-015-9611-0]
35. Zhao Y, Ni Q, Zhou R. What factors influence the mobile health service adoption? a meta-analysis and the moderating role of age. Int J Inf Manag 2018 Dec;43:342-350. [doi: 10.1016/j.ijinfomgt.2017.08.006]
36. 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]
37. Akter S, Ray P, D'Ambra J. Continuance of mHealth services at the bottom of the pyramid: the roles of service quality and trust. Electron Mark 2012 Mar;23(1):29-47. [doi: 10.1007/s12525-012-0091-5]
38. Grundy E. Disability in Great Britain: Results From the 1996/97 Disability Follow-Up to the Family Resources survey. Great Britain: Department of Social Security; Jul 31, 1999:94.
39. Timmermann S. The role of information technology in older adult learning. New Dir Adult Cont Educ 1998;77:61-71. [doi: 10.1002/ace.7706]
40. Heart T, Kalderon E. Older adults: are they ready to adopt health-related ICT? Int J Med Inform 2013 Nov;82(11):e209-e231. [doi: 10.1016/j.ijmedinf.2011.03.002] [Medline: 21481631]
41. Xue L, Yen CC, Chang L, Chan HC, Tai BC, Tan SB, et al. An exploratory study of ageing women's perception on access to health informatics via a mobile phone-based intervention. Int J Med Inform 2012 Sep;81(9):637-648. [doi: 10.1016/j.ijmedinf.2012.04.008] [Medline: 22658778]
42. Leithelser RL, Wetherbe JC. Service support levels: an organized approach to end-user computing. MIS Quarterly 1986 Dec;10(4):337-349. [doi: 10.2307/249187]
43. Lee DM. Usage pattern and sources of assistance for personal computer users. MIS Quarterly 1986 Dec;10(4):313-325. [doi: 10.2307/249183]
44. Gallivan MJ, Spitler VK, Koufaris M. Does information technology training really matter? a social information processing analysis of coworkers' influence on IT usage in the workplace. J Manag Inf Syst 2005;22(1):153-192. [doi: 10.1080/07421222.2003.11045830]
45. Compeau DR, Higgins CA. Computer self-efficacy: development of a measure and initial test. MIS Quarterly 1995 Jun;19(2):189-211. [doi: 10.2307/249688]
46. Hoque R, Sorwar G. Understanding factors influencing the adoption of mHealth by the elderly: an extension of the UTAUT model. Int J Med Inform 2017 Dec;101:75-84. [doi: 10.1016/j.ijmedinf.2017.02.002] [Medline: 28347450]
47. Sirdeshmukh D, Singh J, Sabol B. Consumer trust, value, and loyalty in relational exchanges. J Mark 2002 Jan;66(1):15-37. [doi: 10.1509/jmkg.66.1.15.18449]
48. Thong JY, Yap CS, Raman KS. Engagement of external expertise in information systems implementation. J Manag Inf Syst 1994;11(2):209-231. [doi: 10.1080/07421222.1994.11518046]
49. Fornell C, Bookstein FL. Two structural equation models: LISREL and PLS applied to consumer exit-voice theory. J Mark Res 1982 Nov;19(4):440-452. [doi: 10.2307/3151718]
50. Gefen D, Straub DW, Rigdon E. An update and extension to SEM guidelines for administrative and social science research. MIS Quarterly 2011;35(2):iii-xiv. [doi: 10.2307/23044042]
51. Hair JF, Ringle CM, Sarstedt M. PLS-SEM: indeed a silver bullet. J Market Theor Pract 2011;19(2):139-152. [doi: 10.2753/MTP1069-6679190202]

52. Hair JF, Anderson RE, Tatham RL, Black WC. Multivariate Data Analysis. Fifth edition. New York: Prentice Hall; Mar 23, 1998:768.
53. Hair JJ, Hult GT, Ringle C, Sarstedt M. A primer on partial least squares structural equation modeling (PLS-SEM). Great Britain: Sage Publications; 2016.
54. Fornell C, Larcker DF. Evaluating structural equation models with unobservable variables and measurement error. J Mark Res 1981 Feb;18(1):39-50. [doi: 10.2307/3151312]
55. Chin WW. The partial least squares approach to structural equation modeling. In: Marcoulides GA, editor. Modern Methods for Business Research. Hove, East Sussex: Psychology Press; Jan 1998:295-336.
56. Nunnally JC, Bernstein IH. Psychometric theory. New York: McGraw-Hill; 1978:736.
57. Podsakoff PM, MacKenzie SB, Lee JY, Podsakoff NP. Common method biases in behavioral research: a critical review of the literature and recommended remedies. J Appl Psychol 2003 Oct;88(5):879-903. [doi: 10.1037/0021-9010.88.5.879] [Medline: 14516251]
58. Harman HH. Modern factor analysis. Chicago: University of Chicago Press; Apr 1, 1976:508.
59. Baron RM, Kenny DA. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol 1986 Dec;51(6):1173-1182. [Medline: 3806354]
60. Kuan HH, Bock GW. Trust transference in brick and click retailers: an investigation of the before-online-visit phase. Inform Manage 2007 Mar;44(2):175-187. [doi: 10.1016/j.im.2006.12.002]
61. Helgeson VS, Cohen S. Social support and adjustment to cancer: reconciling descriptive, correlational, and intervention research. Health Psychol 1996 Mar;15(2):135-148. [Medline: 8681922]
62. McKinley CJ. Investigating the influence of threat appraisals and social support on healthy eating behavior and drive for thinness. Health Commun 2009 Dec;24(8):735-745. [doi: 10.1080/10410230903264303] [Medline: 20183382]
63. Yan L, Tan Y. Feeling blue? go online: an empirical study of social support among patients. Inf Syst Res 2014 Dec;25(4):667-891. [doi: 10.1287/isre.2014.0538]

Abbreviations

AVE: average variance extracted
CB: covariance-based
CR: composite reliability
eWOM: e-word of mouth
HIT: health information technology
IS: information system
IT: information technology
mHealth: mobile health
PLS: partial least squares
SEM: structural equation modeling
VIF: variance inflation factor

Edited by G Eysenbach; submitted 20.09.18; peer-reviewed by X Zhang, F Guo, D Leightley; comments to author 11.10.18; revised version received 19.11.18; accepted 09.12.18; published 08.01.19.

Please cite as:
Meng F, Guo X, Peng Z, Lai KH, Zhao X
Investigating the Adoption of Mobile Health Services by Elderly Users: Trust Transfer Model and Survey Study
JMIR Mhealth Uhealth 2019;7(1):e12269
URL: http://mhealth.jmir.org/2019/1/e12269/
doi:10.2196/12269
PMID:30622092

Original Paper

Usability Challenges for Health and Wellness Mobile Apps: Mixed-Methods Study Among mHealth Experts and Consumers

Mei Shan Liew[1*], MBA; Jian Zhang[1*], MBA; Jovis See[1], MBA; Yen Leng Ong[1], MBA

Alliance Manchester Business School, Manchester, United Kingdom

*these authors contributed equally

Corresponding Author:
Mei Shan Liew, MBA
Alliance Manchester Business School
Booth Street West
Manchester, M15 6PB
United Kingdom
Phone: 44 161 820 8344
Email: liewmeishan@gmail.com

Abstract

Background: By 2019, there will be an estimated 4.68 billion mobile phone users globally. This increase comes with an unprecedented proliferation in mobile apps, a plug-and-play product positioned to improve lives in innumerable ways. Within this landscape, medical apps will see a 41% compounded annual growth rate between 2015 and 2020, but paradoxically, prevailing evidence indicates declining downloads of such apps and decreasing "stickiness" with the intended end users.

Objective: As usability is a prerequisite for success of health and wellness mobile apps, this paper aims to provide insights and suggestions for improving usability experience of the mobile health (mHealth) app by exploring the degree of alignment between mHealth insiders and consumers.

Methods: Usability-related major themes were selected from over 20 mHealth app development studies. The list of themes, grouped into 5 categories using the Nielsen usability model, was then used as a framework to identify and classify the responses from mHealth expert (insider) interviews. Responses from the qualitative phase were integrated into some questions for a quantitative consumer survey. Subsequently, categorical data from qualitative mHealth insider interviews and numerical data from a quantitative consumer survey were compared in order to identify common usability themes and areas of divergence.

Results: Of the 5 usability attributes described in Nielsen model, *Satisfaction* ranked as the top attribute for both mHealth insiders and consumers. *Satisfaction* refers to user likability, comfort, and pleasure. The consumer survey yielded 451 responses. Out of 9 mHealth insiders' top concerns, 5 were similar to those of the consumers. On the other hand, consumers did not grade themes such as *Intuitiveness* as important, which was deemed vital by mHealth insiders. Other concerns of the consumers include in-app charges and advertisements.

Conclusions: This study supports and contributes to the existing pool of mixed-research studies. Strengthening the connectivity between suppliers and users (through the designed research tool) will help increase uptake of mHealth apps. In a holistic manner, this will have a positive overall outcome for the mHealth app ecosystem.

(JMIR Mhealth Uhealth 2019;7(1):e12160) doi:10.2196/12160

KEYWORDS

mobile applications; health and wellness mobile applications; mHealth apps; health and wellness mobile applications users; mHealth users; health and wellness mobile applications insiders; mHealth insiders

Introduction

Background

In October 2016, for the first time, internet usage from mobile devices exceeded that from desktop or notebook computers [1]. By 2019, 2.7 billion people will be smartphone users [2], thanks to double-digit growth in China along with emerging markets of Southeast Asia. The convenience of mobile devices with app software capabilities supports the use of such apps as a powerful tool to transform the delivery of medical care and health care. A 2016 estimate puts the total number of medical health apps available over 250,000; however, the overall number of downloads of such apps appears to be decreasing [3].

Assessment of mHealth Apps

It is uncertain whether the plethora of mobile heath (mHealth) apps is truly effective in improving health and wellness outcomes. Despite the attempts by several researchers to establish a systematic evaluation framework, evidence of their efficacy remains sparse [4-6]. Nevertheless, various forms of usability research models that may be broadly applicable have been proposed [7].

Obstacles to efficacy evaluation of mHealth apps tend to be related to the relative newness of the technology and the explosive pace of market growth in the past decade. In a study focused on the evaluation of mHealth measurement methodology, reliability, and validity of data, Kumar et al [8] concluded that the main challenges are the effects of variability on time-intensive data collection and the lack of a gold standard to assess convergent validity. In addition, the number of apps makes it challenging to set common standards for effectiveness. A review of 75 controlled trials of mobile technology-based health interventions found that most interventions were of low quality despite being conducted in high-income countries, especially when there was high variability in the types of measured outcomes [9]. For health care systems already burdened by suboptimal outcomes and excessive costs, premature clinical adoption of these mHealth technologies may detract from, rather than contribute to, what is needed for true overall health improvement.

Usability of mHealth Apps

Broadly speaking, "usability" as a concept is about product quality and user experience. Its use as an assessment criteria has been widely adopted in the software inspection world since its emergence in 1990 [10] and is the predominant development delivery target for mobile apps today [11]. It has been shown that without considering usability, mobile apps are unable to retain users; tracking data has shown that users typically allocate less than 30 seconds to learn how to use the app before abandoning it for alternatives, or simply give up using mobile apps for this purpose altogether.

The scope of usability was defined early on within the International Organization for Standardization standards as "Extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use," outlining *Effectiveness*, *Efficiency*, and *Satisfaction* as measurable attributes [12]. In contrast, the Nielsen model [10] is commonly used in the assessment of mHealth apps due to the suitability of its attributes for the assessment of software products; this model measures *Efficiency*, *Satisfaction*, *Learnability*, *Memorability*, and *Errors* as attributes. Zhang and Adipat [11] also highlighted issues specific to mobile devices, such as the mobile context, hardware limitations, and data-entry challenges as additional considerations not addressed with these models.

Research Focus

Despite the increasing willingness of consumers to try mHealth apps [12], it is important to understand and appeal to the motivations of the users and decrease barriers to "digital adherence." User experience research is thus a critical component to the success of mHealth apps. These can be in the form of focus groups and interviews to identify relevant themes [13-15]. Such tailored approaches provide valuable insights that facilitate extension of the target consumer base to nonconventional populations such as the elderly or patients from lower socioeconomic segments. Response analysis may also be performed within the context of relevant models such as the Health Information Technology Acceptance Model and the Mobile Application Rating Scale [16] to allow for full coverage of technological assessment.

Alternatively, a way to improve upon intrinsic limitations of such post-hoc analyses is to track users' behavior in real time, allowing for instant feedback and improvements [17-19]. The main limitation of this method is potential selection bias in consumers, which is unsuitable for a product designed for wider audiences.

Although there are some mixed-research studies for mHealth apps and wearables [20-23], their focus is mostly at the app level for selected app products or target limited audiences. There are few studies at the systems level that attempt to understand the degree of alignment between the supply and demand with respect to the business environment.

Even a well-designed research approach is still a tedious process where the respondent population size is often a constraint. Consequently, answers might lie somewhere in an area that is not captured within the framework of the designed research approach. Compatibility of the research question to its relevant use of research tools has to be explored. Using the simple tool of 5W1H to illustrate, studies have thus far been dedicated to addressing questions such as "where are the gaps," "who are involved," "how can we improve," and even "why are the gaps present." We would like to contribute content to the area of "what," specifically, what should be the focus area(s) of usability for mHealth app development.

The overall objective of this study was to help improve the ecosystem of digital health care by providing concrete directions for mHealth app companies with regard to consumers' real needs. By examining the priorities and degree of alignment between mHealth insiders and consumers, critical hidden roadblocks are uncovered.

Design and Setting

After reviewing various research approaches described by Creswell [24], a mixed-methods research design was selected to collect both categorical and numerical data. Qualitative data obtained from mHealth app insiders through face-to-face interviews were analyzed and integrated into the quantitative survey of the end consumers. Over the past two decades, mixed methods have become increasingly popular [25], and this study can be a valuable contribution to the literature.

The rationale of using the mixed-methods design was to explore the usability and key concerns highlighted by mHealth app insiders without prejudice. The consumer data will be used to validate the themes identified by mHealth app insiders and provide crucial insights that will have been otherwise neglected by the experts, aiming to better synchronize the developmental roadmap for mHealth apps and clarify the priorities in the

consumers' needs within the mHealth apps market. Three specific objectives addressed in this study are (1) identification of the major concerns of usability with regard to mHealth apps from mHealth insiders (eg, product developers), (2) evaluation of the relative importance of the experts' listed concerns in the mind of consumers, and (3) identification of significant themes on usability from the consumers that may otherwise be neglected by the experts.

Methods

Development and Integration of Usability Themes

Under the International Organization for Standardization code 25010, two components specifically involve the user perspective, namely, functional suitability and usability. Functional suitability is applicable at the product level, where specific product-related questions are needed for execution. On the other hand, usability can be tested at the systems level, providing an opportunity for researchers to design high-level exploratory research. Hence, this research study employed the Nielsen usability model with the aim of performing a system-wide alignment analysis between the mHealth insider and consumer [10].

For this study, usability-related major themes were selected from over 20 mHealth app development studies. The final list comprised 22 unique usability themes (Table 1). The list was then used as a framework to identify and classify the responses from mHealth insider interviews during the coding analysis. A total of 19 usability themes mentioned by mHealth insiders in the qualitative data-collection phase matched the identified usability themes. No additional usability themes were captured from the qualitative phase. To meet the objectives of this study, all developed usability themes were integrated into the consumer's quantitative survey questionnaire.

Participants and Sampling

mHealth insiders were broadly identified from two main industries: Medical Healthcare and Information Technology (IT). Target candidates were experts with more than 10 years of work experience in the domain of health care or software app development for the Asia-Pacific market. A total of 19 experts were shortlisted for the qualitative interviews.

SoJump (WJX) was used for the quantitative survey platform [48]. In addition to the extensive survey features, SoJump offers access to China's market, which is a critical interest for Asia Pacific. The minimum sample size was set at 384 participants in order to achieve 95% confidence level based on the ~4.5 billion population within the Asia Pacific region [49].

Data Collection and Statistical Analysis

All the data for this study were collected within a 6-month period in 2017. Operationally, this mixed-research approach relied on lean design principles (Figure 1). The recruitment process for qualitative and quantitative components was based on selective and snowballing methods, respectively. As the qualitative interview fieldwork was more time consuming, qualified candidates were prescreened. Data-cleaning treatment was applied to assure data quality [50]. For example, for the quantitative online questionnaire, two eligibility questions were built into the survey and data from disqualified participants were removed after survey closure.

Qualitative Study

Each interviewee in the panel of mHealth insiders signed a consent form that includes publication rights to the shared content. Written survey questions were provided to the interviewee during the process, which was also audio recorded. During each interview, all questions were sequentially asked, including the open-ended questions.

For the coding analysis of the transcripts, the authors adopted the Bengtsson method [51]. This 4-step data analysis process includes decontextualization, recontextualization, categorization, and compilation. To ensure scientific adequacy, both transcription of the audio file to text and coding analysis for each transcript were verified twice. A total of over 100 man-hours were spent in the qualitative interview data-collection phase. An industry expert assigned by the Alliance Manchester Business School, The University of Manchester, also validated the entire coding analysis processes and findings.

Quantitative Study

After coding analysis and theme identification from the qualitative study (Figure 2), relevant usability-related elements were incorporated into the consumer questionnaire. A pilot test with 10 consumers was performed for fine tuning.

The survey consisted of 34 questions [52] with a profiling section of 14 demographic questions, including qualifying questions to ensure that the respondents are from the Asia Pacific region and using smartphone devices. In addition, the survey platform captured participants' location and device used, which helped in data processing. The questions were arranged in chronological order of consumers' touch-point cycle, from initial awareness to engagement and postuse feedback. As this study focuses on mobile devices, the survey links were sent via mobile device-compatible messaging apps such as Whatsapp (Whatsapp Inc, Mountain View, CA), Line (Line Corporation, Tokyo, Japan), and Wechat (Tencent, Shenzhen, China). A total of 466 respondents completed the online survey.

The choice of input for usability perception questions follows the 5-point Likert scale, from least concern to most concern, or least important to most important. Likert data falls into the ordinal data type, and descriptive statistics include a mode or median for central tendency and frequencies for variability. To statistically claim that each concern is real for all samples, hypothesis testing for a proportion test with the frequency was conducted with a definition of null hypothesis and alternative hypothesis [53]. The one-sample Z-test was chosen for data analysis, as it is most compatible for ordinal data and the purpose is to test the population mean of frequency and determine the statistical significance.

Z-test follows the following formula [54]:

$$z_i=(P_i-P_o)/((P_o*[1-P_o]/n)\text{^}0.5)$$

where *Pi* is the percentage of the population that shows concern or strong concern of 4 and 5 points, respectively, on the Likert scale.

Table 1. Usability themes. Some studies cover multiple themes. For simplicity, only the representative study is stated for each theme.

Usability category	Theme	Author, year	Study subject	Quantitative or qualitative
Learnability	Intuitive, users' gestures are intuitive	Lin et al, 2017 [26]	Usability of data integration and visualization software for multidisciplinary pediatric intensive care	Quantitative
Satisfaction	Provide incentives	Rosario et al, 2012 [27]	A study in usability: Redesigning a health sciences library	Qualitative
Memorability	Familiarity, interface feels familiar and comfortable	Chevalier et al, 2014 [28]	The influence of the search complexity and familiarity with the website on the subjective appraisal of esthetics, mental effort, and usability	Qualitative
Memorability	Notification, utilizes useful notification alerts	Zeitz et al, 2016 [29]	Speed isn't enough: Usability and adoption of an optimized alert notification system	Qualitative
Efficiency	Lean design, data allow seamless sharing across operating system devices	Bosse and Kelly, 2016 [30]	Improving EHR[a] usability using lean methodology, studies in health technology and informatics	Qualitative
Efficiency	Efficiency, responsive and run smoothly	François et al, 2017 [31]	Digital, analogue, or redundant speedometers for truck driving: Impact on visual distraction, efficiency and usability	Qualitative
Efficiency	Actionable insights	Rose et al, 2017 [32]	Evaluating the usability of health insurance information with immigrant patients	Qualitative
Efficiency	IT[b] compatibility, compatible with mobile device and required limited bandwidth	Juslstrom et al, 2011 [33]	Telecoil-mode hearing aid compatibility performance requirements for wireless and cordless handsets: Magnetic signal-to-noise	Quantitative
Efficiency	Responsiveness, regular updates in response to consumer needs	Green and Pearson, 2011 [34]	Integrating website usability with the electronic commerce acceptance model	Qualitative
Learnability	Integration ability - Technology, paired with latest technologies	Lee and Coughlin, 2015 [35]	An integrated approach to identifying determinants and barriers	Quantitative
Learnability	Integration ability - Lifestyle, usage integrated into daily life	Mishuris et al, 2016 [36]	Online diabetes-prevention program	Qualitative
Satisfaction	Enjoyable, fun and interesting to use	Putrino et al, 2017 [37]	Game-based therapy in stroke	Qualitative
Satisfaction	Functional deliverable, contributes to health objectives	Kawamoto et al, 2009 [38]	Enabling a semantically interoperable service-oriented architecture for healthcare	Quantitative
Satisfaction	Match expectation, understands targeted health concerns and key needs	Reed et al, 2016 [39]	Novel personal health technology to support early palliative care	Quantitative
Satisfaction	Addresses specific needs, consumer pain point	Stjernswärd and Hansson, 2017 [40]	Web-based mindfulness intervention for families living with mental health problems	Qualitative
Satisfaction	Active engagement, interactive and engaging	Jimison et al, 2008 [41]	Barriers and drivers of health information technology use for the elderly, chronically ill, and underserved	Qualitative
Errors	Health care experts' involvement, advice from healthcare professionals	Kaipio et al, 2017 [42]	Usability problems do not heal by themselves: National survey on physicians' experiences with EHRs in Finland	Qualitative
Errors	Data accuracy	Ehrler et al, 2015 [43]	Usability of six data entry mobile interfaces for caregivers	Quantitative
Errors	Error free	Andreasen et al, 2017 [44]	Error-free text typing performance of an inductive intra-oral tongue computer interface	Quantitative
Satisfaction	Targets my demographic group	Armbrüster et al, 2007 [45]	The usability of track point and touchpad for middle-aged adults	Qualitative

Usability category	Theme	Author, year	Study subject	Quantitative or qualitative
Satisfaction	New features added frequently	Wolpin et al, 2015 [46]	Record title: Development and usability testing of a web-based cancer symptom and quality-of-life support intervention	Quantitative
Satisfaction	Progression analytics, providing visible progression on how much improvement, etc	Miah et al, 2017 [47]	Extending the framework for mobile health information systems research	Qualitative

[a]EHR: electronic health record.

[b]IT: information technology.

Figure 1. Qualitative and quantitative survey process map.

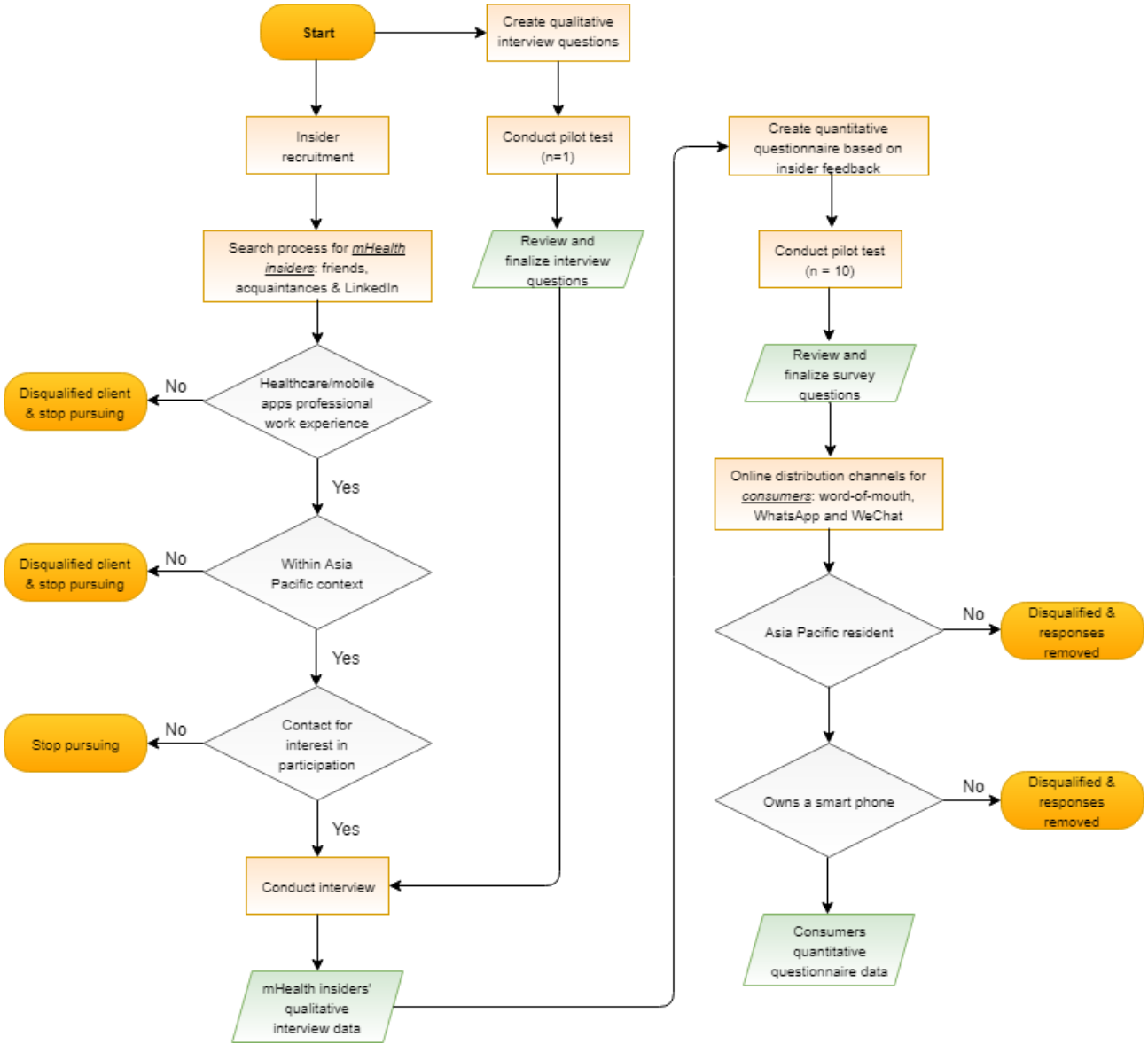


Figure 2. Qualitative study of the interview questions.

The *P* value is the probability of observing a sample statistic as extreme as the test statistic. Normal distribution cumulative probability was used to assess the probability associated with the computed Z-score. P_o is the hypothesized probability from the null hypothesis. The *P* value was then compared with the statistical confidence level, which was set as 95% [55]. The hypotheses that were accepted were considered statistically important themes and will be used as the basis for later discussion.

Results

Qualitative Interviews With mHealth Insiders

Qualitative interviews with mHealth insiders consisted mainly of open-ended questions to encourage them to express and elaborate their views. A total of 14 subjects were interviewed. This number falls within the proposed range of 5 to 25 participants for a phenomenological qualitative study [24,56]. One of the 14 candidates was disqualified due to insufficient experience in the Asia Pacific region.

The Bengtsson qualitative content-analysis method [51] was adopted and modified with a four-step analysis (Figure 3). The first step was to use decontextualization and recontextualization to identify all the common themes from coding analysis, followed by a typical content-analysis process, where the qualitative data were coded and the frequency of codes were counted and analyzed [57]. Next was the selection of themes only related to usability. Of the 32 common themes mentioned, 19 were usability related and the remaining 13 were mostly business and marketing related. The third step was to classify the usability themes into five categories based on the Nielsen definition [10], as described in Usability of mHealth Apps section. Finally, individual attribute priority was ranked based on the total number of mentions and the number of mHealth insiders who mentioned it.

Saturation testing was performed after completion of the qualitative study and content analysis. There is no common definition of saturation, but the generally agreed-upon principles and concepts are: no new themes, no new data or coding, and ability to replicate the study [58]. In this dataset, saturation for mHealth insiders' interviews was tested based on two criteria. The first is at which point all 19 themes were mentioned; we observed that this was reached by the 4th interview (Figure 4). The second is based on frequency of mentions in each of the five usability categories, to observe when saturation of alignment of the proportion of mentions in the categories was reached. We found that the frequency of mentions after the 8th interview was aligned with overall response (Figure 5), with differences within 2% for each category. With the two abovementioned criteria met, we concluded that the number of interviewees was sufficient, and the data collected reached saturation.

Figure 3. The 4-step data analysis flow for consumer interview.

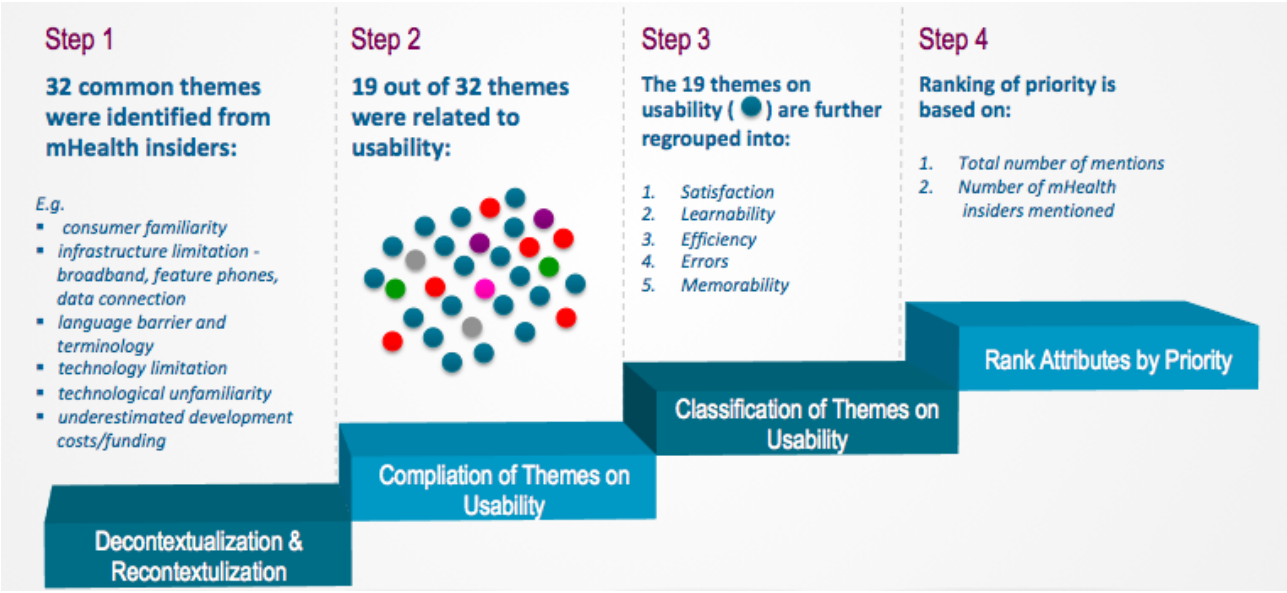


Figure 4. First observation to the 19 obtained usability themes.

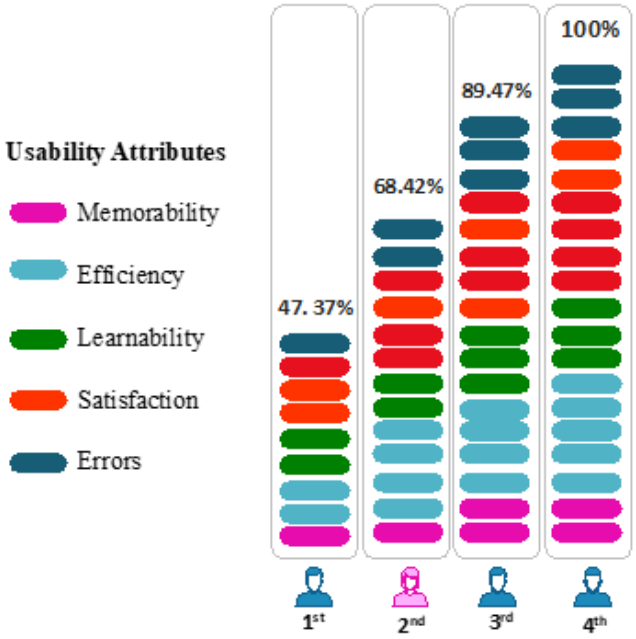
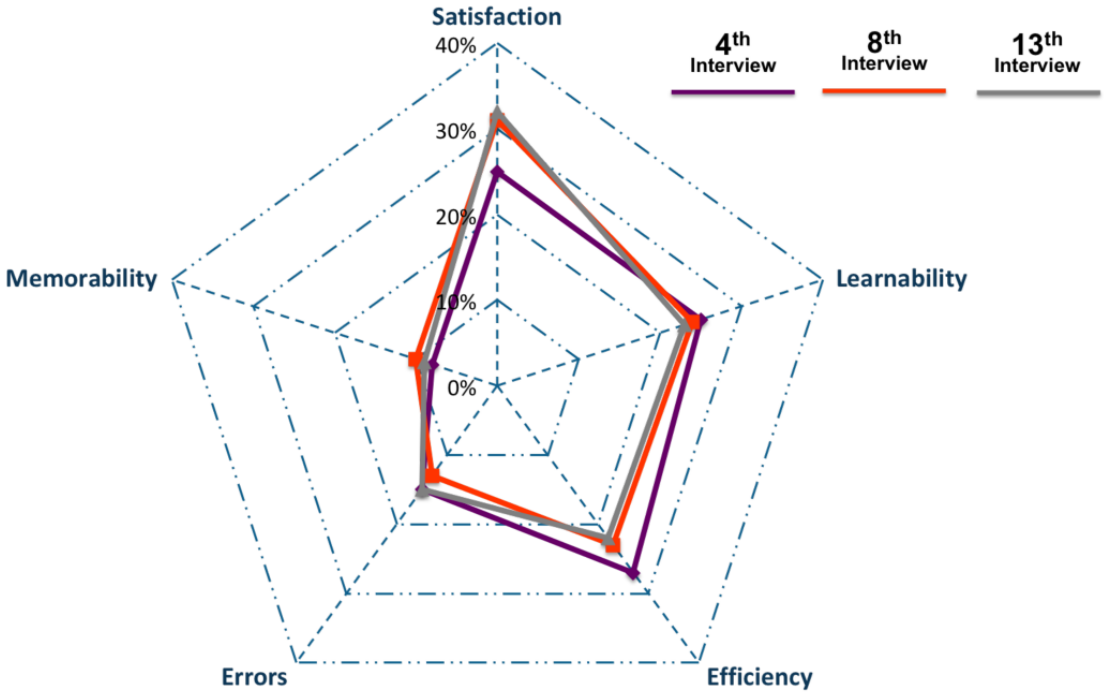


Figure 5. Second observation by frequency of mention.



Emerging Themes of Focus for mHealth Apps

Frequency of mentions was interpreted as an indicator of how the mHealth insiders perceived the importance or worthiness of a theme. The number of interviewees who mentioned a particular theme was interpreted as an indicator of how it has become a common understanding. The 19 usability themes were then summarized into an overview categorized by attribute (Figure 6). *Satisfaction* ranked as the overall most important attribute based on the number of total mentions, with 12 interviewees mentioning themes in this area. Additionally, although the total mentions were fewer, all insiders also discussed themes related to *Learnability* and *Efficiency*. Notably, questions were designed to avoid bias in favor of any particular attribute, which is why there is an unequal number of themes between each of the five usability attributes.

In the fourth step of data analysis, the 19 themes were prioritized based on the total number of mentions and the number of interviewees who mentioned them. Two synthetic conditions were then defined. First, the total number of mentions for that

particular theme should be equal to or greater than the average number of mentions from all interviewees (≥10). Second, the number of insiders who mentioned the theme was equal to or greater than half of all respondents (≥7). With both conditions met, the themes were classified as "important." A total of nine concerns met both criteria, with "lean design thinking" ranking as the top concern (Figure 7). Examples of quotes from the interviews that support the classification for the top concerns are also shown in Figure 8. The remaining 10 themes that did not meet the two criteria were classified as those of lower concern, with "notifications/getting attention" ranking the lowest in usability concern.

Quantitative Survey Results of Consumers

A total of 466 responses were received in the quantitative survey phase, with 451 valid samples after data cleaning. Of these, 53% respondents were men, 97% were Apple and Android operating system users, and more than 90% had bachelor's degrees or higher education. Respondents were from 11 Asian countries, with Singapore and China combined representing 90.9% of the total participants.

Three scenarios were evaluated: Threshold points set as 0.6, 0.7, and 0.8, where only 60%, 70%, or 80% of the population, respectively, agreed that it is a strong concern. All three scenarios passed the precondition check; thus, the Z-test could be used to determine whether the hypothesized population proportion differs significantly from the observed sample proportion [59].

Figure 6. Usability overview from mHealth insiders.

Figure 7. mHealth insiders' top concerns. Asterisk for top concern indicates total mentions≥10 (average) and number of people mentions≥7 (average). mHealth: mobile health.

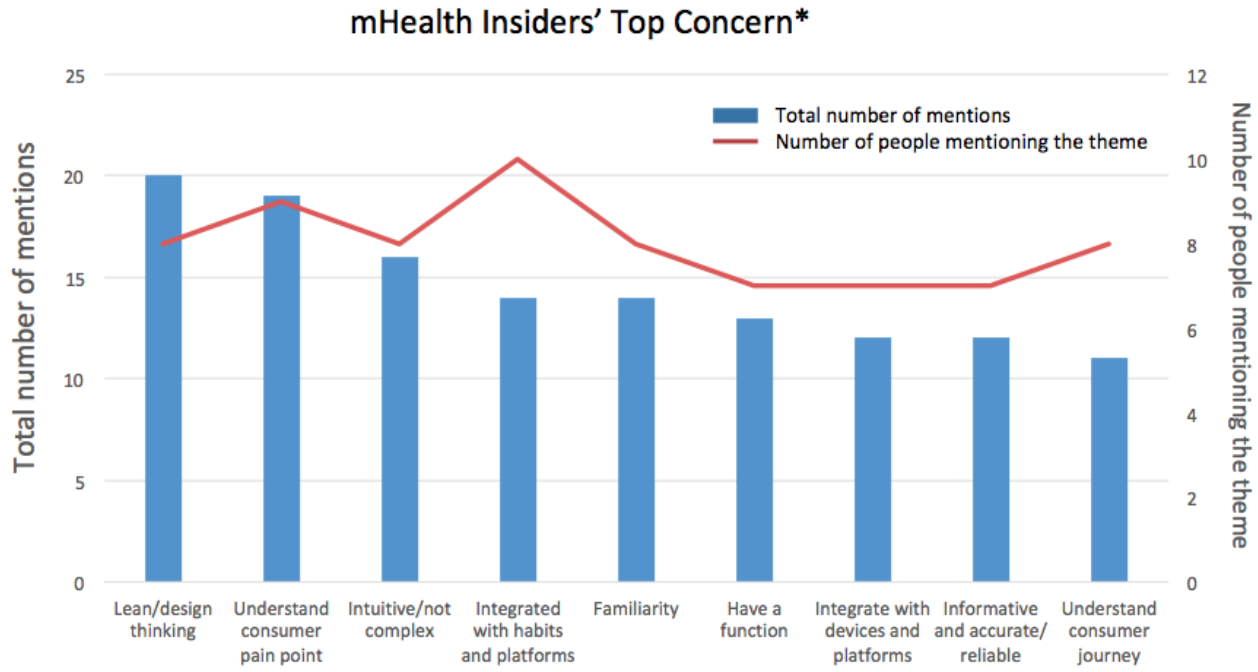
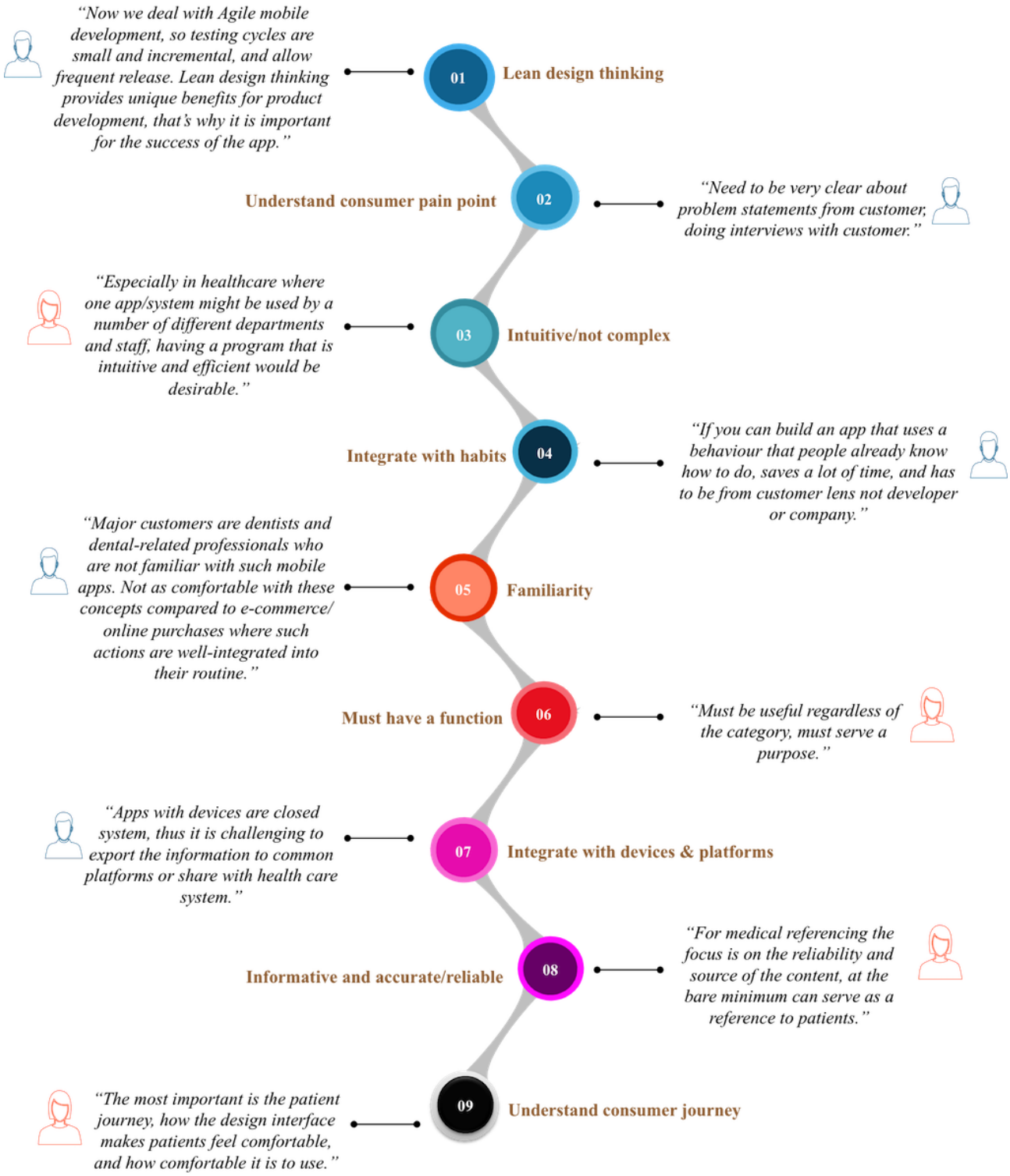


Figure 8. mHealth insider quotes for top concerns.



As shown in Figure 9, when P_o=0.6, 80% (18/22) of the themes across all five usability attributes were ranked important, which was insufficient to differentiate high versus relatively low concerns. Conversely, when P_o=0.8, only 25% of the themes (5/22) from three attributes were considered important. Although the top priority is standing out, it may also eliminate concerns that are worthy of study. Under these constraints, a moderate P_o=0.7 was selected to balance the opposing factors. Herein, the top concerns from mHealth insiders (12/22, with P_o=0.7) were classified into their respective usability attributes (Figure 10).

Given the heavy skewing of respondents' country, further analysis was performed to compare *overall response*, response from *China and Singapore combined*, response from *China only*, and response from *Singapore only*. To check if responses from respondents outside China and Singapore had any impact on the overall findings, the hypothesis test results from the *overall response* was compared against the response from *China and Singapore combined*. One item—*compatible with mobile*

device and requires limited bandwidth—was not found to be of particular significance to the latter population. Diving deeper into this point, respondents from *China only* also did not consider this item as critical; however, it was considered critical for respondents from *Singapore only*. It is interesting to note that Singapore has the cheapest and fastest average bandwidth in Asia [60], suggesting that the lack of importance to China's consumers may speak more towards the confidence in the quality of telecommunications access rather than just the cost consideration alone. The hypothesis test results of China versus Singapore versus the rest of Asia Pacific countries are summarized in Figure 11.

Ten themes common to China and Singapore and overall rated as high concerns and two themes—*compatible with mobile devices* and *healthcare experts' involvement*—were high concerns in Singapore and overall. Furthermore, there were five top concerns specific to Singapore (Area #2 in Figure 11). One possible factor for these Singapore-specific concerns is that there were >25 respondents who were over 50 years old and 30 respondents with only high school and equivalent education as compared to only one respondent from China in each category. For these demographics, there is potentially a higher level of tension regarding digital technology in health care and a stronger need for more engaging apps. This could be confirmed with further studies in the future, with a sufficient sample size.

Figure 9. Sensitivity analysis for comparison of themes.

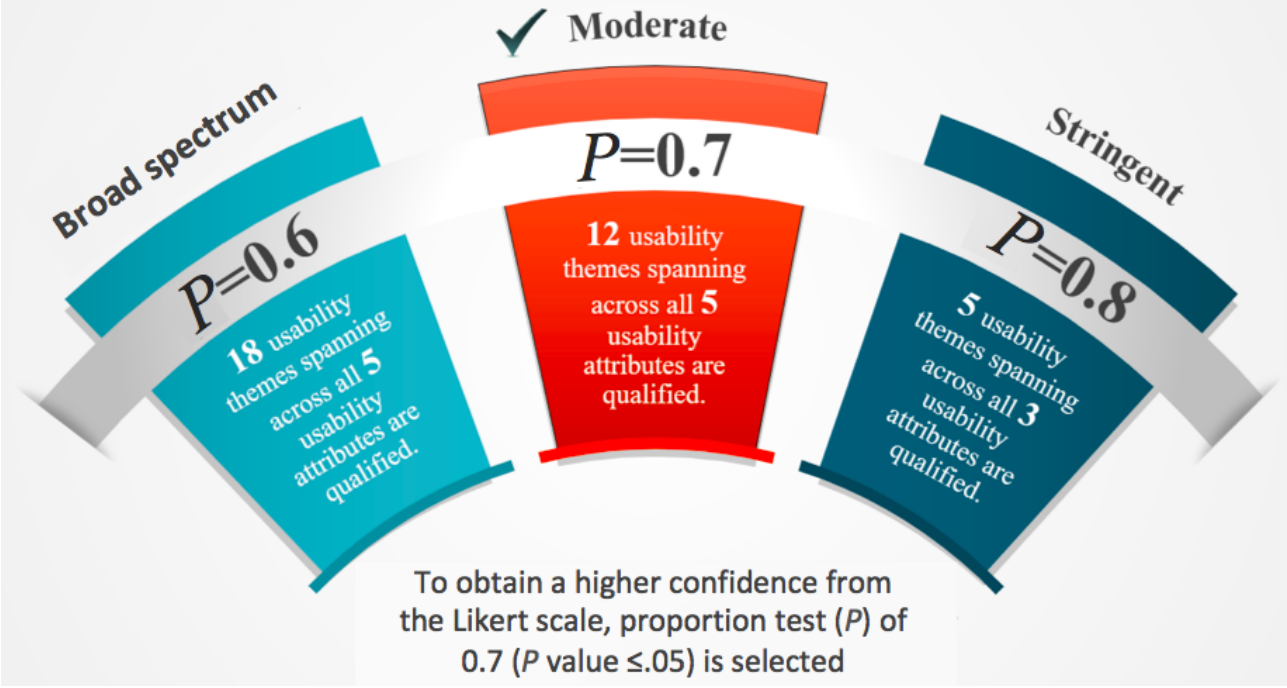
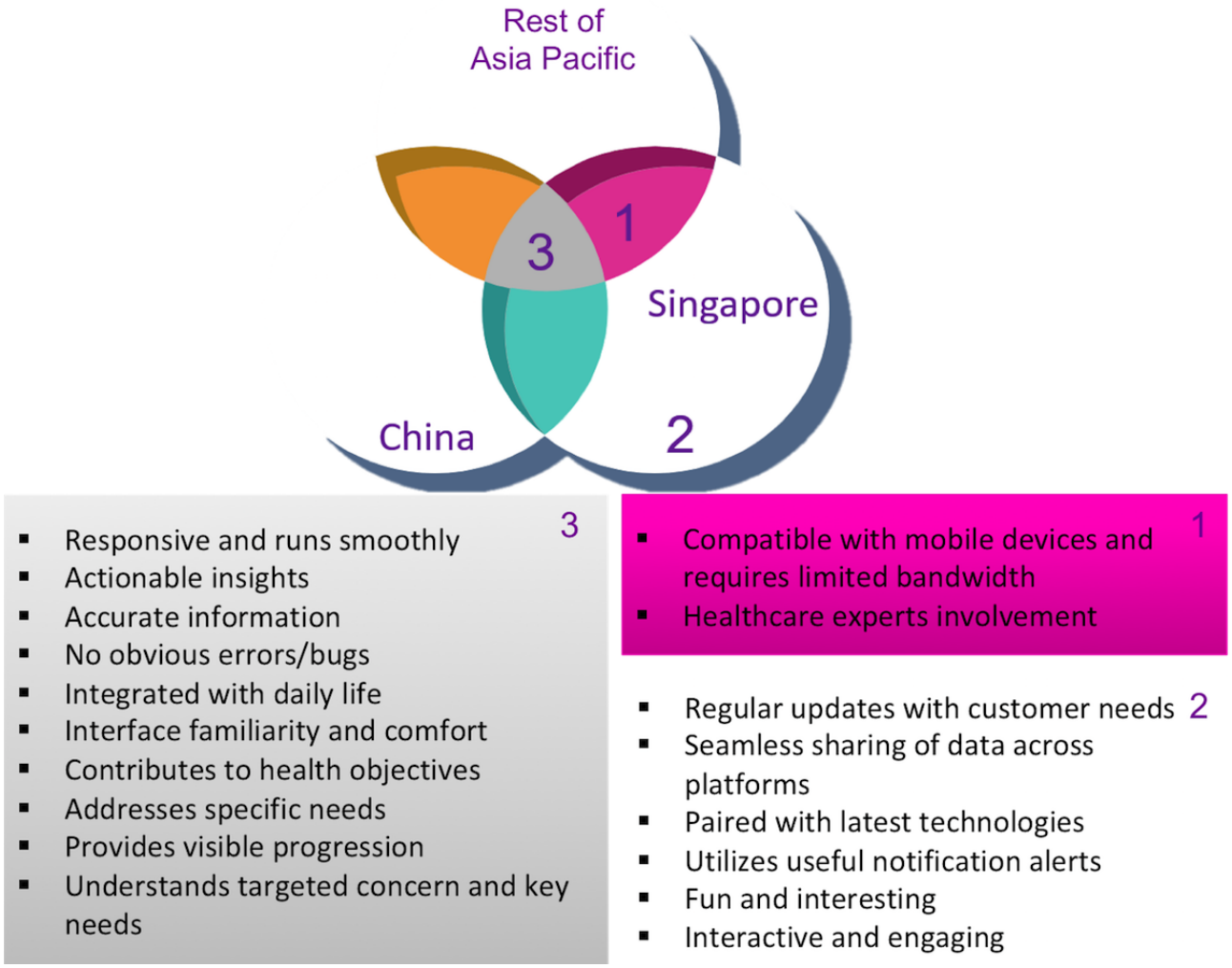


Figure 10. Number of top concerns from consumers.

Figure 11. Demographic comparisons of mHealth insiders' top concerns.



Alignment of Priorities Between the mHealth Apps Insiders and Consumers

Methods were merged by transforming qualitative data into categorical counts and variables and then integrated with the quantitative database, called *integration with data transformation* [61]. The number of total mentions from insiders was normalized into percentages for easier comparison with the statistically significant themes selected by the consumers at the 95% confidence level. Figure 12 summarizes this comparison in a graphical manner.

There was a high degree of alignment between mHealth insiders and consumers in their top concerns. *Satisfaction*, *Learnability*, and *Efficiency* were the top three attributes for both groups, with *Satisfaction* ranking the highest and *Errors* and *Memorability* ranking lower. In terms of divergence, the degree of concern for *Satisfaction* was different, where stronger concern was observed from the consumers. For insiders, the concerns of *Learnability* and *Efficiency* were similar; however, consumers' concerns of *Learnability* were less than those of *Efficiency*. One possible reason could be that once the *Efficiency* is met at a design level (eg, the three-click rule), there will be less reliance on *Learnability* for consumers. Hence, mHealth apps could be mastered with a gentler learning curve. The alignment of concerns in each theme is summarized in Table 2.

Figure 12. mHealth insider and consumer alignment study.

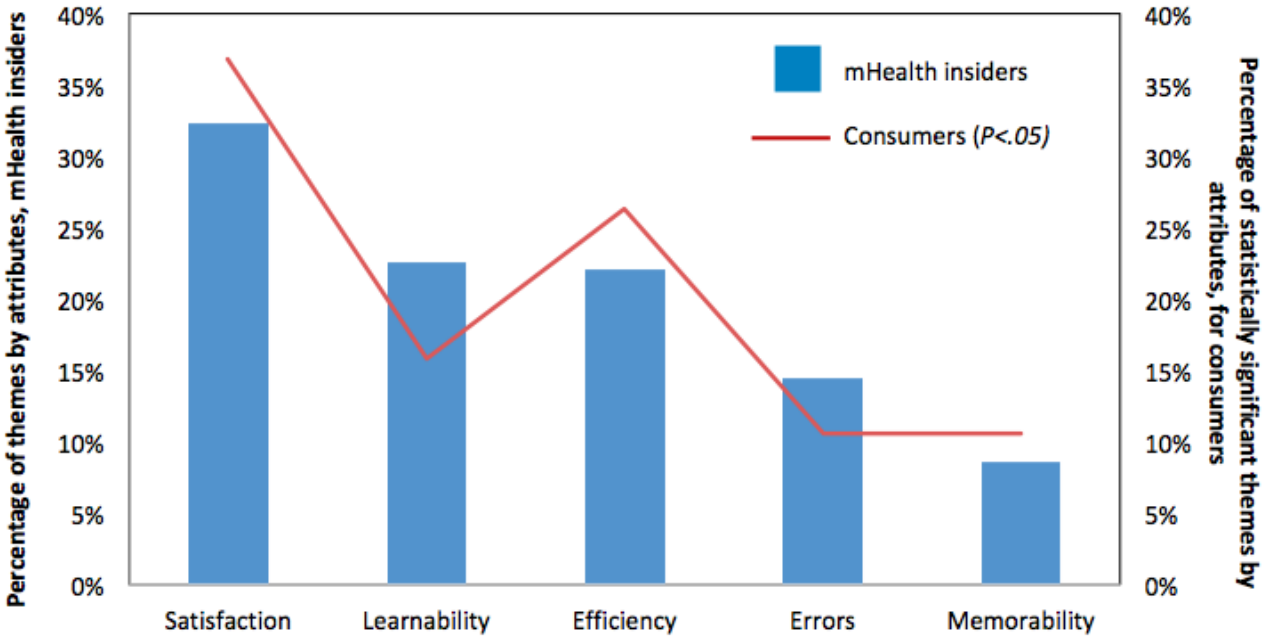
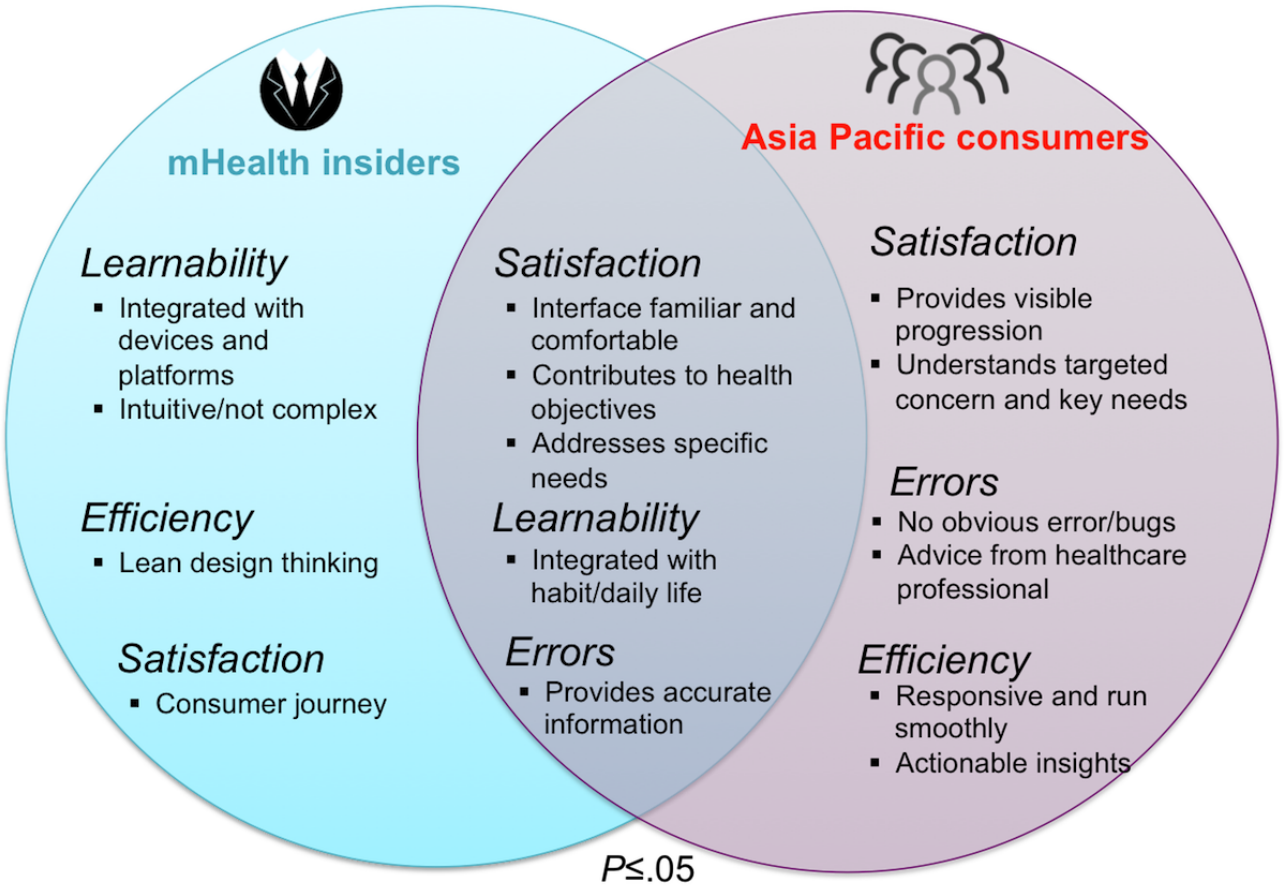


Table 2. Mobile health insiders' top concerns versus consumer's view.

Category	Subcategory	Mobile health insiders' concern level	Consumers (α<0.05)	
			P_O[a]=0.7	P_O=0.6
Efficiency	Lean/design thinking	High	Reject	Accept
Satisfaction	Understand consumer pain point	High	Accept	Accept
Learnability	Intuitive/not complex	High	Reject	Reject
Learnability	Integrated with habits and platforms	High	Accept	Accept
Memorability	Familiarity	High	Accept	Accept
Satisfaction	Have a function	High	Accept	Accept
Learnability	Integrate with devices and platforms	High	Reject	Accept
Errors	Informative and accurate/reliable	High	Accept	Accept
Satisfaction	Understand consumer journey	High	Reject	Accept

[a]Hypothesized probability from the null hypothesis.

Figure 13. Comparison of top concerns from mHealth insiders versus Asia Pacific consumers.



Five of the nine mHealth insider's top themes were also top rated for consumers ($P<.05$, $P_o=0.7$), which also indicated a high level of alignment. *Intuitive functionality* was not statistically significant for consumers, possibly because these respondents already possessed a high level of technology awareness and comfort. Over 60% of the consumer population considered three themes— *lean design thinking*, *integrate with devices and platforms*, and *understand consumer journey* —a high priority ($P<.05$, $P_o=0.6$). However, if we use 70% ($P<.05$, $P_o=0.7$) as the gating threshold, these priorities are not as aligned. This suggests that they are still considered important, but not critical must-haves.

Consumers' Unique Perspective

Asia Pacific consumers were exposed to 22 known usability-related themes summarized from the literature, using a Likert scale rating. Of the 19 usability-related themes mentioned by mHealth insiders, 12 were statistically significant from the consumer's point of view and 7 other themes were paid less attention. Interestingly, one particular theme—*to provide progression analytics* —highlighted by the mHealth consumers was not mentioned in the qualitative study.

The top consumer themes were whether the mHealth app contributes to their health objectives and whether it provides accurate information. This was verified by additional questions in the consumer journey section of the survey, where the top three reasons for them to stop using mHealth apps were: *not as useful as it claims*, *not user-friendly*, and *error or bugs affecting functionality*.

The overlap between the respondent populations is summarized in Figure 13. Commonly emphasized themes were mainly *Satisfaction* related. Some *Learnability*-related themes were ranked as top concern by mHealth insiders but not by the consumers, such as *integration with devices*, which may be a more technical consideration that consumers may already take for granted. In contrast, top concerns from consumers but not from the insiders were *Satisfaction*-related themes, followed by *Errors*- and *Efficiency*-related themes; several of these are indicative of various stages of the consumer journey, such as advice from health care professionals and provision of actionable insights.

Other Important Themes Beyond Usability

Although unrelated to usability, the most-discussed theme from mHealth insiders by the total number of mentions is *the need to have a clear business model and plan for revenue stream*. This is nonetheless critical for the success of any mHealth app and serves as a major motivating factor driving app development in the first place. Thus, to test consumers' willingness to pay for the use of such apps, a similar hypothesis study was done with different P_o, H_o stating that consumers are willing to pay, and H_1 stating that consumers are not willing to pay. The results indicated that fewer than 20% of consumers were willing to pay for both lifestyle/fitness apps and fewer than 35% were willing to pay for medical apps (Table 3).

Table 3. Population proportion test for willingness to pay.

Willingness to pay for a mobile app	P_O[a]=0.4	P_O=0.35	P_O=0.25	P_O=0.20
Lifestyle/fitness	Reject	Reject	Reject	Accept H_o[b]
Medical	Reject	Accept H_o	Accept H_o	Accept H_o

[a]P_O: Hypothesized probability from the null hypothesis.

[b]H_o: Null hypothesis

The finding suggests that although there is a need for revenue stream, it may not be a good choice to charge for the app itself or through in-app advertising. mHealth business owners will need to consider other options, for example, commission from institutions by introducing clients or developing monetization strategies using the consumer data collected.

Finally, other important themes rated important by the mHealth consumers included *adherence to governmental regulations* and *privacy protection*. To illustrate further, with P_o=0.8, the r*espect my privacy* theme was significant, which suggests that privacy is a critical concern for more than 80% of the population. However, this theme was not highly mentioned by mHealth insiders, suggesting that developers may have underestimated the importance and impact of customer data protection and privacy. Although personal data protection legal enforcement is already an integral part of governmental regulations, consumer awareness and education of the existence and contents of such legal protections need to be enhanced. This can be achieved by putting efforts into marketing or enriching the in-app infographics for better illustration.

Discussion

Principal Results and Limitations

In this study, we utilized a mixed-research approach to show that mHealth insiders and consumers are aligned in assigning high priority to a number of usability aspects, particularly with regard to *Satisfaction*, *Learnability*, and *Efficiency* of mHealth apps, although they diverged in terms of the degree of concern for the attributes. The consumer quantitative study also highlighted unique perspectives that were not mentioned by the insiders. Although mHealth is merely the context for the study, we believe that, in general, early involvement and alignment of product development using customized mixed-research and synchronization between categorical and numerical data can result in specific, actionable insights.

The data presented in this paper are representative of a single point-in-time measurement. We speculate that longitudinal research spanning multiple time periods, coupled with real-life mHealth app usage and interaction, would provide deeper insights into individual user perspectives, especially within the fast-evolving digital environment.

The lack of access to a broad insider demographic may have also narrowed the scope of conclusions. Of the 13 mHealth insiders, only two had past experience straddling both health care and IT/app development. Despite a semistructured interview strategy, insiders were found to broadly discuss other topics (eg, sharing personal insights) that were out of the scope of study or provided a subjective view. However, the attributes identified were fairly aligned, and the study outcome was not too negatively impacted.

Separately, although the survey was released on an online platform accessible by the entire Asia Pacific region, we relied largely on personal connections for raising awareness and promotion. As such, respondents from Singapore and China constituted 88% of the total number of submissions, and conclusions made in the study are broadly applicable to these specific markets. Harzing [62] outlines several challenges of primary data collection facing researchers who pose cross-cultural empirical research, including language and translation barriers, cultural barriers, geographical distance, and the liability of foreignness. We anticipated and attempted to address some of these pitfalls, including the use of pilot testing in each country, translations by native speakers, and employing rankings where possible. Nevertheless, the bias of results is a factor that cannot be ignored.

The age range of survey participants was mostly between 31 and 40 years, with a majority of respondents holding a bachelor's degree or higher education. This indicates a fairly young, well-educated population familiar with technology. The potential concerns here are twofold: the responses may not be captured for the specific case of mHealth apps that addressed chronic diseases of old age and the usability concerns of these digital natives may not be coherent with digital immigrants where the learnability bar may be set higher. For example, digital natives are already familiar with a new set of terminology, learning through interactive experience, and concepts of gamification [63], which are widely utilized tools in mobile apps. With a similar argument, only 22% of respondents indicated past usage of medical apps; therefore, survey responses overall may be viewed through experience with lifestyle wellness or fitness mobile apps only.

Comparison With Prior Work

Compared with recent publications, there was general agreement on several aspects. Liu et al (2011) observed that mHealth apps that offered tracking, data visualization, or integration with Internet of Things devices are generally rated higher by users, despite these subcategories being in the minority of the apps on offer [64]. Similar to the findings of Cafazzo et al [65] and Boudreaux [66], consumers from this study were open minded about participating in app testing (35%) and giving feedback (42%) as part of the iterative improvement process.

Studies have also attempted to address potential differences in the level of acceptance correlated with socioeconomic status and health or digital literacy [67-70]. This is one aspect that we

were unable to adequately address, as the reach of the survey did not encompass a broad spectrum of educational and socioeconomic levels or familiarity with technology (reflected in mobile usage, number of apps installed, etc). On the other hand, a qualitative study [14] showed that both digital natives and digital immigrants had comparable awareness and acceptance of mHealth apps and did not differ in their positive impressions of the value of apps.

Surprisingly, the most obvious divergence in our findings with the literature is in the relative importance placed by consumers on the opinions of health care providers during the decision phase of their consumer journey. Conventionally, medical professionals and formal bodies (eg, governmental agencies) have a strong influence on consumer choice [13], along with the impact of strategic initiatives put in place by national and global health organizations [66,71]. However, our study indicated that consumers still rely heavily on word-of-mouth recommendations from familiar sources that are typically not health care experts (65%). This difference indicates that use of mHealth apps is a consequence of subjective decision making rather than a balanced consideration of objective pros and cons.

Application Relevance and Implication of Our Findings

There are two categories of applications in the direction of data maximization (extrapolating insights from current data into other applications) and use of the learning from this mixed-methods approach to address overall business initiatives (ie, supporting sustainable business and market research).

One attractive feature of sequential data collection is that it can also be interpreted separately. Potential new market segments can be identified from the consumer data. For example, there is a slight but noticeable shift in the *comfortability with using mHealth Apps*: Singapore's consumers (12.4%) rated this attribute higher than China's consumers (10.1%). Additionally, Singapore's consumers (17.1%) were far more willing to respond to incentives to download apps than consumers in China (7.1%). For targeting China, which is driving the growth of mobile app usage in Asia, capturing these differentiating characteristics can be essential for success.

The mHealth insiders' dataset can also be dissected to understand the viewpoint differences between experts from different backgrounds. This aspect is particularly significant in this study, as the expertise overlap between health care and software development is small and the impact of various misalignments on the business performance could be further investigated. As discussed, the assumption of a single point of perspective where experts from IT and health care are operating with parallel thoughts needs to be reformed.

The second potential application of this research work is in support of other business initiatives. Business sustainability is one such concern. mHealth apps are currently positioned as a value-added product or companion software tool for connected devices [72]. However, mobile phones today contain processors, sensors, and cameras powerful enough to collect and compute various physiological measurements. This implies that mHealth companies can make use of such alignment studies to develop and reposition mHealth apps as a stand-alone product. This reduces the reliance on electronic connected devices (eg, wristbands and imaging), which is also one of the biggest generators of excessive environmental waste, eliminating the need for the consumers to purchase companion devices and reducing their healthcare financial burden. Furthermore, this lowers the social stress of mastering both the hardware and software apps where there is no obvious value creation.

Another initiative is to maximize the dataset in the direction of generating a predictive modelling tool. Although relatively sparse, the existing dataset can serve as a framework to guide the design of additional studies, or an expansion of the respondent population, to multiply the size of the database. Large datasets will possess sufficient statistical power for creation of various models of consumer behavior and preferences at the discovery stage and increase the probability of creating winning product concepts. Considering the cost of mobile app development (from US $10,000 to >$100,000) [73,74] and the increasing market size [75], a modelling tool can help manage the risk and mitigate the negative costs.

Overall, the research data from this mixed-research study has served to develop ideas, initiate new inquires, and expand the range of inquiry. From concrete to speculative benefits, any future work that can target these applications are strongly encouraged.

Conclusions

The goal of this research was to suggest directions for improvement in the design of mHealth apps for long-term benefits including improved mHealth app download rates, "stickiness," and value within the entire digital mHealth ecosystem. A specific opportunity identified from the literature review is to assist health care companies to better focus on usability aspects rated to be of top importance by mHealth insiders and consumers.

Although the mixed methods used in this study were a more comprehensive and rigorous approach for comparing key interests of mHealth insiders and consumers, there are some challenges in data integration. Hence, there is a need to examine the space between mobile app developers, subject-matter experts, and consumers with regard to the parameters and relative priorities. We have strived to distil usability themes from qualitative interviews into the quantitative survey and vice versa. Alternative interpretations were explored, such as culture; reference standards for importance ranking; and demographic factors such as age, education, and tech-savviness influencing the survey results.

The major findings from this study not only addressed the three outlined research objectives, but also proposed an avenue of structured methodology for researchers of mHealth app development. Taken together, to the best of our knowledge, this is one of the first mixed-research alignment studies that has concretely identified gaps in both the theme and priority of usability concerns between mHealth insiders and consumers.

To conclude, although there is already good alignment of mHealth insiders and consumers in usability, much more could be done to better understand the needs and motivations of consumers with differing concerns. Areas have been identified

where mHealth developers could place greater emphasis in their product-development cycle. Although other factors such as health care regulation and sustainable business models were not considered in this research study, they are important for the digital health care ecosystem to meet the demands in the Asia Pacific region and beyond.

Acknowledgments

This paper is based on the final year project that the authors completed in 2017. We would like to express our sincere thanks to the entire academic team of the Alliance Manchester Business School (AMBS) for the life-changing learning journey. To all the school professors who we met in the workshops and online tutorials, thank you for challenging us to think deeper, broaden our horizons, and use new approaches through different perspectives.

Conflicts of Interest

None declared.

References

1. Heisler Y. BGR. 2016. Mobile internet usage surpasses desktop usage for the first time in history URL: https://bgr.com/2016/11/02/internet-usage-desktop-vs-mobile/ [accessed 2018-08-16] [WebCite Cache ID 71iCYzZ22]
2. Statista. 2016. Number of smartphone users worldwide from 2014 to 2020 (in billions) URL: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ [WebCite Cache ID 71h1joDYA]
3. Research 2 Guidance. The mHealth app market is getting crowded reaching the 259,000 apps URL: https://research2guidance.com/mhealth-app-market-getting-crowded-259000-mhealth-apps-now/ [accessed 2019-01-22] [WebCite Cache ID 71iCj7Rfn]
4. BinDhim NF, Hawkey A, Trevena L. A systematic review of quality assessment methods for smartphone health apps. Telemed J E Health 2015 Feb;21(2):97-104. [doi: 10.1089/tmj.2014.0088] [Medline: 25469795]
5. Huckvale K, Car M, Morrison C, Car J. Apps for asthma self-management: a systematic assessment of content and tools. BMC Med 2012;10:144 [FREE Full text] [doi: 10.1186/1741-7015-10-144] [Medline: 23171675]
6. Chan S, Torous J, Hinton L, Yellowlees P. Towards a Framework for Evaluating Mobile Mental Health Apps. Telemed J E Health 2015 Jul 14. [doi: 10.1089/tmj.2015.0002] [Medline: 26171663]
7. Seffah A, Donyaee M, Kline RB, Padda HK. Usability measurement and metrics: A consolidated model. Software Qual J 2006 Jun;14(2):159-178. [doi: 10.1007/s11219-006-7600-8]
8. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
9. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
10. Nielsen J, Mack R. Usability Inspection Methods. New York, NY: John Wiley & Sons; 1994.
11. Zhang D, Adipat B. Challenges, Methodologies, and Issues in the Usability Testing of Mobile Applications. International Journal of Human-Computer Interaction 2005 Jul;18(3):293-308. [doi: 10.1207/s15327590ijhc1803_3]
12. Fox S, Duggan M. Pew Research Center Internet & Technology. 2012. Mobile Health 2012 URL: http://www.pewinternet.org/2012/11/08/mobile-health-2012/ [accessed 2018-08-15] [WebCite Cache ID 71gyloO6A]
13. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016 Nov 14;16(1):1158 [FREE Full text] [doi: 10.1186/s12889-016-3808-0] [Medline: 27842533]
14. Martinez P. Digital Access to Scholarship at Harvard. 2015. A qualitative study on patient perceptions towards mHealth technology among high risk, chronic disease patients URL: https://dash.harvard.edu/handle/1/17295915 [accessed 2018-08-15] [WebCite Cache ID 71h17SR7M]
15. Ant OA, Wu H, Garrido M, Pronovost PJ, Gurses AP. Usability and perceived usefulness of Personal Health Records for preventive health care: a case study focusing on patients' and primary care providers' perspectives. Appl Ergon 2014 May;45(3):613-628. [doi: 10.1016/j.apergo.2013.09.005] [Medline: 24119975]
16. Anderson K, Burford O, Emmerton L. Mobile Health Apps to Facilitate Self-Care: A Qualitative Study of User Experiences. PLoS One 2016;11(5):e0156164 [FREE Full text] [doi: 10.1371/journal.pone.0156164] [Medline: 27214203]
17. Grindrod KA, Li M, Gates A. Evaluating user perceptions of mobile medication management applications with older adults: a usability study. JMIR Mhealth Uhealth 2014;2(1):e11 [FREE Full text] [doi: 10.2196/mhealth.3048] [Medline: 25099993]
18. Georgsson M, Staggers N. An evaluation of patients' experienced usability of a diabetes mHealth system using a multi-method approach. J Biomed Inform 2016 Feb;59:115-129 [FREE Full text] [doi: 10.1016/j.jbi.2015.11.008] [Medline: 26639894]
19. Mattila E, Pärkkä J, Hermersdorf M, Kaasinen J, Vainio J, Samposalo K, et al. Mobile diary for wellness management--results on usage and usability in two user studies. IEEE Trans Inf Technol Biomed 2008 Jul;12(4):501-512. [doi: 10.1109/TITB.2007.908237] [Medline: 18632330]

20. Schnall R, Rojas M, Bakken S, Brown W, Carballo-Dieguez A, Carry M, et al. A user-centered model for designing consumer mobile health (mHealth) applications (apps). J Biomed Inform 2016 Apr;60:243-251. [doi: 10.1016/j.jbi.2016.02.002] [Medline: 26903153]
21. Årsand E, Frøisland DH, Skrøvseth SO, Chomutare T, Tatara N, Hartvigsen G, et al. Mobile health applications to assist patients with diabetes: lessons learned and design implications. J Diabetes Sci Technol 2012 Sep;6(5):1197-1206 [FREE Full text] [Medline: 23063047]
22. Kalz M, Lenssen N, Felzen M, Rossaint R, Tabuenca B, Specht M, et al. Smartphone apps for cardiopulmonary resuscitation training and real incident support: a mixed-methods evaluation study. J Med Internet Res 2014;16(3):e89 [FREE Full text] [doi: 10.2196/jmir.2951] [Medline: 24647361]
23. Mercer K, Giangregorio L, Schneider E, Chilana P, Li M, Grindrod K. Acceptance of Commercially Available Wearable Activity Trackers Among Adults Aged Over 50 and With Chronic Illness: A Mixed-Methods Evaluation. JMIR Mhealth Uhealth 2016 Jan 27;4(1):e7 [FREE Full text] [doi: 10.2196/mhealth.4225] [Medline: 26818775]
24. Creswell J. Qualitative Inquiry and Research Design : Choosing among Five Traditions. Thousand Oaks, CA: Sage Publications, Inc; 1997.
25. Cameron R. A sequential mixed model research design: Design, analytical and display issues. International Journal of Multiple Research Approaches 2014 Dec 17;3(2):140-152. [doi: 10.5172/mra.3.2.140]
26. Lin YL, Guerguerian A, Tomasi J, Laussen P, Trbovich P. "Usability of data integration and visualization software for multidisciplinary pediatric intensive care: a human factors approach to assessing technology". BMC Med Inform Decis Mak 2017 Aug 14;17(1):122 [FREE Full text] [doi: 10.1186/s12911-017-0520-7] [Medline: 28806954]
27. Rosario J, Ascher MT, Cunningham DJ. A study in usability: redesigning a health sciences library's mobile site. Med Ref Serv Q 2012;31(1):1-13. [doi: 10.1080/02763869.2012.641481] [Medline: 22289091]
28. Chevalier A, Maury A, Fouquereau N. The influence of the search complexity and the familiarity with the website on the subjective appraisal of aesthetics, mental effort and usability. Behaviour & Information Technology 2013 Aug 25;33(2):117-132. [doi: 10.1080/0144929X.2013.819936]
29. Zeitz K, Marchany R, Tront J. Speed isn't Enough: Usability and Adoption of an Optimized Alert Notification System. IEEE Technol Soc Mag 2016 Mar;35(1):47-55. [doi: 10.1109/MTS.2015.2494398]
30. Bosse C, Kelly K. Improving EHR Usability Using LEAN Methodology. Stud Health Technol Inform 2016;225:870. [Medline: 27332384]
31. François M, Crave P, Osiurak F, Fort A, Navarro J. Digital, analogue, or redundant speedometers for truck driving: Impact on visual distraction, efficiency and usability. Appl Ergon 2017 Nov;65:12-22. [doi: 10.1016/j.apergo.2017.05.013] [Medline: 28802430]
32. Rose E, Racadio R, Wong K, Nguyen S, Kim J, Zahler A. Community-Based User Experience: Evaluating the Usability of Health Insurance Information with Immigrant Patients. IEEE Trans Profess Commun 2017 Jun;60(2):214-231. [doi: 10.1109/TPC.2017.2656698]
33. Julstrom S, Kozma-Spytek L, Isabelle S. Telecoil-mode hearing aid compatibility performance requirements for wireless and cordless handsets: magnetic signal-to-noise. J Am Acad Audiol 2011 Sep;22(8):528-541. [doi: 10.3766/jaaa.22.8.5] [Medline: 22031677]
34. Green D, Pearson J. Integrating website usability with the electronic commerce acceptance model. Behaviour & Information Technology 2011 Mar;30(2):181-199. [doi: 10.1080/01449291003793785]
35. Lee C, Coughlin J. PERSPECTIVE: Older Adults' Adoption of Technology: An Integrated Approach to Identifying Determinants and Barriers. J Prod Innov Manag 2014 Jun 03;32(5):747-759. [doi: 10.1111/jpim.12176]
36. Mishuris RG, Yoder J, Wilson D, Mann D. Integrating data from an online diabetes prevention program into an electronic health record and clinical workflow, a design phase usability study. BMC Med Inform Decis Mak 2016 Dec 11;16:88 [FREE Full text] [doi: 10.1186/s12911-016-0328-x] [Medline: 27401606]
37. Putrino D, Zanders H, Hamilton T, Rykman A, Lee P, Edwards D. Patient Engagement Is Related to Impairment Reduction During Digital Game-Based Therapy in Stroke. Games Health J 2017 Oct;6(5):295-302 [FREE Full text] [doi: 10.1089/g4h.2016.0108] [Medline: 28910162]
38. Kawamoto K, Honey A, Rubin K. The HL7-OMG Healthcare Services Specification Project: motivation, methodology, and deliverables for enabling a semantically interoperable service-oriented architecture for healthcare. J Am Med Inform Assoc 2009;16(6):874-881 [FREE Full text] [doi: 10.1197/jamia.M3123] [Medline: 19717796]
39. Bell J, Reed S, Kim K, Ngo V, Wait W, Apesoa-Varano E, et al. Usability and acceptance of novel personal health technology to support early palliative care for patients with cancer and caregivers. JCO 2016 Oct 09;34(26_suppl):73-73. [doi: 10.1200/jco.2016.34.26_suppl.73]
40. Stjernswärd S, Hansson L. User value and usability of a web-based mindfulness intervention for families living with mental health problems. Health Soc Care Community 2017 Dec;25(2):700-709. [doi: 10.1111/hsc.12360] [Medline: 27189237]
41. Jimison H, Gorman P, Woods S, Nygren P, Walker M, Norris S, et al. Barriers and drivers of health information technology use for the elderly, chronically ill, and underserved. Evid Rep Technol Assess (Full Rep) 2008 Nov(175):1-1422. [Medline: 19408968]

42. Kaipio J, Lääveri T, Hyppönen H, Vainiomäki S, Reponen J, Kushniruk A, et al. Usability problems do not heal by themselves: National survey on physicians' experiences with EHRs in Finland. Int J Med Inform 2017 Dec;97:266-281 [FREE Full text] [doi: 10.1016/j.ijmedinf.2016.10.010] [Medline: 27919385]
43. Ehrler F, Haller G, Sarrey E, Walesa M, Wipfli R, Lovis C. Assessing the Usability of Six Data Entry Mobile Interfaces for Caregivers: A Randomized Trial. JMIR Hum Factors 2015 Dec 15;2(2):e15 [FREE Full text] [doi: 10.2196/humanfactors.4093] [Medline: 27025648]
44. Andreasen Struijk LNS, Bentsen B, Gaihede M, Lontis E. Error-Free Text Typing Performance of an Inductive Intra-Oral Tongue Computer Interface for Severely Disabled Individuals. IEEE Trans Neural Syst Rehabil Eng 2017 Dec;25(11):2094-2104. [doi: 10.1109/TNSRE.2017.2706524] [Medline: 28541213]
45. Armbrüster C, Sutter C, Ziefle M. Notebook input devices put to the age test: the usability of trackpoint and touchpad for middle-aged adults. Ergonomics 2007 Mar;50(3):426-445. [doi: 10.1080/00140130601127885] [Medline: 17536778]
46. Wolpin SE, Halpenny B, Whitman G, McReynolds J, Stewart M, Lober WB, et al. Development and usability testing of a web-based cancer symptom and quality-of-life support intervention. Health Informatics J 2015 Mar;21(1):10-23. [doi: 10.1177/1460458213495744] [Medline: 24406906]
47. Miah S, Gammack J, Hasan N. Extending the framework for mobile health information systems Research: A content analysis. Information Systems 2017 Sep;69:1-24. [doi: 10.1016/j.is.2017.04.001]
48. SoJump (WJX). URL: https://www.wjx.cn/ [accessed 2019-01-18] [WebCite Cache ID 75WNU9ZGj]
49. Worldometers. Asia population (live) URL: http://www.worldometers.info/world-population/asia-population/ [accessed 2018-08-16] [WebCite Cache ID 71iCrYSZX]
50. Osborne JW. Best Practices In Data Cleaning: A Complete Guide To Everything You Need To Do Before And After Collecting Your Data. Thousand Oaks, CA: Sage Publications, Inc; 2019.
51. Bengtsson M. How to plan and perform a qualitative study using content analysis. NursingPlus Open 2016;2:8-14. [doi: 10.1016/j.npls.2016.01.001]
52. Healthcare Mobile Application Survey. URL: https://www.wjx.cn/jq/16823139.aspx [accessed 2019-01-18] [WebCite Cache ID 75WSsvJZ2]
53. Cyrus NW. Statistics For Engineers And Scientists. New York, NY: Mcgraw-hill Education; 2019.
54. Massey A, Miller S. Tests of hypotheses using statistics. 2006. URL: https://web.williams.edu/Mathematics/sjmiller/public_html/BrownClasses/162/Handouts/StatsTests04.pdf [WebCite Cache ID 71h1IVh9I]
55. Rossi RJ. Applied Biostatistics For The Health Sciences. Hoboken, NJ: Wiley; 2019.
56. Morse JM. Designing funded qualitative research. In: Denzin NK, Lincoln YS, editors. Handbook of qualitative research. Thousand Oaks, CA: Sage Publications; 1994:220-235.
57. Krippendorff KK. Content Analysis: An Introduction To Its Methodology. Thousand Oaks, CA: Sage Publications, Inc; 2019.
58. Guest G, Bunce A, Johnson L. How Many Interviews Are Enough? Field Methods 2016 Jul 21;18(1):59-82. [doi: 10.1177/1525822X05279903]
59. MacDonald D. Dustin K MacDonald. Z-Test hypothesis testing URL: http://dustinkmacdonald.com/z-test-hypothesis-testing/ [accessed 2018-08-15] [WebCite Cache ID 71h0u0zuH]
60. Teo J. Are You Missing Out on a Faster Internet Connection?. 2015 URL: https://www.singsaver.com.sg/blog/are-you-missing-out-on-a-faster-internet-connection [accessed 2018-08-15] [WebCite Cache ID 71h1ySCeL]
61. Fetters MD, Curry LA, Creswell JW. Achieving integration in mixed methods designs-principles and practices. Health Serv Res 2013 Dec;48(6 Pt 2):2134-2156 [FREE Full text] [doi: 10.1111/1475-6773.12117] [Medline: 24279835]
62. Harzing AW, Reiche BS, Pudelko M. Challenges in international survey research: a review with illustrations and suggested solutions for best practice. EJIM 2013;7(1):112. [doi: 10.1504/EJIM.2013.052090]
63. Prensky M. Digital Natives, Digital Immigrants Part 1. On the Horizon 2001 Sep;9(5):1-6. [doi: 10.1108/10748120110424816]
64. Liu C, Zhu Q, Holroyd KA, Seng EK. Status and trends of mobile-health applications for iOS devices: A developer's perspective. Journal of Systems and Software 2011 Nov;84(11):2022-2033. [doi: 10.1016/j.jss.2011.06.049]
65. 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]
66. Boudreaux ED, Waring ME, Hayes RB, Sadasivam RS, Mullen S, Pagoto S. Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. Transl Behav Med 2014 Dec;4(4):363-371 [FREE Full text] [doi: 10.1007/s13142-014-0293-9] [Medline: 25584085]
67. Caburnay CA, Graff K, Harris JK, McQueen A, Smith M, Fairchild M, et al. Evaluating diabetes mobile applications for health literate designs and functionality, 2014. Prev Chronic Dis 2015;12:E61 [FREE Full text] [doi: 10.5888/pcd12.140433] [Medline: 25950568]
68. Hawn C. Take two aspirin and tweet me in the morning: how Twitter, Facebook, and other social media are reshaping health care. Health Aff (Millwood) 2009;28(2):361-368 [FREE Full text] [doi: 10.1377/hlthaff.28.2.361] [Medline: 19275991]
69. Hofstede Insights. 2017. Country comparison URL: https://www.hofstede-insights.com/country-comparison/china,singapore/ [accessed 2018-08-16] [WebCite Cache ID 71iCMATp4]

70. Tate EB, Spruijt-Metz D, O'Reilly G, Jordan-Marsh M, Gotsis M, Pentz MA, et al. mHealth approaches to child obesity prevention: successes, unique challenges, and next directions. Transl Behav Med 2013 Dec;3(4):406-415 [FREE Full text] [doi: 10.1007/s13142-013-0222-3] [Medline: 24294329]
71. Ricciardi L, Mostashari F, Murphy J, Daniel JG, Siminerio EP. A national action plan to support consumer engagement via e-health. Health Aff (Millwood) 2013 Feb;32(2):376-384. [doi: 10.1377/hlthaff.2012.1216] [Medline: 23381531]
72. Silva BMC, Rodrigues JJPC, de la Torre Diez I, 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]
73. Fu J. CodementorX. 2017. How much does it cost to make an app in 2017? URL: https://www.codementor.io/blog/how-much-does-it-cost-to-make-an-app-in-2017-1nqj6ehste [accessed 2018-08-15] [WebCite Cache ID 71gz6EyHM]
74. Thunderquote. 2016. How much does it cost to develop a mobile app in Singapore? A price guide URL: https://blog.thunderquote.com/2016/08/23/how-much-does-it-cost-to-develop-a-mobile-app-in-singapore/ [WebCite Cache ID 71h2EW2pm]
75. Warman P. Newzoo. 2017. Global Mobile Market Report: App Market to Gross $44.8Bn This Year URL: https://newzoo.com/insights/articles/global-mobile-market-report-app-market-to-gross-44-8bn-this-year/ [accessed 2018-08-15] [WebCite Cache ID 71h2OMGpt]

Abbreviations

IT: information technology
mHealth: mobile health

Edited by G Eysenbach; submitted 10.09.18; peer-reviewed by C Romanzini, AM Navarra; comments to author 22.11.18; revised version received 11.12.18; accepted 12.12.18; published 30.01.19.

Please cite as:
Liew MS, Zhang J, See J, Ong YL
Usability Challenges for Health and Wellness Mobile Apps: Mixed-Methods Study Among mHealth Experts and Consumers
JMIR Mhealth Uhealth 2019;7(1):e12160
URL: http://mhealth.jmir.org/2019/1/e12160/
doi:10.2196/12160
PMID:30698528

Original Paper

Factors for Supporting Primary Care Physician Engagement With Patient Apps for Type 2 Diabetes Self-Management That Link to Primary Care: Interview Study

Julie Ayre[1], BLibSt (Hons); Carissa Bonner[1,2], BCom, BSc (Hons), MPH, PhD; Sian Bramwell[3], RN; Sharon McClelland[3], RN; Rajini Jayaballa[3,4], MBChB, FRACP; Glen Maberly[3,5], BSc (Med), MBBS, FRACP; Kirsten McCaffery[1], BSc (Hon), PhD

[1]Sydney Health Literacy Lab, Sydney School of Public Health, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

[2]Ask, Share, Know: Rapid Evidence for General Practice Decisions Centre for Research Excellence, The University of Sydney, Sydney, Australia

[3]Western Sydney Diabetes, Western Sydney Local Health District, Blacktown, Australia

[4]School of Medicine, Western Sydney University, Blacktown, Australia

[5]Sydney School of Public Health, Faculty of Medicine and Health, The University of Sydney, Sydney, Australia

Corresponding Author:
Kirsten McCaffery, BSc (Hon), PhD
Sydney Health Literacy Lab
Sydney School of Public Health, Faculty of Medicine and Health
The University of Sydney
Edward Ford Building (A27)
The University of Sydney
Sydney,
Australia
Phone: 61 293517220
Email: kirsten.mccaffery@sydney.edu.au

Abstract

Background: The health burden of type 2 diabetes can be mitigated by engaging patients in two key aspects of diabetes care: self-management and regular contact with health professionals. There is a clear benefit to integrating these aspects of care into a single clinical tool, and as mobile phone ownership increases, apps become a more feasible platform. However, the effectiveness of online health interventions is contingent on uptake by health care providers, which is typically low. There has been little research that focuses specifically on barriers and facilitators to health care provider uptake for interventions that link self-management apps to the user's primary care physician (PCP).

Objective: This study aimed to explore PCP perspectives on proposed features for a self-management app for patients with diabetes that would link to primary care services.

Methods: Researchers conducted 25 semistructured interviews. The interviewer discussed potential features that would link in with the patient's primary care services. Interviews were audio-recorded, transcribed, and coded. Framework analysis and the Consolidated Criteria for Reporting Qualitative Research checklist were employed to ensure rigor.

Results: Our analysis indicated that PCP attitudes toward proposed features for an app were underpinned by perceived roles of (1) diabetes self-management, (2) face-to-face care, and (3) the anticipated burden of new technologies on their practice. Theme 1 explored PCP perceptions about how an app could foster patient independence for self-management behaviors but could also increase responsibility and liability for the PCP. Theme 2 identified beliefs underpinning a commonly expressed preference for face-to-face care. PCPs perceived information was more motivating, better understood, and presented with greater empathy when delivered face to face rather than online. Theme 3 described how most PCPs anticipated an initial increase in workload while they learned to use a new clinical tool. Some PCPs accepted this burden on the basis that the change was inevitable as health care became more integrated. Others reported potential benefits were outweighed by effort to implement an app. This study also identified how app features can be positively framed, highlighting potential benefits for PCPs to maximize PCP engagement, buy-in, and uptake. For example, PCPs were more positive when they perceived that an app could facilitate communication and motivation between consultations, focus on building capacity for patient independence, and reinforce rather than replace in-person

care. They were also more positive about app features that were automated, integrated with existing software, flexible for different patients, and included secondary benefits such as improved documentation.

Conclusions: This study provided insight into PCP perspectives on a diabetes app integrated with primary care services. This was observed as more than a technological change; PCPs were concerned about changes in workload, their role in self-management, and the nature of consultations. Our research highlighted potential facilitators and barriers to engaging PCPs in the implementation process.

(JMIR Mhealth Uhealth 2019;7(1):e11885) doi:10.2196/11885

KEYWORDS

diabetes mellitus, type 2; electronic health records; telemedicine; mobile apps; general practitioners; physicians, primary care; self-management; qualitative research; translational medical research

Introduction

For people with diabetes, self-management (including medication adherence, physical activity, healthy diet, and weight management) is a key aspect of care that can mitigate long-term complications of diabetes [1-5]. Because diabetes is a progressive condition, regular interactions with health professionals are important for medical feedback on self-management (such as glycated hemoglobin levels), education, adaptation of the care plan (including adjustment of medications as the condition progresses), and monitoring and treatment of long-term complications [6,7]. As such, self-management and care provided by health professionals are interrelated, and this should ideally also be reflected in clinical diabetes interventions, for example, by fostering ongoing communication between the health care provider and patient to facilitate their respective roles in diabetes care.

As mobile phone ownership increases [8,9], a self-management app that can also be used during consultations could achieve this goal. Mobile phone apps are already available to help people with diabetes engage in self-management [10], and many collect patient data that is highly relevant to the health care provider, including self-monitoring data for blood glucose, physical activity, and diet [11]. Health care providers value self-management apps because they perceive that they encourage patient engagement, provide them with a deeper and more reliable understanding of their patients' behaviors, and improve communication during consultations by providing visualizations of patient data [12-15].

However, despite the potential benefits of online health technologies (including apps), implementation on a large scale remains a key challenge. Research suggests that their effectiveness is often limited by poor uptake and sustained use by health care providers [16,17]. Recent systematic reviews have suggested that key barriers for health care providers are increased workload and disruption to existing clinical processes and staff roles as well as concerns about remuneration, data security, and liability [16,18-20].

Some barriers may specifically relate to self-management apps. This could include the overwhelming complexity of the data that is available to health care providers, provider responsibility to respond to shared self-monitoring data, and health care provider perceptions of poor motivation on the part of patients [12,14]. Overcoming the challenge of poor provider uptake is crucial as strong provider endorsement is in turn a key factor for patient uptake of online tools [21].

Our study aimed to build on these findings by investigating primary care physician (PCP) perspectives on proposed features for a self-management app for people with type 2 diabetes that is linked to their PCP's care plan. This will provide a more specific understanding of how PCPs conceptualize their role in providing care to their patients who have type 2 diabetes and how this role could be better supported by an app.

Methods

Participants

PCPs were recruited from a pool of 50 clinics in Sydney, Australia, that had elected to engage in joint specialist case conferencing, an initiative implemented through the Western Sydney Primary Health Network in an area with culturally and linguistically diverse patient populations. During case conferencing, the PCP discusses diabetes management with the patient in conjunction with an endocrinologist and a credentialed diabetes educator. PCPs were invited to participate in the interview with a view to informing the design of an app developed by a group of collaborating local health authorities called Western Sydney Diabetes. Purposive sampling ensured a diverse range of gender, years of experience, and cultural backgrounds to reflect the broader PCP population in Western Sydney (see Table 1).

Procedure

After providing written consent to participate, JA conducted semistructured interviews for approximately 25 minutes in each PCP's consultation room. Interviews were conducted between November 2017 and June 2018. Questions were based on an interview schedule that was piloted with PCPs prior to this study (Multimedia Appendix 1). Questions pertained to how the PCP currently helps patients to self-manage diabetes and their attitude toward diabetes apps. Participants were also asked for feedback on several specific app features:

- Transfer a patient's individualized care plan into the app
- Export self-monitoring data to PCP software
- Prompt patient to see their PCP (for example, if there is a pattern of high blood glucose readings)
- Send reminders to book cycle of care appointments (for example, PCP check-ups and eye and foot checks)

- Contain bundles of educational material including videos that can be sent to the patient
- Produce a summary report of blood glucose self-monitoring to be used by the PCP during the consultation (see Multimedia Appendix 1)

Interviews were audio-recorded and transcribed verbatim. Ethical approval was obtained from the University of Sydney Human Research Ethics Committees (project number 2017/224) and Western Sydney Local Health District (reference number 5092 AU RED LNR/17/WMEAD/140).

Analysis

Interviews were analyzed using framework analysis, a matrix-based approach to thematic analysis [22], which involved 5 steps: familiarization with the data, indexing, collating similar codes into themes, charting data into a thematic framework, and synthesis and interpretation. Rigor was addressed through indexing a subset of data across 2 researchers, a continuous process of comparing concepts and themes to data, and discussion of potential themes across authors. The project team concluded that theoretical saturation was reached after 25 interviews, where variation in PCP perspectives could be adequately explained through 3 overarching themes (Figure 1).

Table 1. Participant descriptive characteristics.

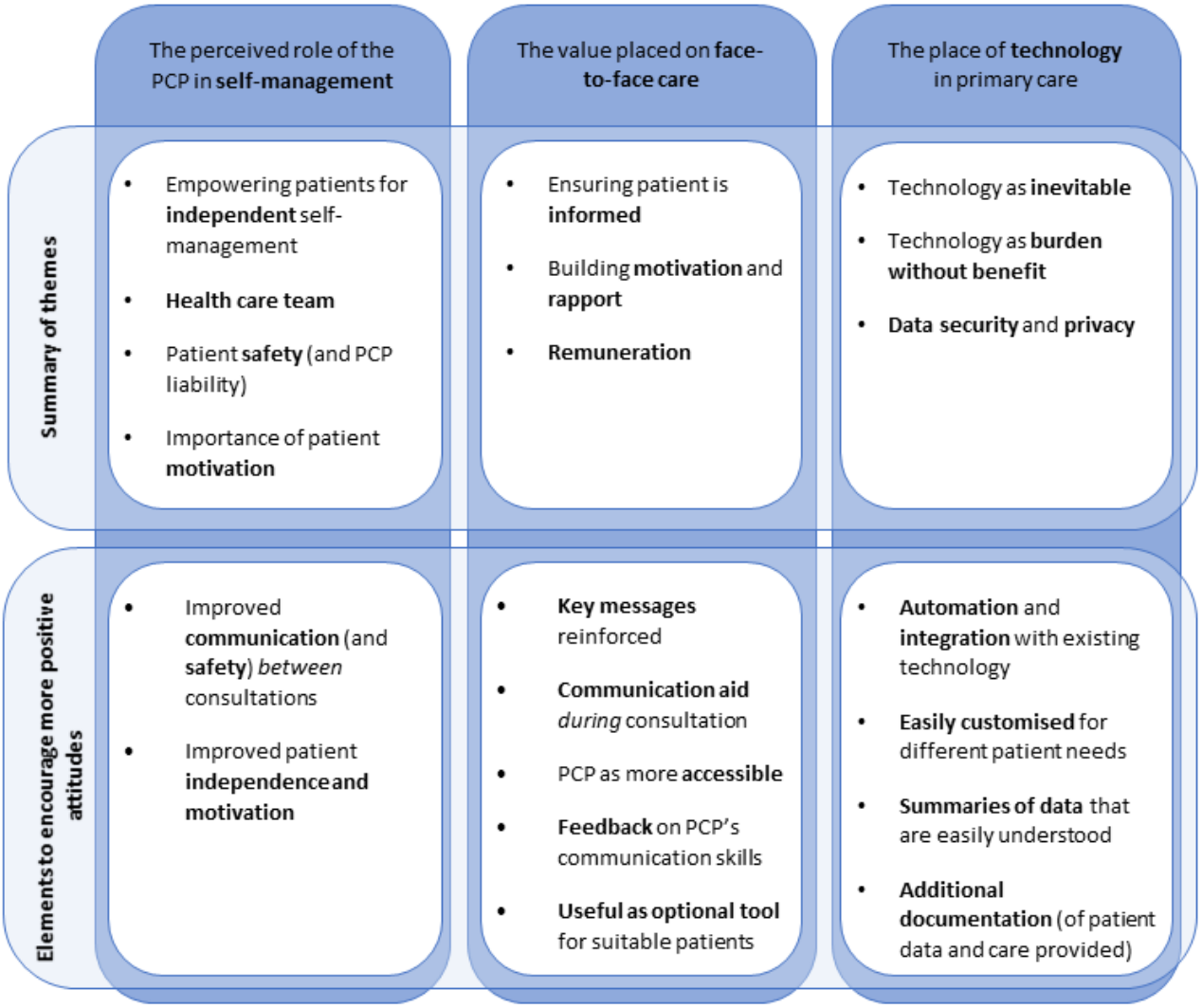
Characteristics	Total, n (%)
Gender (n=25)	
Female	14 (56)
Male	11 (44)
Years qualified as a PCP[a] (n=24)	
<10	5 (20)
10-19	8 (32)
≥20	12 (48)
Country of birth (n=24)	
Australia	6 (25)
Sri Lanka	5 (21)
India	4 (17)
Bangladesh	2 (8)
Philippines	2 (8)
Other[b]	5 (21)
Languages spoken (n=24)	
English only	4 (17)
Tamil	5 (21)
Sinhalese	3 (13)
Chinese	3 (13)
Hindi	3 (13)
Filipino	2 (8)
Other[c]	8 (33)

[a]PCP: primary care physician.

[b]Other includes countries of birth listed by 1 PCP: South Africa, Afghanistan, Malaysia, Fiji, United Kingdom.

[c]Other includes languages spoken by 1 PCP: Afrikaans, Bangla, Bengali, Dari, Kannada, Malay, Swahili.

Figure 1. Mind map of themes. PCP: primary care physician.



Results

Overview

Of the PCPs from 12 clinics who were interviewed, 83% (20/24) spoke a language other than English and 80% (20/25) had been qualified as a PCP for at least 10 years (Table 1).

Most PCPs were open to using a mobile phone app for diabetes self-management in their clinic. Each theme below constitutes a set of beliefs that contributed to these attitudes toward proposed features for a self-management app (Figure 1):

- Perceived role of the PCP in self-management of type 2 diabetes
- Value placed on face-to-face care
- Place of technology in primary care

PCPs often shared similar beliefs and values across these 3 themes regardless of their cultural backgrounds. However, there was substantial variation in perceptions of whether an app would support them in their work. Therefore, each theme also highlights the positive mindsets of PCPs who perceived an app would support the care they currently provide.

Theme 1. Perceived Role of the Primary Care Provider in Self-Management of Type 2 Diabetes

Primary Care Physician Perspective

PCPs emphasized that ongoing self-management was a key aspect of care for their patients with diabetes. However, many believed that this was outside their control and that ultimately the patient must take responsibility for self-management.

> *Well I think it's very important to think medical people want to do what they can, but it's really what people do 24 hours of the day, 7 days a week.* [PCP07, male, speaks a language other than English (LOTE), practicing ≥20 years]

Instead, PCPs perceived that their role was to increase patient capacity for safe and independent self-management. They built this capacity by providing medical advice, general self-management education, and specific feedback (for example, on patterns of blood glucose levels). Several PCPs saw more detailed self-management education as the role of practice nurses, dieticians, or diabetes educators.

App features that expanded PCP responsibilities related to patient self-management raised concerns about liability. For

example, many PCPs did not endorse an app feature that would notify them in real time about patients' blood glucose readings. PCPs perceived that this feature might risk patient safety, particularly if it gave patients the impression that their doctor was actively monitoring their readings. As a result, PCPs anticipated that this would create moral or legal obligations to respond in a timely manner.

> *...probably I don't want to receive the data on a regular basis. I think there is a chance that the clinician might miss it. Or if they've gone on a holiday or if they haven't checked the data or anything. So there's a risk that they could be missed.* [PCP13, male, speaks a LOTE, practicing 10 to 19 years]

Crucially, some PCPs also perceived that a key limitation of an app was that it did not overcome the initial challenge of persuading patients to take responsibility for their self-management. They argued that an app would only be useful for patients who were already independent and therefore would not target those who most needed to engage in self-management care.

> *My problem with these apps...we are already selecting a group of people who are going to be motivated enough to put this data into the app...whereas the majority of patients, or the people we have trouble with...who are not bothered to exercise and things, I'm not sure this is going to...convince them to do it...if they can't walk for half an hour will they take the trouble of putting all this in the mobile app? That's where my skepticism is.* [PCP17, female, speaks a LOTE, practicing ≥20 years]

Facilitating Features

Despite these concerns, several PCPs perceived that some app features could improve patient safety and independence. For example, PCPs were more positive about an app when they saw it as an opportunity to improve patient safety between consultations. This was particularly important for patients they only saw intermittently.

> *I think we can deal with the problem before it becomes out of hand and then we can fix the problem on the patient's point of view or modify the medications if needed to, and so it's a win, win situation for both.* [PCP19, male, speaks a LOTE, practicing ≥20 years]

PCPs also preferred features that placed the onus on the patient to take action. This included notifications to patients about patterns in blood glucose readings (where notifications were generated automatically by algorithms) and automated reminders for check-ups.

Last, PCPs who were more positive about using an app perceived that most of their patients engaged in self-management, at least to some degree. As such, they anticipated that an app could benefit patients by increasing motivation between consultations and capitalizing on transient moments of greater motivation.

> *...So when they've got a little bit of motivation in one perspective, or like for a particular problem, you want to try and foster that as soon as it comes on board.* [PCP04, male, speaks a LOTE, practicing <10 years]

Theme 2. Value Placed on Face-to-Face Care

Primary Care Physician Perspective

All PCPs greatly preferred face-to-face care to online care. Many perceived that face-to-face care enabled them to make sure patients actually took in the information from educational materials. They perceived that they could tailor the information for the patient more easily in person (for example, by presenting information in small amounts, emphasizing the most important points, and checking understanding). PCPs also believed it was a more effective platform to ensure that the patient was at the very least exposed to appropriate information, whereas links could easily be ignored or forgotten.

> *Because you can send patients the link online...you can give them a lot of information. Don't know how much of it they've actually looked at. That's probably the biggest limitation, is not knowing exactly what they've looked at overall.* [PCP15, male, other demographic data missing]

Some PCPs also expressed concern that an app would undermine the patient-physician relationship. They argued that face-to-face meetings were important for developing rapport with the patient. One PCP explained how he personalized care with his patients by discussing their motivation to engage in their health.

> *...so I think it's important to ask the patient, what matters to you? Well, I don't want to end up like my Mum...Or I don't want to go blind...Or I don't want my kidneys to fail. So...ok, so how is it day-to-day what we can take steps to manage that? So you've got to find out what motivates the patient.* [PCP02, male, speaks only English, practicing <10 years]

A small number of PCPs (n=4) also identified that there was a risk that they may not be paid for time spent delivering care online because there was no existing process.

> *...we give good care but we also like to be acknowledged for that care and remunerated, appropriately.* [PCP01, female, speaks a LOTE, practicing ≥20 years]

Facilitating Features

Conversely, PCPs were more positive when they perceived that an app would be a welcome adjunct to face-to-face care that could address existing challenges. For example, PCPs were interested in brief educational materials that would reinforce the key messages discussed during the consultation. This was particularly important when patients found it difficult to take in information during the consultation because they were distressed or overwhelmed.

> *...I realize you can tell people something and it just goes straight over their head. Or you've given them some bad news and then you tell them something else, it just hasn't registered...I guess in one sense, anything that helps information being given to the*

> *patient is good...* [PCP03, male, speaks only English, practicing ≥20 years]

Others identified that an app could help overcome difficulties conveying information during the consultation itself. PCPs discussed how resources that use pictures or that are available in the patient's first language and simple graphs of blood glucose levels could help to convey messages during the consultation and open up discussions about barriers to lifestyle change and the complications of diabetes.

> *So I think things that are visual are good 'cause patients, they don't like all these numbers... So things like this [pie chart] is helpful. So if I can say, all right so the green bit is you in the normal range, but the blue bit is when your readings are too high...we need to try and get your green bits to be a bigger portion...* [PCP18, female, speaks a LOTE, practicing <10 years]

In contrast to the perception that an app would weaken rapport, some PCPs argued that an app that is linked to the patient's health care provider could foster a stronger connection with the patient. They discussed how it would create a sense of accessibility to the PCP and encourage patients with patterns of high or low blood glucose readings to see their doctor more frequently.

> *...I've got a lot of patients who I have a lot of trouble... getting in with their numbers. And I think those patients would benefit from something like this...So, yeah, if it's something they're engaging in and they're getting those numbers and they see it and it prompts them to come in and discuss it with me, I think that'd be a good thing.* [PCP16, female, speaks a LOTE, practicing <10 years]

A few PCPs also saw an app as a tool to improve their communication skills. These PCPs anticipated that some patients may respond to notifications about high or low blood glucose levels by inundating the clinic with inquiries. They perceived that this could be a reflection of poor communication with the patient, particularly in terms of setting clear expectations, providing information and checking understanding.

> *...if they have more questions it's probably because they've not been given the correct information in the first place. So it's almost an aid to [PCP] doing their job properly....is how I view it.* [PCP14, male, speaks a LOTE, practicing 10 to 19 years]

Last, PCPs were more positive when they perceived an app as an optional tool which patients could elect to use. Many PCPs perceived an app could be very useful for patients who were younger or more familiar with mobile phones but would be of limited use with others, primarily older patients who did not regularly use apps or who had significant vision problems (a common complication of diabetes).

Theme 3. Place of Technology in Primary Care

Primary Care Physician Perspective

Most PCPs anticipated that an app would increase the burden of clinical care. They believed that there would be increased workload initially while they learned to use an app, as well as ongoing time required to provide care remotely. Interestingly, despite this shared awareness about the burden of a new clinical technology, PCPs varied substantially in their attitudes. Several PCPs accepted the burden, believing that clinical practice was inevitably shifting towards digital health and mobile phone technologies.

> *I mean that's a bridge we're actually going to have to cross anyway. That's the reality. So, I mean, the way I'd phrase it is, that's a burden we're going to have to take on board, given how technology's going.* [PCP04, male, speaks a LOTE, practicing <10 years]

Those with negative attitudes perceived that potential benefits would be outweighed by the effort needed to implement an app. Some also perceived that an app would add unnecessary complexity to their work. For many of these PCPs the current appointment reminder systems and blood glucose logbooks were sufficient.

> *...they're not that computer savvy, especially the older generations. In which case it's more confusing for them and for me, because then I won't know under what circumstances there was a high BSL. Maybe it was another reason...like it's almost more confusing sometimes.* [PCP14, male, speaks a LOTE, practicing 10 to 19 years]

> *...every time there's a little bit of fatigue when ...[local health district] or, when [health body] comes up with yet another initiative, it's like well I've always been doing it this way, I've trained on this way and now my life is about to get even more complex.* [PCP02, male, speaks only English, practicing <10 years]

Last, only two PCPs raised issues about data security and privacy. They perceived that they would be held personally responsible if they endorsed and used an app that did not adhere to government privacy policies.

Facilitating Features

PCPs were more positive when they perceived that burden would be mitigated. For example, PCPs valued app functions that were largely automated and that were integrated with existing technology and practices.

> *...[my existing online interface that integrates consultation bookings and patient reminders with practice software] is really good because it actually talks to my software, even though it's a Web app, it'll actually talk to my software, then a recall reminder has been sent. So then it doesn't mean I have to go to something, type it in, go to something else, type it in.* [PCP02, male, speaks only English, practicing <10 years]

It is worth noting that PCPs encouraged complexity to ensure that variability in patient characteristics and goals was accommodated. For example, PCPs reported that care plans were different for patients who were newly diagnosed, those who were transitioning to insulin, and those who switched to a new medication. As such, many PCPs preferred options that allowed them to customize the frequency of reminders and the

presentation of patient data. One PCP even highlighted that the colors used for high and low blood glucose readings should reflect whether care was focused on preventing hypoglycemia or preventing long-term complications.

> *...But I think that's more practitioner-dependent and how you practice your medicine... If the overall aim of the app is to improve diabetic control then the highs go red. If the overall aim of the app is to prevent complications from diabetes, especially hypoglycemia, then the hypoglycemia end up red.* [PCP04, male, speaks a LOTE, practicing <10 years]

PCPs also viewed apps more favorably when they perceived additional benefits that an app could provide. For example, PCPs valued features that would analyze patient data and summarize the relevant information so that it could be understood quickly. One PCP discussed that if summary data were available before the consultation, she could make better use of her time with the patient.

> *I also think if we only get [patient data] in the consult there's time constraint. We only get 15 to 20 minutes and we won't necessarily have time to go through all the results and discuss it with the patient and come up with a management plan. So I think if...we can go through [summary data] before we see the patient and we can also have a plan formulated because then we can just discuss it with them and manage it.* [PCP22, female, speaks a LOTE, practicing 10 to 19 years]

Another perceived secondary benefit of an app was improved documentation for patient data and records of care. PCPs perceived that improving patient data would increase patient accountability. They perceived that improving the documentation of care would form a stronger basis for remuneration and improve care that was shared with other health professionals.

> *...So if it's going to be arranged in such a way that the patient uses this app and sends messages to the doctor and the doctor can use that particular opportunity to say that "look this is the care that I've given" and use that as an outcome-based visit without seeing the patient face-to-face.* [PCP01, female, speaks a LOTE, practicing ≥20 years]

Discussion

Principal Findings

This study explored PCP perspectives on proposed features for a diabetes self-management app that would be linked to their practice software. Our analysis indicated that these attitudes were underpinned by perceived roles of PCPs in diabetes self-management, the role of face-to-face care, and the anticipated burden of new technologies in their practice. This study also identified how app features can be positively framed, highlighting potential benefits for PCPs in order to maximize PCP engagement, buy-in, and uptake.

The barriers and facilitators identified in this study can be incorporated into an implementation framework. For example, normalization process theory [23] suggests that 4 factors lead to successful uptake and sustained use of new technologies. This study identified several strategies related to coherence (how consumers understand a new technology within the context of existing systems), cognitive participation (how consumers engage with and commit to using a new technology), and collective action (perceived impact of the new technology on workflow, workload, roles, responsibility and training) (Table 2). The fourth factor, reflexive monitoring, relates to technologies that have already been implemented.

Comparison With Prior Work

The findings in this study identified three key barriers that are specific to uptake of diabetes self-management apps that link to the health care provider. First, there was a clear tension between avoiding an increase in workload and the need for app functions and settings that can be customized to the diverse clinical goals of patients with diabetes. For example, PCPs perceived it was important to have different schedules of prompts when introducing new medication compared to regular blood glucose self-monitoring. A balance between these aspects is needed as workload is a key barrier to provider uptake [16].

Second, PCPs challenged the idea of real-time notifications of patient data. PCPs understood the theoretical value of real-time notifications but perceived that this feature would fail in real clinical settings and could actually put patient safety at risk. This has been identified in previous research on digital self-monitoring [14]. Furthermore, many PCPs voiced that this would not support them in their goal to build patient capacity for independent self-management. Alternatives such as automated prompts directing the patient to see their PCP may be a more realistic option.

Third, some PCPs argued that even if an app is effective, the benefits would be severely limited if it were only suitable for patients who were already motivated, had sufficient familiarity with mobile phone apps, and had adequate vision. However, it should also be noted that mobile phone ownership is high in Australia, with the greatest increases in ownership seen in older age groups [8]. Other research has also reported that health care providers tend to underestimate patient motivation for lifestyle management [24]. Regardless, it will be important for future work to establish how to identify patients who will benefit most from an app and whether other interventions might be more suitable for less motivated patients.

The other barriers identified in this study were less specific to diabetes self-management and are consistently reported in research on implementation of online health interventions. For example, increased workload and changes to scope of practice are common factors for poor uptake of new health technologies [12,13,16-20]. PCPs also suggested that face-to-face care was important for developing rapport with the patient and ensuring that the patient had understood important information. Similar concerns were highlighted in a review of telehealth interventions for patients with heart failure [18]. In addition, face-to-face care is often valued by health professionals because it is perceived as the main method of remuneration [12,24]. However, in our study only four of the PCPs raised this issue.

Table 2. Summary of themes and how suggested strategies relate to normalization process theory.

Theme	Description	Examples of suggested strategies by normalization process theory component		
		Coherence	Cognitive participation	Collective action
Theme 1. Perceived role of the PCP[a] in self-management of type 2 diabetes	• PCP goal is to facilitate independent self-management for patients with diabetes • Care is shared across practice staff • Patients aren't motivated to self-manage	• Explain where the goals of the app overlap or are likely to differ from PCP goals to support patients • Explain any medicolegal risks, particularly in terms of remote monitoring of blood glucose • Explain how staff can continue existing roles through the app • Explain how to identify patients who have enough baseline motivation/independence to use the app	• Explain how an app can address existing challenges • Explain how the app can bolster patient motivation between consults	• Explain that the Intervention will be available to various staff including nurses
Theme 2. Value placed on face-to-face care	• Face-to-face care is valuable • PCPs are remunerated primarily through face-to-face care • Patients don't use mobile phones	• Explain how the app is an optional additional tool; it does not replace face-to-face care. Provide guidance on how to best identify patients who are suited to the app • Be explicit about whether/how work conducted through the app will be remunerated	• Explain how app can improve efficiency of analysis of self-monitoring data • Explain how app can facilitate communication during consultation and promote the take home message • Explain how app can prompt patient to visit doctor	• Not applicable
Theme 3. Place of technology in primary care	• This is just another thing we have to learn to use (with little added benefit) • It will take a lot of time to learn to use the app • Patients are not one-size-fits-all • Data security and privacy	• Be explicit about whether/how work conducted through the app will be remunerated • Be explicit about the implications of data security and privacy issues for the PCP	• Explain how app can improve documentation of care • Must also be flexible enough to accommodate different patient goals and care plans	• Minimize workflow disruption and avoid unnecessary increase in workload through automation and integration with existing technology

[a]PCP: primary care physician.

Limitations

There are several limitations to this study. First, PCPs were drawn from clinics that were already voluntarily engaged with a regional public health body that aimed to improve the efficiency and effectiveness of primary care. These clinics were therefore likely to be more receptive to public health initiatives (including the app, which would be delivered through a collaborative body that includes the local health district). As such, the PCPs in this study may be more positive than other PCPs in the community. This is particularly important regarding sensitive issues such as remuneration and may explain why this was raised by so few PCPs.

The app was intended to first roll out in Western Sydney, a suburban region with a highly culturally diverse population and areas of socioeconomic disadvantage. As such, recruitment focused on PCPs in that area and results are more likely to reflect the perspectives of PCPs working in that kind of setting. In addition, although the research was carried out by an associate of the team that would eventually develop the app, efforts were made to ensure that PCPs understood the interviewer's independence and that they did not feel pressure to provide positive responses about apps.

Second, PCPs discussed their attitudes toward hypothetical app features rather than an actual app. As such, these findings reflect more abstract preconceptions and assumptions about apps. This is useful for anticipating potential barriers and engaging PCPs. However, this approach may have also overemphasized PCP openness to changes in workload as it is more likely to reflect aspirational goals of care.

Conclusions

Diabetes self-management apps that are linked to the patient's PCP have the potential to be highly effective. However, in reality these interventions are often limited by poor health care provider uptake. This study investigated PCP perspectives on

a diabetes app that was integrated with primary care services. PCPs perceived this as more than a technological change; they were concerned about changes in workload, their role in self-management, and the nature of consultations. This research highlighted potential facilitators and barriers to engaging PCPs in the implementation process.

Acknowledgments

This research is supported by an Australian Government Research Training Program Scholarship.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview schedule, example summary report presented to primary care physicians during the interview, and Consolidated Criteria for Reporting Qualitative Research checklist.

[PDF File (Adobe PDF File), 125KB - mhealth_v7i1e11885_app1.pdf]

References

1. Boulé NG, Haddad E, Kenny GP, Wells GA, Sigal RJ. Effects of exercise on glycemic control and body mass in type 2 diabetes mellitus: a meta-analysis of controlled clinical trials. JAMA 2001 Sep 12;286(10):1218-1227. [Medline: 11559268]
2. Norris SL, Lau J, Smith SJ, Schmid CH, Engelgau MM. Self-management education for adults with type 2 diabetes: a meta-analysis of the effect on glycemic control. Diabetes Care 2002 Jul;25(7):1159-1171. [Medline: 12087014]
3. Anderson JW, Kendall CW, Jenkins DJ. Importance of weight management in type 2 diabetes: review with meta-analysis of clinical studies. J Am Coll Nutr 2003 Oct;22(5):331-339. [Medline: 14559925]
4. Rhee MK, Slocum W, Ziemer DC, Culler SD, Cook CB, El-Kebbi IM, et al. Patient adherence improves glycemic control. Diabetes Educ 2005;31(2):240-250. [doi: 10.1177/0145721705274927] [Medline: 15797853]
5. Inzucchi SE, Bergenstal RM, Buse JB, Diamant M, Ferrannini E, Nauck M, et al. Management of hyperglycaemia in type 2 diabetes: a patient-centered approach. Position statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetologia 2012 Jun;55(6):1577-1596. [doi: 10.1007/s00125-012-2534-0] [Medline: 22526604]
6. Royal Australian College of General Practitioners. General practice management of type 2 diabetes: 2016-2018 URL: https://www.racgp.org.au/FSDEDEV/media/documents/Clinical%20Resources/Guidelines/Diabetes/General-practice-management-of-type-2-diabetes_1.pdf [accessed 2018-11-12] [WebCite Cache ID 73rvDERpd]
7. National Institute for Health and Care Excellence. 2015 Dec 02. NICE guideline (NG28)—type 2 diabetes in adults: management URL: https://www.nice.org.uk/guidance/ng28 [accessed 2018-11-12] [WebCite Cache ID 73rvVOEXt]
8. Deloitte Touche Tohmatsu. 2017. Mobile Consumer Survey 2017: the Australian cut URL: https://www2.deloitte.com/au/mobile-consumer-survey [accessed 2018-11-12] [WebCite Cache ID 71D7tLcR6]
9. Deloitte Development LLC. 2017. 2017 Global Mobile Consumer Survey: US edition URL: https://www2.deloitte.com/content/dam/Deloitte/us/Documents/technology-media-telecommunications/us-tmt-2017-global-mobile-consumer-survey-executive-summary.pdf [accessed 2018-02-20] [WebCite Cache ID 71D89kiBP]
10. Garabedian LF, Ross-Degnan D, Wharam JF. Mobile phone and smartphone technologies for diabetes care and self-management. Curr Diab Rep 2015 Dec;15(12):109. [doi: 10.1007/s11892-015-0680-8] [Medline: 26458380]
11. Hood M, Wilson R, Corsica J, Bradley L, Chirinos D, Vivo A. What do we know about mobile applications for diabetes self-management? A review of reviews. J Behav Med 2016 Dec;39(6):981-994. [doi: 10.1007/s10865-016-9765-3] [Medline: 27412774]
12. 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 [FREE Full text] [doi: 10.2196/humanfactors.7709] [Medline: 28974481]
13. 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]
14. Gabriels K, Moerenhout T. Exploring entertainment medicine and professionalization of self-care: interview study among doctors on the potential effects of digital self-tracking. J Med Internet Res 2018 Jan 12;20(1):e10 [FREE Full text] [doi: 10.2196/jmir.8040] [Medline: 29330140]
15. Mishuris RG, Yoder J, Wilson D, Mann D. Integrating data from an online diabetes prevention program into an electronic health record and clinical workflow: a design phase usability study. BMC Med Inform Decis Mak 2016 Dec 11;16:88 [FREE Full text] [doi: 10.1186/s12911-016-0328-x] [Medline: 27401606]

16. Granja C, Janssen W, Johansen MA. Factors determining the success and failure of ehealth interventions: systematic review of the literature. J Med Internet Res 2018 May 01;20(5):e10235 [FREE Full text] [doi: 10.2196/10235] [Medline: 29716883]
17. Greenhalgh T, Wherton J, Papoutsi C, Lynch J, Hughes G, A'Court C, et al. Beyond adoption: a new framework for theorizing and evaluating nonadoption, abandonment, and challenges to the scale-up, spread, and sustainability of health and care technologies. J Med Internet Res 2017 Nov 01;19(11):e367 [FREE Full text] [doi: 10.2196/jmir.8775] [Medline: 29092808]
18. Greenhalgh T, A'Court C, Shaw S. Understanding heart failure; explaining telehealth: a hermeneutic systematic review. BMC Cardiovasc Disord 2017 Dec 14;17(1):156 [FREE Full text] [doi: 10.1186/s12872-017-0594-2] [Medline: 28615004]
19. Irizarry T, DeVito DA, Curran CR. Patient portals and patient engagement: a state of the science review. J Med Internet Res 2015;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]
20. Cresswell K, Sheikh A. Organizational issues in the implementation and adoption of health information technology innovations: an interpretative review. Int J Med Inform 2013 May;82(5):e73-e86. [doi: 10.1016/j.ijmedinf.2012.10.007] [Medline: 23146626]
21. Irizarry T, Shoemake J, Nilsen ML, Czaja S, Beach S, DeVito DA. Patient portals as a tool for health care engagement: a mixed-method study of older adults with varying levels of health literacy and prior patient portal use. J Med Internet Res 2017 Mar 30;19(3):e99 [FREE Full text] [doi: 10.2196/jmir.7099] [Medline: 28360022]
22. Ritchie J, Spencer L, O'Connor W. Carrying out qualitative analysis. In: Ritchie J, Lewis J, editors. Qualitative Research Practice: A Guide for Social Science Students and Researchers. Thousand Oaks: Sage Publishers; 2003:219-262.
23. May C, Finch T. Implementing, embedding, and integrating practices: an outline of normalization process theory. Sociology 2009 Jun 15;43(3):535-554. [doi: 10.1177/0038038509103208]
24. Brandt CJ, Søgaard GI, Clemensen J, Sndergaard J, Nielsen JB. General practitioners' perspective on ehealth and lifestyle change: qualitative interview study. JMIR Mhealth Uhealth 2018 Apr 17;6(4):e88 [FREE Full text] [doi: 10.2196/mhealth.8988] [Medline: 29666045]

Abbreviations

LOTE: language other than English
PCP: primary care physician

Edited by G Eysenbach; submitted 12.08.18; peer-reviewed by J Nielsen, E Sezgin; comments to author 21.09.18; revised version received 03.10.18; accepted 10.10.18; published 16.01.19.

Please cite as:
Ayre J, Bonner C, Bramwell S, McClelland S, Jayaballa R, Maberly G, McCaffery K
Factors for Supporting Primary Care Physician Engagement With Patient Apps for Type 2 Diabetes Self-Management That Link to Primary Care: Interview Study
JMIR Mhealth Uhealth 2019;7(1):e11885
URL: http://mhealth.jmir.org/2019/1/e11885/
doi:10.2196/11885
PMID:30664468

Original Paper

Variability in Doctors' Usage Paths of Mobile Electronic Health Records Across Specialties: Comprehensive Analysis of Log Data

Ji Yeong Soh[1*], RN; Sang-Hyuk Jung[1*], BE; Won Chul Cha[1,2,3*], MD; Mira Kang[1,3,4], MD, PhD; Dong Kyung Chang[1,3,5], MD, PhD; Jaegon Jung[6], PhD; JeanHyoung Lee[3], ME; Jong Soo Choi[1,3], PhD; Kyunga Kim[1,7], PhD

[1]Department of Digital Health, Samsung Advanced Institute for Health Sciences & Technology, Sungkyunkwan University, Seoul, Republic of Korea
[2]Department of Emergency Medicine, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea
[3]Health Information and Strategy Center, Samsung Medical Center, Seoul, Republic of Korea
[4]Center for Health Promotion, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea
[5]Department of Gastroenterology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea
[6]Department of Computer Engineering, Seoul Digital University, Seoul, Republic of Korea
[7]Statistics and Data Center, Research Institute for Future Medicine, Samsung Medical Center, Seoul, Republic of Korea
[*]these authors contributed equally

Corresponding Author:
Kyunga Kim, PhD
Department of Digital Health
Samsung Advanced Institute for Health Sciences & Technology
Sungkyunkwan University
81, Irwon-ro
Gangnam-gu
Seoul, 06351
Republic of Korea
Phone: 82 10 8910 5015
Email: kyunga.j.kim@samsung.com

Abstract

Background: With the emergence of mobile devices, mobile electronic health record (mEHR) systems have been utilized by health care professionals (HCPs), including doctors, nurses, and other practitioners, to improve efficiency at the point of care. Although several studies on mEHR systems were conducted, including those analyzing their effects and HCPs' usage frequency, only a few considered the specific workflows of doctors based on their specialties in which the work process differs greatly.

Objective: This study aimed to investigate the differences in mEHR usage paths across clinical specialties.

Methods: We collected the log data of 974 doctors who worked from August 2016 to August 2017 and used the mEHR system at the Samsung Medical Center, one of the biggest hospitals in South Korea. The doctors were classified into 3 groups based on their specialty: the physician, the surgeon, and other hospital-based physician (OHBP) groups. We used various descriptive and visualization methods to understand and compare doctors' usage paths of mEHRs. First, the average numbers of log-ins per day and features used per log-in were examined over different specialties and positions. Second, the number of features used by each doctor was visualized via a heat map to provide an overview of mEHR usage across feature types and doctors' specialties. Third, we conducted a path analysis via a Sankey diagram to describe main usage paths and association rule mining to find frequent paths in mEHR usage.

Results: The physician group logged on most frequently, whereas the OHBP group logged on least frequently. In fact, the number of log-ins per day of residents in the physician group was 4.4 times higher than that of staff members in the other groups. The heat map visualization showed a visible difference among specialty groups. The physician group used more consultation-related features, whereas the surgeon group used more surgery-related features. Generally, 50% of the doctors spent about 15 seconds at a time when using mEHRs. In the Sankey diagram, the physician group showed diverse usage patterns with higher complexity compared with the other 2 groups; in particular, their paths contained more loops, which reflected repetitive checks on multiple patients. The most frequent path included inpatient summary, which means that most users stopped at the point of summary and did not proceed to view more details.

Conclusions: The usage paths of mEHRs showed considerable differences among the specialty groups. Such differences can be accommodated into an mEHR design to enhance the efficiency of care.

(JMIR Mhealth Uhealth 2019;7(1):e12041) doi:10.2196/12041

KEYWORDS

mobile apps; electronic health record; mobile health; mobile electronic health record

Introduction

Background

The advent of mobile phones has accelerated the expansion of mobile health (mHealth) market because they are equipped with various apps and functions such as wireless connectivity and messaging capabilities. Mobile phone–based mHealth apps have emerged as strong tools for patients and health care professionals (HCPs) in the digital health care era [1,2]. The development of mHealth apps for the private market has been promoted as a result of increasing global access to mobile technology [3].

Many institutions seek mobile electronic health record (mEHR) systems that can improve efficiency at the point of care [4]. The mEHRs provide HCPs with ubiquitous access to patient data in real time, and hence, enable them to communicate with others when facilitating a patient's care [5-10]. A well-made mEHR can improve workflow efficiency, thereby lowering costs and reducing the work burden of HCPs [11,12]. Although the overall satisfaction rate of mEHRs has increased [11,12], their benefits and satisfaction may differ among doctors (ie, medical doctors), according to the latter's specialties [13].

Recently, efforts have been made to analyze log data from mEHRs for the evaluation of providers' workflow [14,15]. Even among doctors, the usage paths are likely to differ according to their specialties, resulting in different work processes [13]. However, research on specialty-based paths with an in-depth analysis has not yet been conducted. This is insufficient to reflect the characteristics of its practical use.

Objectives

In this study, we analyzed real mEHR log data of doctors and investigated specialty-based mEHR usage paths. The difference in the usage paths can be reflected in mEHRs to improve their efficiency and usability.

Methods

Mobile Electronic Health Record System

This study was conducted at the Samsung Medical Center (SMC), one of the largest tertiary referral hospitals in South Korea with more than 2000 beds and approximately 1000 doctors. In 2017, the average daily visit was about 8000 and 220 for the outpatient and emergency departments, respectively. The next-generation medical information system, including a new version of an electronic health record (EHR) system known as the Data Analysis & Research Window for Integrated kNowledge (DARWIN), was launched in July 2016. At the same time, the previous mEHR system was majorly revised and launched with a new name, mDARWIN version 2.3.7-2.4.8 (Figure 1). mDARWIN is based on Android 2.3 Gingerbread (Google Inc, California, United States) and has Wi-Fi and 3G capabilities (Figure 2). It comprises a main menu, list-level features, and patient-level features (Figure 3). The app is mainly for the use of doctors. After log-in, on using a certified user's identification number and password, users can choose from the main menu to view a list-level feature or select a function. From each list-level feature, users can choose patient-level features for more activities or leave and move to other list-level features. Each session closes when either a user logs out or it automatically logs out after no activity for a certain amount of time. mDARWIN also supports fingerprint log-in and near-field communication.

Study Subjects and Data Collection

Target subjects were doctors who had logged on to the mEHR system from August 2016 to August 2017. Visiting doctors were excluded because of short usage duration. Doctors who used the system at least once a month were still included in the analysis. To examine the association between usage and specialty, the subjects were categorized into 3 groups based on their specialties. The physician group consisted of internal medicine, family medicine, pediatrics, and critical care. The surgeon group included general surgery, neurosurgery, and otorhinolaryngology. The other hospital-based physician (OHBP) group covered anesthesiology, pathology, and radiology. The subjects were further classified by job position (staff members, clinical fellows, and residents). The log data for all subjects were collected from the mDARWIN server. For each subject, sessions were identified as a series of features used from log-in to logout. The sessions lasting longer than 60 min were filtered out. This study was approved by the institutional review board of the study site (SMC 2017-12-074).

Data Analysis

Overall usage for individual features in the mEHR system was investigated by summarizing usage frequencies of features from the log data. The frequencies were normalized within each specialty group and presented against specialty departments in a heat map visualization [16]. For each specialty group, usage paths were identified in 3 steps. First, all pairs of adjacent features in every session were recognized. Second, we computed the amount of the first-order transition for each feature pair.

Finally, usage paths were constructed as sets of feature pairs with large first-order transition amounts for each specialty group and then visualized using Sankey diagrams [17]. For better visualization, flows with small frequencies were omitted from the diagrams. In addition, we performed association rule mining (ARM) to identify the top 5 usage paths per specialty according to support values [18].

Figure 1. A screenshot of the mDARWIN screen displayed after login. The main menu functions as a portal for specific contents. mDARWIN: Data Analysis & Research Window for Integrated kNowledge.

Figure 2. System architecture of mDARWIN. It was designed to accommodate 2 different network connectivity choices. AP: access point; DB: database; EAI: enterprise application integration; EHR: electronic health record; LTE: long-term evolution; MCI: multi-channel integration; OCS: order communication system; SQL: structured query language.

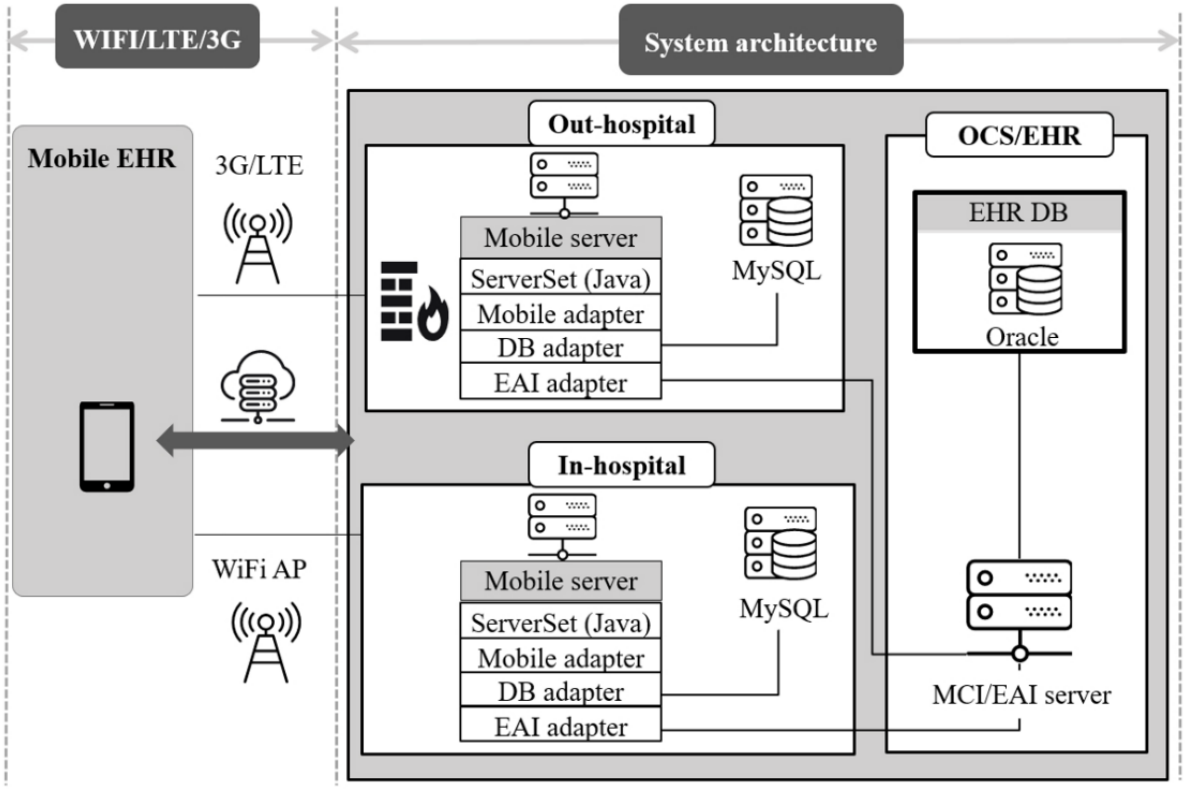
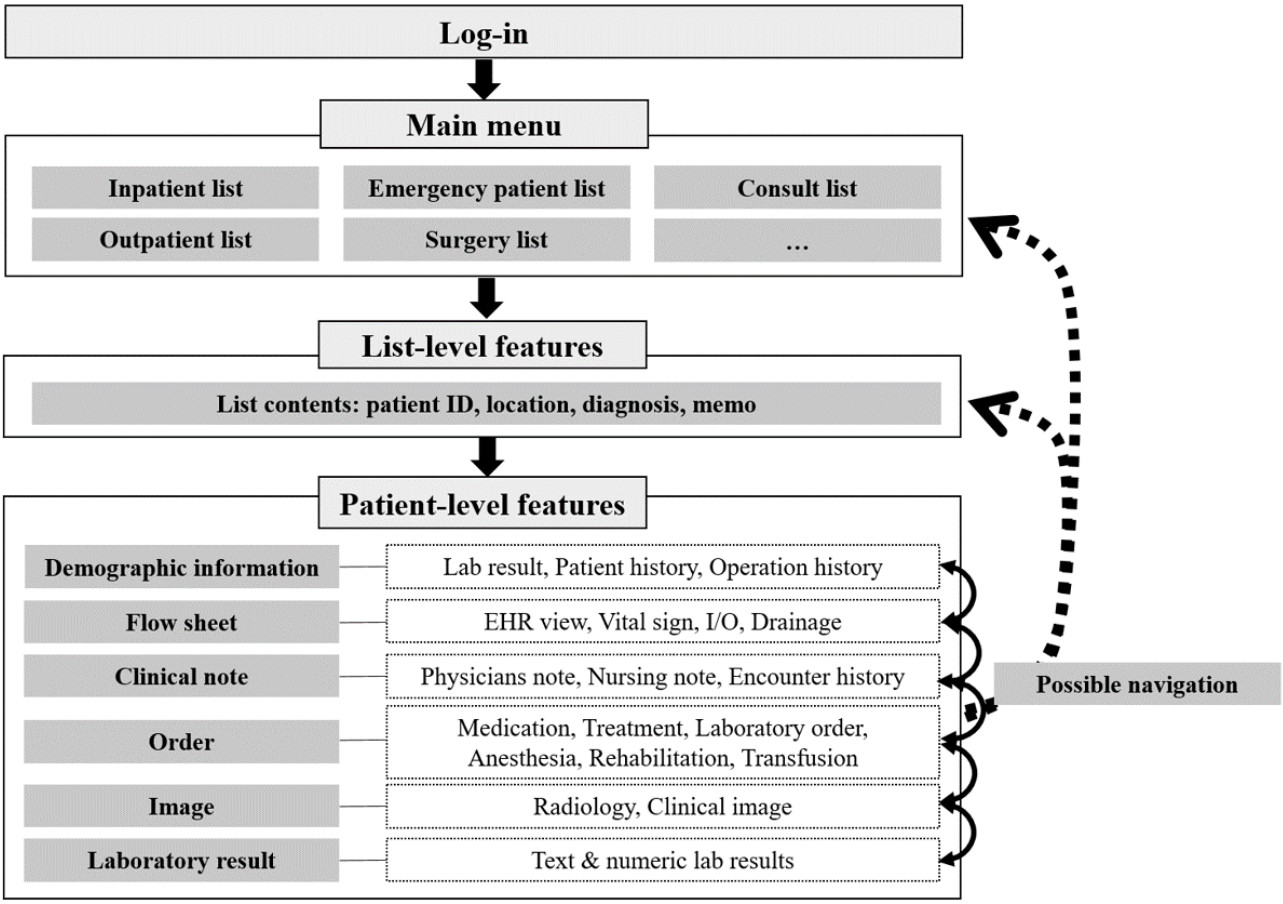


Figure 3. Feature list of mDARWIN. The system consists of 3 levels: main menu, list-, and patient-level features. EHR: electronic health record; ID: identification; I/O: input/output.



All analyses were performed using R software version 3.4.3 (R Foundation for Statistical Computing, Vienna, Austria) [19].

Results

Principal Results

During the study period, 974 unique doctors used mEHRs and generated 2,777,311 event logs, with the following distribution: the physician group, 46.9% (457/974); the surgeon group, 39.4% (384/974); and the OHBP group, 13.7% (133/974). The average number of daily log-ins per user and content per log-in by specialty group and position are shown in Multimedia Appendix 1. There were 24.8% (242/974) staff members, 23.7% (231/974) clinical fellows, and 51.4% (501/974) residents. The average number of daily log-ins per user was 1.4, with an SD of 1.5. After each log-in, users visited 5.5 features (SD 2.8) on average. Doctors in the physician group, especially residents, showed the most frequent log-in activities (2.2 and 3 times more frequent than the surgeon and OHBP groups, respectively). Doctors in the OHBP group, especially staff members, tended to visit diverse features per log-in, compared with the other groups (see Figure 4). Different usage of features was observed among specialties in the heat map (Figure 5). Frequently used features were indicated as hotspots in the heat map and differed across users' specialties. There were hotspots in the consultation-related features for the physician group and in the surgery-related features for the surgeon group. The OHBP group used features evenly, among most features, whereas the other groups used a specific set of features intensively. No distinguished difference was observed in the use of emergency- and outpatient-related features for all 3 groups. Compared with other features, the usage of outpatient features was less frequent in all 3 groups.

Path Analysis

The identified usage paths were specialty-specific, in that they varied across specialties (see Sankey diagrams in Figures 6-8). Compared with the other 2 groups showing heavy flows to surgical features, the physician group showed diverse flows and paths. For instance, they showed repetitive transition patterns among the same features (often called loops), whereas the surgeon and the OHBP groups did not form loops and had more simple paths. The repetitive patterns seemed to reflect physicians' work processes containing repetitive checks on multiple patients.

Among the top 5 paths identified via ARM for each group, most paths included an *inpatient summary* feature, with a high support value of more than 40% (Table 1) [20]. However, the 2-feature path from *inpatient list* to *inpatient summary* was most frequently taken than multiple-feature paths. This finding implied that most users tended to stop at the point of summary and did not proceed to view more details. Regarding frequently used paths, consultation- and emergency-related paths were recognized in the physician group, whereas the operation-related path was identified in the surgeon and the OHBP groups. For all 3 groups, outpatient features were not ranked in the top 5 paths.

Figure 4. The average numbers of daily log-ins per user and features per log-in according to users' specialty and position. The numbers in parentheses stand for standard deviation. OHBP: other hospital-based physician.

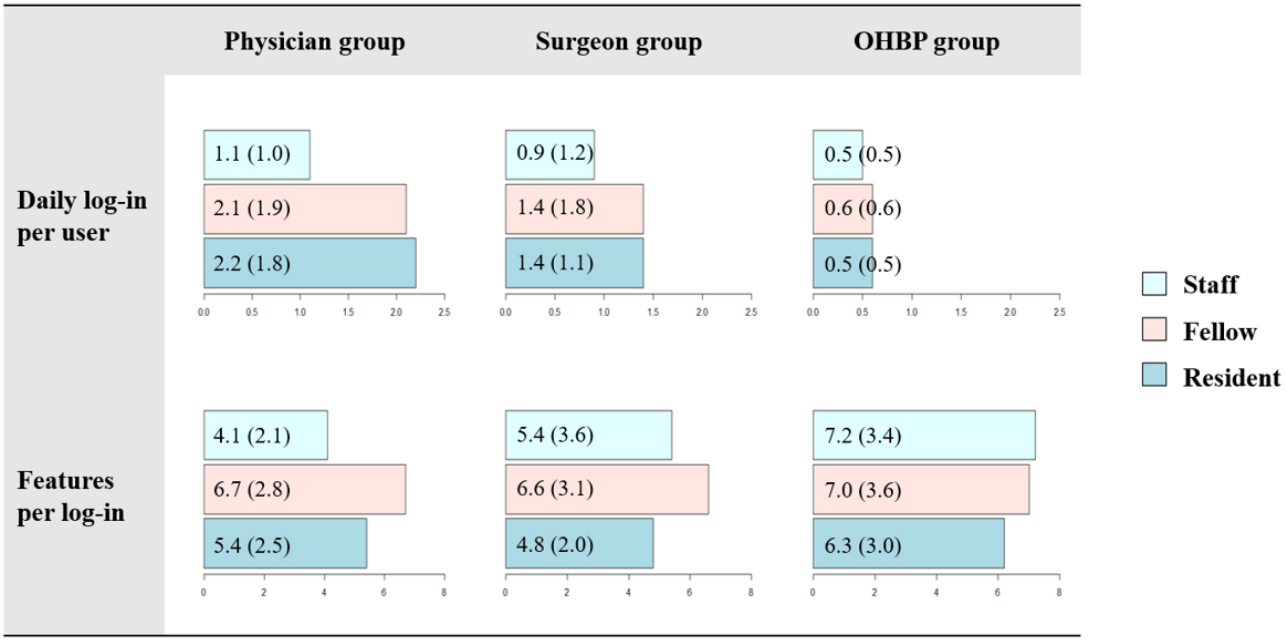


Figure 5. Heat map visualization of feature usage patterns according to users' specialties. Rows and columns stand for specialty departments and individual features, respectively. OHBP: other hospital-based physician.

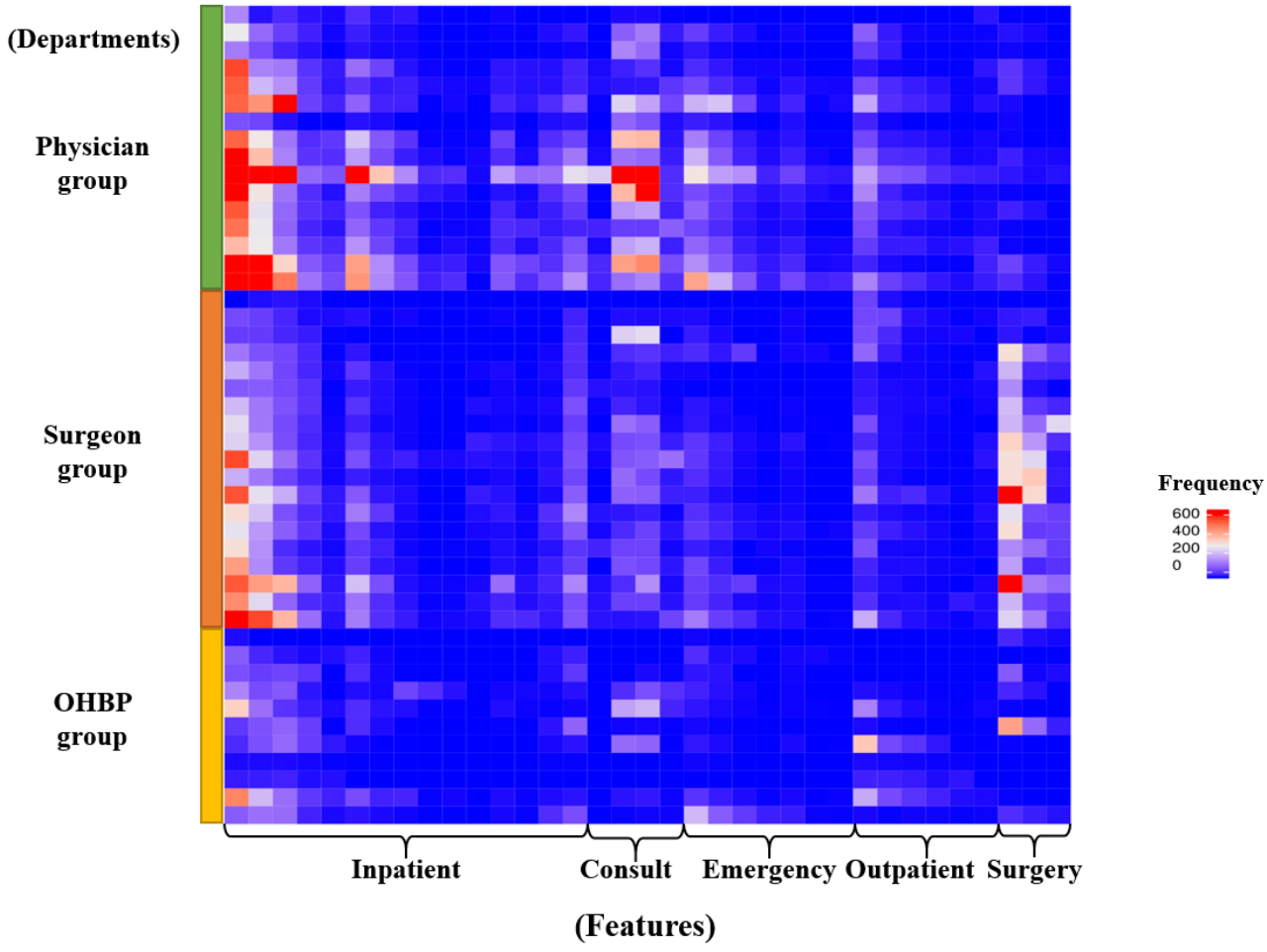


Figure 6. Sankey diagram of usage paths identified for the Physician group. Light blue, light green, light orange, orange and red colors were used to indicate impatient, consult, surgery, outpatient and emergency features, respectively.

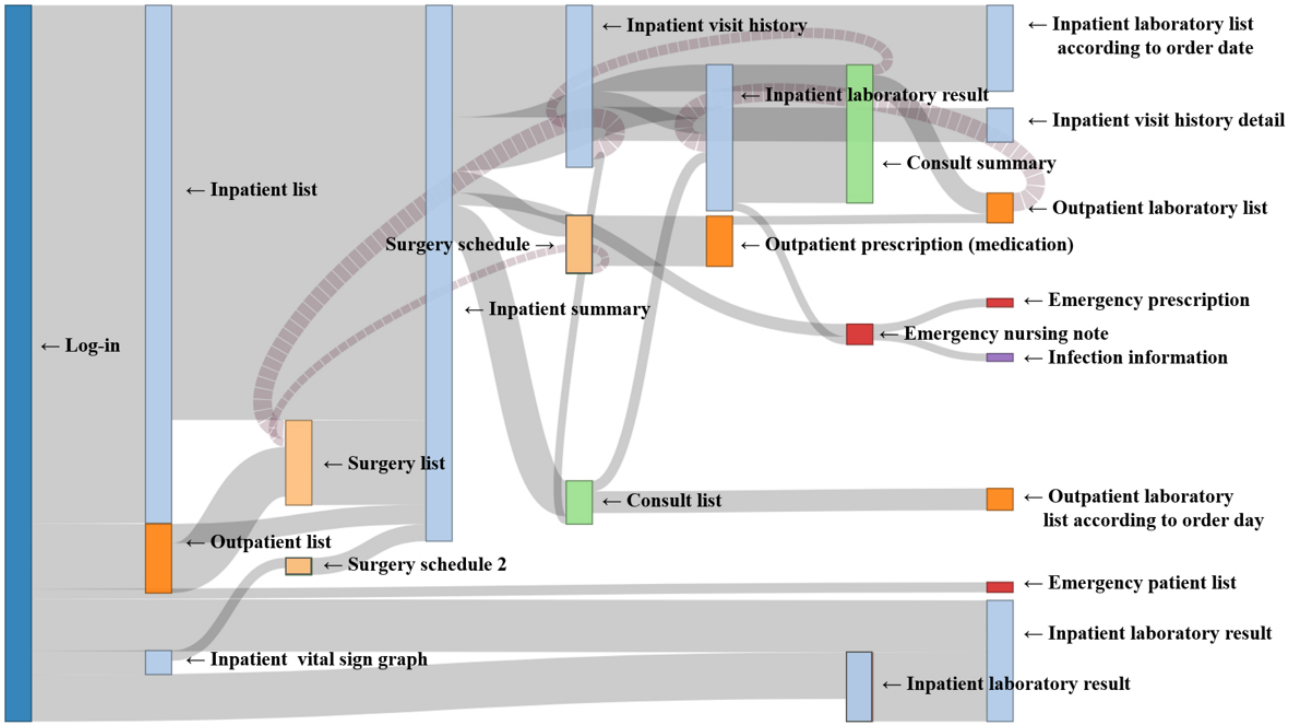


Figure 7. Sankey diagram of usage paths identified for the Surgeon group. Light blue, light green, light orange, orange and red colors were used to indicate impatient, consult, surgery, outpatient and emergency features, respectively.

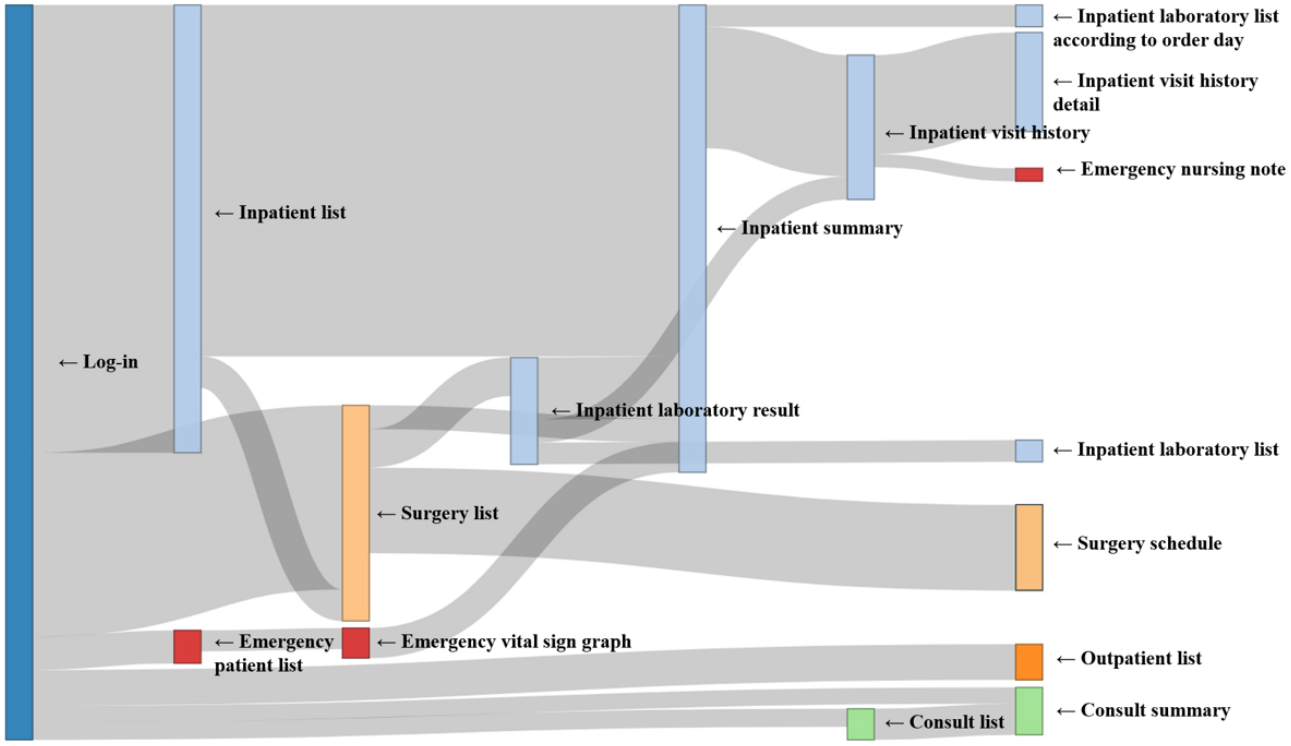


Figure 8. Sankey diagram of usage paths identified for the other hospital-based physicians (OHBP) group. Light blue, light green, light orange, orange and red colors were used to indicate impatient, consult, surgery, outpatient and emergency features, respectively.

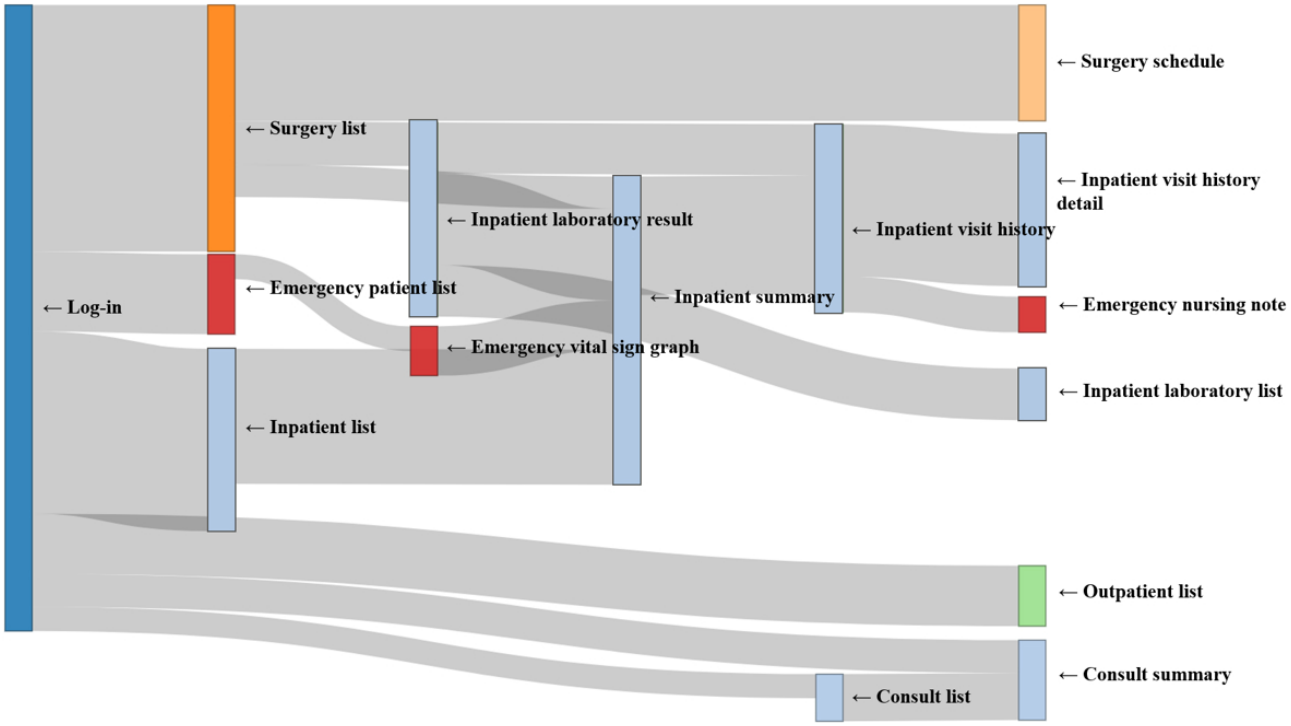


Table 1. Top 5 usage paths identified via association rule mining for each specialty group.

Specialty, rank and path[a]	N	Support (%)	Confidence (%)	Lift
Physician group				
Inpatient list (77.59) → inpatient summary (54.59)	98,590	47.93	61.77	1.13
Consult list (12.48) → consult summary (12.81)	17,594	8.55	66.79	5.35
Inpatient list (77.59) → inpatient summary (54.59) → inpatient visit history (8.5300)	16,594	8.07	14.78	1.73
Emergency patient list (12.86) → emergency vital sign graph (7.11)	14,614	7.10	55.22	7.77
Inpatient list (77.59) → inpatient summary (54.59) → inpatient visit history (8.53) → inpatient visit history detail (4.08)	8348	4.06	47.56	11.66
Surgeon group				
Inpatient list (64.61) → inpatient summary (50.36)	45,777	40.54	80.50	1.25
Inpatient list (64.61) → inpatient summary (50.36) → inpatient visit history (10.63)	11,886	10.53	20.90	1.97
Inpatient list (64.61) → inpatient summary (50.36) → inpatient laboratory result (8.76)	9328	82.61	16.40	1.98
Operation list (29.23) → surgery schedule (8.12)	9056	80.20	27.44	3.38
Inpatient list (64.61) → inpatient summary (50.36) → inpatient visit history (10.63) → inpatient visit history detail (6.27)	7058	6.25	58.78	9.37
Other hospital-based physician group				
Inpatient list (32.69) → inpatient summary (40.46) → inpatient visit history (18.84)	3200	18.68	46.16	2.45
Inpatient list (32.69) → inpatient summary (40.46)	3123	18.23	45.05	1.38
Inpatient list (32.69) → inpatient summary (40.46) → inpatient laboratory result (15.87)	2719	15.87	39.22	2.47
Operation list (40.25) → surgery schedule (14.81)	2532	14.78	36.72	2.48
Inpatient list (32.69) → inpatient summary (40.46) → inpatient visit history (18.84) → inpatient visit history detail (12.22)	2263	13.21	70.13	5.31

[a]Features in each path are listed with their support (ie, usage ratio per session) in parentheses.

Discussion

Principal Findings

Evidence for the effectiveness of mobile apps on health care is increasing [21]. EHRs are also viewed as a means of improving HCPs' decisions and clinical health outcomes [22]. However, the level of evidence on the value of mEHR is relatively low. Analyzing the work process of users, which can be matched later to clinical implication, is necessary to measure the value of any health information technology (IT) system [23].

For the users of IT systems, including mEHRs, it is essential to acquire appropriate information with the least number of click-throughs, such as log-ins, transitions, and navigation. The types and amount of information must be more tailored and intuitively visualized for the user's intent [24]. If IT solutions are not refined enough, they would increase the burden on the workflow [23]. The problem lies in the fact that HCPs do not have the answer to optimization before using them in the field or even after using them for a while. Therefore, the understanding of current usage patterns is a crucial part in system refinement and optimization.

In this study, we conducted a comprehensive analysis of the mEHR log data to investigate doctors' usage patterns using multiple analytic tools, such as a heat map, a Sankey diagram, and ARM. The heat map showed a cross-sectional volumetric view of the association between users and services (ie, departments and individual features) and hence enabled us to examine overall usage patterns of individual features according to the specialty. The Sankey diagram used the information on the first-order transition between 2 features and presented the sequential characteristics of usage patterns. Examples include frequent transitions, repetitive visits, and loops. ARM assessed co-occurrence of 2 or multiple features with quantitative criteria and identified important paths by searching for frequent if-then relationships among features. These criteria, such as support, confidence, and lift, helped to further characterize the identified paths. Our comprehensive analytic approach can be a good starting point to understand the current usage status of an mEHR and hence reveal direction for better usability (eg, feature development and user-interface modification).

More frequent use of the mEHR was observed for the physician group than for the other groups. In the volumetric view and sequential characteristic analysis, physicians tended to utilize more inpatient features and navigate through multiple low-level features in a repetitive manner. These observations implied that the current mEHR environment is more targeted at physicians who need to look up the system as they conduct inpatient care and daily patient rounds across different locations. Therefore, some improvement can be pursued to make repetitive transitions among frequently used features more efficient.

Compared with physicians, surgeons and OHBPs connected the system less frequently and used a smaller number of features. It is partly due to the shortage of specialty-specific features for them. Surgeons, especially, may benefit from features or tools related to the operating theater. For instance, augmented reality and virtual reality technology focused on the surgery process would be points of improvement for surgical specialties.

Outpatient features showed a very low usage rate in all specialty groups. It seems natural in that a desktop-based system might be more effective where doctors do not need to move around (eg, medical office, examination, and consultation rooms). To improve system efficiency, the mEHR can be modified by removing never-used features and changing the order of appearance of features according to their usage frequencies and so on.

Limitations

First, this is a single-system study with in-house software, which could bear a potential limitation for generalization. However, mEHR systems in most institutes are in the developing status, and no sufficient consensus over its standard is reached. This single-system analysis is still valuable in terms of evidence.

Second, the outcomes and measurements of this study were set only on mobile logs. Neither the practical and clinical purpose nor subjective opinions by doctors were considered. When an in-depth log analysis is combined with an investigation of users' perception, the usability of an mEHR system can be comprehensively evaluated. This comprehensive evaluation can lead to connecting the need of electronic features to clinical process and, thereby, to better system development.

Third, the offline characteristics of the specific department that utilized the features were not reflected. The difference of mEHR and EHR utilization patterns was not considered, which limits the interpretation of results on practical practices.

Conclusions

In this study, a comprehensive analysis of the mEHR log data revealed considerable differences in usage patterns among specialty groups of medical doctors. The usage paths were further characterized for each specialty and demonstrated the need and direction for the improvement of the current system including specialty-specific user interfaces.

Acknowledgments

This study was supported by the National Research Foundation of Korea (NRF). This research was supported by the Basic Science Research Program through the NRF, funded by the Ministry of Education (grant number 2018R1C1B6002877).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Average number of daily log-ins per user and content per log-in by specialty group and position.

[PDF File (Adobe PDF File), 24KB - mhealth_v7i1e12041_app1.pdf]

References

1. Lee JH. Future of the smartphone for patients and healthcare providers. Healthc Inform Res 2016 Jan;22(1):1-2 [FREE Full text] [doi: 10.4258/hir.2016.22.1.1] [Medline: 26893944]
2. Bender JL, Yue RY, To MJ, Deacken L, Jadad AR. A lot of action, but not in the right direction: systematic review and content analysis of smartphone applications for the prevention, detection, and management of cancer. J Med Internet Res 2013 Dec 23;15(12):e287 [FREE Full text] [doi: 10.2196/jmir.2661] [Medline: 24366061]
3. Helbostad JL, Vereijken B, Becker C, Todd C, Taraldsen K, Pijnappels M, et al. Mobile health applications to promote active and healthy ageing. Sensors (Basel) 2017 Mar 18;17(3) [FREE Full text] [doi: 10.3390/s17030622] [Medline: 28335475]
4. Ying A. Mobile physician order entry. J Healthc Inf Manag 2003;17(1):58-63. [Medline: 12553223]
5. 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]
6. Choi W, Park M, Hong E, Kim S, Ahn R, Hong J, et al. Early experiences with mobile electronic health records application in a tertiary hospital in Korea. Healthc Inform Res 2015 Oct;21(4):292-298 [FREE Full text] [doi: 10.4258/hir.2015.21.4.292] [Medline: 26618036]
7. Choi W, Park MA, Hong E, Kim S, Ahn R, Hong J, et al. Development of mobile electronic health records application in a secondary general hospital in Korea. Healthc Inform Res 2013 Dec;19(4):307-313 [FREE Full text] [doi: 10.4258/hir.2013.19.4.307] [Medline: 24523996]
8. Bloomfield RA, Polo-Wood F, Mandel JC, Mandl KD. Opening the Duke electronic health record to apps: implementing SMART on FHIR. Int J Med Inform 2017 Mar;99:1-10. [doi: 10.1016/j.ijmedinf.2016.12.005] [Medline: 28118917]
9. Kim S, Lee KH, Hwang H, Yoo S. Analysis of the factors influencing healthcare professionals' adoption of mobile electronic medical record (EMR) using the unified theory of acceptance and use of technology (UTAUT) in a tertiary hospital. BMC Med Inform Decis Mak 2016 Jan 30;16:12 [FREE Full text] [doi: 10.1186/s12911-016-0249-8] [Medline: 26831123]
10. Wu RC, Orr MS, Chignell M, Straus SE. Usability of a mobile electronic medical record prototype: a verbal protocol analysis. Inform Health Soc Care 2008 Jun;33(2):139-149. [doi: 10.1080/17538150802127223] [Medline: 18604757]
11. Duhm J, Fleischmann R, Schmidt S, Hupperts H, Brandt SA. Mobile electronic medical records promote workflow: physicians' perspective from a survey. JMIR Mhealth Uhealth 2016 Jun 06;4(2):e70 [FREE Full text] [doi: 10.2196/mhealth.5464] [Medline: 27268720]
12. Fleischmann R, Duhm J, Hupperts H, Brandt SA. Tablet computers with mobile electronic medical records enhance clinical routine and promote bedside time: a controlled prospective crossover study. J Neurol 2015 Mar;262(3):532-540 [FREE Full text] [doi: 10.1007/s00415-014-7581-7] [Medline: 25476692]
13. Ben-Assuli O. Electronic health records, adoption, quality of care, legal and privacy issues and their implementation in emergency departments. Health Policy 2015 Mar;119(3):287-297. [doi: 10.1016/j.healthpol.2014.11.014] [Medline: 25483873]
14. Lee Y, Park YR, Kim J, Kim JH, Kim WS, Lee JH. Usage pattern differences and similarities of mobile electronic medical records among health care providers. JMIR Mhealth Uhealth 2017 Dec 13;5(12):e178. [doi: 10.2196/mhealth.8855] [Medline: 29237579]
15. Kim J, Lee Y, Lim S, Kim JH, Lee B, Lee JH. What clinical information is valuable to doctors using mobile electronic medical records and when? J Med Internet Res 2017 Oct 18;19(10):e340 [FREE Full text] [doi: 10.2196/jmir.8128] [Medline: 29046269]
16. Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 2016 Sep 15;32(18):2847-2849. [doi: 10.1093/bioinformatics/btw313] [Medline: 27207943]
17. Yetman CJ, Allaire JJ, Ellis P, Gandrud C. The Comprehensive R Archive Network. 2017. networkD3: D3 JavaScript Network Graphs from R URL: https://cran.r-project.org/web/packages/networkD3/ [accessed 2018-12-21] [WebCite Cache ID 74pKKUJVJ]
18. Hahsler M, Chelluboina S, Hornik K, Buchta C. The arules R-package ecosystem: analyzing interesting patterns from large transaction data sets. J Mach Learn Res 2011 Jun;12:2021-2025 [FREE Full text]
19. R Core Team. Global Biodiversity Information Facility. 2018. R: A language and environment for statistical computing URL: https://www.gbif.org/tool/81287/r-a-language-and-environment-for-statistical-computing [accessed 2018-12-23] [WebCite Cache ID 74tcTnjV1]
20. Hahsler M, Gruen B, Hornik K. GitHub. 2018 Dec 03. Mining Association Rules and Frequent Itemsets URL: https://github.com/mhahsler/arules [accessed 2018-12-21] [WebCite Cache ID 74pmCOax3]
21. Burdette SD, Herchline TE, Oehler R. Surfing the web: practicing medicine in a technological age: using smartphones in clinical practice. Clin Infect Dis 2008 Jul 01;47(1):117-122. [doi: 10.1086/588788] [Medline: 18491969]

22. Blumenthal D, Tavenner M. The "meaningful use" regulation for electronic health records. N Engl J Med 2010 Aug 05;363(6):501-504. [doi: 10.1056/NEJMp1006114] [Medline: 20647183]
23. Institute of Medicine. Crossing the Quality Chasm: A New Health System for the 21st century. Washington, DC: National Academies Press; Nov 17, 2001.
24. Soh JY, Cha WC, Chang DK, Hwang JH, Kim K, Rha M, et al. Development and validation of a multidisciplinary mobile care system for patients with advanced gastrointestinal cancer: interventional observation study. JMIR Mhealth Uhealth 2018 May 07;6(5):e115 [FREE Full text] [doi: 10.2196/mhealth.9363] [Medline: 29735478]

Abbreviations

ARM: association rule mining
DARWIN: Data Analysis & Research Window for Integrated kNowledge
EHR: electronic health record
HCP: health care professional
IT: information technology
mEHR: mobile electronic health record
mHealth: mobile health
NRF: National Research Foundation of Korea
OHBP: other hospital-based physician
SMC: Samsung Medical Center

Edited by G Eysenbach; submitted 26.08.18; peer-reviewed by Y Park, L Garvin, KM Kuo; comments to author 21.09.18; revised version received 27.10.18; accepted 09.12.18; published 17.01.19.

Please cite as:
Soh JY, Jung SH, Cha WC, Kang M, Chang DK, Jung J, Lee J, Choi JS, Kim K
Variability in Doctors' Usage Paths of Mobile Electronic Health Records Across Specialties: Comprehensive Analysis of Log Data
JMIR Mhealth Uhealth 2019;7(1):e12041
URL: http://mhealth.jmir.org/2019/1/e12041/
doi:10.2196/12041
PMID:30664473

Original Paper

Assessment of Physical Activity by Wearable Technology During Rehabilitation After Cardiac Surgery: Explorative Prospective Monocentric Observational Cohort Study

Isabeau Thijs[1*], MSc; Libera Fresiello[2,3*], MSc, PhD; Wouter Oosterlinck[1], MD, PhD; Peter Sinnaeve[4], MD, PhD; Filip Rega[1], MD, PhD

[1]Research Unit of Cardiac Surgery, Department of Cardiovascular Sciences, University Hospitals Leuven, Leuven, Belgium

[2]Department of Cardiac Surgery, Katholiek Universiteit Leuven, Leuven, Belgium

[3]Institute of Clinical Physiology, National Research Council, Pisa, Italy

[4]Research Unit of Cardiology, Department of Cardiovascular Sciences, University Hospitals Leuven, Leuven, Belgium

[*]these authors contributed equally

Corresponding Author:
Isabeau Thijs, MSc
Research Unit of Cardiac Surgery
Department of Cardiovascular Sciences
University Hospitals Leuven
Herestraat 49, Bus 7003
Leuven, 3000
Belgium
Phone: 32 +3216344219
Email: isabeau.thijs@outlook.com

Abstract

Background: Wearable technology is finding its way into clinical practice. Physical activity describes patients' functional status after cardiac surgery and can be monitored remotely by using dedicated trackers.

Objective: The aim of this study was to compare the progress of physical activity in cardiac rehabilitation by using wearable fitness trackers in patients undergoing coronary artery bypass surgery by either the conventional off-pump coronary artery bypass (OPCAB) or the robotically assisted minimally invasive coronary artery bypass (RA-MIDCAB). We hypothesized faster recovery of physical activity after RA-MIDCAB in the first weeks after discharge as compared to OPCAB.

Methods: Patients undergoing RA-MIDCAB or OPCAB were included in the study. Each patient received a Fitbit Charge HR (Fitbit Inc, San Francisco, CA) physical activity tracker following discharge. Rehabilitation progress was assessed by measuring the number of steps and physical activity level daily. The physical activity level was calculated as energy expenditure divided by the basic metabolic rate.

Results: A total of 10 RA-MIDCAB patients with a median age of 68 (min, 55; max, 83) years and 12 OPCAB patients with a median age of 69 (min, 50; max, 82) years were included. Baseline characteristics were comparable except for body mass index (RA-MIDCAB: 26 kg/m²; min, 22; max, 28 versus OPCAB: 29 kg/m²; min, 27; max, 33; $P<.001$). Intubation time ($P<.05$) was significantly lower in the RA-MIDCAB group. A clear trend, although not statistically significant, was observed towards a higher number of steps in RA-MIDCAB patients in the first week following discharge.

Conclusions: RA-MIDCAB patients have an advantage in recovery in the first weeks of revalidation, which is reflected by the number of steps and physical activity level measured by the Fitbit Charge HR, as compared to OPCAB patients. However, unsupervised assessment of daily physical activity varied widely and could have consequences with regard to the use of these trackers as research tools.

(JMIR Mhealth Uhealth 2019;7(1):e9865) doi:10.2196/mhealth.9865

KEYWORDS

fitness trackers; coronary artery bypass; cardiac surgery; cardiac rehabilitation; postoperative care; wearable; physical activity; exercise

Introduction

Kolesov V (1964) performed the first coronary artery bypass graft (CABG) using the internal mammary artery to treat a patient with ischemic myocardial heart disease [1]. Off-pump coronary artery bypass (OPCAB) surgery was developed to reduce potential adverse effects induced by the use of cardiopulmonary bypass and cardioplegic arrest [2]. In this approach, the harvesting of the internal mammary arteries and anastomoses are performed on a beating heart through a median sternotomy.

Robotically assisted minimally invasive direct coronary artery bypass (RA-MIDCAB) aims to further reduce the invasiveness of the OPCAB approach by avoiding midline sternotomy. In RA-MIDCAB surgery, the internal mammary arteries are prelevated via a thoracoscopy using robotic assistance. The grafting of the bypass is performed in a second stage via a small (4-5 cm) left anterolateral minithoracotomy. Similar to OPCAB, the anastomosis is performed on a beating heart, without the use of cardiopulmonary bypass. The RA-MIDCAB approach reduces morbidity, length of hospital stay, need for blood transfusion, and wound infections [3-5]. The time for recovery is controversial among studies [3-5]: Some indicate an earlier recovery to full physical activity [3,6].

After cardiac surgery, patients need structured support to improve functional capacity and restore their quality of life. Phase II cardiac rehabilitation programs are developed to deliver comprehensive support such as monitoring in physical and psychological conditions and education of patients on healthy long-term routines. Phase II cardiac rehabilitation is suggested as a class I recommendation in the treatment of cardiac diseases by the European Society of Cardiology, the American Heart Association, and the American College of Cardiology [7-9]. Adherence could potentially be tracked by the use of remote monitoring systems.

Physical activity or fitness trackers are wearable sensors, often worn as a wristband or embedded in a smartwatch or mobile phone, that collect data on one's daily physical activity. Most of these commercially available trackers include an accelerometer to assess step counts; distance walked; and intensity, duration, and type of movement (eg, walking, running, and jogging). Thus, users can have direct access to their personal data and take an active role in monitoring their health [10,11].

These trackers are also of use in clinical practices and research. Accelerometry data can be derived noninvasively and in unsupervised, free-living conditions, which provides an opportunity to better describe patients' activity of daily living and health status in terms of mobility, behavioral pattern, and functional ability. Consequently, these data can contribute to more comprehensive, relevant, and high-quality clinical research data [10]. In clinical practices, home telemonitoring trials show favorable results in pulmonary and cardiac patients [12]. In cardiac rehabilitation, multiple cardiac telecare trials have shown a noninferiority or superiority of telemonitoring and telecoaching of patients in a cardiac rehabilitation program compared to conventional center-based supervised cardiac rehabilitation programs [12,13,14]. These physical activity trackers have the ability to encourage exercise and lifestyle behavior and monitor and share progress [11,12,15]. As such, wearables could potentially have a future in at-home management and remote monitoring of patients with chronic diseases and in secondary preventive care after cardiac surgery.

At the University Hospital of Leuven, an explorative clinical observational study was performed to evaluate physical activity in patients after coronary artery bypass (CAB) surgery. This study aimed to quantify physical fitness at particular time points and investigate whether minimal-access surgical procedures can assure faster recovery and better outcome than the conventional, more invasive surgical procedure.

Methods

Study Design

Protocol

An explorative prospective monocentric observational cohort study was conducted at the Cardiac Surgery Department of the University Hospitals of Leuven. The clinical protocol conformed to the principles outlined in the Declaration of Helsinki and was approved by the ethical committee of the University Hospitals Leuven. All patients provided written informed consent prior to inclusion in the study.

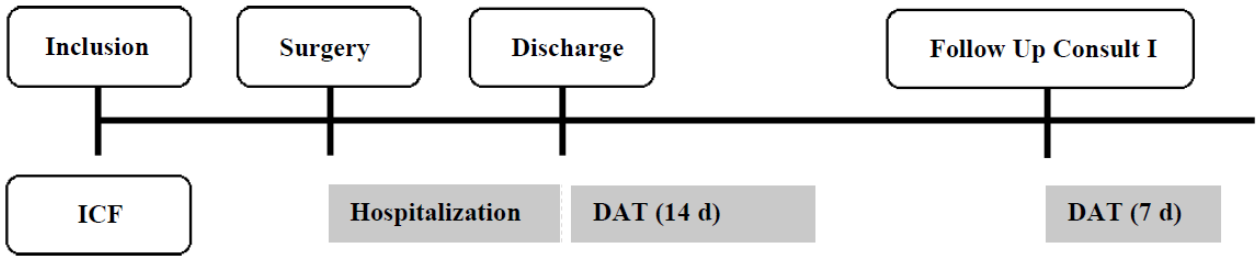
Patients

The study included patients with coronary artery disease who were eligible for elective surgical revascularization according to the most recent guidelines of the European Society of Cardiology [16]. They were scheduled to undergo either an RA-MIDCAB or OPCAB procedure. In both approaches, grafts are anastomosed on the diseased vessels without the use of cardiopulmonary bypass—the so-called off-pump technique. In OPCAB, a full sternotomy is performed, whereas in RA-MIDCAB, the anastomosis is performed through a small left anterolateral thoracotomy. Additionally, in RA-MIDCAB, internal mammary arteries (used as grafts) are harvested using robotic assistance from the Da Vinci Surgical System (Intuitive Surgical Inc, Sunnyvale, CA).

Eligibility Criteria

The main exclusion criteria were urgently scheduled and on-pump procedures, mobility problems that could interfere with physical activity, and the presence of cognitive impairment that prevented subjects from fully understanding the protocol. An overview of the inclusion and exclusion criteria is provided in Multimedia Appendix 1.

Figure 1. Timeline of the study design constructed in three evaluation time points: a preoperative inclusion, a 14-day Fitbit-wearing period after discharge, and a 7-day Fitbit-wearing period after a follow-up consult, 4 weeks after discharge. ICF: Informed Consent Form; DAT: Daily Activity Tracking.



Evaluation Design

The protocol was organized in three evaluation time points: a preoperative baseline assessment and two periods of wearing a physical activity tracker (Figure 1). At discharge, patients received a Fitbit Charge HR (Fitbit Inc, San Francisco, CA) and were asked to wear the wearable device for 14 consecutive days. Four weeks later, a follow-up consultation was scheduled, and the patient was again asked to wear the device for 7 days. The patients were asked to return the tracker to the hospital by mail. Additionally, throughout the hospital stay, clinical data were collected. An overview of the variables for which data were collected is shown in Multimedia Appendix 2.

Physical Fitness Assessment

Subjects' daily physical activity was described by parameters recorded by the Fitbit Charge HR. This activity tracker is a wrist band with an interface through which patients can monitor their real-time progress. The tracker was set according to the height, weight, and age of the subject. Subjects were instructed to wear the wristband as much as possible during the day. The tracker records the daily number of steps, the distance walked, the flights of stairs taken, the intensity and duration of exercise, the estimated energy expenditure, the sleeping pattern and the heart rate variation using the Pure Pulse technology (FitBit Inc).

Data Analysis

Data from the Fitbit Charge HR were analyzed by calculating the weekly average step counts and energy expenditure (kcal). For energy expenditure, the physical activity level was calculated by dividing the total caloric expenditure by the basic metabolic rate. This physical activity level represents the physical activity adjusted for weight, height, and age (included in the basic metabolic rate). For every subject, the first (discharge) and last (return) day of Fitbit wearing were excluded, since biased results were expected. Furthermore, continuous heart rate was evaluated to check periods when patients did not wear the device. Up to 2 hours a day of non-wearing time were neglected; if the non-wearing period was longer, that day was excluded. This time loss could be due to battery charging or activities such as bathing or showering.

Outcome Measures

Primary Outcome

The primary outcome was the objective physical activity score described by the Fitbit activity tracker data during the two periods mentioned above in Evaluation Design. Weekly average number of steps and weekly average physical activity level were used to quantify physical activity.

Secondary Outcomes

Secondary outcomes included observational data including demography, cardiac and noncardiac history, operative variables, and postoperative complications until 4 weeks after discharge.

Statistical Analyses

Statistical analyses were carried out using IBM SPSS Statistics, Version 24 (IBM Corp, Armonk, NY). Since only a small sample size was included, differences in continuous variables were analyzed using two-sided Mann Whitney *U* tests and reported as median and range with minimal and maximal values. For dichotomous variables, a Fisher exact test was performed, and for categorical variables, a Mann Whitney *U* test was performed. For repeated measures analysis, the nonparametric Friedman *t* test was used. Statistical significance was considered for *P* values<.05.

Results

Patient Recruitment

Patients were recruited from January 2017 to April 2017. In total, 25 patients were enrolled, of which 11 were RA-MIDCAB patients and 14 were OPCAB patients. Three patients were excluded after the surgery: one patient was excluded from the RA-MIDCAB group due to a prolonged hospital stay as a consequence of acute on chronic kidney failure, and two patients dropped out after surgery in the OPCAB group (one withdrew from the study and one died). Furthermore, after the 14-day Fitbit-wearing period, three patients in the OPCAB group and two patients in the RA-MIDCAB group dropped out. A study flowchart is presented in Figure 2. Results of the baseline characteristics are listed in Table 1.

The two groups did not significantly differ in age (*P*=.79), gender (*P*=>.99), height (*P*=.79), and weight (*P*=.07), but OPCAB patients had a significantly higher body mass index (29 kg/m²) than RA-MIDCAB patients (26 kg/m²; *P*<.001).

Heart failure distribution according to the New York Heart Association class was not significantly different in both groups (*P*=.89), and most subjects belonged to class I and II. The median left ventricular ejection fraction was 60% (min, 45; max, 78) in the RA-MIDCAB group and 60% (min, 40; max, 78) in the OPCAB group. Furthermore, both groups showed a

similar distribution in the European System for Cardiac Operative Risk Evaluation (EuroSCORE II; *P*=.21), left ventricular ejection fraction (*P*=.69), history of arrhythmias (*P*=.46), history of myocardial infarction (*P*=.90), and mitral regurgitation (*P*>.99).

There was a trend towards the presence of hypercholesterolemia in OPCAB patients compared to RA-MIDCAB patients (92% and 50%, respectively, *P*=.06). No significant differences were observed in the presence of comorbidities (Table 1).

Intubation time was significantly higher in the OPCAB group (*P*<.05), with 15 hours 30 minutes (min, 8 hours 2 minutes; max, 21 hours 50 minutes) in contrast to 8 hours 45 minutes (min, 5 hours 49 minutes; max, 23 hours) in the RA-MIDCAB group (Table 1). In addition, the number of grafts was significantly higher in the OPCAB group (*P*<.005), and the operation duration was higher in the OPCAB group (*P*=.10). The median duration was 5 hours 15 minutes (min, 3 hours 7 minutes; max, 6 hours 58 minutes), whereas the median duration in the RA-MIDCAB group was 4 hours 40 minutes (min, 3 hours 18 minutes; max, 5 hours 26 min).

Postoperatively, five patients in the RA-MIDCAB group and one patient in the OPCAB group were transferred to the postanesthetic care unit; a trend towards significance was observed in this parameter (*P*=.06). The length of stay at the postoperative care units (including the postanesthetic care unit and intensive care unit) was significantly lower in the RA-MIDCAB group (*P*<.001), with a median stay of 20 hours 45 minutes (min, 15 hours 30 minutes; max, 45 hours) in this group and 30 hours 45 minutes (min, 18 hours 30 minutes; max, 77 hours) in the OPCAB group. Furthermore, the overall hospital stay was similar in both groups (*P*=.21).

Figure 2. Study flowchart. CABG: coronary artery bypass graft; RA-MIDCAB: robotically assisted minimally invasive direct coronary artery bypass; OPCAB: off-pump coronary artery bypass; AKI: acute kidney injury.

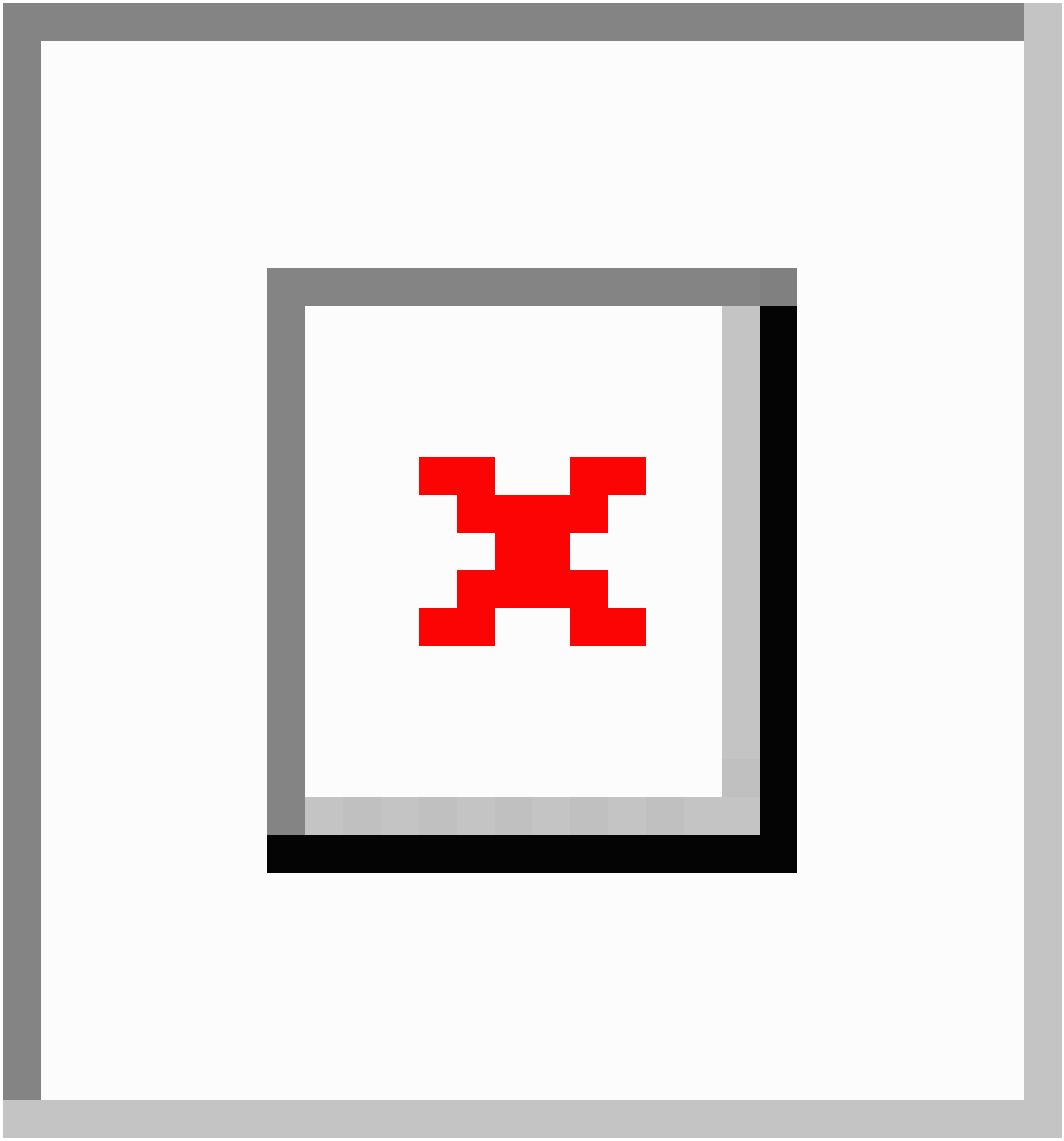


Table 1. Baseline characteristics and postoperative parameters.

Characteristics	RA-MIDCAB[a] (n=10)	OPCAB[b] (n=12)	*P* value
Demographics			
Sex (male), n (%)	9 (90)	10 (78)	>.99
Age (y), median (min, max)	68 (55, 83)	69 (50, 82)	.79
Height (cm), median (min, max)	172 (154, 178)	171 (155, 178)	.79
Weight (kg), median (min, max)	77 (57, 90)	83 (75, 100)	.07
Body mass index (kg/m^2), median (min, max)	26 (22, 28)	29 (27, 33)	<.005[c]
Cardiac history			
NYHA[d] Class, n			.89
I	4	5	
II	4	5	
III	2	2	
Presence of unstable angina, n	2	3	.79
Left ventricular ejection fraction (%), median (min, max)	60 (45, 78)	60 (40, 78)	.69
EuroSCORE II[e] (%), median (min, max)	1.7 (0.6, 5.5)	1.1 (0.6, 2.8)	.21
History of arrhythmias, n	1	0	.46
History of myocardial infarction, n			.9
Non-ST-segment elevated myocardial infarction, n	2	3	
ST-segment elevated myocardial infarction, n	1	0	
Mild mitral regurgitation, n	5	5	>.99
Non-cardiac history			
Smoking status, n			.64
Ex-smoker since >1 month	5	6	
Smoker	3	1	
Diabetes mellitus type II, n	3	2	.62
Arterial hypertension, n	7	10	.62
Pulmonary hypertension, n[f]	1	1	>.99
Hypercholesterolemia, n	5	11	.06
Renal impairment, n	9	6	.83
Peripheral vascular diseases, n	1	2	>.99
Chronic obstructive pulmonary disease, n	1	2	>.99
Operative data, median (min, max)			
Operation duration, min	281 (198, 326)	316 (187, 418)	.1
Intubation time, min	527 (349, 1380)	931 (482, 1310)	<.05[c]
Number of grafts, n	2 (1, 2)	3.5 (2, 4)	<.005[c]
Postoperative data			
Postanesthetic care unit stay, n	5	1	
Postoperative care unit length of stay (min), median (min, max)	1245 (930, 2700)	1440 (1110, 4620)	<.001[c]
Hospital length of stay (d), median (min, max)	6 (4, 12)	7 (5, 15)	.21
Complications, n			
Wound infections	0	2	.48

Characteristics	RA-MIDCAB[a] (n=10)	OPCAB[b] (n=12)	*P* value
Pulmonary infections	0	2	.48
Pleural effusions	2	3	>.99
New arrhythmias	2	1	.57
Hypokalemia	4	3	.65
Pericarditis	2	0	.2

[a]RA-MIDCAB: robotically assisted minimally invasive direct coronary artery bypass.

[b]OPCAB: off-pump coronary artery bypass.

[c]*P* values are significant.

[d]NYHA: New York Heart Association functional classification of heart failure.

[e]EuroSCORE II: European System for Cardiac Operative Risk Evaluation.

[f]Pulmonary hypertension was defined as pulmonary pressure>25 mmHg.

Table 2. Step count analysis recorded by Fitbit Charge HR in RA-MIDCAB and OPCAB patients.

Average step counts	RA-MIDCAB[a]	Percentage[b]	OPCAB[c]	*P* value
Overall analysis, median (min, max), n				
Week 1	3715 (1637, 6720), 10	335	1110 (739, 10,195), 11	.06
Week 2	4357 (1415, 7671), 10	238	1832 (856, 11,282), 10	.33
Week 5	6012 (3473, 11579), 8	105	5719 (2128, 11,948), 9	.7
Analysis without dropouts, median (min, max), n				
Week 1	3715 (1734, 6720), 8	371	1001 (739, 10,195), 9	.07
Week 2	4357 (1512, 7286), 8	459	949 (856, 11,282), 9	.17
Week 5	6012 (3473, 11579), 8	105	5719 (2128, 11,948), 9	.7
Repeated measures Friedman *t* test				
Chi-square	28		30	—[d]
P value	<.001		<.001	—

[a]RA-MIDCAB: robotically assisted minimally invasive direct coronary artery bypass.

[b]Percentage of number of steps in RA-MIDCAB patients compared to OPCAB patients.

[c]OPCAB: off-pump coronary artery bypass.

[d]Not applicable.

Primary Outcome

Step counts

Data were corrected for periods when the physical activity tracker was not worn. In week 1, a total of 3 days for one patient were excluded from the analysis. In week 2, a total of 9 days, distributed over four patients, were excluded. In week 5, 2 days, distributed over two patients, were excluded.

In the first week, the RA-MIDCAB group showed a higher average number of steps than the OPCAB group, a result almost statistically significant (*P*=.06). Similarly, in the second week after discharge, RA-MIDCAB patients took more steps, but no significant difference was observed between the groups (*P*=.33). In week 5, the OPCAB group bridged the gap in the number of steps, and the average number of steps was similar between the two groups (*P*=.70; Table 2; Figures 3 and 4).

A nonparametric Friedman *t* test was performed to analyze the repeated measures analysis for the number of steps over time. A significant change over time was observed in the RA-MIDCAB group (28 steps; *P*<.001) and the OPCAB group (30 steps; *P*<.001; Table 2).

Physical Activity Level

With regard to the physical activity level, no significant differences were observed in weeks 1, 2, and 5 between the RA-MIDCAB and OPCAB groups (*P*=.36, *P*=.36, and *P*=.50, respectively). However, the physical activity level was higher in the RA-MIDCAB group than in the OPCAB group in all weeks (Table 3).

Figure 3. Weekly average number of steps in robotically assisted minimally invasive direct coronary artery bypass (RA-MIDCAB) and off-pump coronary artery bypass (OPCAB) patients plotted over time. Weekly average step count is plotted as median over time. n indicates the number of patients included in the cohort result. *P* value is for the Mann Whitney *U* test for the difference between the two groups at that time point. X^2 results of repeated measures Friedman *t* test.

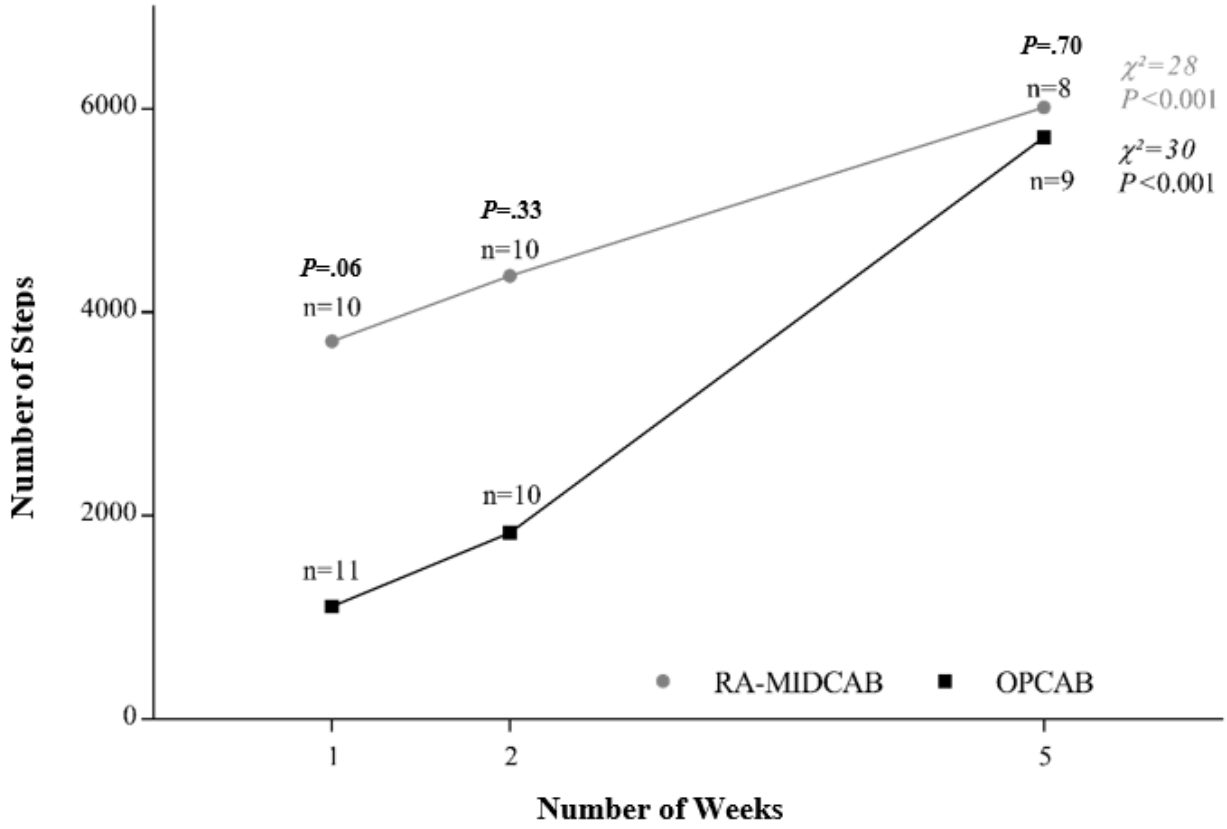


Figure 4. Boxplots of weekly average number of steps in robotically assisted minimally invasive direct coronary artery bypass (RA-MIDCAB) and off-pump coronary artery bypass (OPCAB) patients. Weekly average step counts is shown as box and whisker plots, presenting medians, 25% and 75% quartiles, minimums, and maximums. n indicates the number of patients included in the cohort result. M_W1: MIDCAB result in week 1; O_W1: OPCAB results in week 1; M_W2: MIDCAB results in week 2; O_W2: OPCAB results in week 2; M_W5: MIDCAB results in week 5; O_W5: OPCAB results in week 5.

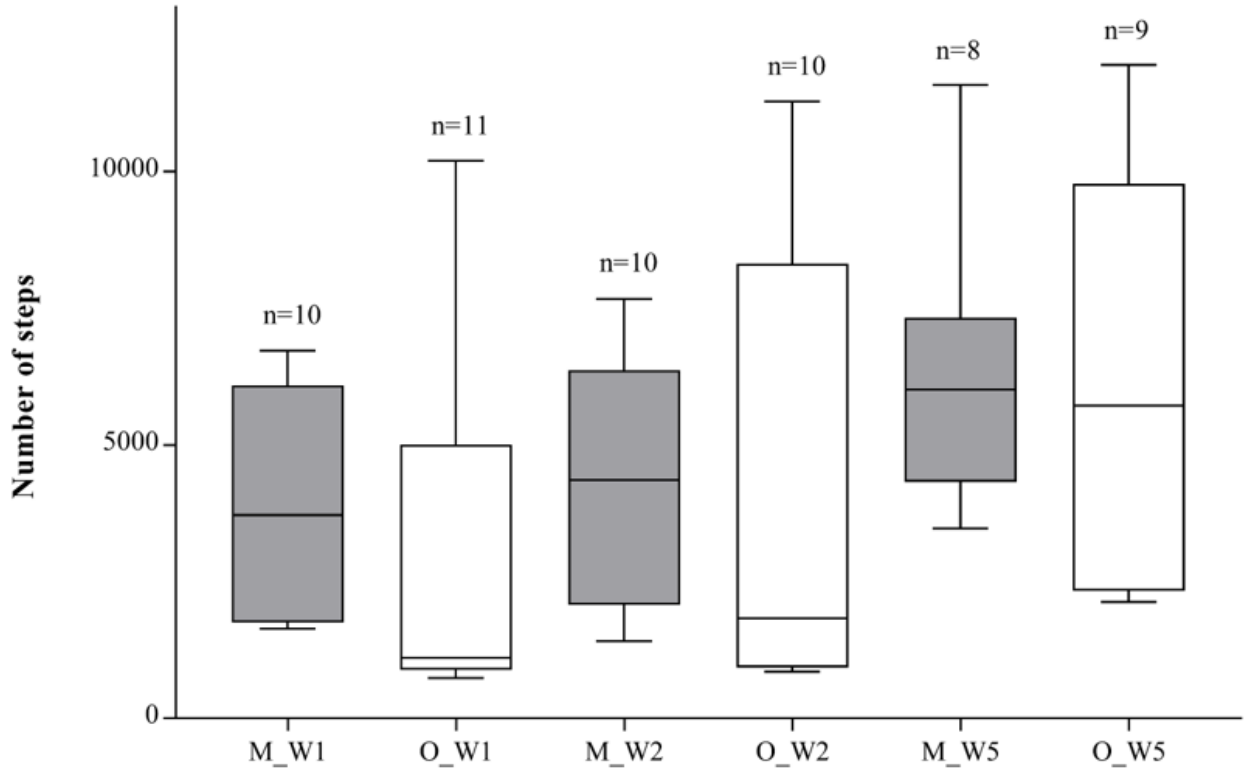


Table 3. Physical activity level[a] analysis recorded by Fitbit Charge HR in RA-MIDCAB and OPCAB patients.

Time point	RA-MIDCAB[b]		OPCAB[c]		*P* value
	median (min, max)	n	median (min, max)	n	
Overall analysis					
Week 1	1.39 (1.05, 1.71)	10	1.29 (1.08, 1.59)	11	.36
Week 2	1.41 (1.04, 1.63)	10	1.32 (1.04, 1.60)	10	.36
Week 5	1.52 (1.13, 1.90)	8	1.44 (1.16, 1.80)	9	.5
Analysis without dropouts					
Week 1	1.39 (1.05, 1.71)	8	1.26 (1.08, 1.59)	9	.4
Week 2	1.41 (1.04, 1.63)	8	1.23 (1.04, 1.60)	9	.42
Week 5	1.52 (1.13, 1.90)	8	1.44 (1.16, 1.80)	9	.5

[a]Physical activity level calculated as total energy expenditure divided by basic metabolic rate.

[b]RA-MIDCAB: robotically assisted minimally invasive direct coronary artery bypass.

[c]OPCAB: off-pump coronary artery bypass; data are reported as median (min, max).

Discussion

Principal Findings

We evaluated physical activity in cardiac rehabilitation by using the Fitbit Charge HR tracker device after conventional and minimally invasive CAB surgery. A clear trend was observed towards a higher physical activity level in RA-MIDCAB patients than in OPCAB patients, which was reflected in the number of steps and physical activity level, although statistical significance was not reached.

Value of Wearable Activity Trackers in Surgical Outcome Research

To our knowledge, this is the first study to use wearable activity tracking in a clinical environment to compare the outcome of

two types of cardiac surgery interventions. The Fitbit Charge HR provided useful information about patients' physical activity in this study. Wearable activity trackers are finding their way in research and medical practices [10,15,17]. An important limitation, however, is that commercially available activity trackers are often not thoroughly validated for their accuracy and reliability. Studies showed that step count accuracy is dependent on gait patterns in healthy volunteers [18,19]. In the elderly and chronically ill, a negative correlation was found between the gait pattern and step count accuracy, as assessed by other commercially available activity trackers [18,20]. Postsurgical patients who are still in recovery and probably walk at a slow speed will create a bias in the number of steps counted. In addition, the Fitbit Charge HR is still rarely used in research. However, older-generation models have been tested for their step count accuracy [21-24] and energy expenditure estimation [25,26].

Taking into account the limitations of the Fitbit technology, it is worth highlighting that our data were not analyzed as exact results but were only used to compare the RA-MIDCAB and OPCAB groups. Any error in the step count technology would affect RA-MIDCAB and OPCAB patients in a similar way without impacting the comparative analysis performed in this work. No baseline references are available, and Fitbit does not disclose algorithms or mean error values. Hence, it was not possible to apply mean error corrections. However, it is distinctively true that further research is needed before commercially available self-monitoring wearables can be used in clinical applications.

Besides the lack of validation, wearable activity trackers have a promising future. Activity trackers provide the possibility for patients to monitor their activity patterns and share their progress with physicians, friends, and family members. Therefore, these trackers can be used as motivational tools to reach and maintain a healthy active lifestyle [15]. In this study, subjects were nearly always compliant and motivated to wear the tracker and monitor their own progress. Physical activity is one of the most health-enhancing practices, especially in primary and secondary prevention of cardiovascular risk factors. Physical activity counselling (by use of wearable activity monitors) has been shown to improve healthy lifestyles [27-29]. Savage et al (2008) found significant correlations between the daily number of steps in the first weeks of Phase II cardiac rehabilitation and cardiovascular risk factors [30].

The future for this wearable activity trackers is still unknown, but its implementation in medical practice would provide many benefits, for instance, in cardiac rehabilitation to overcome barriers to cardiac rehabilitation programs. Multiple trials have shown that wearable activity tracking and digital health devices encourage patients to improve their physical behavior and are therefore useful tools in cardiac rehabilitation. The Telerehab III trial showed that telerehabilitation, through use of a commercially available accelerometer, provided substantial and persistent health benefits and novel, cost-efficient care [31].

Comparison of Minimally Invasive and Conventional Coronary Artery Bypass

Our study showed that both groups were comparable. No significant differences were found in baseline characteristics, except for body mass index and the number of anastomosed grafts. EuroSCORE II calculation showed no significant difference in the predicted operative mortality. OPCAB patients have a more pronounced disease and may be considered more unwell, which can be reflected in the body mass index. In this group of patients, however, it seems that body mass index does not have a significant influence on physical activity and performance. Spearman correlation analysis showed that weight/body mass index was not related with the number of steps at any time point (Multimedia Appendix 3). Surgical and postsurgical data reflect the difference between both procedures. RA-MIDCAB entails a shorter intubation time and a shorter stay at postoperative care units, with more patients transferred to the postanesthetic care unit postoperatively (fast-track treatment). These findings are in line with previous studies [4,32]. Hospital length of stay, however, showed no significant difference between the two study groups; this is in contrast to the findings of other studies [33-35], which could be due to the small sample size of the present study. However, this variable might be dependent on institutional protocols and decision making of physicians and surgeons [36].

Despite the similar length of stay and baseline characteristics between the two groups, a clear trend was observed toward higher physical activity reflected in step counts and physical activity level in RA-MIDCAB patients in the first weeks, although statistical significance was not reached. Step counts depict the actual daily walking of patients during the day. The physical activity level depicts the energy expenditure as a result of activity, adjusted for individual basic metabolic rate. Therefore, both parameters interpret physical activity in a different manner and should be interpreted accordingly. It is harder to reach significance in the physical activity level analysis due to the smaller scale of variations. Owing to its explorative nature, this study is probably underpowered to detect smaller differences and is influenced greatly by outliers. Both groups showed significant changes in the number of steps over time (Table 2). Together with the differences between the groups, this could indicate that RA-MIDCAB patients advance in the early stages and OPCAB patients need some time to catch up. It must be noted that physical activity varied greatly among subjects, which could be due to the accuracy levels and algorithms of the device itself. In addition, physical activity is dependent on personal habits and character, referred to as self-efficacy [37,38], and the motivational support from the environment (relatives and friends). Patients who are sedentary before the surgery would likely abide by this lifestyle after surgery. Patients who are regularly active would probably be more motivated to achieve their prior level of fitness before the disease became symptomatic. This was illustrated in the Telerehab III trial where patients partially relapsed after telerehabilitation was stopped [31]. However, this effect would influence both groups similarly. Furthermore, inclusion in a clinical study and the intervention for monitoring activity by use of a tracker could be motivating factors. The tracker makes

it possible for patients to monitor their own progress and activity. These factors could possibly also contribute to achieving a higher level of activity.

Both groups of patients walked about 5000 to 6000 steps a day at steady state in the fifth week after discharge. The American Heart Association recommends that healthy subjects walk 10,000 steps a day for overall better health outcomes, including cardiovascular outcome [39]. The official recommendation by the American Heart Association and World Health Organization is 150 minutes of moderate-intensity aerobic physical activity a day for 5 days a week [40], equivalent to 7000-8000 steps a day. Prior studies evaluating the number of steps in patients with coronary artery disease in secondary prevention proposed a target of 7500 steps a day to correlate with improved condition in terms of lipid profiles, muscle endurance, and body mass index [41,42].

Limitations

This explorative observational study has multiple limitations. The lack of validation for this wearable technology was already described in the Discussion section.

Subjects were scheduled for either OPCAB or RA-MIDCAB surgery based on the coronary anatomy (number of grafts), comorbidities, and endovascular options. Therefore, they were matched according to baseline characteristics, cardiac history, and comorbidity profile, but the groups were found to be significantly different for body mass index, which was higher in the OPCAB group. Although body mass index was not significantly correlated to the number of steps at any point of the study, it might still be an influencing factor for physical activity. However, the difference was not taken into account in further analysis.

As stated above, physical activity is influenced by other factors in addition to the impact of a surgical intervention. Not all patients are equally active in nature and the differences may depend on self-efficacy, the choice for physiotherapy, and cardiac rehabilitation thereafter. Physiotherapy in the first weeks after surgery and the following Phase II cardiac rehabilitation may significantly influence the progress in physical activity.

Subjects were recommended to wear the activity tracker all day and to take it off only for charging or while showering. To verify if patients were constantly wearing the device, we checked for any missing data in continuous heart rate monitoring. Apart from this measure, it was difficult to supervise the wearing time. In an ideal setting, these patient cohorts would be analyzed in a randomized controlled trial. In addition, the present study is based on a small sample size. Hence, the results should be interpreted with caution, and further investigations should be carried out before outlining definitive conclusions.

Conclusions

This research aimed to evaluate postsurgical cardiac rehabilitation progress by using commercially available wearable technology. We confirm our hypothesis that RA-MIDCAB patients have an advantage over OPCAB patients with regard to revalidation. Although not statistically significant, the RA-MIDCAB patient cohort showed a clear trend towards higher physical activity level in the first weeks after surgery. The exact hinge point must be confirmed with a larger number of patients. This work highlighted the feasibility of the use of wearable technology for physical activity monitoring in a clinical setting. Further research should be conducted to evaluate the accuracy and reliability of wearable technology before it serves clinical applications, especially in nonhealthy subjects with an altered gait pattern.

Conflicts of Interest

None of the authors have any conflict of interest to declare with regard to the design or execution of this study. The Fitbit trackers were purchased by the Department of Cardiac Surgery of the University Hospitals Leuven.

Multimedia Appendix 1

Inclusion and exclusion criteria.

[PDF File (Adobe PDF File), 15KB - mhealth_v7i1e9865_app1.pdf]

Multimedia Appendix 2

Baseline characteristics and demographics.

[PDF File (Adobe PDF File), 11KB - mhealth_v7i1e9865_app2.pdf]

Multimedia Appendix 3

Spearman correlations with the number of steps in weeks 1, 2, and 5.

[PDF File (Adobe PDF File), 13KB - mhealth_v7i1e9865_app3.pdf]

References

1. Cheng TO. History of coronary artery bypass surgery--half of a century of progress. Int J Cardiol 2012 May 17;157(1):1-2. [doi: 10.1016/j.ijcard.2012.02.018] [Medline: 22425221]

2. Soylu E, Harling L, Ashrafian H, Attaran S, Athanasiou C, Punjabi P, et al. A systematic review of the safety and efficacy of distal coronary artery anastomotic devices in MIDCAB and TECAB surgery. Perfusion 2016 Oct;31(7):537-543. [doi: 10.1177/0267659115618004] [Medline: 26590391]
3. Ruel M, Une D, Bonatti J, McGinn J. Minimally invasive coronary artery bypass grafting: is it time for the robot? Curr Opin Cardiol 2013 Nov;28(6):639-645. [doi: 10.1097/HCO.0b013e3283653fd1] [Medline: 24077608]
4. Rogers C, Pike K, Angelini G, Reeves B, Glauber M, Ferrarini M, et al. An open randomized controlled trial of median sternotomy versus anterolateral left thoracotomy on morbidity and health care resource use in patients having off-pump coronary artery bypass surgery: the Sternotomy Versus Thoracotomy (STET) trial. J Thorac Cardiovasc Surg 2013 Aug;146(2):306-16.e1 [FREE Full text] [doi: 10.1016/j.jtcvs.2012.04.020] [Medline: 22944093]
5. Magee M, Mack M. Robotics and coronary artery surgery. Curr Opin Cardiol 2002 Nov;17(6):602-607. [Medline: 12466701]
6. Poston RS, Tran R, Collins M, Reynolds M, Connerney I, Reicher B, et al. Comparison of economic and patient outcomes with minimally invasive versus traditional off-pump coronary artery bypass grafting techniques. Ann Surg 2008 Oct;248(4):638-646 [FREE Full text] [doi: 10.1097/SLA.0b013e31818a15b5] [Medline: 18936577]
7. Niebauer J. Is There a Role for Cardiac Rehabilitation After Coronary Artery Bypass Grafting? Treatment After Coronary Artery Bypass Surgery Remains Incomplete Without Rehabilitation. Circulation 2016 Jun 14;133(24):2529-2537. [doi: 10.1161/CIRCULATIONAHA.116.021348] [Medline: 27297345]
8. Mendes M. Is There a Role for Cardiac Rehabilitation After Coronary Artery Bypass Grafting? There is No Role for Cardiac Rehabilitation After Coronary Artery Bypass Grafting. Circulation 2016 Jun 14;133(24):2538-2543. [doi: 10.1161/CIRCULATIONAHA.115.017800] [Medline: 27297346]
9. European Association of Cardiovascular Prevention Rehabilitation Committee for Science Guidelines, EACPR, Corrà U, Piepoli MF, Carré F, Heuschmann P, Document Reviewers, et al. Secondary prevention through cardiac rehabilitation: physical activity counselling and exercise training: key components of the position paper from the Cardiac Rehabilitation Section of the European Association of Cardiovascular Prevention and Rehabilitation. Eur Heart J 2010 Aug;31(16):1967-1974. [doi: 10.1093/eurheartj/ehq236] [Medline: 20643803]
10. Munos B, Baker P, Bot B, Crouthamel M, de Vries G, Ferguson I, et al. Mobile health: the power of wearables, sensors, and apps to transform clinical trials. Ann N Y Acad Sci 2016 Dec;1375(1):3-18. [doi: 10.1111/nyas.13117] [Medline: 27384501]
11. Kas K. Nooit Meer Ziek. Antwerpen: Van Halewyck Nv; 2015.
12. Piwek L, Ellis D, Andrews S, Joinson A. The Rise of Consumer Health Wearables: Promises and Barriers. PLoS Med 2016 Feb;13(2):e1001953 [FREE Full text] [doi: 10.1371/journal.pmed.1001953] [Medline: 26836780]
13. Huang K, Liu W, He D, Huang B, Xiao D, Peng Y, et al. Telehealth interventions versus center-based cardiac rehabilitation of coronary artery disease: A systematic review and meta-analysis. Eur J Prev Cardiol 2015 Aug;22(8):959-971. [doi: 10.1177/2047487314561168] [Medline: 25488550]
14. Frederix I, Vanhees L, Dendale P, Goetschalckx K. A review of telerehabilitation for cardiac patients. J Telemed Telecare 2015 Jan;21(1):45-53. [doi: 10.1177/1357633X14562732] [Medline: 25475219]
15. Kaiser D, Harrington R, Turakhia M. Wearable Fitness Trackers and Heart Disease. JAMA Cardiol 2016 Dec 01;1(2):239 [FREE Full text] [doi: 10.1001/jamacardio.2016.0354] [Medline: 27437902]
16. Windecker S, Kolh P, Alfonso F, Collet J, Cremer J, Falk V, Grupa Robocza Europejskiego Towarzystwa Kardiologicznego (ESC), Europejskie Stowarzyszenie Chirurgii Serca i Klatki Piersiowej (EACTS) do spraw rewaskularyzacji mięśnia sercowego, European Association for Percutaneous Cardiovascular Interventions (EAPCI). [2014 ESC/EACTS Guidelines on myocardial revascularization]. Kardiol Pol 2014;72(12):1253-1379. [doi: 10.5603/KP.2014.0224] [Medline: 25524605]
17. Eapen Z, Peterson E. Can Mobile Health Applications Facilitate Meaningful Behavior Change?: Time for Answers. JAMA 2015;314(12):1236-1237. [doi: 10.1001/jama.2015.11067] [Medline: 26393844]
18. Leth S, Hansen J, Nielsen O, Dinesen B. Evaluation of Commercial Self-Monitoring Devices for Clinical Purposes: Results from the Future Patient Trial, Phase I. Sensors (Basel) 2017 Jan 22;17(1) [FREE Full text] [doi: 10.3390/s17010211] [Medline: 28117736]
19. Fokkema T, Kooiman T, Krijnen W, VAN DER Schans CP, DE Groot M. Reliability and Validity of Ten Consumer Activity Trackers Depend on Walking Speed. Med Sci Sports Exerc 2017 Dec;49(4):793-800. [doi: 10.1249/MSS.0000000000001146] [Medline: 28319983]
20. Lauritzen J, Muñoz A, Luis Sevillano J, Civit A. The usefulness of activity trackers in elderly with reduced mobility: a case study. Stud Health Technol Inform 2013;192:759-762. [Medline: 23920659]
21. Wong C, Mentis H, Kuber R. The bit doesn't fit: Evaluation of a commercial activity-tracker at slower walking speeds. Gait Posture 2018 Dec;59:177-181. [doi: 10.1016/j.gaitpost.2017.10.010] [Medline: 29049964]
22. Kooiman T, Dontje M, Sprenger S, Krijnen W, van der Schans CP, de Groot M. Reliability and validity of ten consumer activity trackers. BMC Sports Sci Med Rehabil 2015;7:24 [FREE Full text] [doi: 10.1186/s13102-015-0018-5] [Medline: 26464801]
23. Bai Y, Welk GJ, Nam YH, Lee JA, Lee JM, Kim Y, et al. Comparison of Consumer and Research Monitors under Semistructured Settings. Med Sci Sports Exerc 2016 Jan;48(1):151-158. [doi: 10.1249/MSS.0000000000000727] [Medline: 26154336]

24. An H, Jones G, Kang S, Welk G, Lee JM. How valid are wearable physical activity trackers for measuring steps? Eur J Sport Sci 2017 Apr;17(3):360-368. [doi: 10.1080/17461391.2016.1255261] [Medline: 27912681]
25. Evenson K, Goto M, Furberg R. Systematic review of the validity and reliability of consumer-wearable activity trackers. Int J Behav Nutr Phys Act 2015 Dec 18;12:159 [FREE Full text] [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]
26. Diaz K, Krupka D, Chang M, Peacock J, Ma Y, Goldsmith J, et al. Fitbit: An accurate and reliable device for wireless physical activity tracking. Int J Cardiol 2015 Apr 15;185:138-140 [FREE Full text] [doi: 10.1016/j.ijcard.2015.03.038] [Medline: 25795203]
27. Berra K, Rippe J, Manson J. Making Physical Activity Counseling a Priority in Clinical Practice: The Time for Action Is Now. JAMA 2015;314(24):2617-2618. [doi: 10.1001/jama.2015.16244] [Medline: 26662069]
28. Lin JS, O'Connor E, Evans CV, Senger CA, Rowland MG, Groom HC. Behavioral counseling to promote a healthy lifestyle in persons with cardiovascular risk factors: a systematic review for the U.S. Preventive Services Task Force. Ann Intern Med 2014 Oct 21;161(8):568-578. [doi: 10.7326/M14-0130] [Medline: 25155549]
29. Chiauzzi E, Rodarte C, DasMahapatra P. Patient-centered activity monitoring in the self-management of chronic health conditions. BMC Med 2015 Apr 09;13:77 [FREE Full text] [doi: 10.1186/s12916-015-0319-2] [Medline: 25889598]
30. Savage P, Ades P. Pedometer step counts predict cardiac risk factors at entry to cardiac rehabilitation. J Cardiopulm Rehabil Prev 2008;28(6):370-7; quiz 378. [doi: 10.1097/HCR.0b013e31818c3b6d] [Medline: 19008690]
31. Frederix I, Solmi F, Piepoli M, Dendale P. Cardiac telerehabilitation: A novel cost-efficient care delivery strategy that can induce long-term health benefits. Eur J Prev Cardiol 2017 Dec;24(16):1708-1717. [doi: 10.1177/2047487317732274] [Medline: 28925749]
32. Gong W, Cai J, Wang Z, Chen A, Ye X, Li H, et al. Robot-assisted coronary artery bypass grafting improves short-term outcomes compared with minimally invasive direct coronary artery bypass grafting. J Thorac Dis 2016 Mar;8(3):459-468 [FREE Full text] [doi: 10.21037/jtd.2016.02.67] [Medline: 27076941]
33. Rabindranauth P, Burns J, Vessey T, Mathiason M, Kallies K, Paramesh V. Minimally invasive coronary artery bypass grafting is associated with improved clinical outcomes. Innovations (Phila) 2014;9(6):421-426. [doi: 10.1097/IMI.0000000000000110] [Medline: 25438111]
34. Lapierre H, Chan V, Sohmer B, Mesana T, Ruel M. Minimally invasive coronary artery bypass grafting via a small thoracotomy versus off-pump: a case-matched study. Eur J Cardiothorac Surg 2011 Oct;40(4):804-810. [doi: 10.1016/j.ejcts.2011.01.066] [Medline: 21393011]
35. Halkos M, Vassiliades T, Myung R, Kilgo P, Thourani V, Cooper W, et al. Sternotomy versus nonsternotomy LIMA-LAD grafting for single-vessel disease. Ann Thorac Surg 2012 Nov;94(5):1469-1477. [doi: 10.1016/j.athoracsur.2012.05.049] [Medline: 22776082]
36. Birla R, Patel P, Aresu G, Asimakopoulos G. Minimally invasive direct coronary artery bypass versus off-pump coronary surgery through sternotomy. Ann R Coll Surg Engl 2013 Oct;95(7):481-485 [FREE Full text] [doi: 10.1308/003588413X13629960047119] [Medline: 24112493]
37. Blanchard C, Arthur H, Gunn E. Self-efficacy and outcome expectations in cardiac rehabilitation: associations with women's physical activity. Rehabil Psychol 2015 Feb;60(1):59-66. [doi: 10.1037/rep0000024] [Medline: 25602362]
38. Sweet S, Fortier M, Strachan S, Blanchard C, Boulay P. Testing a Longitudinal Integrated Self-Efficacy and Self-Determination Theory Model for Physical Activity Post-Cardiac Rehabilitation. Health Psychol Res 2014 Jan 13;2(1):1008 [FREE Full text] [doi: 10.4081/hpr.2014.1008] [Medline: 26973926]
39. Nocon M, Hiemann T, Müller-Riemenschneider F, Thalau F, Roll S, Willich S. Association of physical activity with all-cause and cardiovascular mortality: a systematic review and meta-analysis. Eur J Cardiovasc Prev Rehabil 2008 Jun;15(3):239-246. [doi: 10.1097/HJR.0b013e3282f55e09] [Medline: 18525377]
40. Global recommendations on physical activity for health Internet. World Health Organization URL: http://www.who.int/dietphysicalactivity/factsheet_recommendations/en/ [accessed 2018-03-02] [WebCite Cache ID 6xcElUr36]
41. Bäck M, Cider A, Gillström J, Herlitz J. Physical activity in relation to cardiac risk markers in secondary prevention of coronary artery disease. Int J Cardiol 2013 Sep 20;168(1):478-483. [doi: 10.1016/j.ijcard.2012.09.117] [Medline: 23041099]
42. Houle J, Valera B, Gaudet-Savard T, Auclair A, Poirier P. Daily steps threshold to improve cardiovascular disease risk factors during the year after an acute coronary syndrome. J Cardiopulm Rehabil Prev 2013;33(6):406-410. [doi: 10.1097/HCR.0000000000000021] [Medline: 24104407]

Abbreviations

CAB: coronary artery bypass
CABG: coronary artery bypass grafting
EuroSCORE II: European System for Cardiac Operative Risk Evaluation
NYHA: New York Heart Association functional classification of heart failure
OPCAB: off-pump coronary artery bypass
RA-MIDCAB: robotically assisted minimally invasive direct coronary artery bypass
STEMI: ST-segment elevated myocardial infarction

Edited by G Eysenbach; submitted 17.01.18; peer-reviewed by J Goris, F Vogt, J Pietilä; comments to author 15.02.18; revised version received 13.03.18; accepted 10.12.18; published 31.01.19.

Please cite as:
Thijs I, Fresiello L, Oosterlinck W, Sinnaeve P, Rega F
Assessment of Physical Activity by Wearable Technology During Rehabilitation After Cardiac Surgery: Explorative Prospective Monocentric Observational Cohort Study
JMIR Mhealth Uhealth 2019;7(1):e9865
URL: http://mhealth.jmir.org/2019/1/e9865/
doi:10.2196/mhealth.9865
PMID:30702433

Original Paper

Mobile App for Improved Self-Management of Type 2 Diabetes: Multicenter Pragmatic Randomized Controlled Trial

Payal Agarwal[1,2], MD; Geetha Mukerji[1,3,4], MSc, MD; Laura Desveaux[1,4,5], PhD; Noah M Ivers[1,2,4,5], MD, PhD; Onil Bhattacharyya[1,2,5], MD, PhD; Jennifer M Hensel[1,5,6,7], MD, MSc; James Shaw[1,5], PhD; Zachary Bouck[1], MPH; Trevor Jamieson[1,3,8], MBI, MD; Nike Onabajo[1], MSc; Madeline Cooper[1], BAH; Husayn Marani[4], MSc; Lianne Jeffs[9], RN, PhD, FAAN; R Sacha Bhatia[1,3,5], MBA, MD

[1]Women's College Hospital Institute for Health System Solutions and Virtual Care, Women's College Hospital, Toronto, ON, Canada
[2]Department of Family and Community Medicine, University of Toronto, Toronto, ON, Canada
[3]Department of Medicine, University of Toronto, Toronto, ON, Canada
[4]Institute for Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada
[5]Women's College Research Institute, Women's College Hospital, Toronto, ON, Canada
[6]Department of Psychiatry, University of Manitoba, Winnipeg, MB, Canada
[7]Department of Psychiatry, University of Toronto, Toronto, ON, Canada
[8]Division of General Internal Medicine, St Michael's Hospital, Toronto, ON, Canada
[9]Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, ON, Canada

Corresponding Author:
Payal Agarwal, MD
Women's College Hospital Institute for Health System Solutions and Virtual Care
Women's College Hospital
76 Grenville Street
Toronto, ON, M5S 1B2
Canada
Phone: 1 416 323 6400
Email: payal.agarwal@wchospital.ca

Abstract

Background: As the increasing prevalence of type 2 diabetes mellitus has put pressure on health systems to appropriately manage these patients, there have been a growing number of mobile apps designed to improve the self-management of diabetes. One such app, BlueStar, has been shown to significantly reduce hemoglobin A_{1c} (HbA_{1c}) levels in small studies and is the first app in the United States to receive Food and Drug Administration approval as a mobile prescription therapy. However, the impact of the app across *real-world* population among different clinical sites and health systems remains unclear.

Objective: The primary objective of this study was to conduct a pragmatic randomized controlled trial of the BlueStar mobile app to determine if app usage leads to improved HbA_{1c} levels among diverse participants in real-life clinical contexts. We hypothesized that this mobile app would improve self-management and HbA_{1c} levels compared with controls.

Methods: The study consisted of a multicenter pragmatic randomized controlled trial. Overall, 110 participants randomized to the immediate treatment group (ITG) received the intervention for 6 months, and 113 participants randomized to the wait-list control (WLC) group received usual care for the first 3 months and then received the intervention for 3 months. The primary outcome was glucose control measured by HbA_{1c} levels at 3 months. Secondary outcomes assessed intervention impact on patient self-management, experience of care, and self-reported health utilization using validated scales, including the Problem Areas in Diabetes, the Summary of Diabetes Self-Care Activities, and the EuroQol-5D. Intervention usage data were collected directly from the app.

Results: The results of an analysis of covariance controlling for baseline HbA_{1c} levels did not show evidence of intervention impact on HbA_{1c} levels at 3 months (mean difference [ITG−WLC] −0.42, 95% CI −1.05 to 0.21; *P*=.19). Similarly, there was no intervention effect on secondary outcomes measuring diabetes self-efficacy, quality of life, and health care utilization behaviors. An exploratory analysis of 57 ITG participants investigating the impact of app usage on HbA_{1c} levels showed that each additional day of app use corresponded with a 0.016-point decrease in participants' 3-month HbA_{1c} levels (95% CI −0.03 to −0.003). App

usage varied significantly by site, as participants from 1 site logged in to the app a median of 36 days over 14 weeks (interquartile range [IQR] 10.5-124); those at another site used the app significantly less (median 9; IQR 6-51).

Conclusions: The results showed no difference between intervention and control arms for the primary clinical outcome of glycemic control measured by HbA_{1c} levels. Although there was low usage of the app among participants, results indicate contextual factors, particularly site, had a significant impact on overall usage. Future research into the patient and site-specific factors that increase app utilization are needed.

Trial Registration: Clinicaltrials.gov NCT02813343; https://clinicaltrials.gov/ct2/show/NCT02813343 (Archived by WebCite at https://clinicaltrials.gov/ct2/show/NCT02813343)

(JMIR Mhealth Uhealth 2019;7(1):e10321) doi:10.2196/10321

KEYWORDS

mobile apps; diabetes mellitus, type 2; self-management; blood glucose self-monitoring; randomized controlled trial; pragmatic clinical trial

Introduction

The worldwide burden of type 2 diabetes mellitus (T2DM) continues to increase, with almost 9% of the global population expected to have T2DM by 2035 [1]. The increasing prevalence of T2DM will put pressure on health systems to appropriately manage these patients to avoid diabetic complications. Optimizing self-management of glycemic control and other risk factors in conjunction with pharmacologic therapy may be an efficient way to improve patient outcomes [2-5]. Although self-management is traditionally offered through in-person educational programs, this is resource intensive, and advances in mobile technology provide the opportunity to deliver effective self-management support to patients that is convenient and potentially cost-effective [6-9].

There are a growing number of mobile apps designed to improve the self-management of T2DM patients [10-12], although few have been rigorously evaluated. One diabetes management app, called BlueStar, a smartphone-enabled app that is designed to serve as a virtual coach for patients, has been shown to significantly reduce hemoglobin A_{1c} (HbA_{1c}) levels in T2DM patients, seen by primary care physicians [13]. As a result, BlueStar is the first app in the United States to be given Food and Drug Administration approval as a mobile prescription therapy [14]. Previous studies using BlueStar were small however and conducted on a relatively homogenous patient population [13]. As a result, it remains unknown whether the result of these studies would be generalizable to a diverse *real-world* population across different clinical sites. In addition, multiple studies of mobile apps for chronic diseases have highlighted the importance of contextual and implementation factors, including clinician training, integration into existing workflows, and ongoing clinician engagement with the patient as important influencers of clinical outcomes [15], yet previous studies were not designed to assess these factors.

The purpose of this study was to conduct a pragmatic randomized controlled trial of the BlueStar mobile app on T2DM patients with poorly controlled blood sugar to determine if the use of the app would lead to improved HbA_{1c} levels compared with controls in real-life clinical contexts. We hypothesized that this mobile app would improve patient self-management, and ultimately, patients with the app would have improved HbA_{1c} levels compared with controls.

Methods

Settings

Participants were recruited from 3 hospital-based diabetes education programs (DEPs) in Ontario, Canada. Most health services in Ontario, Canada, are financed through the publicly funded Ontario Health Insurance Program (OHIP), which covers medically necessary services delivered by physicians, including primary, specialty, and emergency care. Patients with T2DM typically get most of their diabetes care in short visits from family physicians who may or may not have additional multidisciplinary support. In addition, OHIP covers services provided by DEPs, which are multidisciplinary, nonphysician -led programs designed to deliver self-management education of diabetes and self-management support [16]. The 3 recruitment sites included (1) a DEP located in an urban area in a large city center (>2 million people), (2) 1 located in a midsize city in a remote area of the province (<150,000 people), and (3) 1 located in a semiurban area surrounding a large city center (<600,000 people). These sites serve a diverse range of patients including a large immigrant community, rural patients, and a large Aboriginal population. The services of these programs are complementary to primary care delivered through the patients' primary care provider (PCP) and usually do not include medication titration.

Trial Design

The study consisted of a multicenter, pragmatic randomized controlled trial with blinded outcome assessment designed to evaluate the effectiveness of the BlueStar app. A full description of the protocol has been previously published [17]. Participants with an HbA_{1c} level higher than 8.0% were recruited from the 3 DEPs, where they received support for diabetes management and randomized in a ratio of 1:1 to 2 groups: (1) immediate treatment group (ITG) or (2) wait-list control (WLC) group. The ITG received the intervention immediately for a total duration of 6 months. The WLC group received usual care for the first 3 months, at which point they received the intervention and used the app for a total of 3 months. Outcomes were measured at baseline as well as 3 and 6 months.

Participants

Participants were eligible for inclusion in the study if they met the following criteria: (1) adults aged older than 18 years; (2) obtaining care for T2DM at a participating DEP; (3) HbA_{1c} ≥8.0% (and at least 1% above the participant's target level) on most recent laboratory report within the last 3 months; (4) currently using an active email address or able and willing to obtain one; and (5) able to read the English language (self-reported). Patients were excluded if they have type 1 diabetes, were on continuous glucose monitoring, had an insulin pump, were on dialysis, pregnant, or are unable to use a computer or mobile phone because of severe mental or physical impairment.

Recruitment Process

Potential participants were identified by a clinician at each site during their regular scheduled appointments at a participating DEP. Those wanting more information met with the site coordinator and were given a brochure on the intervention and a copy of study consent form to review. If interested, the site coordinator would facilitate a phone call between the participant and study research assistant to obtain verbal consent. Participants were then randomized to 1 of 2 arms. Baseline questionnaires were completed over the phone by the research assistant at that time or within 2 weeks of randomization. Patients randomized to the ITG would meet with the site coordinator to receive the phone loaded with the BlueStar app along with a training session designed by the Ontario Telemedicine Network. Participants in the WLC group would arrange an appointment with the site coordinator in 3 months to receive their intervention and training.

Allocation

Randomization was done in a centralized fashion by the Applied Health Research Centre (AHRC) at the Li Ka Shing Knowledge Institute of St. Michael's Hospital in Toronto, Canada. Subject randomization was computer generated and stratified by site, using block sizes of 2 or 4, through REDCap [18], a Web-based electronic data entry system at the AHRC. Once the participant completed a baseline questionnaire, the centralized research assistant accessed the randomization sequence and informed the patient of their allocation to receive 1 of 2 treatments with a 1:1 randomization scheme (ITG or WLC).

Intervention

The intervention was the BlueStar mobile app, designed to act as a virtual coach for patients with T2DM. The app was preloaded onto a cellular network–connected Samsung smartphone (with all other features disabled). The phone was connected to a cellular data plan for internet connectivity and was able to connect to local Wi-Fi networks. If participants used the app without an internet connection, the information was saved and uploaded to the secure server when the phone regained an internet connection. Patients could enter information related to T2DM management into the app, including baseline health, daily blood glucose readings, exercise activity, and food intake (see Multimedia Appendix 1). The app used this information to deliver customized, evidence-based messages in real time that aim to impact motivation, behavior, and education. The messages, based on the Transtheoretical Model of Behavior Change, included educational and affirmational content to encourage sustained behavior changes. Educational messages were aligned with the American Association of Diabetes Educators 7 Standard of Care [19]. The app also facilitated the transfer of data to the user's clinician through *Smart Visit* reports that provide a clinical overview of current diabetes management including recent blood sugar readings.

Patients in the WLC group received usual diabetes care by the DEP and their primary care physician for the first 3 months of the study. To align with the principles of pragmatic trials, the usual care received was not standardized among participants [20].

Outcomes and Data Collection

The primary outcome for the trial was glucose control measured by HbA_{1c} levels at 3 months. Secondary outcomes assessed intervention impact on patient self-management, experience of care, and self-reported health utilization using patient-reported outcomes measures and patient-reported experience measures. This included patient self-efficacy measured using 2 validated scales for diabetes, the Problem Areas in Diabetes [21] and the Summary of Diabetes Self-Care Activities [22], as well as quality of life measures using the EuroQol-5D (EQ-5D) [23].

Data were collected centrally by research assistants and inputted into the REDCap database. All outcomes were assessed at 3 and 6 months. Intervention usability, an additional secondary outcome, was evaluated by an adapted version of the Mobile App Rating Scale. App utilization data were routinely collected through the app. Utilization measures include the mean number of engagements per week and the frequency of use of each feature per week.

Statistical Analysis

Patient characteristics and baseline HbA_{1c} levels were summarized using descriptive statistics, including means and SD for continuous variables and proportions for categorical variables. Data were analyzed according to the intention-to-treat principle. Primary analyses used analysis of covariance (ANCOVA) with all complete cases. A secondary analysis adjusting for study site, length of diabetes diagnosis, ethnicity, and length in DEP was also conducted. A sensitivity analysis to explore the impact of missing data was conducted by identifying all characteristics that significantly differed between those included and not included and then adding these to the primary model with the assumption that the data are at least missing at random. Self-reported health utilization data including hypoglycemic episodes, visits to a primary care physician, visits to a specialist, visits to the emergency department, and hospital admission were converted to binary outcomes (event vs no event) and analyzed using a logistic regression model.

After 6 months, HbA_{1c} levels among those participants in the ITG were compared using a paired *t* test to look for sustained impact of the intervention. App utilization data were analyzed descriptively, including frequency of use (mean uses per week) by site and feature. An exploratory analysis to assess the impact of app usage on 3-month HbA_{1c} and Problem Areas in Diabetes

(PAID) scale values was conducted using general linear models that controlled for baseline values.

Power was determined assuming an ANCOVA analysis with an estimate correlation between baseline and follow-up HbA_{1c} measurements of 0.80. The power to detect a difference of 0.7% in HbA_{1c} levels using an SD of 2% between treatment groups at 3 months is 99.7% at a significance level of 5%, based on a sample size of 255 (which assumes a dropout rate of 15% from the target sample size of 300 participants).

Results

Study Participants

Potential participants were identified based on the study criteria and enrolled in the study between June and December 2016. We invited 463 patients; of those, 145 were not interested, 74 were unreachable for follow-up, and 5 did not complete baseline questionnaires (Figure 1). Randomization was completed on 240 participants, but 17 were excluded (8 in the WLC group and 9 in ITG) because of an eligibility HbA_{1c} <8.0%. Thus, 223 participants were included in the study. On follow-up, 77.1% (172/223) of participants completed a baseline HbA_{1c} value, whereas 65.5% (146/223) completed the primary outcome (HbA_{1c} levels at 3 months). A comparison of baseline characteristics shows no significant differences among those who completed the primary outcome versus those who did not, except that nonwhite were less likely to have a 3-month HbA_{1c} value (Multimedia Appendix 2). In total, 120 participants (63 in the WLC group and 57 in the ITG) had both baseline and 3-month HbA_{1c} values completed.

Figure 1. Flowchart of enrollment. HbA_{1c}: hemoglobin A_{1c}.

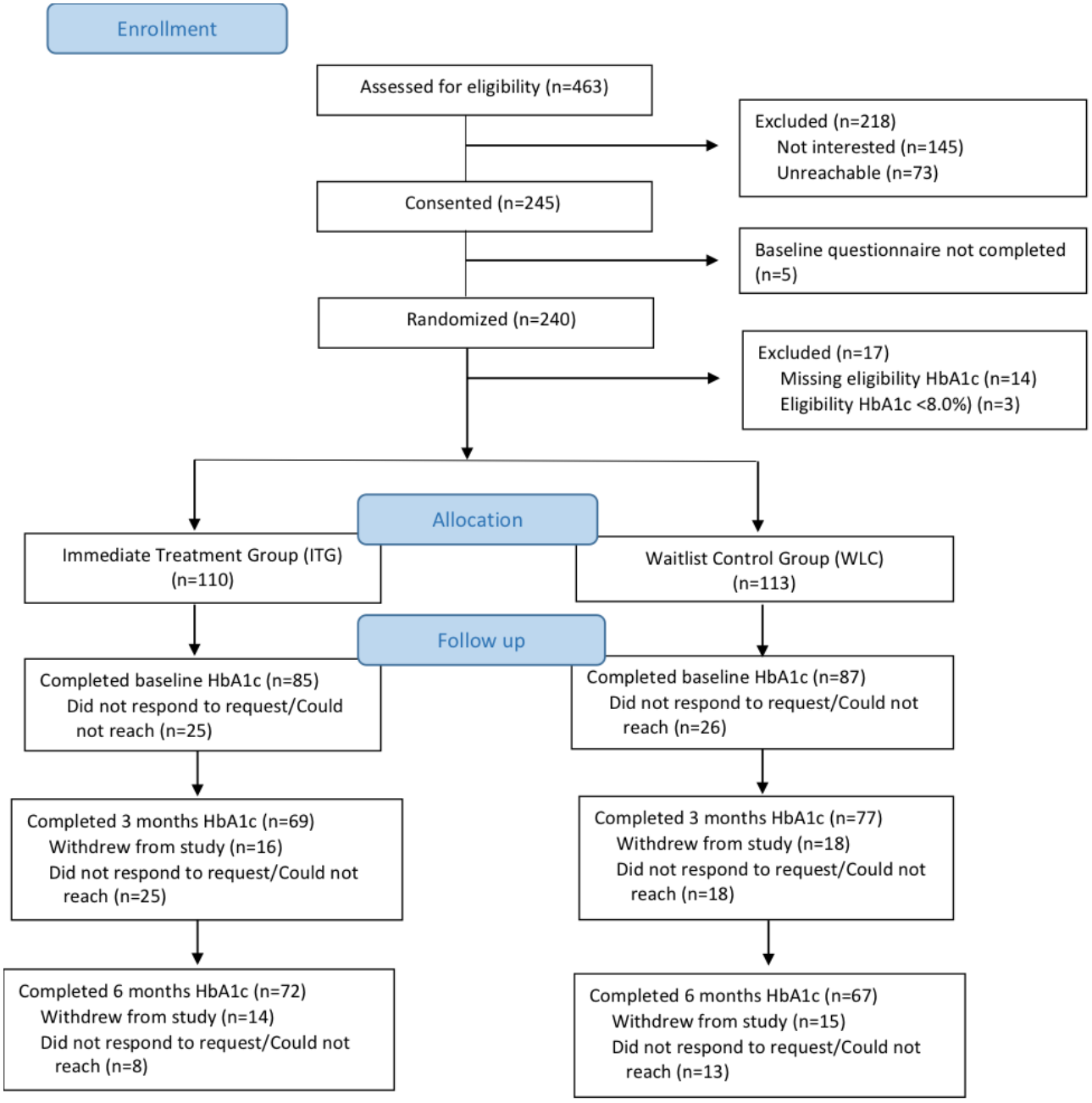


Table 1. Baseline characteristics of participants.

Variable	Immediate treatment group (n=110), n (%)	Wait-list control (n=113), n (%)	Overall (N=223), n (%)
Age[a] (years), mean (SD)	51.5 (10.6)	52.1 (10.7)	51.8 (10.7)
Sex, n (%)			
Male	61 (55.0)	55 (49.0)	116 (52.0)
Female	48 (44.0)	58 (51.0)	106 (48.0)
Not specified	1 (0.9)	0 (0.0)	1 (0.0)
Ethnicity, n (%)			
Caucasian	46 (41.8)	50 (44.3)	96 (43.1)
Non-Caucasian	64 (58.2)	60 (53.0)	124 (55.6)
Refuse to answer	0 (0.0)	2 (1.8)	2 (0.9)
Missing	0 (0.0)	1 (0.9)	1 (0.5)
Education, n (%)			
High school or less	32 (29.1)	37 (32.7)	69 (31.0)
College degree or diploma	49 (46.6)	44 (38.9)	93 (41.7)
Undergraduate university degree	11 (10.0)	14 (12.4)	25 (11.2)
Postgraduate degree	5 (4.6)	6 (5.3)	11 (5.0)
Other	4 (3.6)	8 (7.1)	12 (5.3)
Not applicable	2 (1.8)	1 (0.9)	3 (1.3)
Refuse to answer	6 (5.4)	2 (1.8)	8 (3.6)
Missing	1 (0.9)	1 (0.9)	2 (0.9)
Household income (Can $), n (%)			
<$35,000	30 (27.3)	24 (21.2)	54 (24.3)
$35,000-$50,000	10 (9.1)	24 (21.2)	34 (15.1)
>$50,000-$80,000	23 (20.9)	17 (15.0)	40 (18.0)
>$80,000-$150,000	17 (15.5)	21 (18.6)	38 (17.0)
>$150,000	6 (5.5)	5 (4.4)	11 (5.0)
Not applicable	9 (8.2)	4 (3.5)	13 (5.8)
Refuse to answer	15 (13.6)	16 (14.2)	31 (13.9)
Missing	0 (0.0)	2 (1.8)	2 (0.9)
Time since diabetes diagnosis, n (%)			
0-6 months	16 (14.6)	24 (21.2)	40 (18)
>6 months to 2 years	25 (22.7)	27 (23.9)	52 (23)
>2-5 years	26 (23.6)	13 (11.5)	39 (18)
5+ years	41 (37.3)	47 (41.6)	88 (40)
Unsure	1 (0.9)	2 (1.8)	3 (1)
Missing	1 (0.9)	0 (0.0)	1 (0.0)
Baseline value for HbA_{1c}[b], mean (SD)	8.89 (1.82)	9.03 (1.53)	8.96 (1.68)
Time in diabetes education, n (%)			
New patient	35 (31.8)	41 (363)	76 (34.1)
1-6 months	15 (13.6)	22 (19.5)	37 (16.6)
>6-12 months	22 (20.0)	19 (16.8)	41 (18.4)
1+ years	36 (32.7)	31 (27.4)	67 (30.1)

Variable	Immediate treatment group (n=110), n (%)	Wait-list control (n=113), n (%)	Overall (N=223), n (%)
Unsure	1 (0.9)	0 (0)	1 (0.4)
Missing	1 (0.9)	0 (0)	1 (0.4)
Insulin use, n (%)			
Yes	50 (45.0)	60 (53.0)	110 (49.0)
No	60 (55.0)	53 (47.0)	113 (51.0)

[a]N=222.

[b]HbA_{1c}: hemoglobin A_{1c}, N=172.

Table 1 summarizes the demographic characteristics of the study population. There were no significant differences in patient characteristics including age, gender, ethnicity, education, and household income. About 18.0% (40/223) of participants were diagnosed with T2DM within the last 6 months, whereas 39.5% (88/223) had a diagnosis of T2DM for over 5 years. The average HbA_{1c} level for the study population was 8.96% (SD 1.68) and was similar between the 2 study arms, and the use of insulin was similar between the 2 groups. Additional clinical features, including baseline medication usage and comorbidities, were similar across study arms (Multimedia Appendix 3).

Outcomes

Primary Outcome

Figure 2 shows the HbA_{1c} levels for patients in the ITG and WLC group at baseline, 3 months, and 6 months. At 3 months, the unadjusted mean HbA_{1c} values were 8.22% for the ITG and 8.41% for the WLC group. The results of an ANCOVA controlling for baseline values of 120 participants (63 WLC and 57 ITG) did not show evidence of impact on HbA_{1c} levels at 3 months for those in the ITG (mean difference [ITG-WLC] −0.42, 95% CI −1.05 to 0.21; *P*=.19). This nonsignificant difference between groups persisted after adjustment for study site, length of diabetes diagnosis, ethnicity, and length of time spent in the DEP (mean difference [ITG−WLC] −0.12, 95% CI −0.71 to 0.47).

Baseline characteristics were compared between the 120 participants included in the above model with the 103 participants who had incomplete HbA_{1c} data and were excluded to determine whether the 2 subgroups differed systematically from one another. After adjusting the main ANCOVA model for all covariates found to be associated with complete versus incomplete HbA_{1c} data (ie, site, time since diabetes diagnosis, ethnicity, antidepressant use, dyslipidemia, and obesity), the effect of treatment on 3-month HbA_{1c} levels remained statistically insignificant (least squares adjusted mean difference −0.33, 95% CI −0.99 to 0.34).

An exploratory analysis of ITG participants investigating the impact of app usage on 3-month HbA_{1c} levels while adjusting for baseline HbA_{1c} levels was conducted using an ANCOVA. Only 57 participants were complete cases and included in the regression. Each additional day of app use corresponded with a 0.016-point decrease in participant's 3-month HbA_{1c} levels (95% CI −0.03 to −0.003; *P*=.02). In other words, 25 days of additional use of the app corresponded with an HbA_{1c} reduction of 0.4%. A correlation matrix of this analysis (Multimedia Appendix 4) found a weak correlation between increased use of the exercise feature with lower HbA_{1c} levels at 3 months (ρ_s=−0.33; *P*=.01). An analysis of ITG participants, using a paired *t* test, did not show a statistically significant difference in HbA_{1c} levels between 3 and 6 months (mean difference 0.16, 95% CI −0.48 to 0.81).

Secondary Outcomes

Overall, there was no difference in patient-reported diabetes self-care behaviors (measured by PAID and Summary of Diabetes Self-Care Activities-6) or general health status (measured by EQ-5D) at 3 months between intervention arms in both the unadjusted and adjusted models (Multimedia Appendix 5). Furthermore, there was no difference in health care utilization at 3 months between groups (Table 2). An exploratory analysis of 63 ITG participants investigating the impact of app usage on PAID score levels at 3 months, adjusting for baseline scores, did not show evidence of significance (95% CI −0.28 to 0.091; *P*=.32).

Figure 2. Mean HbA_{1c} (hemoglobin A_{1c}) values for intervention and control groups from baseline to 6 months.

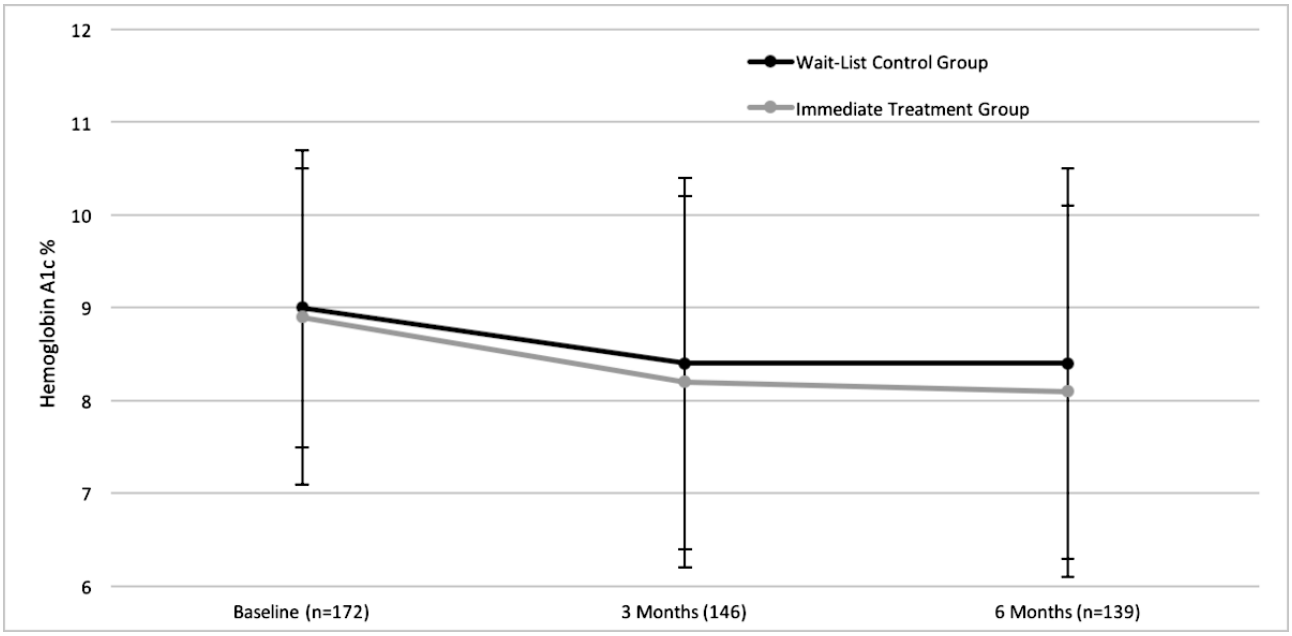


Table 2. Health service utilization.

Outcome (N=baseline/3 month)	ITG[a] (% with event)		WLC[b] (% with event)		Odds ratio (95% CI)	*P* value
	Baseline, n (%)	3 months, n (%)	Baseline, n (%)	3 months, n (%)		
Emergency deparment visits (223/139)	21 (19.0)	5 (7.5)	12 (10.6)	5 (6.8)	1.11 (0.03-0.18)	0.86
Hypoglycemic episodes (223/139)	32 (29.0)	21 (31.8)	25 (22.1)	15 (20.5)	1.80 (0.84-3.89)	0.13
Hospital admission (223/139)	23 (20.9)	9 (13.6)	16 (14.1)	4 (5.4)	2.72 (0.78-9.91)	0.11
Visit to primary care provider (222/138)	95 (86.3)	57 (86.3)	103 (91.1)	64 (87.6)	0.79 (0.29-2.19)	0.65
Visit to specialist (223/139)	78 (70.9)	37 (56.0)	70 (61.9)	46 (63.0)	0.75 (0.38-1.48)	0.40

[a]ITG: immediate treatment group.

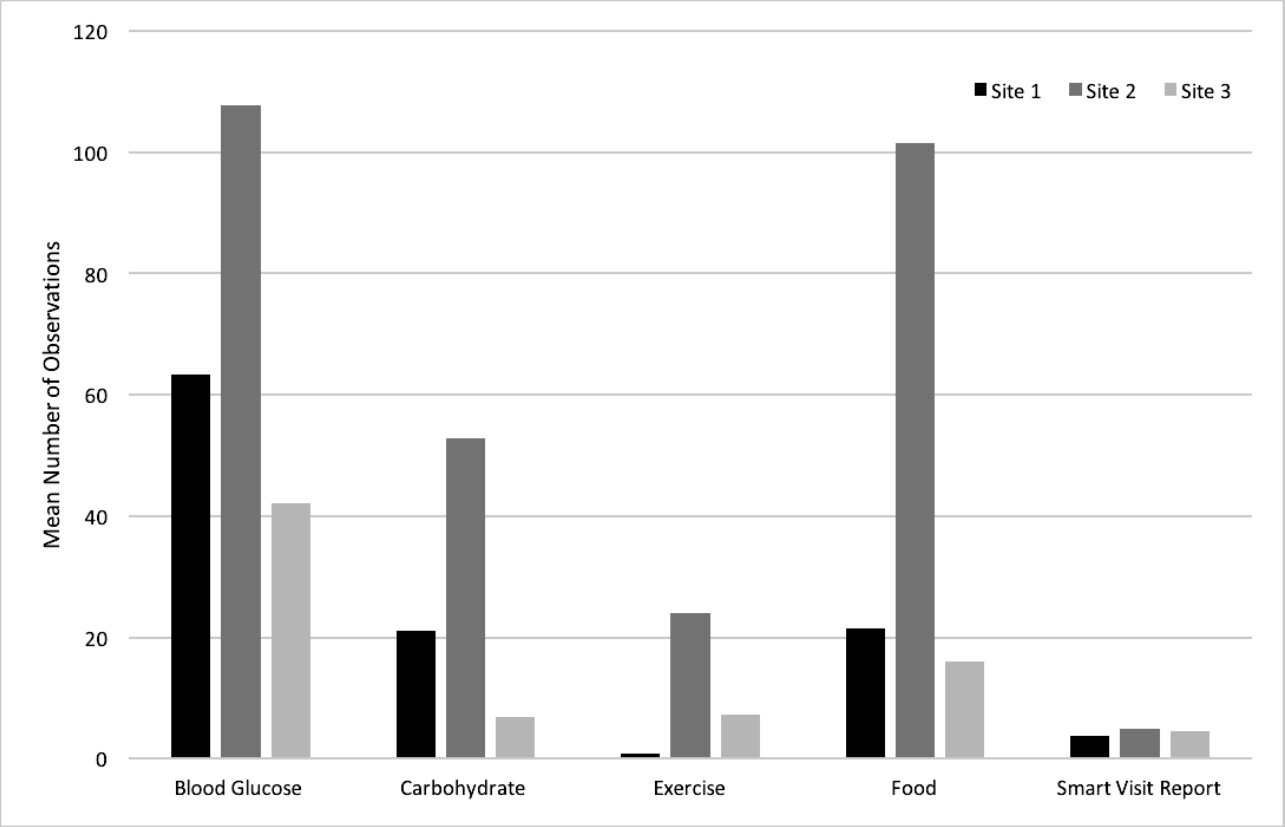
[b]WLC: wait-list control.

Mobile App Utilization and Satisfaction

Overall, there was low app utilization among ITG participants with a mean number of log-in days of 42.4 (SD 52.1) over 26 weeks, of which 46.4% (51/110) of participants used the app for 10 days or less. There was a small percentage of high users, with 18.2% (20/110) of participants using the app 100 days or more over a 182-day period. Multimedia Appendix 6 shows average number of log-in days among ITG participants over 26 weeks, showing significant decreasing mean usage over time. Blood glucose tracking was the most utilized feature with an average of 76.6 entries over 14 weeks (SD 96.59), whereas exercise tracking was the least utilized (mean 26.7 [SD 53.4]; see Figure 3). Of note, this graph also shows high variability in usage by site. Over the first 14 weeks, site 2 showed the highest number of log-in days by participants (median 36; interquartile range [IQR] 10.5-124), whereas participants from site 3 used the app significantly less (median 9; IQR 6-51). Site 1 has intermediate usage (median 17; IQR 7-72). Users with a diagnosis of diabetes in the last 6 months were the most engaged as assessed by days of log-in (median 24.5; IQR 8.7-73.5), whereas those with a diagnosis for over 5 years also had high engagement (median 18; IQR 8-86).

User ratings were completed by 105 participants to assess satisfaction with the app. Almost half of those who responded (45.7, 48/105) stated they would recommend the app to all people like them. Moreover, 41.0% (43/105) stated they would use the app 50 times or more if they continued to have access to it. About half (53.3%, 56/105) gave the app a rating of 4 to 5 stars of 5, whereas 39.0% (41/105) gave the app a rating of 3 stars. When asked if they would be willing to pay for the app, the majority of participants (55.2%, 58/105) stated they would not.

Figure 3. Mean number of observations recorded by feature and site for immediate treatment group (ITG) participants over 14 weeks.



Discussion

Principal Findings

The aim of this study was to evaluate the clinical impact of the BlueStar app for diabetes self-management in a real-world multisite implementation. The results showed no difference between intervention and control arms for the primary clinical outcome of glycemic control as measured by HbA_{1c}. Furthermore, we found no intervention effect on secondary outcomes measuring diabetes self-efficacy, quality of life, and health care utilization behaviors. Of note, there was relative low use of the app overall, with almost half of intervention group users having minimal engagement with the app. Many app features were poorly utilized, including diet and exercise tracking, which have previously shown to play an important role in T2DM self-management [10]. There was a small number of highly engaged users, and exploratory analysis suggests a correlation with app usage, and improvement in HbA_{1c} levels at 3 months analysis suggests that 25 days of usage associated with an improvement in HbA_{1c} level by 0.4%, a clinically significant change [24].

To our knowledge, this is the largest pragmatic multisite trial of evaluation of a mobile app for self-management of T2DM, and the results are in contrast to prior published studies of mobile app for diabetes self-management. These studies of mobile apps for T2DM largely consist of small single-site studies with a homogeneous population [8,25]. A meta-analysis of 10 studies of T2DM apps reported a medium reduction in HbA_{1c} level of 0.55% among those using an app, with all studies reporting some positive benefit. However, these tended to have small study populations, and 8 of the 10 studies included additional ongoing feedback from the PCP as part of the intervention [8]. Similarly, a previous study of the same mobile app, which showed significant decrease in HbA_{1c} level among intervention participants, was conducted with only 30 participants. Moreover, in that study, the intervention arm received the mobile app plus multiple follow-up interactions from the research team to the physician and patient [13]. This large multisite study likely represents a more realistic assessment of impact for a diabetes health app across a health system than smaller, higher touch, single-site studies.

Our findings suggest that when evaluating a mobile app for chronic disease management, it is important to ask not only if the app works but also in what context, for which patients, and how to promote ongoing engagement of use. Overall, there was low usage of the app among participants. However, results indicate that contextual factors, particularly site, had a significant impact on overall usage of the app. App usage overall and across features was almost twice as high among site 2 compared with site 3. Despite comprehensive implementation protocols, there were substantial differences in time spent training clinicians, time training patients, and ongoing engagement with patients between clinical test sites, with the highest use site spending the greatest time and resources on implementation. In addition, it is increasingly evident that digital health apps designed to improve chronic disease self-management require ongoing patient engagement as a key determinant of clinical impact [26-30]. Therefore, a successful implementation and evaluation of these apps require careful consideration of factors that impact patient app utilization [30]. In this study, patients with a new diagnosis of T2DM had

significantly higher usage than those who were diagnosed more than 6 months prior. Previous studies have shown patient factors including age, internal motivation, and personal values impact utilization of mobile health technologies [31]. This aligns with the results of a qualitative evaluation conducted with a subset of patients from this study. It found that perceived self-efficacy, competing priorities, and beliefs about the usefulness of virtual solutions had a significant impact on app utilization [17].

A recent systematic review of factors that impact engagement with digital health interventions highlighted the importance of both patient factors and engagement and recruitment methods [32]. Several recent studies of apps for T2DM have emphasized the importance of an implementation that includes a strong clinical endorsement and ongoing clinical support to increase overall usage [33,34]. A qualitative study of patients who dropped out of a study evaluating a self-management app for T2DM cited lack of clinician support as the primary reason for leaving. Our complementary qualitative study found participants with high app utilization identified the health care provider and/or site coordinator as a significant source of support in app adoption. These align with our quantitative findings that variation in app usage across sites was at least in part driven by variation in implementation. Future implementations of digital health apps would benefit from a clear effort to include factors that improve engagement, including a strong clinical endorsement, ongoing physician involvement, and patient reminders [35].

Limitations

Several limitations to this study warrant discussion. Importantly, the study was underpowered to detect small but potentially still important differences in HbA_{1c} levels. The studies' high dropout rate of 34.5% (77/223), while in line with prior electronic health (eHealth) studies, may have led to an underestimation of the clinical impact among participants [36,37]. There were several study design factors that likely contributed to the low app usage and lack of a detected intervention effect. Instead of downloading the app, participants were given the intervention on a second phone they used for the duration of the study in an attempt to standardize implementation by the funder. However, the use of a second phone to deliver eHealth interventions has been a noted barrier to usage in previous studies, and future mobile app evaluations would likely benefit from allowing participants to use their own smartphones when possible [28]. Given previous evidence on the benefits of strong primary care participation in diabetes self-management apps, the use of DEPs as the primary site of recruitment likely had a negative impact on enrollment, usage, and clinical impact [38,39]. Clinicians at the selected DEPs did not have regular communication with PCPs, and therefore, there was no robust pathway to report use of the app or possible treatment enhancements to the PCP. Future implementations of this, or similar apps, would likely benefit from strong primary care involvement throughout the study who can support self-management through direct treatment changes including medication titration. Finally, as discussed previously, significant variations in implementation across sites likely also had significant impact on site usage and overall ability to detect a clinical effect.

Conclusions

In this large real-world evaluation of a mobile app for diabetes self-management, we found no significant difference in HbA_{1c} levels between the intervention and control groups. Future research into the patient and site-specific factors that would increase app utilization would be warranted.

Acknowledgments

The research team would like to acknowledge Susan Franchi and Lorrie McKevitt (St. Joseph Care Group, Thunder Bay); Kevin Edwards, Ana Offenheim, and Richard Hundert (North York General Hospital); and Tina Parmar and Zita Maingi-Michire (William Osler Health System), as well as the site leads, project coordinators, and research assistants for their roles in project planning, recruitment, and data collection.

Conflicts of Interest

PA received a stipend to work on a project, unrelated to this study, from the Ontario Telemedicine Network who partially supported this evaluation. NMI received an honorarium to act as cochair of the Dissemination and Implementation Committee for the national guidelines developed by Diabetes Canada.

Multimedia Appendix 1

BlueStar mobile app interface.

[PDF File (Adobe PDF File), 363KB - mhealth_v7i1e10321_app1.pdf]

Multimedia Appendix 2

Comparison of baseline characteristics among those with and without an HbA_{1c} (hemoglobin A_{1c}) value at 3 months.

[PDF File (Adobe PDF File), 35KB - mhealth_v7i1e10321_app2.pdf]

Multimedia Appendix 3

Baseline clinical characteristics of participants.

[PDF File (Adobe PDF File), 27KB - mhealth_v7i1e10321_app3.pdf]

Multimedia Appendix 4

Spearman correlation coefficients of immediate treatment group (ITG) participants with complete cases (n=57).

[PDF File (Adobe PDF File), 29KB - mhealth_v7i1e10321_app4.pdf]

Multimedia Appendix 5

Patient-reported outcome and experience measures.

[PDF File (Adobe PDF File), 29KB - mhealth_v7i1e10321_app5.pdf]

Multimedia Appendix 6

Mean app usage among immediate treatment group (ITG) users over time.

[PNG File, 31KB - mhealth_v7i1e10321_app6.png]

Multimedia Appendix 7

CONSORT‐EHEALTH checklist (V.1.6.1).

[PDF File (Adobe PDF File), 2MB - mhealth_v7i1e10321_app7.pdf]

References

1. Guariguata L, Whiting DR, Hambleton I, Beagley J, Linnenkamp U, Shaw JE. Global estimates of diabetes prevalence for 2013 and projections for 2035. Diabetes Res Clin Pract 2014 Feb;103(2):137-149. [doi: 10.1016/j.diabres.2013.11.002] [Medline: 24630390]
2. Norris S, Lau J, Smith SJ, Schmid CH, Engelgau MM. Self-management education for adults with type 2 diabetes: a meta-analysis of the effect on glycemic control. Diabetes Care 2002 Jul;25(7):1159-1171. [Medline: 12087014]
3. Gaede P, Lund-Andersen H, Parving HH, Pedersen O. Effect of a multifactorial intervention on mortality in type 2 diabetes. N Engl J Med 2008 Feb 07;358(6):580-591. [doi: 10.1056/NEJMoa0706245] [Medline: 18256393]
4. Powers M, Bardsley J, Cypress M, Duker P, Funnell MM, Fischl AH, et al. Diabetes self-management education and support in type 2 diabetes: a joint position statement of the American Diabetes Association, the American Association of Diabetes Educators, and the Academy of Nutrition and Dietetics. Clin Diabetes 2016 Apr;34(2):70-80 [FREE Full text] [doi: 10.2337/diaclin.34.2.70] [Medline: 27092016]
5. Worswick J, Wayne SC, Bennett R, Fiander M, Mayhew A, Weir MC, et al. Improving quality of care for persons with diabetes: an overview of systematic reviews - what does the evidence tell us? Syst Rev 2013 May 07;2:26 [FREE Full text] [doi: 10.1186/2046-4053-2-26] [Medline: 23647654]
6. Fitzner K, Moss G. Telehealth--an effective delivery method for diabetes self-management education? Popul Health Manag 2013 Jun;16(3):169-177. [doi: 10.1089/pop.2012.0054] [Medline: 23216062]
7. Pal K, Eastwood SV, Michie S, Farmer AJ, Barnard ML, Peacock R, et al. Computer-based diabetes self-management interventions for adults with type 2 diabetes mellitus. Cochrane Database Syst Rev 2013;3:CD008776. [doi: 10.1002/14651858.CD008776.pub2] [Medline: 23543567]
8. Hou C, Carter B, Hewitt J, Francisa T, Mayor S. Do mobile phone applications improve glycemic control (HbA1c) in the self-management of diabetes? A systematic review, meta-analysis, and GRADE of 14 randomized trials. Diabetes Care 2016 Dec;39(11):2089-2095. [doi: 10.2337/dc16-0346] [Medline: 27926892]
9. Saffari M, Ghanizadeh G, Koenig HG. Health education via mobile text messaging for glycemic control in adults with type 2 diabetes: a systematic review and meta-analysis. Prim Care Diabetes 2014 Dec;8(4):275-285. [doi: 10.1016/j.pcd.2014.03.004] [Medline: 24793589]
10. Årsand E, Frøisland DH, Skrøvseth SO, Chomutare T, Tatara N, Hartvigsen G, et al. Mobile health applications to assist patients with diabetes: lessons learned and design implications. J Diabetes Sci Technol 2012 Sep 01;6(5):1197-1206 [FREE Full text] [doi: 10.1177/193229681200600525] [Medline: 23063047]
11. Bovbjerg M, Lee J, Wolff R, Bangs B, May MA. Pilot study for managing complex chronic care medicaid patients with diabetes using a mobile health application achieves “Triple Aim” improvement in a primary care setting. Clin Diabetes 2017 Oct;35(4):227-231 [FREE Full text] [doi: 10.2337/cd17-0006] [Medline: 29109612]

12. Whitehead L, Seaton P. The effectiveness of self-management mobile phone and tablet apps in long-term condition management: a systematic review. J Med Internet Res 2016 May 16;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
13. Quinn C, Clough SS, Minor JM, Lender D, Okafor MC, Gruber-Baldini A. WellDoc mobile diabetes management randomized controlled trial: change in clinical and behavioral outcomes and patient and physician satisfaction. Diabetes Technol Ther 2008 Jun;10(3):160-168. [doi: 10.1089/dia.2008.0283] [Medline: 18473689]
14. Waltz E. IEEE Spectrum (2014). BlueStar, the Frist Prescription-Only App URL: https://spectrum.ieee.org/biomedical/devices/bluestar-the-first-prescriptiononly-app [accessed 2018-11-12] [WebCite Cache ID 73sd2hKA7]
15. Gude W, van der Veer SN, de Keizer NF, Coiera E, Peek N. Optimizing digital health informatics interventions through unobtrusive quantitative process evaluations. Stud Health Technol Inform 2016;228:594-598. [Medline: 27577453]
16. Siminerio L, Ruppert KM, Gabbay RA. Who can provide diabetes self-management support in primary care? Findings from a randomized controlled trial. Diabetes Educ 2013;39(5):705-713. [doi: 10.1177/0145721713492570] [Medline: 23782622]
17. Desveaux L, Agarwal P, Shaw J, Hensel JM, Mukerji G, Onabajo N, et al. A randomized wait-list control trial to evaluate the impact of a mobile application to improve self-management of individuals with type 2 diabetes: a study protocol. BMC Med Inform Decis Mak 2016 Dec 15;16(1):144 [FREE Full text] [doi: 10.1186/s12911-016-0381-5] [Medline: 27842539]
18. Harris P, 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]
19. American Association of Diabetes Educators. 2010. AADE7 Self-Care Behaviors URL: https://www.diabeteseducator.org/living-with-diabetes/aade7-self-care-behaviors [accessed 2018-11-12] [WebCite Cache ID 73sdJ76l0]
20. Thorpe K, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. J Clin Epidemiol 2009 May;62(5):464-475. [doi: 10.1016/j.jclinepi.2008.12.011] [Medline: 19348971]
21. Welch G, Jacobson AM, Polonsky WH. The problem areas in diabetes scale. An evaluation of its clinical utility. Diabetes Care 1997 May;20(5):760-766. [Medline: 9135939]
22. Toobert D, Hampson SE, Glasgow RE. The summary of diabetes self-care activities measure: results from 7 studies and a revised scale. Diabetes Care 2000 Jul;23(7):943-950 [FREE Full text] [Medline: 10895844]
23. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res 2011 Dec;20(10):1727-1736 [FREE Full text] [doi: 10.1007/s11136-011-9903-x] [Medline: 21479777]
24. Food and Drug Administration. U.S. Department of Health and Human Services. 2008. Diabetes Mellitus: Developing Drugs and Therapeutic Biologics for Treatment and Prevention URL: https://www.fda.gov/downloads/Drugs/.../Guidances/ucm071624.pdf [accessed 2018-11-12] [WebCite Cache ID 73sdr2g3t]
25. McMillan K, Kirk A, Hewitt A, MacRury S. A systematic and integrated review of mobile-based technology to promote active lifestyles in people with type 2 diabetes. J Diabetes Sci Technol 2017 Dec;11(2):299-307 [FREE Full text] [doi: 10.1177/1932296816656018] [Medline: 27334301]
26. Hartin P, Nugent CD, McClean SI, Cleland I, Tschanz JT, Clark CJ, et al. The empowering role of mobile apps in behavior change interventions: the gray matters randomized controlled trial. JMIR Mhealth Uhealth 2016 Aug 02;4(3):e93 [FREE Full text] [doi: 10.2196/mhealth.4878] [Medline: 27485822]
27. Morrison L, Hargood C, Lin SX, Dennison L, Joseph J, Hughes S, et al. Understanding usage of a hybrid website and smartphone app for weight management: a mixed-methods study. J Med Internet Res 2014 Oct 22;16(10):e201 [FREE Full text] [doi: 10.2196/jmir.3579] [Medline: 25355131]
28. Goyal S, Nunn CA, Rotondi M, Couperthwaite AB, Reiser S, Simone A, et al. A mobile app for the self-management of type 1 diabetes among adolescents: a randomized controlled trial. JMIR Mhealth Uhealth 2017 Jun 19;5(6):e82 [FREE Full text] [doi: 10.2196/mhealth.7336] [Medline: 28630037]
29. Andrews S, Sperber NR, Gierisch JM, Danus S, Macy SL, Bosworth HB, et al. Patient perceptions of a comprehensive telemedicine intervention to address persistent poorly controlled diabetes. Patient Prefer Adherence 2017;11:469-478 [FREE Full text] [doi: 10.2147/PPA.S125673] [Medline: 28424543]
30. Alvarado M, Kum HC, Gonzalez Coronado K, Foster MJ, Ortega P, Lawley MA. Barriers to remote health interventions for type 2 diabetes: a systematic review and proposed classification scheme. J Med Internet Res 2017 Dec 13;19(2):e28 [FREE Full text] [doi: 10.2196/jmir.6382] [Medline: 28193598]
31. Scott A, Alore EA, Naik AD, Berger DH, Suliburk JW. Mixed-methods analysis of factors impacting use of a postoperative mHealth app. JMIR Mhealth Uhealth 2017 Feb 08;5(2):e11 [FREE Full text] [doi: 10.2196/mhealth.6728] [Medline: 28179215]
32. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Dec 15;16(1):120 [FREE Full text] [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]

33. Lie S, Karlsen B, Oord ER, Graue M, Oftedal B. Dropout from an eHealth intervention for adults with type 2 diabetes: a qualitative study. J Med Internet Res 2017 May 30;19(5):e187 [FREE Full text] [doi: 10.2196/jmir.7479] [Medline: 28559223]
34. Cui M, Wu X, Mao J, Wang X, Nie M. T2DM self-management via smartphone applications: a systematic review and meta-analysis. PLoS One 2016;11(11):e0166718 [FREE Full text] [doi: 10.1371/journal.pone.0166718] [Medline: 27861583]
35. Attwood S, Morton KL, Mitchell J, Van Emmenis M, Sutton S, VBI Programme Team. Reasons for non-participation in a primary care-based physical activity trial: a qualitative study. BMJ Open 2016 Dec 23;6(5):e011577 [FREE Full text] [doi: 10.1136/bmjopen-2016-011577] [Medline: 27217288]
36. Eysenbach G. The law of attrition. J Med Internet Res 2005 Mar 31;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
37. Donkin L, Christensen H, Naismith SL, Neal B, Hickie IB, Glozier N. A systematic review of the impact of adherence on the effectiveness of e-therapies. J Med Internet Res 2011 Aug 05;13(3):e52 [FREE Full text] [doi: 10.2196/jmir.1772] [Medline: 21821503]
38. Chomutare T, Fernandez-Luque L, Arsand E, Hartvigsen G. Features of mobile diabetes applications: review of the literature and analysis of current applications compared against evidence-based guidelines. J Med Internet Res 2011 Sep 22;13(3):e65 [FREE Full text] [doi: 10.2196/jmir.1874] [Medline: 21979293]
39. Katz R, Mesfin T, Barr K. Lessons from a community-based mHealth diabetes self-management program: "it's not just about the cell phone". J Health Commun 2012;17(Suppl 1):67-72. [doi: 10.1080/10810730.2012.650613] [Medline: 22548601]

Abbreviations

AHRC: Applied Health Research Centre
ANCOVA: analysis of covariance
DEP: diabetes education program
eHealth: electronic health
EQ-5D: EuroQol-5D
HbA_{1c}: hemoglobin A_{1c}
IQR: interquartile range
ITG: immediate treatment group
OHIP: Ontario Health Insurance Program
PAID: Problem Areas in Diabetes
PCP: primary care provider
T2DM: type 2 diabetes mellitus
WLC: wait-list control

Edited by G Eysenbach; submitted 06.03.18; peer-reviewed by C Mamillapalli, L Berard, A Gilani; comments to author 17.07.18; revised version received 08.09.18; accepted 14.09.18; published 10.01.19.

Please cite as:
Agarwal P, Mukerji G, Desveaux L, Ivers NM, Bhattacharyya O, Hensel JM, Shaw J, Bouck Z, Jamieson T, Onabajo N, Cooper M, Marani H, Jeffs L, Bhatia RS
Mobile App for Improved Self-Management of Type 2 Diabetes: Multicenter Pragmatic Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(1):e10321
URL: https://mhealth.jmir.org/2019/1/e10321/
doi:10.2196/10321
PMID:30632972

Original Paper

Capturing Daily Disease Experiences of Adolescents With Chronic Pain: mHealth-Mediated Symptom Tracking

Chitra Lalloo[1], PhD; Amos Hundert[1], BSc; Lauren Harris[1], OT; Quynh Pham[2], MSc; Fiona Campbell[3], MD; Jill Chorney[4], MD; Bruce Dick[5], PhD; Mark Simmonds[5], MD; Joseph Cafazzo[6], PEng, PhD; Jennifer Stinson[7], RN, PhD

[1]iOUCH Pain Lab, Department of Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, ON, Canada

[2]Centre for Global eHealth Innovation, Toronto General Hospital, University Health Network, Toronto, ON, Canada

[3]Department of Anesthesia and Pain Medicine, The Hospital for Sick Children, Toronto, ON, Canada

[4]Department of Anesthesia, Pain Management & Perioperative Medicine, Dalhousie University, Halifax, NS, Canada

[5]Department of Anesthesiology & Pain Medicine, University of Alberta, Edmonton, AB, Canada

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

[7]Department of Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, ON, Canada

Corresponding Author:
Chitra Lalloo, PhD
iOUCH Pain Lab
Department of Child Health Evaluative Sciences
The Hospital for Sick Children
555 University Avenue
Toronto, ON, M5G 1X8
Canada
Phone: 1 416 813 7654
Email: chitra.lalloo@sickkids.ca

Abstract

Background: Chronic pain is a common problem in adolescents that can negatively impact all aspects of their health-related quality of life. The developmental period of adolescence represents a critical window of opportunity to optimize and solidify positive health behaviors and minimize future pain-related disability and impaired work productivity. This research focuses on the development and evaluation of a smartphone-based pain self-management app for adolescents with chronic pain.

Objective: The objectives of this study were to characterize (1) the feasibility of deploying a mobile health (mHealth) app (*iCanCope*) to the personal smartphones of adolescent research participants; (2) adherence to daily symptom tracking over 55 consecutive days; (3) participant interaction with their symptom history; and (4) daily pain-related experiences of adolescents with chronic pain.

Methods: We recruited adolescents aged 15-18 years from 3 Canadian pediatric tertiary care chronic pain clinics. Participants received standardized instructions to download the *iCanCope* app and use it once a day for 55 days. Detailed app analytics were captured at the user level. Adherence was operationally defined as per the relative proportion of completed symptom reports. Linear mixed models were used to examine the trajectories of daily symptom reporting.

Results: We recruited 60 participants between March 2017 and April 2018. The mean age of the participants was 16.4 (SD 0.9) years, and 88% (53/60) of them were female. The app was deployed to 98% (59/60) devices. Among the 59 participants, adherence was as follows: low (4, 7%), low-moderate (14, 24%), high-moderate (16, 27%), and high (25, 42%). Most (49/59, 83%) participants chose to view their historical symptom trends. Participants reported pain intensity and pain-related symptoms of moderate severity, and these ratings tended to be stable over time.

Conclusions: This study indicates that (1) the *iCanCope* app can be deployed to adolescents' personal smartphones with high feasibility; (2) adolescents demonstrated moderate-to-high adherence over 55 days; (3) most participants chose to view their symptom history; and (4) adolescents with chronic pain experience stable symptomology of moderate severity.

Trial Registration: ClinicalTrials.gov NCT02601755; https://clinicaltrials.gov/ct2/show/NCT02601755 (Archived by WebCite at http://www.webcitation.org/74F4SLnmc)

(JMIR Mhealth Uhealth 2019;7(1):e11838) doi:10.2196/11838

KEYWORDS

adherence; adolescents; chronic pain; disease experience; feasibility; mHealth; self-report; smartphones; symptom monitoring; mobile phone

Introduction

Chronic pain in adolescents is a common problem that can negatively impact all aspects of health-related quality of life [1,2]. A significant proportion of these adolescents continue to experience pain that persists into adulthood [3,4]. However, the developmental period of adolescence also represents a critical window of opportunity to optimize and solidify positive health behaviors and minimize future pain-related disability and impaired work productivity [5,6].

This research focuses on the development and evaluation of a smartphone-based pain self-management app (*iCanCope*) for adolescents with chronic pain. The *iCanCope* app was developed through a phased, user-centered design approach. In phase 1A, a qualitative needs assessment study was completed with a sample of adolescents with chronic pain (n=23; age 14-18 years) and health care providers (n=7) [7]. Participants took part in focus group or individual interviews to identify their self-management needs and how an app could be designed to meet these needs. In phase 1B, a scoping review was completed to identify and characterize publicly available "pain apps" [8]. In a systematic search of the Apple, Android, Windows, and BlackBerry stores, 279 apps were identified. However, no single app was comprehensive in terms of pain self-management content. In addition, only 8.2% of apps involved a health care professional in the development process; patient engagement was limited, and no apps provided a theoretical rationale. In phase 2A, group design sessions were held with end users (adolescents with chronic pain), app designers, and members of the research team. These design sessions were intended to better understand (1) a typical "day in the life" of a young person with pain; (2) the various points when pain interfered with their function; and (3) how a pain self-management app could be designed to fit into their life. In phase 2B, a prototype app was designed by a team of professional designers and human factors specialists. The prototype then underwent iterative cycles of usability testing with a sample of 15 young people with chronic pain to ensure that it was easy to use and perceived as valuable [9]. In phase 3, a pilot randomized controlled trial (RCT) was conducted to evaluate trial feasibility with a sample of 60 adolescents. It is recommended that electronic health (eHealth) evaluations should use reasonable comparison groups rather than no-treatment or usual care. Thus, participants were randomized to receive one of two possible versions of the *iCanCope* app. Versions A and B included an identical symptom reporting function called the "Check-in." Version B included additional self-management content related to goal setting, pain coping, and social support. In the context of the pilot RCT, we sought to compare app user groups in terms of adherence to symptom tracking while offering all participants a similar study condition (ie, a smartphone app) and a comparable amount of attention from the study team. In phase 4 (future), the pilot RCT will be scaled up to a definitive trial to evaluate the effectiveness of the *iCanCope* program for improving health outcomes. In the context of the definitive RCT, we will compare app user groups in terms of health outcomes over time. These health outcome data will be captured via validated questionnaires to be administered at baseline, 2 month, and 6 month timepoints.

The *iCanCope* symptom tracker applies the principles of ecological momentary assessment (EMA), which refers to a collection of methods that gather longitudinal, real-time data from individuals in their everyday environments [10]. Use of EMA has been shown to improve data quality by minimizing the potential biases associated with retrospective self-report data (eg, memory and self-concept biases) [10,11]. Mobile administration of EMA on devices such as smartphones can markedly improve patient adherence with daily diary reports compared with paper-based approaches [12]. It can also facilitate the capture of time- and date-stamped data, provide users with multiple response options, and embed branching logic for survey questions [12,13]. Thus, EMA can provide a data-rich window into the daily experiences of individuals across a variety of backgrounds and settings. While classical EMA studies are designed to collect dense, contemporaneous data for research use, the purpose of the *iCanCope* symptom tracker is to empower adolescents to track their symptoms, visualize trends, and communicate this information with people of their choosing (eg, caregivers, health care providers).

In a systematic review of pediatric studies that applied mobile-based EMA methodologies, Heron et al [13] identified 24 unique studies published from inception to May 2016 and found that EMA can be successfully implemented with children as young as age 7. In addition, they identified gaps in the existing literature to be addressed in future pediatric EMA studies. Specifically, they recommended that (1) researchers should evaluate the feasibility of youth using their own smartphones to participate in EMA studies; (2) EMA methods should be used to obtain a more complete picture of youth's daily experiences with chronic medical conditions, including disease-related symptoms; and (3) self-report pediatric EMA measures should use pictorial response options instead of traditional Likert scales to optimize comprehension and engagement.

To begin addressing these identified knowledge gaps, this paper focuses on the symptom tracking data from the *iCanCope* phase 3 pilot RCT. Data related to intervention effectiveness will be published separately once the phase 4 trial is complete. The specific research questions (RQs) to be addressed in this paper are as follows:

> RQ1: Is it feasible for a symptom-monitoring app to be remotely deployed to the personal smartphones (iOS or Android) of adolescent research participants?
>
> RQ2: How adherent are 15-18-year olds with chronic pain to a regimen of daily symptom tracking with automated reminders over 55 consecutive days?

RQ3: Over a period of 55 days, how often do 15-18-year olds with chronic pain choose to view their history of self-report symptom data?

RQ4: What are the daily pain-related experiences of 15-18-year olds with chronic pain as per their self-report of pain intensity, pain interference, mood, physical activity, sleep quality, and energy over 55 days?

Methods

This study was approved by the locally responsible Research Ethics Boards. A 2-arm, parallel-group RCT design with 1:1 group allocation was used (Multimedia Appendix 1) [14]. As per recommendations for feasibility studies, a sample size of 20-30 participants per group was targeted [15]. Adolescents were recruited from 3 pediatric tertiary care chronic pain clinics across Canada. Individuals were eligible if they were aged 15-18 years, were diagnosed with chronic pain, were English speaking, and owned a compatible smartphone (ie, iPhone 5 or later or Android device running operating system 4.4.2 or later). Chronic pain was defined as pain that had persisted or recurred for at least 3 months [5]. Individuals were excluded if they had moderate-to-severe cognitive impairment as per their health care provider. Participants were randomized to use one of two possible versions of the *iCanCope* app for a period of 55 days. Versions A and B included an identical symptom reporting function called the "Check-in." Version B also included content related to goal setting, pain coping, and social support. The *iCanCope* app was downloadable from the Canadian Google Play (Android) or App Store (iOS). Each participant completed a guided orientation with research staff to download and learn how to use the app. This orientation was completed over Skype or over the telephone and took approximately 10-15 minutes. During orientation, participants received standardized instructions to download the *iCanCope* app onto their personal device and to sign-in with unique log-in credentials. Participants were asked to complete a symptom Check-in and access the History feature during orientation. These data (ie, first Check-in and History access) were excluded from all analyses because they were created by participants under the direction of research staff, rather than independently. In addition, participants were shown how to customize the time of the daily Check-in notification. Successful deployment was operationally defined as a participant downloading the app, logging in, and setting up his or her app profile.

Participants were instructed to complete one symptom Check-in per day for a period of 55-days following their orientation. The 55-day study period was chosen on the basis of precedent from other Web-based self-management programs for adolescents with chronic pain, which found this duration to be associated with acceptable program adherence and effectiveness [16]. The daily Check-in feature used adolescent-friendly language and pictorial response options to optimize participant engagement (Figure 1). The pain intensity was self-reported on a 0-10 numerical rating scale with the anchors "no pain" and "worst pain." Other symptom categories (pain interference, mood, physical activity, sleep quality, and energy) were captured via individual 5-point scales where a lower score indicated better function. Participants received daily push notification reminders at a time of their choice. Each study participant received a Can $15 gift card in recognition of their time and effort. In addition, each participant received a gift card valued at Can $40 as compensation for using their personal smartphone and data plan during the study.

Participants could access the History function within the app at any time (Figure 2). This function allowed participants to view all of their previous Check-in data. The interface was designed as an dynamic calendar with a transposed "heat map" where different colors correspond to different symptoms (eg, pain intensity). The app is designed such that when the users open the History section, they are shown the pain intensity-specific heat map by default. Each day on the monthly calendar represented a potential Check-in day. If a participant had completed a Check-in on a particular day, that calendar day would be filled with color. More severe ratings were denoted by a darker color shade, which enabled users to examine patterns at the macro calendar-month level. Participants could click on any colored box within the calendar to view their exact numerical rating for that symptom. Furthermore, they could switch between different symptom categories using central filter buttons. The home screen of the app featured a central "banner" that displayed revolving messages to the user, such as "welcome home." Users could swipe on the banner to generate a new message. A reminder related to the History function was one of the revolving banner messages periodically displayed to all users. However, participants did not receive any push notification reminders specific for the History function.

Detailed app usage analytics were captured at the individual level. Stata Version 15 (StataCorp LLC) software was used for all analyses [17]. The study team was equipped to centrally track any technical issues encountered during app deployment. User adherence was operationally defined as the relative proportion of symptom Check-ins that were completed over the 55-day study period: "low adherence,"≤24% (<13/55 reports); "low-moderate adherence," 25%-49% (14/55 to 27/55 reports); "high-moderate adherence," 50%-75% (28/55 to 41/55 reports); "high adherence," 76%-100% (42/55 to 55/55 reports). Descriptive statistics were used to summarize the data for RQs 1-3, broken down by an assigned version of the *iCanCope* app. These data were analyzed to assess measures of central tendency (mean, median) and dispersion (SD, interquartile range). For RQ4, linear mixed models using an independent covariance structure and allowing for random slope and intercept were used to separately examine trajectories of each daily symptom over 55 days [18]. The estimated overall and user-level regression lines over the study period were plotted.

Figure 1. Example screenshots of the *iCanCope* daily symptom Check-in. From left: introductory screen; lowest anchor of pain intensity scale; mid-anchor of physical; activity scale; high-anchor of energy scale.

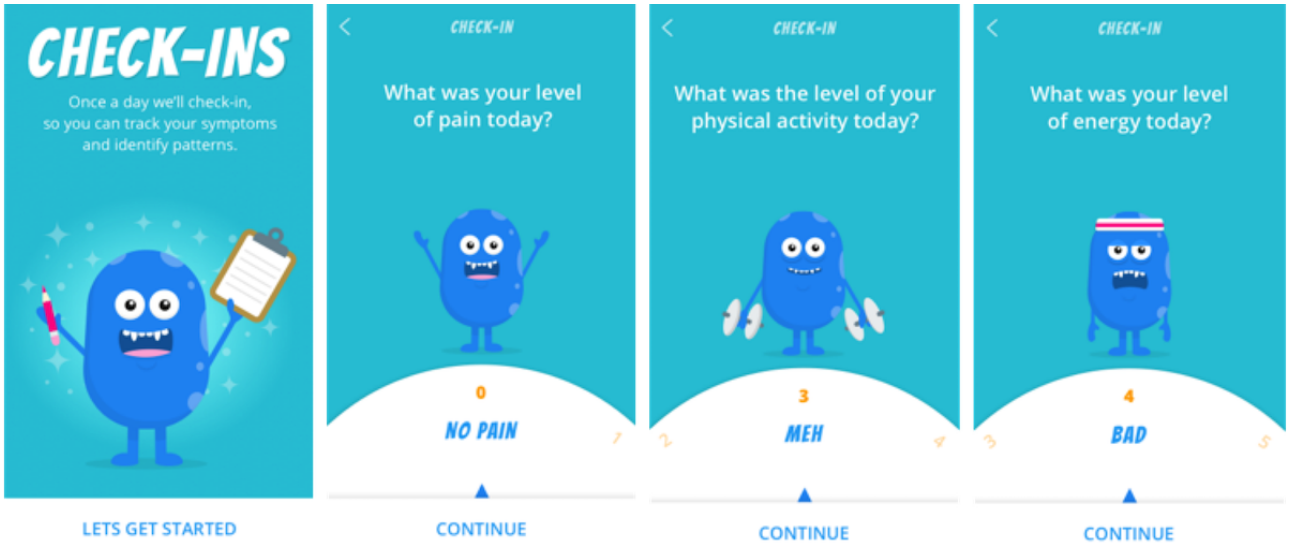
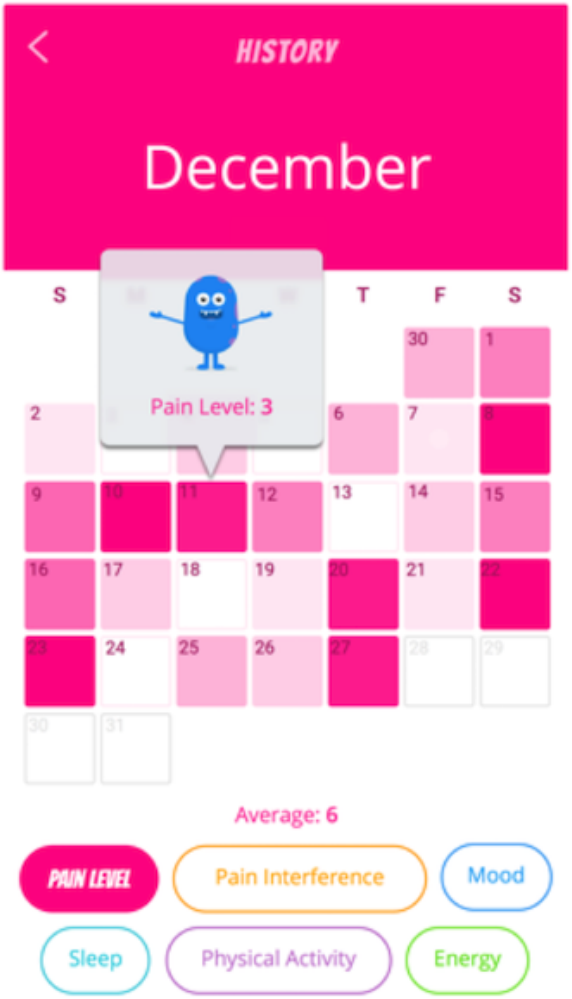


Figure 2. Example screenshot of the *iCanCope* History feature.



Results

Participants

This study was approved by the locally responsible Research Ethics Boards. A sample of 60 adolescents was recruited between March 13, 2017, and April 4, 2018; participants were recruited from clinics in Edmonton, Alberta (18, 30%); Halifax, Nova Scotia (15, 25%); and Toronto, Ontario (27, 45%). The mean age of participants was 16.4 (SD 0.9) years, and 88% (55/60) participants were female. Of the 60 participants, the majority (45, 75%) were iPhone users, with the remainder (15, 25%) being Android users; the breakdown of the *iCanCope* app assignment was as follows: version A (28, 47%) and version B (32, 53%). Table 1 shows additional participant demographic information.

Research Question 1: Feasibility of the App Deployment

No technical issues were encountered during deployment. The *iCanCope* app was successfully deployed to 98% (59/60) devices. The single participant who did not receive the app completed the telephone orientation. However, this participant did not complete the steps required for app setup and also failed to log-in throughout the study. No technical issues were noted in this case. The analytics data presented for RQs 2-4 are drawn from 59 participants.

Research Question 2: Participant Adherence to Regimen of Daily Symptom Tracking

The mean number of completed daily Check-ins was 36.0 (SD 13.9) for version A participants and 33.8 (SD 13.6) for version B participants. As per the operational definitions of adherence, version A participants (n=27) were distributed as follows: low (2, 7%), low-moderate (5, 19%), high-moderate (7, 26%), and high (13, 48%). Version B participants (n=32) were distributed as follows: low (2, 6%), low-moderate (9, 28%), high-moderate (9, 28%), and high (12, 38%). Figure 3 displays the total number of users who completed a Check-in as a function of time, broken down by the app version.

Research Question 3: Participant Interaction With History of Symptom Check-in Data

Overall, 83% (49/59) participants accessed the History function at least once during the 55-day study period. Figure 4 displays the breakdown of views for each symptom category within History according to assigned the app version.

The app is designed such that when users open the History section, they are shown the pain intensity heat map by default. The total view count includes users who opened the History section multiple times in the same day or filtered between different symptoms within the same viewing session.

Table 1. Demographic and chronic pain characteristics of the study sample (N=60).

Characteristic	Participants, n (%)
Age (years)	
15	11 (18)
16	23 (38)
17	18 (30)
18	8 (13)
Type of pain[a]	
Abdominal	17 (28)
Facial	6 (10)
Headache	24 (40)
Low back	22 (37)
Musculoskeletal	18 (30)
Neuropathic	11 (18)
Pelvic	10 (17)
Other	20 (33)
Duration of pain (years)	
<1	2 (3)
1-5	32 (54)
≥5	24 (40)
Missing data	2 (3)

[a]Participants were able to report more than one type of pain.

Figure 3. Total number of users who completed a symptom Check-in as a function of time (N=59).

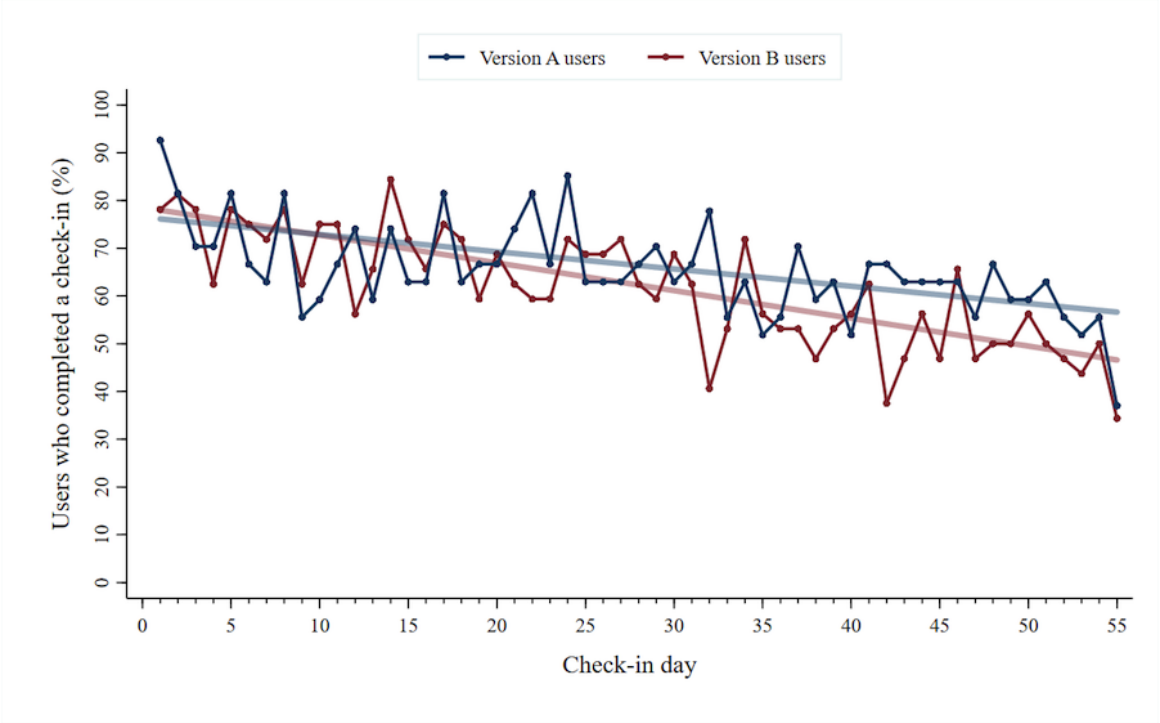
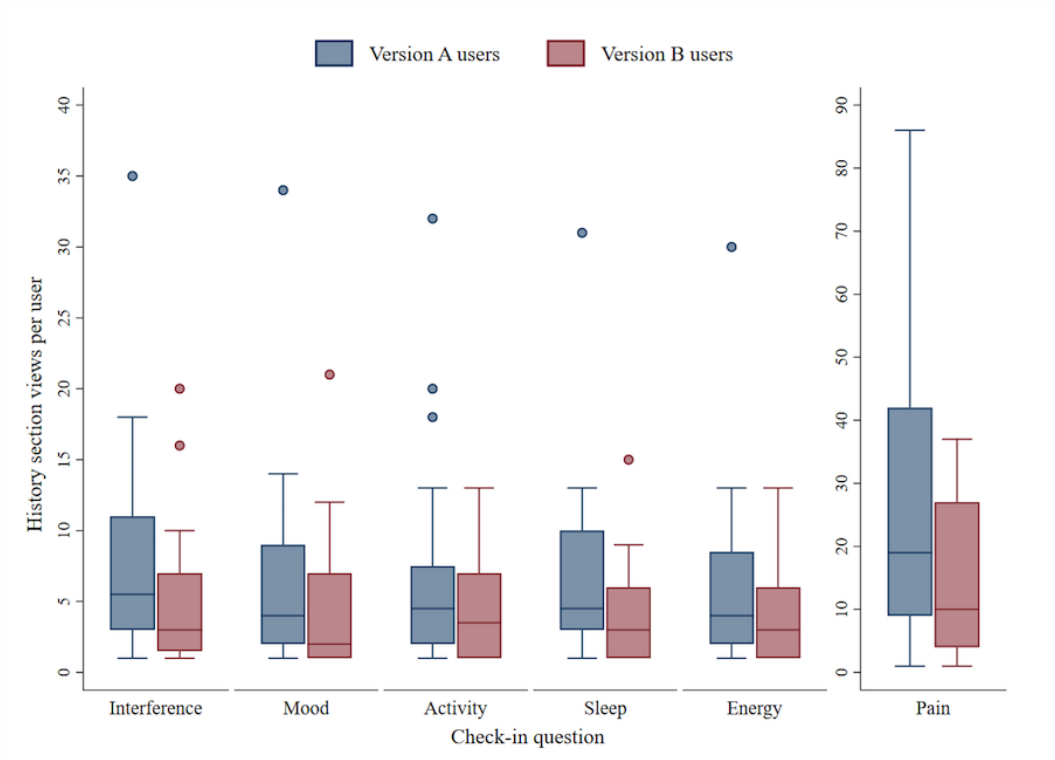


Figure 4. Interaction with the *iCanCope* History function to view symptom trends. The central line within each box denotes the median view count; the lower and upper box hinges denote the 25th and 75th percentiles, respectively; adjacent lines of the whiskers represent the lower and upper adjacent values, respectively; data points above each box represent outlier values.



Research Question 4: Daily Pain-Related Experiences of 15-18-Year Olds With Chronic Pain

A total of 2053 unique data points were analyzed for each symptom category across users. The mean reported pain intensity, captured on a 0-10 numerical rating scale, was 5.5 (SD 2.4). The mean scores for the other symptom categories, captured on individual 1-5 pictorial Likert scales, were as follows: pain interference, 2.9 (SD 1.0); mood, 2.6 (SD 1.0); physical activity, 2.8 (SD 1.1); sleep quality, 2.8 (SD 1.1); and energy, 2.9 (SD 1.0). Figures 5-10 present trajectories of each daily symptom over the 55 days.

Figure 5. Self-reported pain intensity across 55 days using the *iCanCope* daily Check-in.

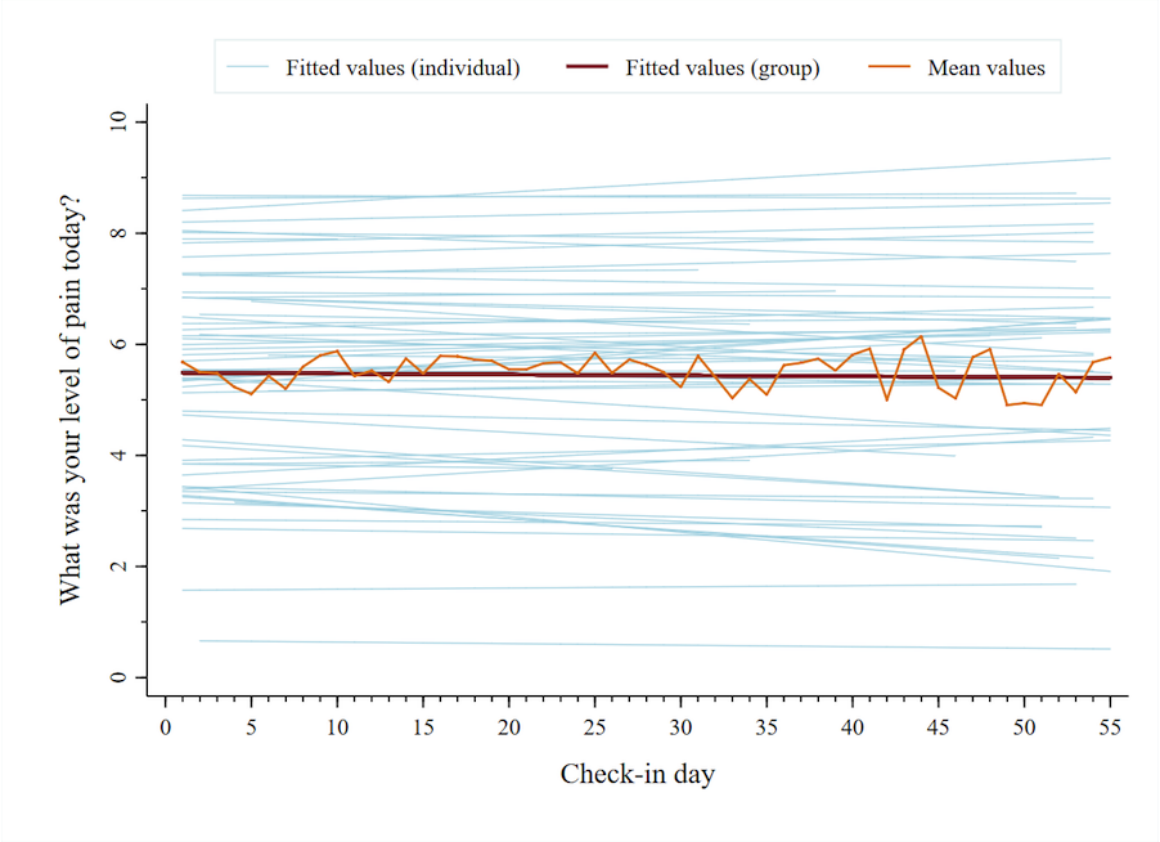


Figure 6. Self-reported pain interference across 55 days using the *iCanCope* daily Check-in.

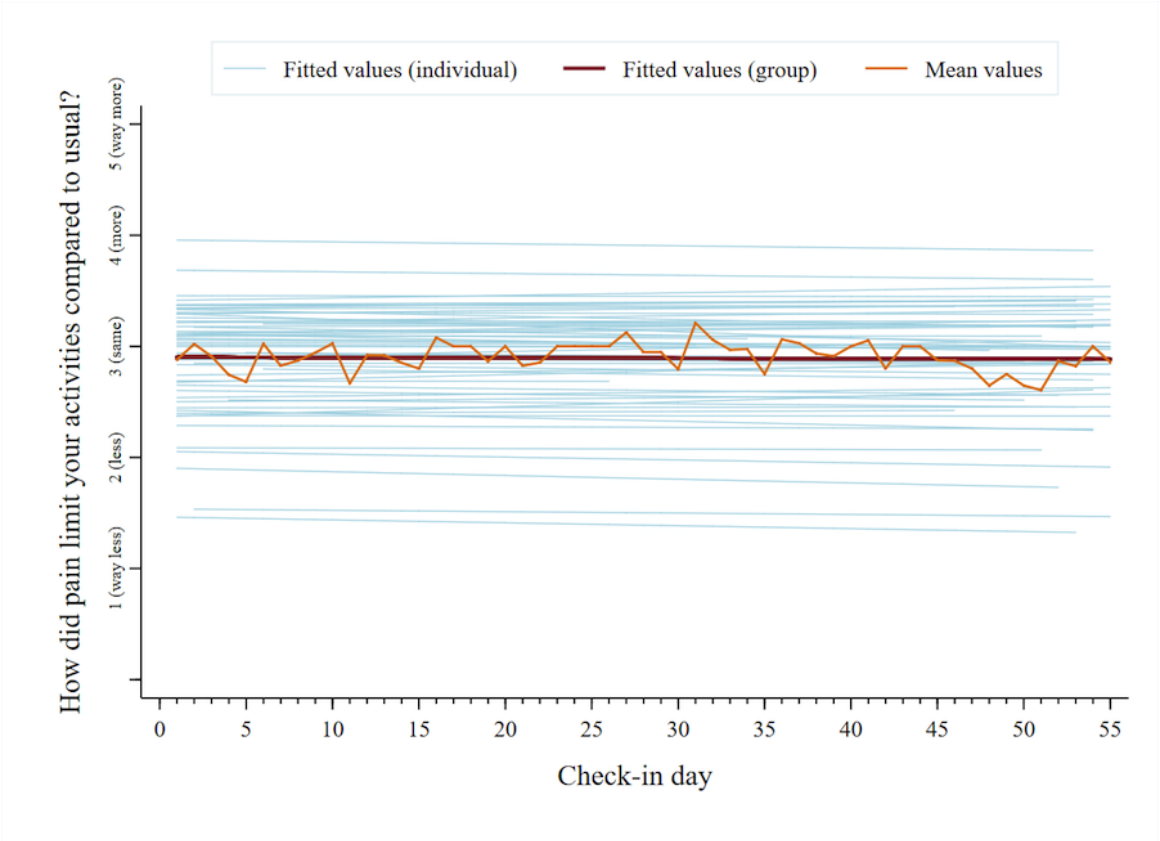


Figure 7. Self-reported mood across 55 days using the *iCanCope* daily Check-in.

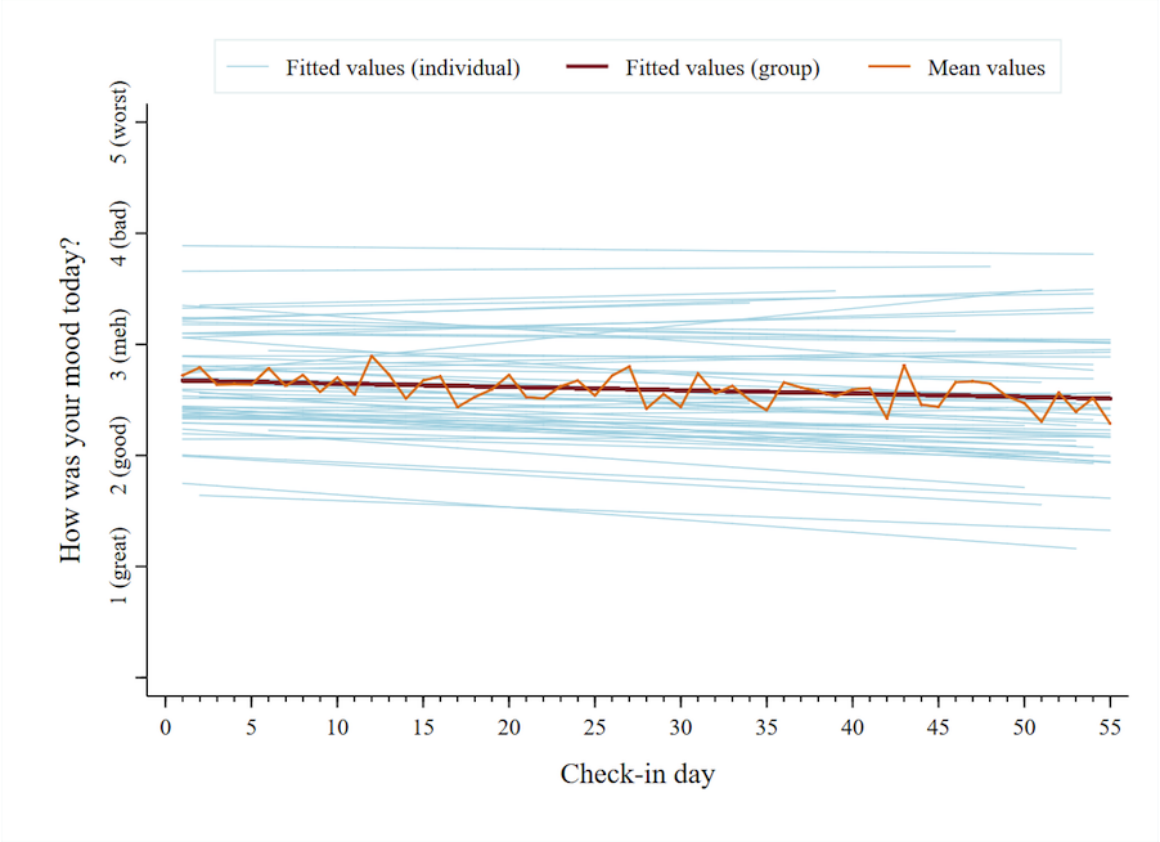


Figure 8. Self-reported physical activity across 55 days using the *iCanCope* daily Check-in.

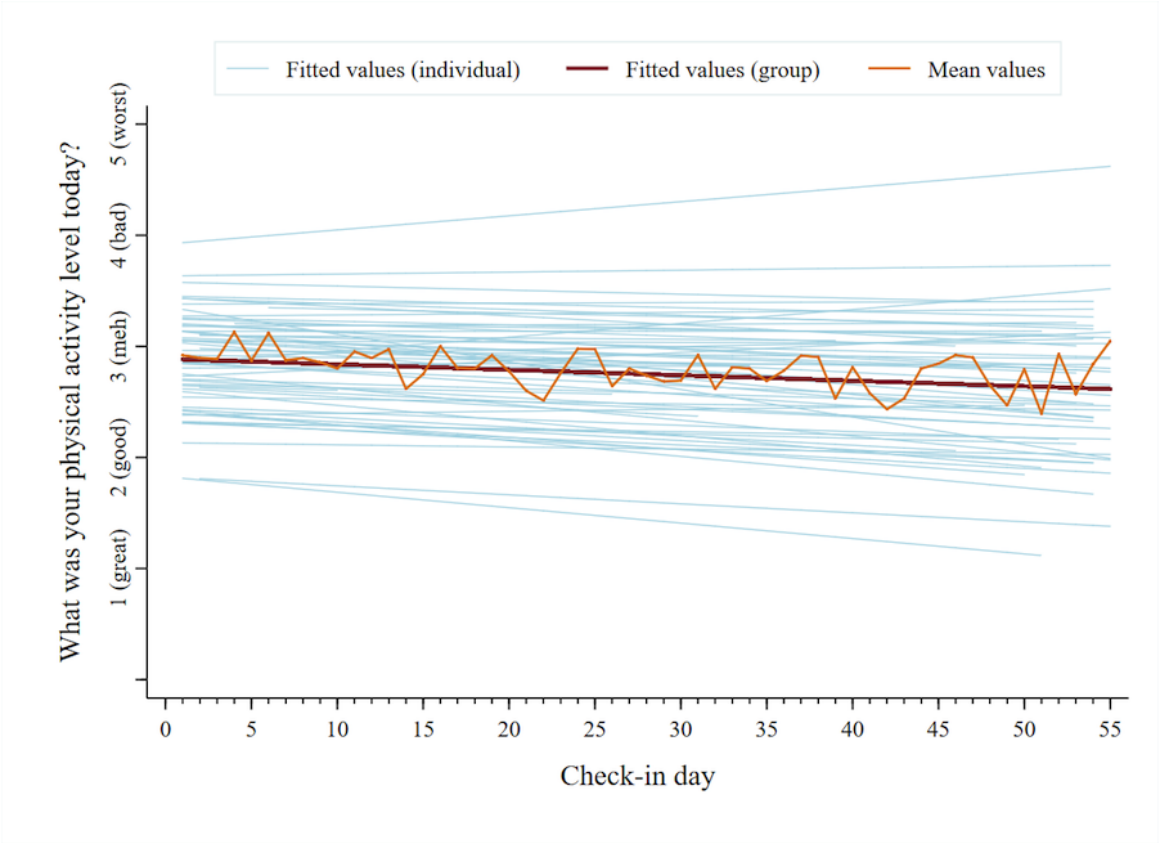


Figure 9. Self-reported sleep quality across 55 days using the *iCanCope* daily Check-in.

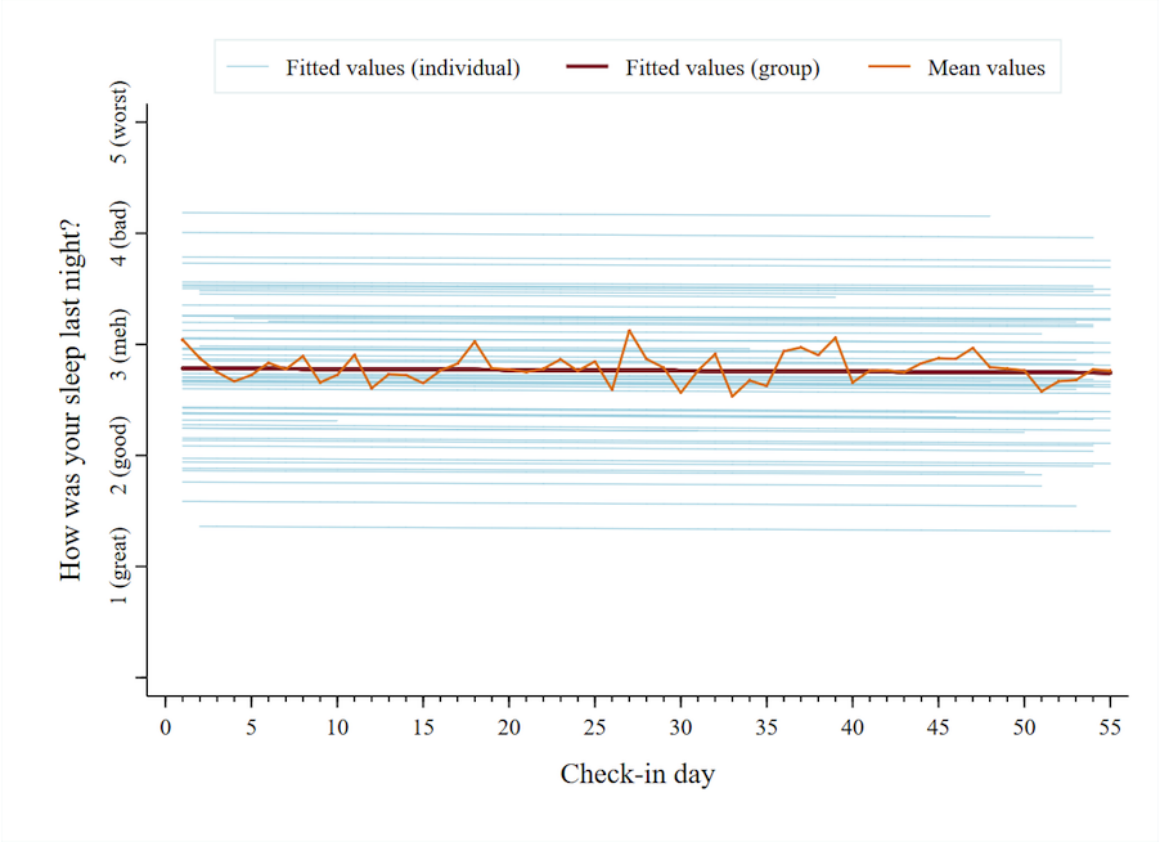
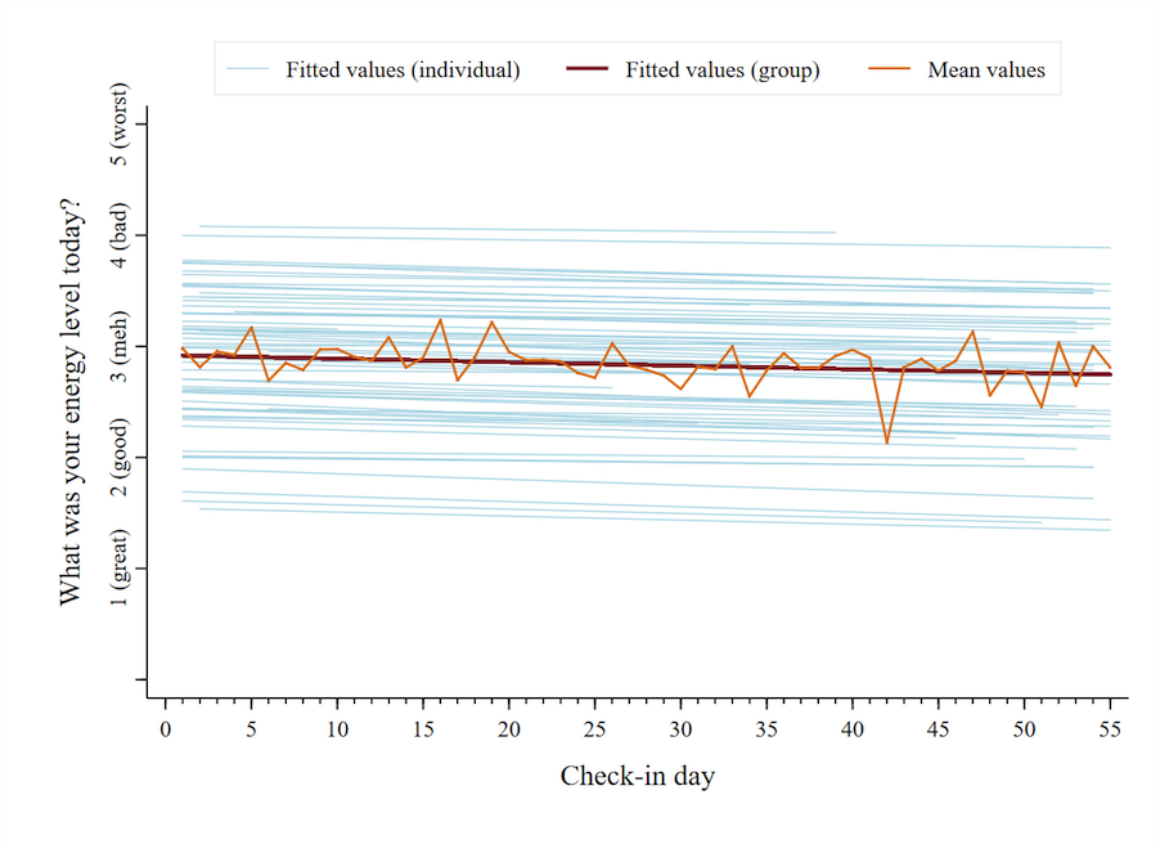


Figure 10. Self-reported energy across 55 days using the *iCanCope* daily Check-in.



Discussion

Principal Findings

Our data demonstrate that a symptom-monitoring app can be remotely deployed to the personal smartphones of adolescents using the infrastructure of the public app stores with a high degree of feasibility (98%, 59/60). The majority of participants exhibited either "high-moderate" or "high" adherence to a regimen of daily symptom tracking. Most (49/59, 83%) participants chose to view and interact with their symptom data through the History function. On average, participants who received version A of the app were more engaged with the symptom tracking feature than those who received version B. Adolescents with chronic pain reported pain intensity and pain-related symptoms of moderate severity, and these reports of their daily disease experience tended to be stable over 55 days.

Comparison With Previous Work

As per the 2 existing systematic reviews of pediatric EMA research, no single study has focused on the daily disease experiences of adolescents with chronic pain [13,19]. However, one EMA study focused on the symptomology of youth with juvenile idiopathic arthritis, a condition that is associated with persistent pain [20]. In a sample of 59 individuals (aged 8-18 years) with arthritis, participants completed a mobile EMA protocol with a sampling frequency of 3 times daily over 28 days. The mean self-reported pain intensity was 36 (SD 23) on a 0-100 scale, which was characterized in the mild-to-moderate range. A subgroup of children (13/59, 22%) reported pain intensity in the high range (>40 of 100) [20]. In a 5-year retrospective study of 2249 patients presenting at a tertiary care pediatric pain clinic in Germany, the sample was characterized by moderate-to-high functional impairment and a mean recalled pain intensity of 6.4 (SD 2.1) over the past 4 weeks [21]. In this study of Canadian adolescents from a tertiary care chronic pain setting, participants reported a mean pain intensity of 5.5 (SD 2.4) and pain-related symptoms of moderate severity that tended to be stable over time. These comparisons suggest that our group of participants is similar to the German sample in terms of chronic pain intensity and impact on the function.

Considerations for Future Pediatric Mobile Ecological Momentary Assessment Studies

Study-Issued Phone Versus Personal Phone

In the systematic review by Heron et al, 11 (46%) studies used smartphones for EMA administration [13]. In all of these studies, participants were issued a smartphone for the study duration rather than being required to download the required software onto their own device. In addition, most studies took steps to "lock down" the study devices by limiting their technical capabilities, such as blocking users from accessing other apps or disabling the phone function. This approach offers researchers with a high level of control over the EMA deployment process and the manner in which participants can interact with the study device. A potential disadvantage of requiring participants to carry a secondary device is that it may disrupt their typical routine ("ecology") and potentially influence their reports. In this study, we chose to deploy the *iCanCope* app to the personal smartphones of study participants. While ceding some control over the deployment process, this approach was intended to encourage adolescents to incorporate the app into their daily routine, including their smartphone-related habits. Given the high penetration of mobile technology in this age group, we also sought to avoid the inconvenience of participants being required to carry multiple devices (personal and study-issued) for 55 days. Indeed, recent pediatric mobile health (mHealth) studies have cautioned against the use of secondary devices, as participants frequently left their study-issued device at home and, thus, missed report notifications [22,23].

Deployment Strategies and Future Scalability

We chose to use the existing infrastructure of publicly accessible app stores (iOS, Android) rather than a mobile device management system (eg, *MobileIron*, *AirWatch*) due to the lower burden for study participants and greater potential for scalability once *iCanCope* is publicly released. By carefully codifying the process of deployment, including both electronic manuals and telephone support from research staff, we were able to install the app onto participant devices with a high success rate. Upon public release of *iCanCope*, we anticipate that app deployment will be remotely supported through Web-based manuals, instructional videos, and email technical support, rather than the individualized telephone orientations used in the pilot RCT. We will apply our operational definition of successful deployment (user downloading the app, logging in, and setting up his or her profile) to measure the effectiveness of these self-guided strategies compared with telephone orientation. In addition, we will monitor user engagement with the future public app compared with the app evaluated through the pilot RCT. Differences between these user groups will include access to monetary compensation (ie, honoraria for study participants only), direct contact with the research team for study participants only, and potential duration of usage (ie, 55 days for study participants vs unlimited access for public users).

Benchmarks for User Adherence

In a systematic review focused on pediatric adherence to mobile EMA protocols, Wen et al identified 42 unique studies that included participants from clinical (16, 38%) and nonclinical (26, 62%) settings [19]. Adherence was typically defined as the proportion of prompts to which participants responded. Among the clinical studies, the average adherence was significantly lower in studies that prompted participants 2-3 times (73.5%) or 4-5 times (66.9%) daily compared with studies with a higher sampling frequency (>6 times; 89.3%). Stone and Shiffman have recommended that researchers should aim to achieve EMA adherence rates of ≥80% [24]. However, as the *iCanCope* app aims to provide useful data to adolescents about their symptomology, rather than to collect research or clinical data, the threshold for "success" is less defined. For instance, if a particular patient experiences little or no change in their daily pain intensity, they may not perceive value in tracking it daily for 55 days. In comparison with most studies identified in the 2 recent pediatric systematic reviews, this study implemented a lower sampling frequency (once vs 2-9 times daily) over a

longer sampling duration (55 days vs 2-42 days) [13,19]. The decision regarding sampling frequency was informed by the conceptualization of *iCanCope* as a program for adolescents and based on the recommendations of patient partners during phase 1. Specifically, these collaborators recommended that we minimize the daily report burden, while also allowing users to create additional *ad hoc* reports if they wished. The decision regarding sampling duration was a function of the phase 3 pilot RCT design.

Design Considerations for Pediatric Studies

Most (63%) pediatric-focused EMA studies have reported on specific design considerations for children, including the use of youthful survey language [13]. In keeping with this trend, the *iCanCope* app design was informed by several core principles, which were developed in collaboration with patient partners during phase 1: (1) *keep it simple*; (2) *help me support my life, not just my pain*; and (3) *a safe and friendly space for me.* Based on these principles, the app was designed to include adolescent-friendly language and pictorial response options on the Check-in (see Figure 1). The specific symptoms tracked by the app were also chosen on the basis of recommendations of adolescents with chronic pain. We posit that these user-informed design choices may have contributed to the moderate-to-high adherence observed in this study and recommend this approach for future pediatric EMA studies.

Considerations for User Engagement With App Symptom Tracking

Different Versions of the iCanCope App

In this study, participants were randomly assigned to use one of the two versions of the *iCanCope* app. Group-level analysis of the daily Check-in completion illustrated that participants who received version A were more adherent than participants who received version B (see Figure 3). One possible reason for this observed difference is that version A participants received a simpler app that was focused on symptom tracking. In contrast, version B participants received a more complex app with additional self-management content. The presence of these extra features may have diverted the attention of some participants away from the symptom tracking function. It is important to note, however, that the pilot RCT group sizes (n=28 for version A and n=32 for and version B) are limited for discerning the importance of this observed trend. The future phase 4 RCT will generate a larger pool of data to more definitively examine whether there are meaningful differences in symptom tracking adherence between the groups.

Access to Historical Symptom Tracking Data

Given that most EMA studies feature a high sampling density, participants are not typically granted access to their submitted reports due to the complexity of aggregating large volumes of data into digestible output in near real time. However, as *iCanCope* collects a manageable volume of data and is meant to empower adolescents, it was important to provide users with the ability to view their symptom trends. In general, participants accessed History multiple times over the course of the study, suggesting that they found value in this feature. During the phase 1 studies, adolescents indicated interest in using the History function to communicate with their health care team during clinic appointments [9]. This user requirement was taken into consideration when designing the History feature. For instance, a calendar interface was chosen so that users could access a bird's-eye view of their symptom trends in response to common clinician queries about their pain and function since the last clinic visit. During app orientation, participants were shown how to use their History to communicate symptoms with their health care providers. Conceivably, some study participants did choose to use their data in this way during the study, although the research team did not track specific modes of use.

Limitations

Some limitations of this study should be noted. The unique methodological characteristics of our study (eg, sampling density and duration, purpose of data collection) must be considered when making direct comparisons with traditional EMA studies. The low sampling density may have failed to capture daily fluctuations in pain and related symptoms. The app Check-in was the sole source of collected symptom data and was reliant on participant self-report. No additional symptom data sources were included such as wearable accelerometers or parent report. It was not feasible to track if and how participants chose to share their symptom History with their health care providers.

Conclusions

This paper begins to address identified knowledge gaps in the field of adolescent EMA research through an mHealth app in pediatric chronic pain. We suggest that future research should extend our work by (1) evaluating the feasibility of deploying EMA apps to younger children; (2) experimenting with protocols of different sampling densities and durations; (3) triangulating self-report data with passive ambulatory data collection methods; and (4) examining other chronic disease groups.

Acknowledgments

We thank the entire *iCanCope* development and support team (Centre for Global eHealth Innovation, University Health Network) for their ongoing partnership. Joseph Charles Victor (Institute for Clinical Evaluative Sciences) provided valued advice on our data analysis methods. Participant recruitment was overseen by Tamara Do Amaral (The Hospital for Sick Children), Adena Cox and Jillian Banfield (IWK Health Centre), and Daniella San Martin-Feeney and Michelle Verrier (Stollery Children's Hospital). We also thank Carley Ouellette (McMaster University; Chronic Pain Network, Strategy for Patient-Oriented Research) for providing feedback on the manuscript draft.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist (V 1.6.2).

[PDF File (Adobe PDF File), 537KB - mhealth_v7i1e11838_app1.pdf]

References

1. Gold JI, Yetwin AK, Mahrer NE, Carson MC, Griffin AT, Palmer SN, et al. Pediatric chronic pain and health-related quality of life. J Pediatr Nurs 2009 Apr;24(2):141-150. [doi: 10.1016/j.pedn.2008.07.003] [Medline: 19268235]
2. King S, Chambers CT, Huguet A, MacNevin RC, McGrath PJ, Parker L, et al. The epidemiology of chronic pain in children and adolescents revisited: a systematic review. Pain 2011 Dec;152(12):2729-2738. [doi: 10.1016/j.pain.2011.07.016] [Medline: 22078064]
3. Brattberg G. Do pain problems in young school children persist into early adulthood? A 13-year follow-up. Eur J Pain 2004 Jun;8(3):187-199. [doi: 10.1016/j.ejpain.2003.08.001] [Medline: 15109969]
4. Fearon P, Hotopf M. Relation between headache in childhood and physical and psychiatric symptoms in adulthood: national birth cohort study. BMJ 2001 May 12;322(7295):1145 [FREE Full text] [Medline: 11348907]
5. American Pain Society. 2012. Assessment and Management of Children with Chronic Pain: A Position Statement from the American Pain Society URL: http://americanpainsociety.org/uploads/get-involved/pediatric-chronic-pain-statement.pdf [accessed 2018-11-28] [WebCite Cache ID 74GPfFl3Q]
6. Sawyer SM, Aroni RA. Self-management in adolescents with chronic illness. What does it mean and how can it be achieved? Med J Aust 2005 Oct 17;183(8):405-409. [Medline: 16225444]
7. Stinson JN, Lalloo C, Harris L, Isaac L, Campbell F, Brown S, et al. iCanCope with Pain™: User-centred design of a web- and mobile-based self-management program for youth with chronic pain based on identified health care needs. Pain Res Manag 2014;19(5):257-265 [FREE Full text] [Medline: 25000507]
8. Lalloo C, Jibb LA, Rivera J, Agarwal A, Stinson JN. "There's a Pain App for That": Review of Patient-targeted Smartphone Applications for Pain Management. Clin J Pain 2015 Jun;31(6):557-563. [doi: 10.1097/AJP.0000000000000171] [Medline: 25370138]
9. Stinson J, Lalloo C, Harris L, Isaac L, Campbell F, Brown S, et al. iCanCope with Pain™: User-centred design of a web- and mobile-based self-management program for youth with chronic pain based on identified health care needs. Pain Res Manag 2014;19(5):257-265 [FREE Full text] [Medline: 25000507]
10. Smyth JM, Stone AA. Ecological Momentary Assessment Research in Behavioral medicine. Journal of Happiness Studies 2003;4(1):35-52. [doi: 10.1023/A:1023657221954]
11. Stone AA, Broderick JE, Shiffman SS, Schwartz JE. Understanding recall of weekly pain from a momentary assessment perspective: absolute agreement, between- and within-person consistency, and judged change in weekly pain. Pain 2004 Jan;107(1-2):61-69. [Medline: 14715390]
12. Palermo TM, Valenzuela D, Stork PP. A randomized trial of electronic versus paper pain diaries in children: impact on compliance, accuracy, and acceptability. Pain 2004 Feb;107(3):213-219. [Medline: 14736583]
13. Heron KE, Everhart RS, McHale SM, Smyth JM. Using Mobile-Technology-Based Ecological Momentary Assessment (EMA) Methods With Youth: A Systematic Review and Recommendations. J Pediatr Psychol 2017 Nov 01;42(10):1087-1107. [doi: 10.1093/jpepsy/jsx078] [Medline: 28475765]
14. Eysenbach G, CONSORT-EHEALTH Group. CONSORT-EHEALTH: improving and standardizing evaluation reports of Web-based and mobile health interventions. J Med Internet Res 2011 Dec 31;13(4):e126 [FREE Full text] [doi: 10.2196/jmir.1923] [Medline: 22209829]
15. Hertzog MA. Considerations in determining sample size for pilot studies. Res Nurs Health 2008 Apr;31(2):180-191. [doi: 10.1002/nur.20247] [Medline: 18183564]
16. Palermo TM, Wilson AC, Peters M, Lewandowski A, Somhegyi H. Randomized controlled trial of an Internet-delivered family cognitive-behavioral therapy intervention for children and adolescents with chronic pain. Pain 2009 Nov;146(1-2):205-213 [FREE Full text] [doi: 10.1016/j.pain.2009.07.034] [Medline: 19695776]
17. StataCorp. Stata Statistical Software: Release 13. College Station, TX: StataCorp LP; 2013.
18. McCulloch C, Searle S, Neuhaus JM. Generalized, Linear, and Mixed Models. Hoboken, NJ, USA: John Wiley & Sons, Inc; 2000.
19. Wen CKF, Schneider S, Stone AA, Spruijt-Metz D. Compliance With Mobile Ecological Momentary Assessment Protocols in Children and Adolescents: A Systematic Review and Meta-Analysis. J Med Internet Res 2017 Dec 26;19(4):e132 [FREE Full text] [doi: 10.2196/jmir.6641] [Medline: 28446418]

20. Bromberg MH, Connelly M, Anthony KK, Gil KM, Schanberg LE. Self-reported pain and disease symptoms persist in juvenile idiopathic arthritis despite treatment advances: an electronic diary study. Arthritis Rheumatol 2014 Feb;66(2):462-469 [FREE Full text] [doi: 10.1002/art.38223] [Medline: 24504820]
21. Zernikow B, Wager J, Hechler T, Hasan C, Rohr U, Dobe M, et al. Characteristics of highly impaired children with severe chronic pain: a 5-year retrospective study on 2249 pediatric pain patients. BMC Pediatr 2012 May 16;12:54 [FREE Full text] [doi: 10.1186/1471-2431-12-54] [Medline: 22591492]
22. Goyal S, Nunn CA, Rotondi M, Couperthwaite AB, Reiser S, Simone A, et al. A Mobile App for the Self-Management of Type 1 Diabetes Among Adolescents: A Randomized Controlled Trial. JMIR Mhealth Uhealth 2017 Jun 19;5(6):e82 [FREE Full text] [doi: 10.2196/mhealth.7336] [Medline: 28630037]
23. Jibb LA, Stevens BJ, Nathan PC, Seto E, Cafazzo JA, Johnston DL, et al. Implementation and preliminary effectiveness of a real-time pain management smartphone app for adolescents with cancer: A multicenter pilot clinical study. Pediatr Blood Cancer 2017 Oct;64(10). [doi: 10.1002/pbc.26554] [Medline: 28423223]
24. Stone AA, Shiffman S. Capturing momentary, self-report data: a proposal for reporting guidelines. Ann Behav Med 2002;24(3):236-243. [doi: 10.1207/S15324796ABM2403_09] [Medline: 12173681]

Abbreviations

EMA: ecological momentary assessment
mHealth: mobile health
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 20.08.18; peer-reviewed by L McCann, K Heron; comments to author 27.09.18; revised version received 17.10.18; accepted 04.11.18; published 17.01.19.

Please cite as:
Lalloo C, Hundert A, Harris L, Pham Q, Campbell F, Chorney J, Dick B, Simmonds M, Cafazzo J, Stinson J
Capturing Daily Disease Experiences of Adolescents With Chronic Pain: mHealth-Mediated Symptom Tracking
JMIR Mhealth Uhealth 2019;7(1):e11838
URL: http://mhealth.jmir.org/2019/1/e11838/
doi:10.2196/11838
PMID:30664472

Original Paper

A Patient-Centered Mobile Health System That Supports Asthma Self-Management (breathe): Design, Development, and Utilization

Plinio Pelegrini Morita[1,2,3], BSc, MSc, PhD; Melanie S Yeung[3], MHSc; Madonna Ferrone[4], RRT, CRE; Ann K Taite[5,6]; Carole Madeley[7], RRT, CRE, MAppSC; Andrea Stevens Lavigne[7], MBA; Teresa To[8,9,10], PhD; M Diane Lougheed[5,6,10], MD, MSc; Samir Gupta[11,12], MD, MSc; Andrew G Day[5], MSc; Joseph A Cafazzo[2,3,13], PEng, PhD; Christopher Licskai[14,15], MD

[1]School of Public Health and Health Systems, Faculty of Applied Health Sciences, University of Waterloo, Waterloo, ON, Canada
[2]Institute of Health Policy, Management, and Evaluation, University of Toronto, Toronto, ON, Canada
[3]Centre for Global eHealth Innovation, Techna Institute, University Health Network, Toronto, ON, Canada
[4]Hotel Dieu Grace Healthcare, Windsor, ON, Canada
[5]Kingston General Health Research Institute, Kingston, ON, Canada
[6]Asthma Research Unit, Kingston Health Sciences Center, Kingston General Hospital Research Institute, Kingston, ON, Canada
[7]Ontario Lung Association, Toronto, ON, Canada
[8]Child Health Evaluative Sciences, The Hospital for Sick Children, Toronto, ON, Canada
[9]Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada
[10]Institute for Clinical Evaluative Sciences, Toronto, ON, Canada
[11]Division of Respirology, Department of Medicine, University of Toronto, Toronto, ON, Canada
[12]Li Ka Shing Knowledge Institute, St. Michael's Hospital, Toronto, ON, Canada
[13]Institute of Biomaterials and Biomedical Engineering, University of Toronto, Toronto, ON, Canada
[14]Western University, London, ON, Canada
[15]London Health Sciences, Victoria Hospital, London, ON, Canada

Corresponding Author:
Plinio Pelegrini Morita, BSc, MSc, PhD
School of Public Health and Health Systems
Faculty of Applied Health Sciences
University of Waterloo
200 University Avenue West
Waterloo, ON, N2L 3G1
Canada
Phone: 1 5198884567 ext 31372
Email: plinio.morita@uwaterloo.ca

Abstract

Background: Uncontrolled asthma poses substantial negative personal and health system impacts. Web-based technologies, including smartphones, are novel means to enable evidence-based care and improve patient outcomes.

Objective: The aim of this study was to design, develop, and assess the utilization of an asthma collaborative self-management (CSM) platform (*breathe*) using content based on international evidence-based clinical guidelines.

Methods: We designed and developed *breathe* as a Web-based mobile health (mHealth) platform accessible on smartphones, tablets, or desktop with user-centered design methods and International Organization for Standardization–certified quality development processes. Moreover, *breathe* was envisioned as a multifunctional, CSM mHealth platform, with content based on international clinical practice guidelines and compliant with national privacy and security specifications. The system enabled CSM (patient, provider, and *breathe*) and self-monitoring of asthma patients through (1) assessment of asthma control, (2) real-time access to a dynamic asthma action plan, (3) access to real-time environmental conditions, and (4) risk-reduction messaging. The data collection protocol collected user data for 12 months, with clinic visits at baseline and 6 and 12 months. Utilization outcomes included user interactions with the platform, user impressions, self-reported medication use, asthma symptom profile, reported peak flow measurement, and the delivery and impact of email reminders.

Results: We enrolled 138 patients with a mean age of 45.3 years to receive the *breathe* intervention. Majority were female (100/138, 72.5%), had a smartphone (92/138, 66.7%), and had a mean Asthma Control Test score of 18.3 (SD 4.9). A majority reported that *breathe* helped in the management of their asthma. Moreover, *breathe* scored 71.1 (SD 18.9) on the System Usability Scale. Overall, 123 patients had complete usage analytics datasets. The platform sent 7.96 reminder emails per patient per week (pppw), patients accessed *breathe* 3.08 times, journaled symptoms 2.56 times, reported medication usage 0.30 times, and reported peak flow measurements 0.92 times pppw. Furthermore, *breathe* calculated patients' action plan zone of control 2.72 times pppw, with patients being in the green (well-controlled) zone in 47.71% (8300/17,396) of the total calculations. Usage analysis showed that 67.5% (83/123) of the participants used the app at week 4 and only 57.7% (71/123) by week 45. Physician visits, email reminders, and aged 50 years and above were associated with higher utilization.

Conclusions: Individuals with asthma reported good usability and high satisfaction levels, reacted to *breathe* notifications, and had confidence in the platform's assessment of asthma control. Strong utilization was seen at the intervention's initiation, followed by a rapid reduction in use. Patient reminders, physician visits, and being aged 50 years and above were associated with higher utilization.

Trial Registration: ClinicalTrials.gov NCT01964469; https://clinicaltrials.gov/ct2/show/NCT01964469

(JMIR Mhealth Uhealth 2019;7(1):e10956) doi:10.2196/10956

KEYWORDS

smartphone; asthma; self report; self-management; patient compliance; telemedicine; risk reduction behavior; internet; monitoring, physiologic; mobile applications

Introduction

Background

Asthma is a common chronic disease that poses a serious global health problem. In Canada alone, asthma affects 10.8% of Canadians [1]. Globally, the Global Burden of Diseases, Injuries, and Risk Factors Study estimated that 339 million people suffer from asthma, where asthma is the most common chronic disease among children [2,3]. However, 50% of patients with asthma are uncontrolled, leading to substantial personal and health system impacts [4-8]. In Canada, there are 150,000 emergency room visits and 60,000 hospitalizations triggered annually by asthma [9].

Collaborative self-management (CSM) is defined as "a system of coordinated healthcare interventions and communications for populations with conditions in which self-care efforts are significant" [10]. National and international guidelines and systematic review evidence recommend CSM, including a written asthma action plan, patient education, and regular clinical review [11-14]. CSM has been shown to substantially improve important patient and health system challenges, by reducing hospitalizations, emergency room visits, unscheduled visits to a doctor, absenteeism, nocturnal asthma symptoms, and significantly improving quality of life [14]. Moreover, a majority of patients prefer an active or collaborative role in their asthma management, particularly in the context of an asthma exacerbation [15,16]. Despite this strong evidence, these patient preferences, and consistent recommendations in international guidelines [11-13], CSM continues to be available to only a minority of patients (2%-11%) [5,17]. For these reasons, asthma is a chronic disease well suited for an examination of the transformative promise of smartphone mobile health (mHealth) apps in support of CSM.

Smartphones have become ubiquitous, and mHealth apps have the potential to transform elements of chronic disease management [18,19]. mHealth apps offer new opportunities for access to care, disease specific education, monitoring and disease management, personalized goal setting, adherence reminders, and communication. Requisite to the success of smartphone apps as new tools in the management of chronic diseases are a commitment to, and evidence of, user-centered design (UCD); development; and evaluation to ensure privacy, efficacy, and safety. Beyond the requirements of good design and development, the central question of whether patient-facing asthma apps that support CSM are efficacious remains unanswered.

Objectives

We sought to design and develop a multifunctional, CSM mHealth platform for patients with Asthma, based on clinical content from international evidence-based guidelines, following a UCD process and then evaluate its utilization to inform iterative product improvement.

Methods

Overview

The *breathe* development program was structured in 2 main phases: (1) the design and development process for building the *breathe* mHealth platform including architecture, design, platform content, functional elements, user experience, and utilization (University Health Network REB 12-0102-AE and 12-0102-AE_Amendment) and (2) an evaluation of the patient outcomes by randomized controlled trial (RCT; Western University HSREB 102842, Queens University HSREB 6007261, ClinicalTrials.gov NCT01964469) and by a population-based cohort study. The utilization data reported in this manuscript are derived from the intervention (*breathe*) arm of the RCT. The RCT comparing conventional best practice plus the *breathe* platform with conventional best practice has been completed, and the main results are published in abstract form [20]. The focus of this manuscript is to share the design and development of the *breathe* platform, *breathe* utilization,

and the user experience. The results of the RCT will be published in an upcoming manuscript.

Development Specifications of breathe

Specifications were developed collaboratively with Canada Health Infoway and included (1) a user-centered Web-based asthma self-management platform available on any Web-enabled device including mobile phone browsers and standard Web browsers on laptop, desktop, and tablet to ensure equitable access of the app; (2) patient access to their personal health information and electronic health (eHealth) records through connectivity with TELUS health space, which was a localized version of Microsoft HealthVault (Web-based personal health record developed by Microsoft); (3) alignment with national and provincial clinical and eHealth priorities, as per the Canadian Thoracic Society (CTS), Ontario Lung Association (OLA), eHealth Ontario (a provincial agency tasked with the implementation of Ontario's public Electronic Health Record System), and the Ontario Ministry of Health and Long-Term Care; and (4) scalability to the provincial level and ability to be leveraged by other jurisdictions within Canada. Evidence-based best practices from the CTS Asthma guidelines [11] and the Global Initiative for Asthma guidelines [12] guided clinical content development.

The Development Team

breathe was developed by the Centre for Global eHealth Innovation at the University Health Network in collaboration with clinicians; researchers; and scientists from Western University, Queen's University, Hospital for Sick Children, and the University of Toronto. The Centre is certified under International Organization for Standardization 13485, an international quality management system, to ensure the safety and quality of innovations. The mHealth platform development was guided by a 16-member interdisciplinary steering committee including asthma expert respirologists, certified asthma educators, population health scientists, knowledge translation experts, and eHealth experts. These experts were informed by 4 working groups: benefits evaluation, technical, consumer engagement (patients with asthma), and clinical. Working groups comprised a few members of the steering committee, along with additional individuals who contributed specific expertise such as consumers (patients with asthma), information technology professionals, and clinicians.

The Design Process

breathe (Figure 1) was designed using UCD methods [18,21,22], ensuring that the input and requirements of final users of the technology (patients, caregivers, and physicians) were included in the design process. The iterative UCD process included 11 interviews and 5 usability testing and cognitive walk-through cycles [21,22]. The semistructured interviews were conducted with representative end users (adults who have asthma) to test assumptions related to the use of a monitoring system as an intervention to enhance healthy self-management behaviors and disease-related decision making. These interviews employed a qualitative, ethnographic approach. Information was gathered and organized by extracting common themes identified by the participants. This initial user research provided the necessary evidence for the conceptualization and initial prototypes of the intervention, which was subsequently used in usability testing and walk-throughs. This UCD process explored the intuitiveness of the app and identified user preferences and expectations. Multiple cycles of cognitive walk-throughs and usability testing allowed the *breathe* team to improve the design based on user feedback and observed issues, focusing on the needs of the platform's final users and avoiding the paradoxes of expertise [21]. The final design of the platform ensured that functionality was aligned with clinical needs and patient preferences and limitations.

Evaluation of the Patient Experience, Platform Usability, and Utilization

Patient Recruitment

The utilization data reported in this manuscript are derived from the intervention (*breathe*) arm of the RCT designed to evaluate patient outcomes [20]. Participants were recruited from 6 primary care and 2 specialty asthma clinics in Ontario, Canada. A convenience sample of patients was self-identified after viewing posters in the clinic or invited to participate by clinic staff. The participating clinics were geographically distributed—for example, North, East, Southwest, and Central Ontario—with a range of urban and rural communities. All participants randomized to *breathe* had a baseline onboarding clinic visit where they were provided with *breathe* accounts, received a brief orientation, and completed a 6- and 12-month follow-up visit.

Platform Usability, Consumer Satisfaction, and Confidence

Overall, 2 customized consumer satisfaction questionnaires and the standardized System Usability Scale (SUS) [23] were administered at 6 months and 12 months post enrollment.

Measuring Platform Utilization

breathe was designed to collect usage data (in-app analytics) to enable data-driven design and evaluation. Information flowing through the *breathe* data server was logged and used as a part of this evaluation. The *breathe* server tracked medications prescribed to patients, self-reported medication use, actual peak flow compared with personal best or normal, action plan zone of control, general access, and email notifications sent by the system. Each entry to the database was identified with a unique user ID and time stamped to enable further analysis.

Statistical Methods

The statistics reported in this manuscript are primarily descriptive. We reported counts and percentages for categorical variables as well as means and SDs for continuous variables or pseudocontinuous variables derived as means of multiple ordinal questionnaire items. We used the Wilcoxon rank-sum Test to compare the number of weeks with at least one login during the 52 weeks between groups defined by age, college education, smartphone use, and baseline Asthma Control Test (ACT) score. Age groups were defined as less than 50 years versus aged 50 years and above because it approximately divided the population in half.

Figure 1. Examples of the various features designed for breathe. The first row provides examples of the main home screen, the current zone of control that the patient is in, and environment information. The second row provides examples of the journaling feature where users can report symptoms, medication intake, and review entries. The last row has examples of the desktop version of breathe, where the zone of control review and action plans are displayed. These are not actual plans, medications, or patient data but instead, prototypes of the breathe interface.

Results

mHealth Platform Architecture of *breathe*

breathe is a Web-based mHealth platform that utilizes HTML5 and responsive design allowing a single version of the platform to be accessible on any device (smartphone, tablet, or personal computer; see Figure 2). Moreover, *breathe* interfaces with TELUS health space, where it receives up-to-date medication and peak flow ranges from the integration of clinical data repositories (electronic medical records). Furthermore, *breathe* retrieves real-time environmental conditions directly from Environment Canada, which include current and forecasted weather conditions, in addition to the Air Quality Health Index (AQHI) with relevant risk-reduction health messaging from Health Canada. The AQHI is a simple 1 to 10 scale designed to help individuals understand air quality, the impact of poor air quality on their health, and what actions to take to minimize health risks [24].

Functionality of *breathe*

The health care provider developed an *asthma app prescription* in a collaborative triad of patient, provider, and app. The health care provider determined the patient's asthma medications, their individualized action plan by zone, and peak flow ranges for control zone calculations (if applicable). The *breathe* platform did not advise on the selection of medications and did not create the action plan. This remained a physician responsibility. Integrating with TELUS health space offered patients the option to share *breathe* data with family members and other health care providers, which could be accomplished through the health space Web-based profile. The *breathe* features can be seen in Figure 1. Each of these features was designed to engage users and collect relevant data to support self-management, as described below.

Journal

The Journal feature allows patients to track daily symptoms, record reliever and controller medication usage, and log peak flow measurements. The historical review feature allows users to look back at previous journal entries and peak flow values entered.

Your Zone

The journal entries feed an integrated asthma control algorithm at the *breathe* server, based on the CTS Asthma Guidelines [11,12] that analyzes patient inputs and immediately advises the patient of their current zone of control: (1) green zone—in control, (2) yellow zone—uncontrolled, or (3) red zone—dangerously uncontrolled. The zone of control assessment is paired with the actionable recommendations from patients' personalized asthma action plan. The zone of control is dynamic, immediately updated with any new journal entries and resets after the action plan has been executed, ensuring a tailored and customized intervention to the patient [25]. Patients were notified of changes on their zone of control through the app dashboard and in the Your Zone section.

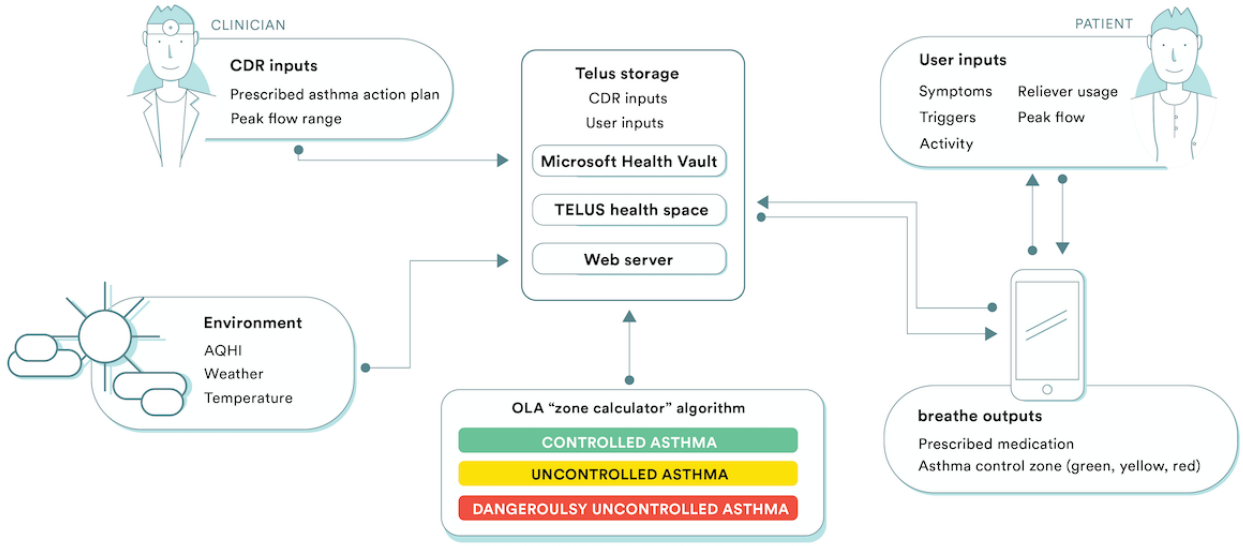
Trends

Data visualization and analysis of several trends, including identified triggers, control zone, and peak flow values, were available to users. An example of the usefulness of this feature is that trigger frequency reported back to patients may enable patient insights into which triggers to avoid in the future.

Environment

This feature provides real-time current and forecast of location-specific (based on users' input about their location) environmental conditions including temperature, humidity index, weather forecast, and the AQHI with specific poor air quality risk-reduction health messaging.

Figure 2. Architecture of the *breathe* platform (Ontario Lung Association [OLA], Air Quality Health Index [AQHI], Clinical Data Repository [CDR]). In cases where peak flow was part of the action plan, peak flow ranges were entered by the provider. Patients were responsible for entering peak flow measurements.



Account

This feature includes a variety of options including changing default (7:00 pm EST) time and email address to receive emailed medication adherence reminders and setting a location for location-specific environmental information. Email adherence reminders were automatically generated based on predefined rule-based logic including a welcome email, *check-in* emails for users not accessing the platform within 7 days, and daily adherence reminders for controller medications. There was no limit to the daily reminder emails; however, we designed the system to try to mitigate fatigue by creating approximately 30 different body messages that were randomly emailed to the user. A demonstration of how the *breathe* platform works can be found in the Canada Health Infoway website, with a detailed description of functionalities and platform capabilities.

Patient Population

A total of 344 patients were recruited into the RCT between October 31, 2012, and March 31, 2014, of whom 171 were allocated to the *breathe* intervention arm. Consent was withdrawn (n=10) or we were unable to find the patient to consent for data transfer (n=23) in 33 patients, leaving 138 patient that could be used in this analysis. Complete platform utilization data were available in 89.1% (123/138) participants, and 12-month usability and satisfaction questionnaires were available for 86.2% (119/138) participants. The majority of the 138 patients were women 72.5% (100/138), mean age 45.3 (SD 15.8) years, and 97.1% (134/138) were Caucasian. Of these participants, 66.7% (92/138) had a smartphone, and the majority 83% (76/92) reported being *comfortable* or *very comfortable* using it. Patients recruited had a mean ACT score of 18.3 (SD 4.9), suggesting well- to somewhat-well-controlled baseline asthma [26].

breathe Usability, Patient Satisfaction, and Confidence (12-Month Data)

Usability was evaluated by the SUS, a validated composite measure, which is scored from 0 to 100, with higher scores representing greater usability (Table 1). The *breathe* system scored 71.1 (SD 19.9) at 12 months indicating good usability, as defined by Bangor et al [27]. The mean of 7 ease-of-use questions scaled from 1-very difficult to 5-very easy was 4.1 (SD 0.9). A majority found *breathe* components useful and were satisfied with the design. (Table 1)

Satisfaction was evaluated using 5-point Likert scale responses, 1-strongly disagree, 3-do not know or neutral, 5-strongly agree. A total of 63.8% (74/116) of patients agreed or strongly agreed that the *breathe* app was helpful in the management of their asthma. Moreover, 65.2% (75/115) of patients were confident that the *breathe* app was correct when it presented the patient's asthma action plan zone of control. Furthermore, 49.6% (58/117) of participants agreed or strongly agreed that they would continue to use the app after the study if it remained available (Table 1).

Actual *breathe* Usage

The 123 patients in the intervention arm with utilization data accessed *breathe* 19,678 times (3.08 times per patient per week, pppw), reported symptoms in their diary 16,357 times (2.56 times pppw), reported medication use 1922 times (combined use of reliever and controller; 0.30 times pppw), and reported peak flow measurements 5864 times (0.92 times pppw). Total counts can include patients accessing the platform multiple times in the same day.

Table 1. Usability questionnaire.

Usability and user satisfaction of *breathe*	Statistics at 12 months
Satisfaction	
The *breathe* app that was provided to me by the clinic is helpful in the management of my asthma, n (%)	
Disagree or strongly disagree	21 (18.1)
Agree or strongly agree	74 (63.8)
I would continue to use the *breathe* app if it were available to me after the study, n (%)	
Disagree or strongly disagree	30 (25.6)
Agree or strongly agree	58 (49.6)
I was confident that when the *breathe* app was correct when it assessed my asthma zone of control, n (%)	
Disagree or strongly disagree	16 (13.9)
Agree or strongly agree	75 (65.2)
System Usability Scale (score range 0-100), mean (SD)	71.1 (19.9)
Evaluation of specific functional components of *breathe* (on a scale of 1-very difficult, 3-don't know or neutral, and 5-very easy)	
Ease of use: mean of 7 questions (n=119)	4.1 (0.9)
Usefulness: mean of 12 questions (n=118)	3.6 (0.9)
Design of components: mean of 12 questions reported (n=119)	4.2 (0.7)

Figure 3. Panel A: Total calculations of zone of control calculations per month of the intervention calculated from enrollment Panel B: Percentage of zone of control calculations per month of the intervention.

breathe calculated patients' action plan zone of control 17,396 times (2.72 times pppw). Patients were most often in the green zone of control (47.71% of calculations, 8300/17,396), followed by yellow zone (23.90%, 4158/17,396) and red zone (6.40%, 1110/17,396). In 22.00% (3826/17,396) of the calculations, *breathe* did not have enough information to return a zone of control back to the patient based on the programmed algorithm in the *breathe* platform (Figure 3).

breathe sent 50,939 emails (7.96 times pppw) to remind participants to take their controller medications or to return to the platform after 7 days of no usage. *breathe* did not log email responses potentially generated by the users.

Tracking patient log-ins to the platform demonstrated a fall in use within the first 4 weeks of initiation and thereafter a standard decay in usage (Figure 4), whereby 67.5% (83/123) of the participants used the platform weekly initially and only 57.7% (71/123) used the platform in week 45. Figure 4 presents our patient log-in data along with Eysenbach attrition curve [28].

Further utilization analysis demonstrated patterns of use that related to patient behavior, *breathe* functionality, or the interaction of both.

- *Time of day:* Analyzing log-ins by time of day revealed 2 periods of increased utilization (Figure 5). First, there was higher platform use between 5:00 am and 10:00 am, which corresponds to the time of the day when most patients are waking up and preparing for their day. Second, there was a dramatic spike in utilization just after 7:00 pm, the default time of day when the *breathe* system email reminders were automatically sent by the app server. This finding was sustained each month over the 12 months of the study (Figure 6).
- *Symptom reporting:* Evaluation of the *Journal* functional element within the platform revealed approximately twice as many reports of good days (a day without symptoms) compared with days with symptoms (Figure 7), which aligns with our expectations for well-controlled asthma.
- *Scheduled physician visits:* Finally, based on controller medication recording, there was an increase in platform utilization in weeks 26 and 52, corresponding to scheduled follow-up visits (Figure 8).

The post hoc analysis of patient factors that may have influenced utilization including age, education level, smartphone use, and asthma control is presented in Table 2. Only age (≥50 years) was associated with higher utilization.

Figure 4. Attrition in breathe use throughout the 12-months of the study, with Eysenbach attrition curve plotted as a reference.

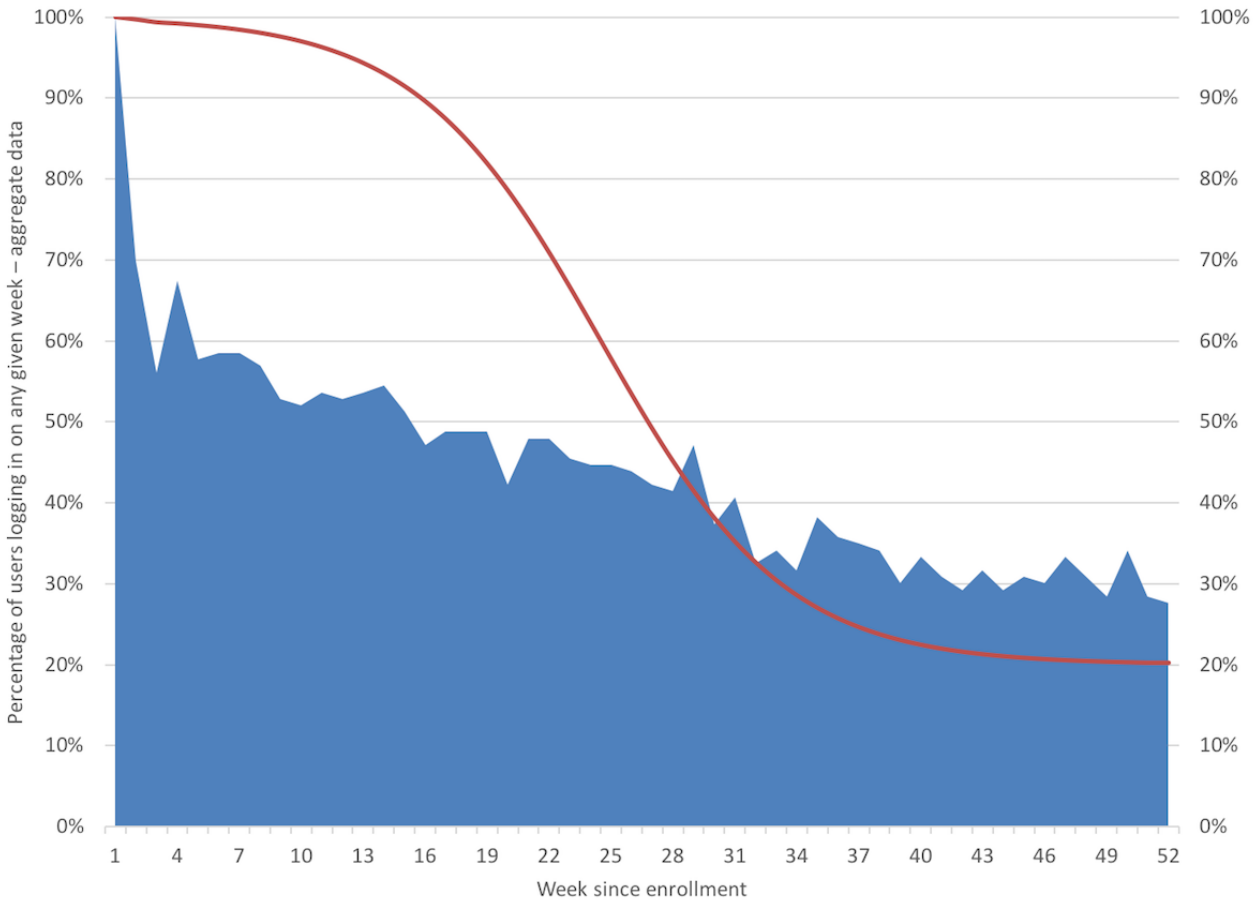


Figure 5. App use tracked by number of logins by time of day exploring the effectiveness of reminders. Note that automatic app reminders are default to send around 7:00 pm.

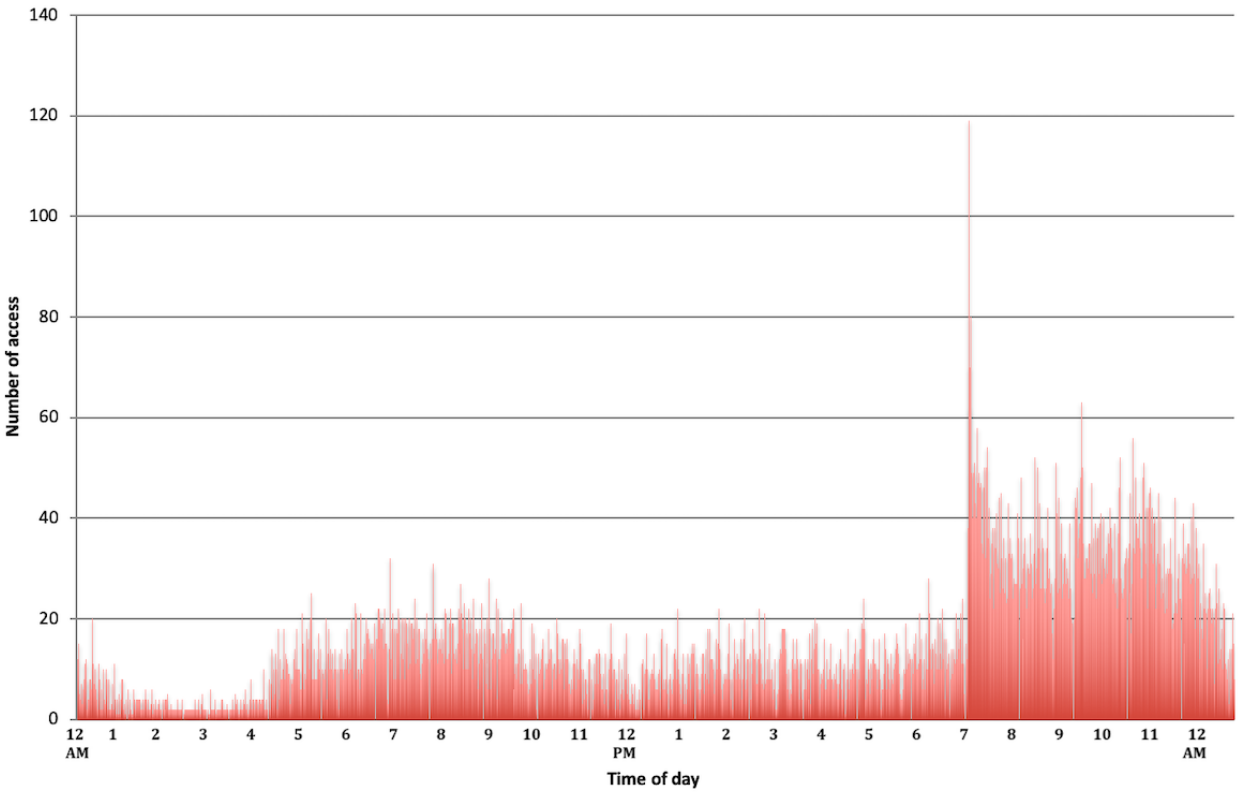


Figure 6. Sustained effect of email reminders on app use over the 12 months of intervention.

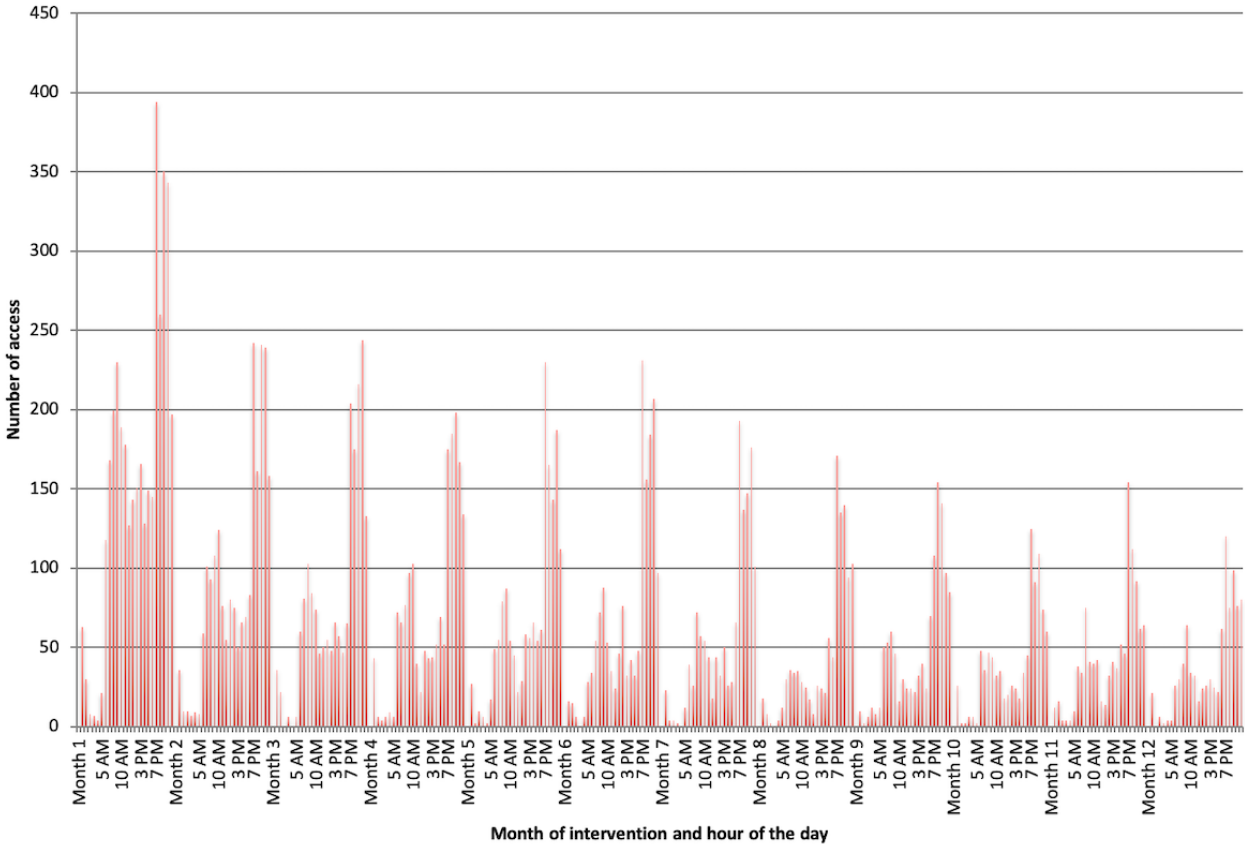


Figure 7. Panel A: Number of reported “good days” (no symptoms) and symptom episodes since enrollment. Panel B: Percentage of reported “good days” (no symptoms) and symptom episodes since enrollment.

Figure 8. Self-reported controller medication use showing the effect of clinic visits (surveillance effect) on self-reporting behavior (clinic visits were scheduled at 6 and 12 months from the beginning of the intervention).

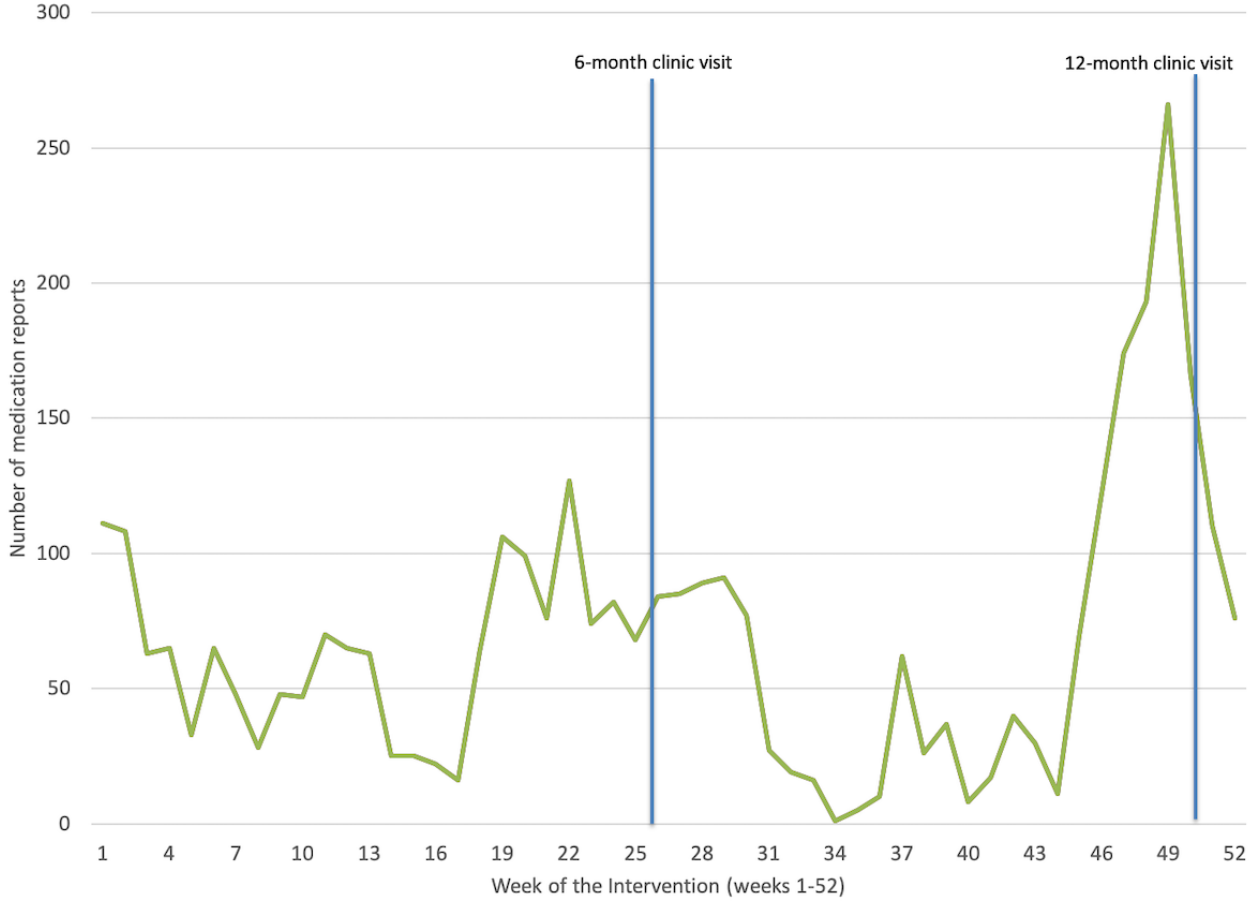


Table 2. Utilization by patient characteristics, indicating the number of weeks with at least one log-in during a 52-week period.

Group	Weeks (n)	Mean (SD)	*P* value[a]
Age (years)			
<50	73	18.2 (17.9)	<.001
≥50	49	30.1 (18.2)	—[b]
College educated			
No	34	21.6 (20.8)	.42
Yes	88	23.5 (18.2)	—
Use smartphone			
No	40	23.8 (19.7)	.66
Yes	82	22.6 (18.6)	—
Baseline Asthma Control Test score			
<20	65	22.7 (18.7)	.97
≥20	57	23.2 (19.2)	—

[a]*P* value from Wilcoxon rank-sum test.

[b]Not applicable.

Discussion

Despite a decade of mHealth app development, there remains a limited body of evidence demonstrating improved health outcomes with apps [29-31]. Currently, there are 3 published RCTs evaluating patient-facing, multifunctional asthma apps developed to support CSM. Liu et al [32] showed increased quality of life, increased use of controller medications, improved lung function (8%), and decreased emergency service use.

Merchant et al [33] demonstrated the effectiveness of the Propeller Health Asthma Platform at reducing the use of short-acting β-agonist (SABA) by 0.41 activations per day (vs 0.31 control), increasing the number of SABA-free days by 21% (vs 17% control). Conversely, Ryan et al [34] found that mobile phone–based monitoring did not improve asthma control or patient self-efficacy compared with a paper-based monitoring system. Our experience with an asthma app prototype in a recent pilot study revealed a high level of satisfaction with the app (80% of users viewed the app positively, with the majority wishing to continue using the app after the study), regular participation in self-management, and improvements in asthma-related quality of life [35]. The central question of whether mHealth platforms that support asthma management such as *breathe* are efficacious remains unanswered. We assert that good platform design is a precondition to posing and answering this question.

Usability, Satisfaction, and Confidence

Patients had a high level of satisfaction with the individual design components of *breathe*. They rated *breathe* usability as good on the validated SUS and high on standard Likert scales. A central function of the *breathe* platform was to present patients with a real-time dynamic action plan based on their symptoms and peak flow entries. The app effectively returned a dynamic zone of control calculation back to patients. *breathe* patients were confident that the calculated assessment was accurate. By objectively measuring control, *breathe* resolves a long-standing barrier to action plan utilization in the community, the barrier of inaccurate control assessment by patients. Patients who overestimate control will not activate their action plan as prescribed and thus not experience the substantial associated clinical benefits [5-8].

The *breathe* mHealth platform was an important facilitator of teachable moments and acted as an unidirectional communication bridge between providers and patients in the community through the delivery of 50,939 reminder email messages and communicating asthma control and care recommendations through the platform 19,678 times. An examination of utilization suggests that patients responded to these notifications by accessing the app after the reception of these emails, and patient questionnaires indicate that they had confidence in the care and control recommendations.

App Usage

The goal of UCD is to create and sustain a certain level of adherence to the platform, as adherence is a prerequisite to positive behavioral change and improved health outcomes. Despite good ratings for ease of use and a high degree of satisfaction with the *breathe* system, actual platform use declined substantially over time, which in general aligns with reviews describing attrition rates in eHealth deployments [28,36-38]. In his seminal viewpoint paper "The Law of Attrition," Eysenbach argues the need for a *science of attrition* and recommends that usage metrics be measured, analyzed, and discussed to identify reasons for attrition [28]. The *breathe* utilization curve differs substantially from Eysenbach specifically related to a dramatic fall in utilization in the first 4 to 6 weeks. We evaluated factors associated with increased or decreased platform utilization.

We considered that decreased utilization (attrition) in this study might have been related to population and design characteristics, including technology savviness, patients with relatively good disease control, infrequent physician monitoring, or because patients achieved their expected outcomes (or the correct *digital dose* of the intervention).

Technology Savviness

All participants had access to either a smartphone or a computer. Although, 55.2% (76/138) of our population had a smartphone and reported being comfortable or very comfortable with its use, one-third did not have a smartphone and therefore accessed the platform by laptop, desktop, or tablet. We considered that the nonsmartphone subset may have been less technology savvy, contributing to the decline in utilization and particularly may have contributed to the sharp decline in the first 4 weeks. However, our post hoc analysis did not find an association between utilization and having a smartphone.

Age and Education Level

We considered that younger age and higher education level might have an impact on utilization. We did not find an association between utilization and education level. In a post hoc analysis, we were able to demonstrate that being aged 50 years and above was significantly associated with higher utilization. Although general app use is normally greater in a younger population, we speculate that our participants over the age of 50 with a chronic disease may have had a higher level of concern about their chronic disease and potentially find more value in health-related apps than a younger population. We observed that increased utilization was associated with time of day, anticipated physician visits, and email reminders.

Good Disease Control

Patients in this study had relatively well controlled asthma as indicated by high baseline ACTs and a high percentage of good days when compared with episode days. We did not have a specific engagement strategy to motivate patients to return to the platform when they were feeling well. Failure to engage the users in moments of disease stability has been described by other authors as a critical factor affecting attrition across diseases [39-41]. However, our post hoc analysis did identify an association between utilization and high versus low scores on the ACT.

Physician Monitoring

In this study, patients were evaluated by a physician only twice after enrollment. Infrequent monitoring may have increased the attrition rate. Increased *breathe* platform utilization was associated with upcoming 6- and 12-month clinic appointments. An increase in eHealth utilization in response to anticipated clinical review has been described by Mohr et al as supportive accountability [42] and by others [43,44] as a strong factor influencing sustained adherence. The surveillance effect has a direct influence on how engaged patients are with the platform and how much they adhere to the intervention. Along the same lines, eHealth platforms that provide some level of feedback

and peer support appear to demonstrate better adherence rates [45]. The need for regular clinical review to motivate platform adherence aligns with the literature supporting written asthma action plans, where efficacy requires regular clinical review [14]. The finding related to increased medication reporting at 6 and 12 months also suggests that for most of the year, medication use was underreported. Self-reported medication use may underreport actual use [46]. New Bluetooth-enabled smart inhalers that automatically log medication use [47] will be considered in the future development of *breathe*.

Patients Achieved Their Expected Outcomes

Patients were satisfied with *breathe*, and 63.8% (74/116) agreed or strongly agreed that "the *breathe* application is helpful in the management of my asthma." Thus, it is possible that after an interval, having achieved their personal goals, patients no longer felt a need to use the platform.

Email Reminders

Increased *breathe* platform utilization was associated in time with email adherence reminders. Others have identified reminders as powerful design features to increase adherence and engagement with eHealth platforms [48], to alert participants of important events [19,49], or to alert them of aspects of the treatment they have missed [45]. Although alarm fatigue has been described in long-term interventions, wherein reminders lose their impact over time [50,51], we demonstrated a sustained effect of reminders over the 12 months.

Usage Analysis Summary

Patterns of usage analysis identified physician visits and email reminders as strongly associated with utilization. A post hoc analysis identified being aged 50 years and above as significantly associated with higher utilization.

Limitations

The population studied was a convenience sample from primary and specialty clinics with a dedicated asthma program, and at the time of enrollment, patients had relatively good asthma control. As such, patients' evaluation of the app and their utilization patterns may not be representative of the general asthma population. Since this project was completed, native apps have largely supplanted Web browser–based apps such as *breathe*. The improved performance of native app platforms may positively impact utilization and reduce attrition.

Conclusions

We followed UCD methods to develop *breathe*, a multifunctional asthma CSM platform with content based on international clinical practice guidelines, compliant with national privacy and security specifications, to support patients as active participants in chronic disease management at home, work, and in the community. *breathe* enabled self-management and self-monitoring of asthma patients through assessment of asthma control, real-time access to a dynamic action plan, environmental conditions display, and air quality risk-reduction messaging. Individuals with asthma reported good usability and high satisfaction levels and had confidence in the platform's assessment of asthma control. We embedded in-platform analytics, evaluated utilization, and examined the utilization patterns in the context of known patient characteristics. We related increased utilization to physician monitoring, email reminders, and age 50 years and above. Looking to the future, embedded app analytics combined with data-driven design will enable real-time evaluation of mHealth platforms, enabling innovators to execute design improvements during the deployment of the technology.

Lessons Learned or Future Considerations

As we iterate development of the *breathe* platform based on lessons learned, we will seek to (1) leverage the surveillance effect of in-platform or in-person patient-physician contact to support utilization, (2) create a specific strategy to engage patients when they are feeling well and to reengage as they become unwell, (3) create a strategy to support adherence specifically for asthma patients aged less than 50 years, (4) integrate automated logging technology (smart or connected inhalers) to capture actual medication utilization, (5) leverage the sustained impact of patient reminders on utilization, (6) create a more interactive experience to enhance platform use, (7) utilize embedded app analytics that provide continuous evaluation of usage to enable the execution of design improvements during platform deployment, and (8) develop the next version of the *breathe* platform with a native iOS or Android app.

Acknowledgments

The OLA through investment from the Canada Health Infoway's Consumer Health Solutions Program and the Ministry of Health and Long-Term Care (Asthma Program) supported the development of *breathe*.

Conflicts of Interest

PPM reports grants from the Natural Sciences and Engineering Research Council of Canada (NSERC), Canadian Institute for Health Research (CIHR), MITACS, and Ontario Centres of Excellence (OCE) during the conduct of this study. PPM is a member of advisory boards at Roche Canada. The work developed in this paper was not funded by any of these companies. MSY reports grants from Canada Health Infoway during the conduct of the study. MF, AKT, and CM have nothing to disclose. ASL reports grants from Canada Health Infoway, from Ontario Ministry of Health and Long-Term Care, during the conduct of the study. TT reports grants from Ontario Ministry of Health and Long-Term Care, grants from Ontario Ministry of the Environment and Climate Change, grants from CIHR, grants from Health Canada, grants from Canadian Respiratory Research Network, outside the submitted work. MDL received honoraria from the Astra Zeneca Severe Asthma PRECISION Program, and funds were paid

directly to Queen's University for participation in multicenter clinical trials from Astra Zeneca, GlaxoSmithKline, Hoffman LaRoche Ltd, Janssen, and Novartis; grants were paid directly to Queen's University from the OLA, the Government of Ontario's Innovation Fund, Allergen NCE, Canadian Institutes of Health Research; and personal fees were paid by the Public Service Occupational Health Program Regions and Programs Bureau Health Canada or Government of Canada for preparation of a report on pollution exposure at post and the role of surveillance spirometry, outside the submitted work. SG has nothing to disclose. AGD reports that his employer was paid from grants from Canada Health Infoway and The Lung Association to cost recover his time spent on this project. JAC reports grants from Canada Health Infoway during the conduct of the study. CL declares that he is a member of advisory boards or equivalent in commercial organizations as AstraZeneca, Novartis, Boehringer Ingelheim, and GlaxoSmithKline as well as receiving funding from commercial organizations as AstraZeneca, Novartis, Boehringer Ingelheim, Pfizer, and Bayer. The work developed in this project was not funded by any of these companies.

References

1. Public Health Agency of Canada. Government of Canada. 2018 Mar. Report from the Canadian Chronic Disease Surveillance System: Mental Illness in Canada, 2015 URL: https://www.canada.ca/en/public-health/services/publications/diseases-conditions/report-canadian-chronic-disease-surveillance-system-mental-illness-canada-2015.html [accessed 2018-11-28] [WebCite Cache ID 74GALXxmV]
2. World Health Organization. 2017 Aug 31. Asthma - Fact Sheet Internet URL: http://www.who.int/en/news-room/fact-sheets/detail/asthma [accessed 2018-11-28] [WebCite Cache ID 74GAWux3B]
3. GBD 2016 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 328 diseases and injuries for 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep 16;390(10100):1211-1259 [FREE Full text] [doi: 10.1016/S0140-6736(17)32154-2] [Medline: 28919117]
4. Chapman KR, Boulet LP, Rea RM, Franssen E. Suboptimal asthma control: prevalence, detection and consequences in general practice. Eur Respir J 2008 Feb;31(2):320-325 [FREE Full text] [doi: 10.1183/09031936.00039707] [Medline: 17959642]
5. FitzGerald JM, Boulet L, McIvor RA, Zimmerman S, Chapman KR. Asthma control in Canada remains suboptimal: the Reality of Asthma Control (TRAC) study. Can Respir J 2006;13(5):253-259 [FREE Full text] [Medline: 16896426]
6. Lai CK, De Guia TS, Kim Y, Kuo S, Mukhopadhyay A, Soriano JB, Asthma Insights and Reality in Asia-Pacific Steering Committee. Asthma control in the Asia-Pacific region: the asthma insights and reality in Asia-Pacific study. J Allergy Clin Immunol 2003 Feb;111(2):263-268. [doi: 10.1067/mai.2003.30] [Medline: 12589343]
7. Rabe KF, Adachi M, Lai CK, Soriano JB, Vermeire PA, Weiss KB, et al. Worldwide severity and control of asthma in children and adults: the global asthma insights and reality surveys. J Allergy Clin Immunol 2004 Jul;114(1):40-47. [doi: 10.1016/j.jaci.2004.04.042] [Medline: 15241342]
8. Sekerel BE, Gemicioglu B, Soriano JB. Asthma insights and reality in Turkey (AIRET) study. Respir Med 2006 Oct;100(10):1850-1854 [FREE Full text] [doi: 10.1016/j.rmed.2006.01.024] [Medline: 16564685]
9. Government of Canada. 2007 Jul 26. Life and Breath: Respiratory Disease in Canada Internet URL: https://www.canada.ca/en/public-health/services/reports-publications/2007/life-breath-respiratory-disease-canada-2007.html [accessed 2018-11-28] [WebCite Cache ID 74GAf8cN3]
10. Windham B, Bennett R, Gottlieb S. Care management interventions for older patients with congestive heart failure. Am J Manag Care 2003 Jun;9(6):447-59; quiz 460 [FREE Full text] [Medline: 12816174]
11. Lougheed MD, Lemiere C, Ducharme FM, Licskai C, Dell SD, Rowe BH, Canadian Thoracic Society Asthma Clinical Assembly. Canadian Thoracic Society 2012 guideline update: diagnosis and management of asthma in preschoolers, children and adults. Can Respir J 2012;19(2):127-164 [FREE Full text] [doi: 10.1155/2012/635624] [Medline: 22536582]
12. Global Initiative for Asthma. 2018. Global Strategy for Asthma Management and Prevention URL: https://ginasthma.org/2018-gina-report-global-strategy-for-asthma-management-and-prevention/ [accessed 2018-11-28] [WebCite Cache ID 74GB90JEP]
13. Lougheed M, Lemière C, Dell S, Ducharme F, Fitzgerald JM, Leigh R, Canadian Thoracic Society Asthma Committee. Canadian Thoracic Society Asthma Management Continuum--2010 Consensus Summary for children six years of age and over, and adults. Can Respir J 2010;17(1):15-24 [FREE Full text] [doi: 10.1155/2010/827281] [Medline: 20186367]
14. Gibson PG, Powell H, Coughlan J, Wilson AJ, Abramson M, Haywood P, et al. Self-management education and regular practitioner review for adults with asthma. Cochrane Database Syst Rev 2003(1):CD001117. [doi: 10.1002/14651858.CD001117] [Medline: 12535399]
15. Caress A, Beaver K, Luker K, Campbell M, Woodcock A. Involvement in treatment decisions: what do adults with asthma want and what do they get? Results of a cross sectional survey. Thorax 2005 Mar;60(3):199-205 [FREE Full text] [doi: 10.1136/thx.2004.029041] [Medline: 15741435]
16. Adams RJ, Smith BJ, Ruffin RE. Patient preferences for autonomy in decision making in asthma management. Thorax 2001 Feb;56(2):126-132 [FREE Full text] [Medline: 11209101]

17. Tsuyuki R, Sin D, Sharpe H, Cowie R, Nilsson C, Man SP, Alberta Strategy to Help Manage Asthma (ASTHMA) Investigators. Management of asthma among community-based primary care physicians. J Asthma 2005 Apr;42(3):163-167. [Medline: 15962871]
18. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. Biomed Instrum Technol 2012;Suppl:49-56. [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
19. Goyal S, Morita P, Lewis GF, Yu C, Seto E, Cafazzo JA. The systematic design of a behavioural mobile health application for the self-management of type 2 diabetes. Can J Diabetes 2016 Feb;40(1):95-104. [doi: 10.1016/j.jcjd.2015.06.007] [Medline: 26455762]
20. Licskai C, Ferrone M, Taite A, Madeley C, Lavigne A, To T, et al. The evaluation of breathe - a patient mobile health (mHealth) app for adult asthma. Am J Respir Crit Care Med 2016;193:a1084.
21. Morita PP, Cafazzo JA. Challenges and paradoxes of human factors in health technology design. JMIR Hum Factors 2016 Mar 01;3(1):e11 [FREE Full text] [doi: 10.2196/humanfactors.4653] [Medline: 27025862]
22. Johnson CM, Johnson TR, Zhang J. A user-centered framework for redesigning health care interfaces. J Biomed Inform 2005 Feb;38(1):75-87 [FREE Full text] [doi: 10.1016/j.jbi.2004.11.005] [Medline: 15694887]
23. Elkin PL, Beuscart-Zephir M, Pelayo S, Patel V, Nøhr C. The usability-error ontology. Stud Health Technol Inform 2013;194:91-96. [Medline: 23941937]
24. Government of Canada. 2016 Jun 07. Air Quality Health Index URL: https://www.canada.ca/en/environment-climate-change/services/air-quality-health-index/about.html [accessed 2018-11-28] [WebCite Cache ID 74G9DL4w9]
25. Oinas-Kukkonen H. Persuasive systems design: key issues, process model, and system features. Commun Assoc Inf Syst 2009;24(1):28. [doi: 10.17705/1CAIS.02428]
26. Nathan RA, Sorkness CA, Kosinski M, Schatz M, Li JT, Marcus P, et al. Development of the asthma control test: a survey for assessing asthma control. J Allergy Clin Immunol 2004 Jan;113(1):59-65. [doi: 10.1016/j.jaci.2003.09.008] [Medline: 14713908]
27. Bangor A, Kortum P, Miller J. An empirical evaluation of the system usability scale. Int J Hum Comput Interact 2008 Jul 30;24(6):574-594. [doi: 10.1080/10447310802205776]
28. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
29. Wicks P, Chiauzzi E. 'Trust but verify'--five approaches to ensure safe medical apps. BMC Med 2015;13:205 [FREE Full text] [doi: 10.1186/s12916-015-0451-z] [Medline: 26404791]
30. Pham Q, Wiljer D, Cafazzo JA. Beyond the randomized controlled trial: a review of alternatives in mHealth clinical trial methods. JMIR Mhealth Uhealth 2016 Sep 09;4(3):e107 [FREE Full text] [doi: 10.2196/mhealth.5720] [Medline: 27613084]
31. Huckvale K, Morrison C, Ouyang J, Ghaghda A, Car J. The evolution of mobile apps for asthma: an updated systematic assessment of content and tools. BMC Med 2015;13:58 [FREE Full text] [doi: 10.1186/s12916-015-0303-x] [Medline: 25857569]
32. Liu W, Huang C, Wang C, Lee K, Lin S, Kuo H. A mobile telephone-based interactive self-care system improves asthma control. Eur Respir J 2011 Feb;37(2):310-317 [FREE Full text] [doi: 10.1183/09031936.00000810] [Medline: 20562122]
33. Merchant RK, Inamdar R, Quade RC. Effectiveness of population health management using the propeller health asthma platform: a randomized clinical trial. J Allergy Clin Immunol Pract 2016;4(3):455-463. [doi: 10.1016/j.jaip.2015.11.022] [Medline: 26778246]
34. Ryan D, Price D, Musgrave SD, Malhotra S, Lee AJ, Ayansina D, et al. Clinical and cost effectiveness of mobile phone supported self monitoring of asthma: multicentre randomised controlled trial. Br Med J 2012;344:e1756 [FREE Full text] [Medline: 22446569]
35. Licskai C, Sands TW, Ferrone M. Development and pilot testing of a mobile health solution for asthma self-management: asthma action plan smartphone application pilot study. Can Respir J 2013 Aug;20(4):301-306 [FREE Full text] [Medline: 23936890]
36. Eysenbach G. The law of attrition. J Med Internet Res 2005;7(1):e11 [FREE Full text] [doi: 10.2196/jmir.7.1.e11] [Medline: 15829473]
37. Christensen H, Griffiths KM, Farrer L. Adherence in internet interventions for anxiety and depression. J Med Internet Res 2009;11(2):e13 [FREE Full text] [doi: 10.2196/jmir.1194] [Medline: 19403466]
38. Lang E, Afilalo M, Vandal AC, Boivin J, Xue X, Colacone A, et al. Impact of an electronic link between the emergency department and family physicians: a randomized controlled trial. CMAJ 2006 Jan 31;174(3):313-318 [FREE Full text] [doi: 10.1503/cmaj.050698] [Medline: 16399880]
39. Kelders SM, Kok RN, Ossebaard HC, Van Gemert-Pijnen JE. Persuasive system design does matter: a systematic review of adherence to web-based interventions. J Med Internet Res 2012;14(6):e152 [FREE Full text] [doi: 10.2196/jmir.2104] [Medline: 23151820]
40. Brouwer W, Kroeze W, Crutzen R, De Nooijer J, De Vries NK, Brug J, et al. Which intervention characteristics are related to more exposure to internet-delivered healthy lifestyle promotion interventions? A systematic review. J Med Internet Res 2011;13(1):e2 [FREE Full text] [doi: 10.2196/jmir.1639] [Medline: 21212045]

41. Schubart JR, Stuckey HL, Ganeshamoorthy A, Sciamanna CN. Chronic health conditions and internet behavioral interventions: a review of factors to enhance user engagement. Comput Inform Nurs 2011 Feb;29(2 Suppl):TC9-T20. [doi: 10.1097/NCN.0b013e3182155274] [Medline: 21372645]
42. 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;13(1):e30 [FREE Full text] [doi: 10.2196/jmir.1602] [Medline: 21393123]
43. Cramer J, Scheyer R, Mattson R. Compliance declines between clinic visits. Arch Intern Med 1990 Jul;150(7):1509-1510. [Medline: 2369248]
44. Osterberg L, Blaschke T. Adherence to medication. N Engl J Med 2005 Aug 04;353(5):487-497. [doi: 10.1056/NEJMra050100] [Medline: 16079372]
45. Logan AG, Irvine MJ, McIsaac WJ, Tisler A, Rossos PG, Easty A, et al. Effect of home blood pressure telemonitoring with self-care support on uncontrolled systolic hypertension in diabetics. Hypertension 2012 Jul;60(1):51-57. [doi: 10.1161/HYPERTENSIONAHA.111.188409] [Medline: 22615116]
46. Voils CI, Hoyle RH, Thorpe CT, Maciejewski ML, Yancy WS. Improving the measurement of self-reported medication nonadherence. J Clin Epidemiol 2011 Mar;64(3):250-254 [FREE Full text] [doi: 10.1016/j.jclinepi.2010.07.014] [Medline: 21194887]
47. Burgess SW, Wilson SS, Cooper DM, Sly PD, Devadason SG. In vitro evaluation of an asthma dosing device: the smart-inhaler. Respir Med 2006 May;100(5):841-845 [FREE Full text] [doi: 10.1016/j.rmed.2005.09.004] [Medline: 16216485]
48. Fry JP, Neff RA. Periodic prompts and reminders in health promotion and health behavior interventions: systematic review. J Med Internet Res 2009;11(2):e16 [FREE Full text] [doi: 10.2196/jmir.1138] [Medline: 19632970]
49. Goyal S, Morita PP, Picton P, Seto E, Zbib A, Cafazzo JA. Uptake of a consumer-focused mHealth application for the assessment and prevention of heart disease: the <30 days study. JMIR Mhealth Uhealth 2016;4(1):e32 [FREE Full text] [doi: 10.2196/mhealth.4730] [Medline: 27012937]
50. Cvach M. Monitor alarm fatigue: an integrative review. Biomed Instrum Technol 2012;46(4):268-277. [doi: 10.2345/0899-8205-46.4.268] [Medline: 22839984]
51. Denkinger CM, Grenier J, Stratis AK, Akkihal A, Pant-Pai N, Pai M. Mobile health to improve tuberculosis care and control: a call worth making. Int J Tuberc Lung Dis 2013 Jun;17(6):719-727. [doi: 10.5588/ijtld.12.0638] [Medline: 23541232]

Abbreviations

ACT: Asthma Control Test
AQHI: Air Quality Health Index
CIHR: Canadian Institute for Health Research
CSM: collaborative self-management
CTS: Canadian Thoracic Society
eHealth: electronic health
mHealth: mobile health
OLA: Ontario Lung Association
pppw: per patient per week
RCT: randomized controlled trial
SABA: short-acting β-agonist
SUS: System Usability Scale
UCD: user-centered design

Edited by G Eysenbach; submitted 18.06.18; peer-reviewed by G Matthias, E Meinert, P Wicks; comments to author 12.07.18; revised version received 02.09.18; accepted 27.09.18; published 28.01.19.

Please cite as:
Morita PP, Yeung MS, Ferrone M, Taite AK, Madeley C, Stevens Lavigne A, To T, Lougheed MD, Gupta S, Day AG, Cafazzo JA, Licskai C
A Patient-Centered Mobile Health System That Supports Asthma Self-Management (breathe): Design, Development, and Utilization
JMIR Mhealth Uhealth 2019;7(1):e10956
URL: http://mhealth.jmir.org/2019/1/e10956/
doi:10.2196/10956
PMID:30688654

Original Paper

Breast, Prostate, and Colorectal Cancer Survivors' Experiences of Using Publicly Available Physical Activity Mobile Apps: Qualitative Study

Anna L Roberts[1], BSc (Hons), MSc; Henry WW Potts[2], BA (Hons), MSc, PhD; Dimitrios A Koutoukidis[3,4], BSc (Hons), MSc, RD, PhD; Lee Smith[5], BSc (Hons), MSc, PhD; Abigail Fisher[1], BSc (Hons), PhD

[1]Research Department of Behavioural Science & Health, University College London, London, United Kingdom

[2]Institute of Health Informatics, University College London, London, United Kingdom

[3]Nuffield Department of Primary Care Health Sciences, University of Oxford, Oxford, United Kingdom

[4]National Institute for Health Research Oxford Biomedical Research Centre, Oxford, United Kingdom

[5]The Cambridge Centre for Sport and Exercise Sciences, Department of Life Sciences, Anglia Ruskin University, Cambridge, United Kingdom

Corresponding Author:
Anna L Roberts, BSc (Hons), MSc
Research Department of Behavioural Science & Health
University College London
Gower Street
London, WC1E 6BT
United Kingdom
Phone: 44 02076791723
Email: anna.roberts.15@ucl.ac.uk

Abstract

Background: Physical activity (PA) can improve a range of outcomes following a cancer diagnosis. These include an improvement in experience of side effects of treatment (eg, fatigue) and management of comorbid conditions. PA might also increase survival and reduce recurrence. Digital interventions have shown potential for PA promotion among cancer survivors, but most in a previous review were Web-based, and few studies used mobile apps. There are many PA apps available for general public use, but it is unclear whether these are suitable as a PA intervention after a cancer diagnosis.

Objective: This study sought posttreatment nonmetastatic breast, prostate, and colorectal cancer survivors' opinions of using smartphone apps to promote PA and gathered their views on existing publicly available PA apps to inform a future intervention.

Methods: Each participant was randomly assigned to download 2 of 4 apps (Human, The Walk, The Johnson & Johnson Official 7 Minute Workout, and Gorilla Workout). Participants used each app for 1 week consecutively. In-depth semistructured telephone interviews were then conducted to understand participants' experiences of using the apps and how app-based PA interventions could be developed for cancer survivors. The interviews were analyzed using thematic analysis.

Results: Thirty-two participants took part: 50% (16/32) had prostate cancer, 25% (8/32) had breast cancer, and 25% (8/32) had colorectal cancer. Three core themes were identified. The first theme was that multiple factors affect engagement with PA apps and this is highly personalized. Factors affecting engagement included participants' perceptions of (1) the advantages and disadvantages of using apps to support PA, (2) the relevance of the app to the user (eg, in terms of cancer-related factors, their PA goals, the difficulty level of the app, the way in which they interact with their mobile phone, and the extent to which the app fits with their self-identity), (3) the quality of the app (eg, usability, accuracy, quality of production, and scientific evidence-base), and (4) the behavior change techniques used to promote PA. In the second theme, participants recommended that apps that promote walking are most appealing, as walking removes many barriers to PA. Finally, the participants suggested that PA apps should be integrated into cancer care, as they valued guidance and recommendations from health care professionals.

Conclusions: This sample of breast, prostate, and colorectal cancer survivors was receptive to the use of apps to promote PA. Although no publicly available PA app was deemed wholly suitable, many suggestions for adaptation and intervention development were provided. The results can inform the development of an app-based PA intervention for cancer survivors. They also highlight the wide-ranging and dynamic influences on engagement with digital interventions, which can be applied to other evaluations of mobile health products in other health conditions and other health behaviors.

(JMIR Mhealth Uhealth 2019;7(1):e10918) doi:10.2196/10918

KEYWORDS

physical activity; health behavior; cancer survivors; mobile apps; mHealth; digital health

Introduction

The number of people diagnosed with cancer continues to increase and is estimated to reach more than 20 million new cases per year worldwide by 2025 [1]. Earlier diagnosis and improvements in treatment mean that rates of cancer survival are also increasing. However, fatigue [2], pain [3], sleep problems [4], weight gain [5,6], and anxiety and depression [7,8] are common among cancer survivors, and 70% have a comorbid chronic condition [9]. Depending on cancer type and the area at which treatment is targeted, particular groups of cancer survivors are at greater risk of more specific late effects. For instance, lymphedema is common among breast cancer survivors [10], and incontinence is common among prostate and colorectal cancer survivors [11,12]. These sequelae can have a profound negative impact on quality of life (QoL) [13], and interventions to improve outcomes in cancer survivors are urgently required.

There is now strong evidence that physical activity (PA) can improve a range of important cancer outcomes for breast, prostate, and colorectal cancer survivors, 3 of the most prevalent cancer types worldwide [1]. Observational evidence shows that PA might reduce cancer-specific and all-cause mortality and cancer recurrence in breast, prostate, and colorectal cancer survivors [14-16]. PA has also been shown to improve overall health-related QoL, emotional well-being, and social functioning and reduce anxiety, fatigue, pain, and sleep problems in cancer survivors [17]. Therefore, cancer survivors are now advised to meet the same PA guidelines as the general adult population. This includes a minimum of 150 min of at least moderate-intensity PA and 2 instances of strength and resistance-based training per week [18-21]. Where this is not achievable, avoiding inactivity is recommended. However, when PA is measured objectively using accelerometers, as few as 16% of breast cancer survivors meet PA recommendations, and those with the highest level of comorbidity are the least active [22]. People diagnosed with cancer are less likely to engage in PA than people who have never received a diagnosis [23]. In a sample of 631 breast cancer survivors, the self-reported proportion meeting PA guidelines declined from 34% at 2 years post diagnosis to 21% at 10 years post diagnosis [24]. Side effects of treatment and fear about what type of PA, when, and how to start or increase PA safely are often reported as barriers to PA after a cancer diagnosis [25-27]. As a result, PA interventions for people affected by cancer are required.

Face-to-face interventions are time- and resource-intensive, and accessibility can be limited [28,29]. Increasing ownership of smartphones provides an avenue for scalable digital behavior change interventions, including those that aim to increase PA. In the context of cancer, for which age is the strongest risk factor [30], smartphone ownership has, in recent years, risen most rapidly among the older age groups. In the United Kingdom specifically, smartphone ownership increased from 32% to 47% in people aged over 55 years from 2015 to 2017 [31,32]. In the United States, smartphone ownership is even higher among older age groups (73% among those aged 50-64 years and 46% in people aged over 65 years) [33].

Digital behavior change interventions have been shown to increase PA in the general adult population [34] and our recent meta-analysis of 15 studies showed that digital interventions have the potential to increase cancer survivors' moderate-vigorous PA by approximately 40 min per week [35]. However, of the studies included in this review, the majority used Web-based interventions, and only 2 small feasibility studies evaluated the use of mobile apps in PA promotion [36,37]. Mobile apps have the benefit of being able to deliver behavior change techniques (BCTs) in real time using a device that is usually switched on, usually carried with the person, and often has inbuilt functions to monitor PA and deliver immediate feedback. There are many health and fitness apps aimed at the general population that are currently available on commercial app stores, which might already be appropriate for cancer survivors or could be adapted to increase their suitability. Exploring cancer survivors' experiences of using different types of existing apps is, therefore, a useful way to understand which types of PA apps might be most appropriate or successful, before making potentially large investments into app or intervention development.

Qualitative research methods provide a rich understanding of people's experiences, thoughts, and opinions and seeking the perspectives of intended users is a critical element of digital intervention development [38,39]. Robertson et al conducted focus groups with breast, prostate, colorectal, and endometrial cancer survivors where feedback was collected for potential PA app features and messages [40]; however, the feedback provided was hypothetical. We suggest that by allowing participants to actually experience using different types of apps and BCTs over a period of time provides greater ecological validity. Since the publication of the meta-analysis, Short et al conducted an experiential mixed-methods study where 10 cancer survivors were referred to one of 15 existing PA apps, which were used for a 1- to 2-week period [41]. Although this study explored the participants' experience and preliminary efficacy of the app referral service, it did not explore participants' opinions of using the apps in detail. We see the value in a deeper understanding of participants' perceptions of their preferences for and influences on engagement with PA apps. For the purposes of this study, we use a broad, integrative definition of engagement comprising "1) the extent (e.g. amount, frequency, duration, depth) of usage and 2) a subjective experience characterized by attention, interest and affect" [42].

Therefore, the aim of this study was to seek breast, prostate, and colorectal cancer survivors' opinions of using apps to promote PA and gather their views on existing publicly available PA apps to inform a future intervention.

Methods

Mobile Apps

During our initial scoping of the smartphone app stores, no apps that were specifically designed to promote PA among cancer survivors were identified. This is in line with a previous Australian study exploring the use of PA apps among cancer survivors [41]. Therefore, the PA apps considered for this study were identified from apps that were featured in the "Health and Fitness" section of the British Apple App Store (iOS), along with other apps that the study authors were aware of from previous work in digital health and that might have been suitable for this study. The following criteria were considered in deciding which apps might be suitable for the study:

- Content: The apps needed to vary from each other in terms of the type of PA, and their format, features, and BCTs to allow comparison between different types of apps.
- Typicality: Although the apps needed to vary in terms of their content, we also felt that the apps chosen should be typical of the various types of popular PA apps that are available (eg, activity trackers and workout programs).
- Suitability: The apps needed to be suitable for people who have undergone cancer treatment and, therefore, needed to have the flexibility to cater for different levels of fitness and familiarity with PA. Given the target group, apps that catered for low levels of fitness/familiarity with PA, but with an option to increase this if required, were of interest. Each app was reviewed for its suitability for use by breast, prostate, and colorectal cancer survivors by a physiotherapist specializing in oncology.
- Stability: The apps were required to have been launched at least 2 years before the study.
- Availability: The apps needed to be available on both iOS and Android devices.

We felt that 4 apps should be included in the study, based on a number of considerations. These included the number of apps required to compare multiple participants' opinions across several different PA apps, the number of participants required for the study, and feasibility of recruitment and data analysis. Given the consideration of all of the above factors, the 4 chosen apps were "Human," "The Walk," "The Johnson & Johnson Official 7 Minute Workout" (J&J), and "Gorilla Workout" (see Table 1 for a description of each of the apps and an assessment of the incorporated BCTs, coded using the BCT Taxonomy (v1) [43] by AR and DK, with discrepancies resolved via discussion). Figures 1-4 show screenshots of the 4 apps.

Recruitment

Participants were recruited via advertisements within community-based cancer support groups (either by verbal descriptions from group leaders at meetings or via posters, flyers, and email mailing lists), Facebook cancer support groups, and charitable organizations (eg, Macmillan Cancer Support's Cancer Voices and Tackle Prostate Cancer). We initially aimed to recruit 32 participants to attempt to ensure sufficient representation from participants diagnosed with each of the 3 cancer types and so that approximately 16 participants would be allocated to use each of the 4 apps throughout the study. If new themes continued to be identified, we would continue recruitment until saturation was achieved.

Table 1. App characteristics.

App (Developer)	Price	Description	Behavior change techniques
Human (Humanco, Inc)	Free	Encourages users to meet daily 30/60/90/120 min goal of walking, running, and/or cycling measured using mobile phone's activity tracker. Delivers push notifications when users have not met their goal or during periods of inactivity. Compares activity levels to other app users nearby	1.1 Goal setting (behavior); 2.2 Feedback on behavior; 2.3 Self-monitoring of behavior; 6.2 Social comparison; 7.1 Prompts/cues; 10.3 Nonspecific reward
The Walk (Six to Start)	£2.29 (iOS); £2.59 (Android)	An interactive story-based game where walking unlocks audio clips to hear the next part to the story and other rewards. Time to complete an episode is based on the users' current physical activity level and walking is measured using the mobile phone's activity tracker	2.2 Feedback on behavior; 10.3 Nonspecific reward; 10.6 Nonspecific incentive
The Johnson & Johnson Official 7 Minute Workout (Johnson & Johnson Health and Wellness Solutions, Inc)	Free	7-min workouts are created to include aerobic and resistance exercises alternating between upper and lower body, core, and total body exercises. The workouts can be tailored to the users' current fitness and motivation levels and are provided with detailed video demonstrations and audio guidance	1.4 Action planning; 2.3 Self-monitoring of behavior; 4.1 Instruction on how to perform behavior; 6.1 Demonstration of the behavior; 7.1 Prompts/cues; 8.7 Graded tasks; 9.1 Credible source
Gorilla Workout (Heckr LLC)	£0.79 (iOS); £0.83 (Android)	The default program is tailored to the users' current fitness level and gradually increases in difficulty. Each exercise has written guidance with an associated video with visual and audio demonstrations. Users can also choose to complete their own selection of exercises (from a list of 43) with the same written/video demonstrations. Daily push notifications are delivered to remind users to complete their workout	4.1 Instruction on how to perform behavior; 6.1 Demonstration of the behavior; 7.1 Prompts/cues; 8.7 Graded tasks

Figure 1. Screenshots of Human.

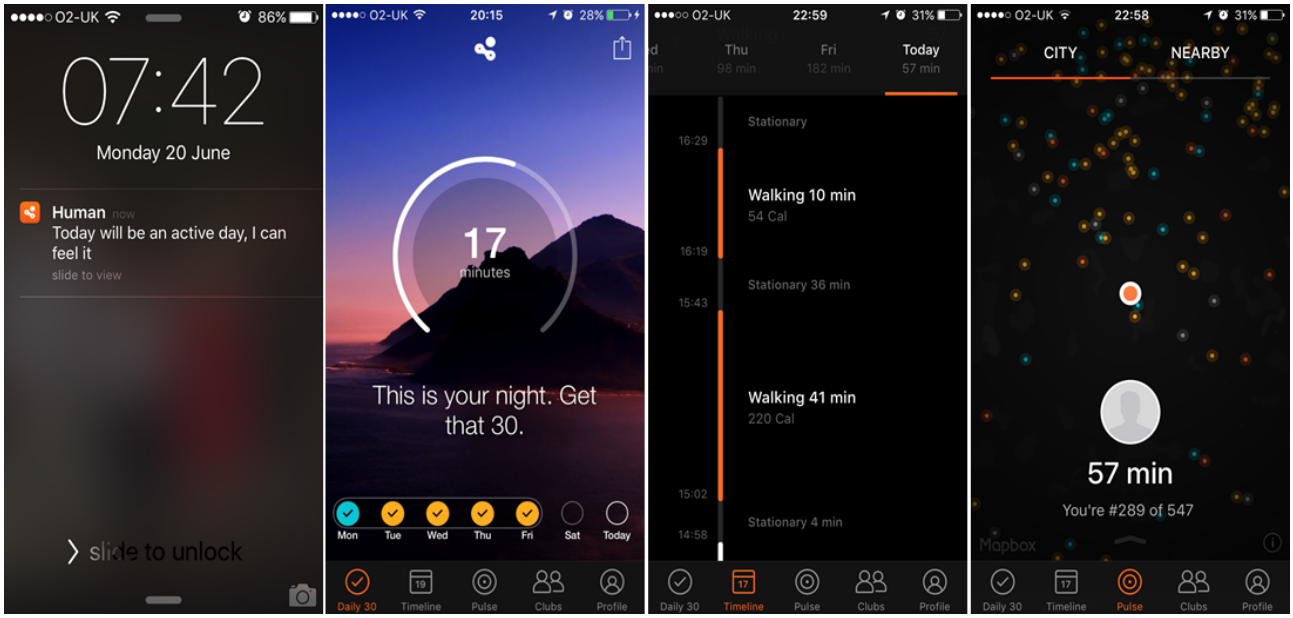
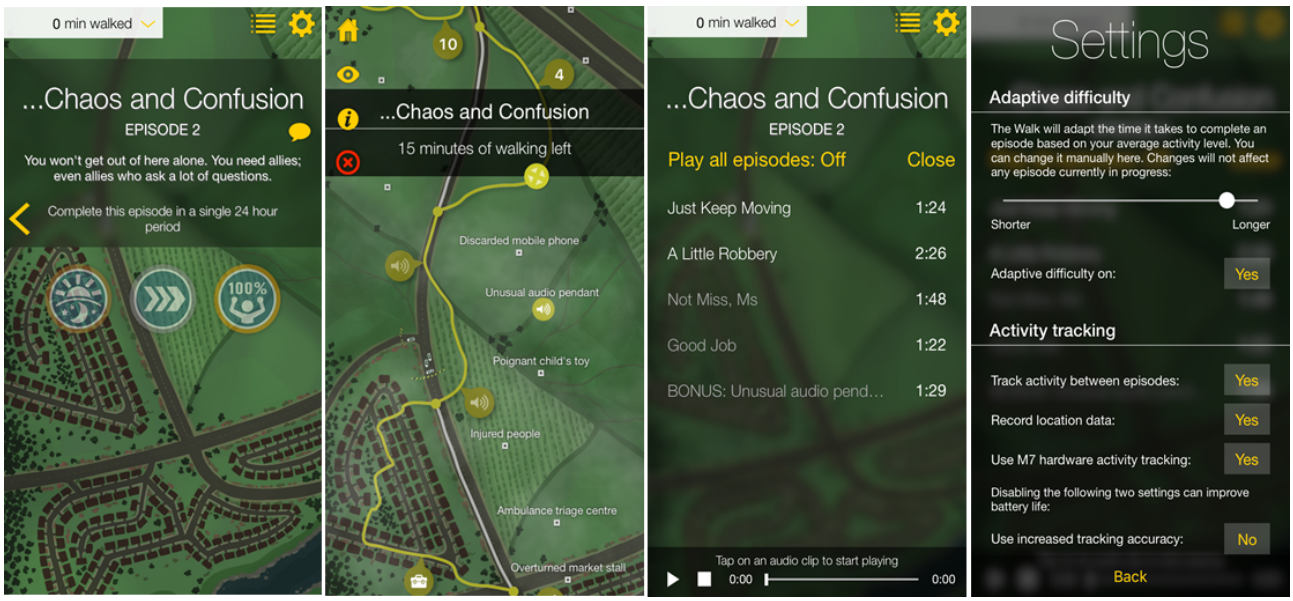


Figure 2. Screenshots of The Walk.



Participants were required to be aged 18 years or older; to have been diagnosed with breast, prostate, or colorectal cancer; to have finished primary curative treatment (as it is likely that individuals still undergoing primary treatment or with metastatic disease might require additional support and monitoring to be active); to have no known impairment or comorbidity that meant a clinician had advised them not to exercise; and to own a smartphone. Although participants were required to have finished primary curative treatment (surgery, radiotherapy, and chemotherapy), participants still taking maintenance hormone therapy or under active surveillance were eligible. Participants were offered a £10 voucher as an incentive for completion of this study and to reimburse the cost incurred if asked to install an app that was not free to download. Ethical approval for this study was granted by the UCL Research Ethics Committee (reference: 7663/001).

Procedure

Participants took part in an initial short semistructured telephone questionnaire that confirmed participants' eligibility and requested details of the participants' sociodemographic information (age, gender, and ethnicity), cancer diagnosis, and their experience of using digital technologies to support PA. Participants were asked to describe their perceptions of their current participation in PA (eg, what types of PA and how frequently). This was asked as an introductory question to build rapport with the participants at the beginning of the study and to provide context. A Web-based random number generator (Randomizer) was used to allocate 2 apps to each participant to allow comparison of app features and content but to minimize participant burden. Guidance in downloading and installing each app was provided, if required. Participants were asked to spend approximately 2 consecutive weeks using the apps, (approximately 1 week using each) and were able to choose the

order in which they used the apps to which they were allocated over the 2-week period. Participants were asked to try to use each app at least three to four times throughout that app's trial week and record any comments or opinions in log sheets provided. After 2 to 3 weeks, each participant completed an audio-recorded semistructured telephone interview, using the interview schedule (Table 2) as a guide.

Analysis

Telephone interviews were conducted by AR and transcribed verbatim by an external company. A partly deductive and partly inductive approach to thematic analysis was adopted using the stepped approach described by Braun and Clarke [44]. The deductive approach to thematic analysis involved using the BCT taxonomy [43] as a framework to code any interview data where participants spoke about app features used to promote behavior change. The rest of the data were analyzed using an inductive approach through an iterative reading and rereading of the data. An initial coding framework was developed by AR and revised in collaboration with DK, with discrepancies agreed via discussion. AR applied the final codes that were then incorporated into themes during discussion between all authors. After analysis of these 32 interviews, no new themes were identified and recruitment was concluded. Data analysis was conducted in NVivo 11.

Figure 3. Screenshots of The Johnson & Johnson Official 7 Minute Workout (J&J).

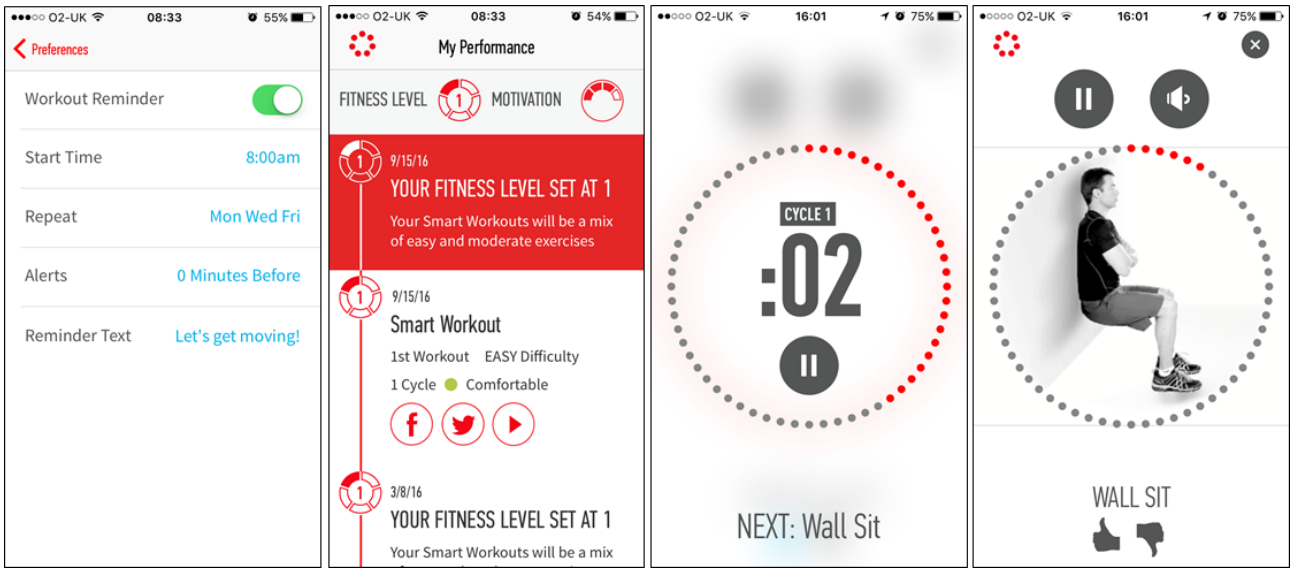


Figure 4. Screenshots of Gorilla Workout.

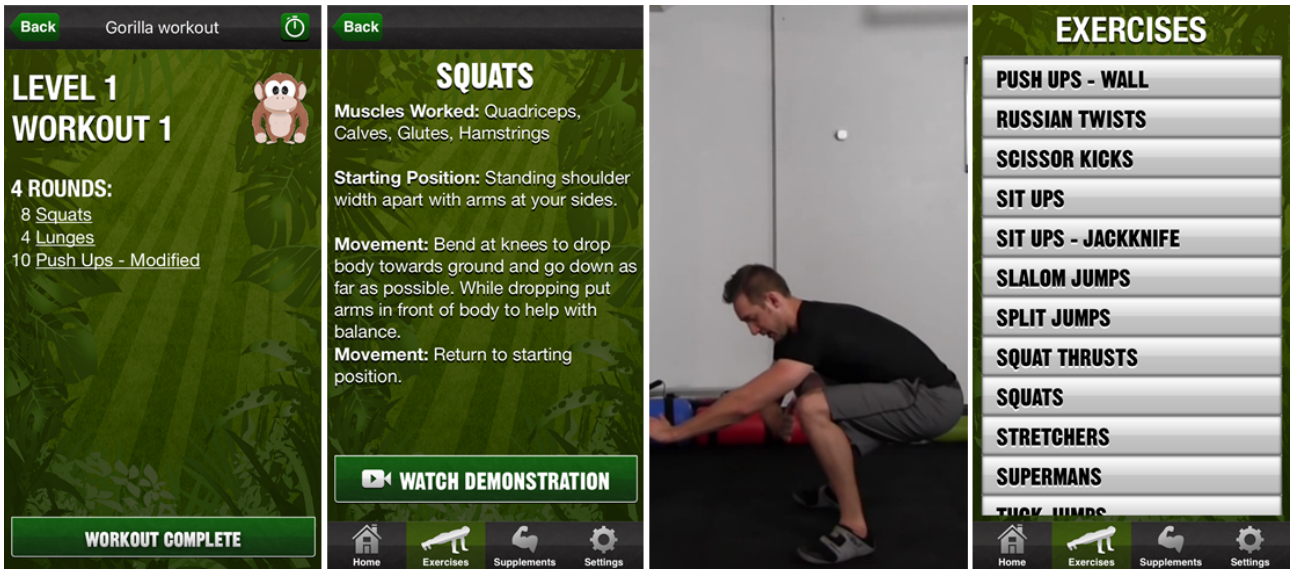


Table 2. Semistructured interview guide.

Discussion point	Details
Recap	Confirm which apps participant was asked to download and try
Download/install	Ask about the participants' ability to find, download, and install each app
First app	Ask participant to start by giving overall opinion of app; Depending on amount of detail provided in overview, ask participant to expand on any points raised in overview, comment on likes/dislikes, comment on specific app features/BCTs[a] (dependent on allocated app)
Second app	Repeat the steps as given for the first app
Appropriateness for cancer	Discuss how appropriate and relevant each of the apps were for their personal circumstances and as a cancer survivor
Adapting for cancer survivors	Discuss how (if at all) the apps could be adapted for cancer survivors. If so, what adaptations/functions to tailor the app would they make
Interest in an app	Discuss participants' interest in a PA[b] app tailored specifically for people who have had cancer
Preferred types of PA	Discuss types of PA that should be promoted to cancer survivors, including intensity, frequency, type of activity, and with relevance to current PA guidelines (ie, 150 min moderate-vigorous PA and 2 sessions of strength and resistance-based exercises per week) and how apps could promote these types of PA (if at all)
Recommendations	Discuss any PA recommendations that were provided to them following cancer diagnosis/treatment and who were they delivered by or where participant looked for them
Intervention communication	Discuss who should direct cancer survivors to a cancer-specific PA app, including when this should be discussed and promoted to patients

[a]BCT: behavior change technique.

[b]PA: physical activity.

Results

A total of 40 participants began the study, and 32 participants completed telephone interviews. Of those who dropped out, lack of time, family circumstances (eg, bereavement), and not wanting to update their smartphone's operating system or register credit card details with Google Play were the listed reasons. Of the 32 participants who completed the study, the mean age was 60 years (range 37-78 years; SD 11 years) and the other sample characteristics are displayed in Table 3.

Broadly, the core themes demonstrate that multiple factors affect engagement with PA apps and this is highly personalized, that apps that promote walking are most appealing for cancer survivors, and that PA apps should be integrated into cancer care.

Multiple Factors Affect Engagement With Physical Activity Apps, and This is Highly Personalized

Key determinants of engagement appeared to be the users' perceptions of (1) the advantages and disadvantages of using apps to support PA, (2) the relevance of the app, (3) the quality of the app, and (4) the BCTs used to promote PA.

Perceived Advantages and Disadvantages of Using Apps to Support Physical Activity

The participants identified a number of advantages of PA apps, which facilitated engagement with the apps. These included the convenience that an app offers in terms of equipment required, cost, and not being required to attend a specific exercise facility:

> *Especially if you can, y'know, the workouts, like the Gorilla workouts that I've looked at so far, they're all just using your own body, where you don't need any special equipment, and all the rest of it…and you don't need to spend £30 a month to join a gym to do it.* [Male, aged 68 years, colorectal cancer]

> *You can just choose when you decide to do it – so you can think, "right, I'm gonna do a little workout now", so y'know, pick your moment, put your phone on and just pick whichever one you want.* [Female, aged 52 years, breast cancer]

They also commented that apps could be useful in building confidence or self-efficacy for PA and how this can be important in relation to side effects:

> *I was left with a lot of tummy problems after my treatment. So in a way you would think that doing a workout at home might suit a lot of people because if their confidence is low, either how they feel about their fitness or that they need to be near the loo or whatever, then being at home should be reassuring, shouldn't it?* [Female, aged 47 years, breast cancer]

Table 3. Sample characteristics (N=32).

Sample characteristics	n (%)
Gender	
Female	10 (31)
Male	22 (69)
Ethnicity	
White British	28 (88)
White-other	1 (3)
Asian/Asian British	2 (6)
Mixed	1 (3)
Cancer type	
Breast	8 (25)
Prostate	16 (50)
Colorectal	8 (25)
Experience of using digital technology to support PA[a]	
Never used	10 (31)
Mobile phone installed PA app (eg, Apple Health and SHealth)	5 (16)
Currently using a PA tracker (eg, pedometer, Fitbit, Garmen, and Strava)	9 (28)
Have used a PA tracker before but not currently using	5 (16)
Using combination of technologies (eg, mobile phone installed PA app + PA tracker)	3 (9)

[a]PA: physical activity.

It was also acknowledged that an app-based PA program could be more effective in comparison with printed materials due to the ubiquity of smartphones and the more engaging nature of interacting with the program in real time:

> *Where apps, of course, have a huge advantage, the days of paper things…exercise sheets, and things which end up in the bottom drawer or in the dustbin, err, apps are better than that, because they're on your phone, and they can be updated, as well...you've always got your phone with you. You haven't always got the list with you.* [Male, aged 69 years, prostate cancer]

> *...it's a bit more interactive and it's there and you can just...I'm gonna press on...whatever this...what's a box jump? For example, and you can press on that and see...see what it is, so it's very, very useful.* [Male, aged 69 years, prostate cancer]

Although only 1 participant mentioned the possible benefit of apps in terms of the level of literacy required to interact with the program, it is important to note that this could improve accessibility to a PA intervention through the visual and interactive features of the apps:

> *Y'know...it's a nice, simple app. You don't need to be that literate.* [Male, aged 60 years, prostate cancer]

However, a number of disadvantages of app-based PA interventions were also raised. These included the possible safety implications of unsupervised PA:

> *...if somebody isn't getting advice from a professional first and they're just picking up an app and...wanted to get a bit more active and doing it at home, I think that something like this could be actually be quite risky.* [Female, aged 43 years, breast cancer]

> *I think you'd have to be careful that people did it properly and that they did it at the right time and didn't…you know, didn't overdo it...some people think, 'ooh, well I'm doing exercise, it must be doing me good,' but it might not be…cause they're doing it too early, or they're doing it wrong...Because there's no supervision, there's no guarantee, is there?...That would be more for strength-based things, really [compared to walking].* [Female, aged 59 years, breast cancer]

Participants also experienced a number of technical issues (eg, impact on battery life, mobile data usage, and smartphone memory):

> *[Human] does drain your battery quite quickly because you have to use, erm, location services all the time...if it was gonna be a regular thing I wouldn't use it every day then just because the fact that it does drain your battery.* [Female, aged 37 years, colorectal cancer]

There were also concerns around data security and access to or usage of personal data:

> *...of course, with the freebies, as we know, what you're doing is you're signing up to allow them to*

> *track your location, other things you might be doing...nothing's really free.* [Male, aged 69 years, prostate cancer]

Perceived Relevance of the App

The participants described a number of factors that influenced their opinions of the perceived relevance of the apps used in this study. The participants described greater engagement with the apps that were perceived as most relevant to them. In relation to cancer, the participants acknowledged that they were a heterogeneous group who will differ in terms of their PA ability and that a successful app must be able to be tailored for this diversity to ensure it feels relevant to the user's ability:

> *Everybody who's had cancer will have a different level of fitness anyway even after cancer, and they'll have a different level of motivation and a different starting point so that's why that 7 minute app is good...you can choose...depending on where your starting point is.* [Female, aged 52 years, breast cancer]

The participants also highlighted that each individual's experience of cancer, treatment, and side effects differs and that a PA app to be used by cancer survivors must acknowledge the potential barriers that patients who have been diagnosed with various types of cancer and experienced different types of treatment and side effects might experience:

> *...depending on what treatment you've had, in terms of, umm, certainly operations, and scars and whether you've got adhesions or...weakened muscles in various places...it's all going to vary, from one cancer to another...there's a lot of variation and, err, that needs to be covered.* [Male, aged 68 years, colorectal cancer]

> *[Gorilla Workout] came up with something like….I can't remember what it said, but something like, "Don't be a slacker, get… you know, get working," or something, and I was like, "Err…hang on a minute." Like, if I'm feeling crap and I'm feeling fatigued, that is not what I want to see.* [Female, aged 38 years, breast cancer]

Furthermore, the participants also described that the types of PA that might feel appropriate or relevant to a cancer survivor could vary depending on where the patient is in their cancer *journey* (eg, diagnosis, treatment, recovery, and survivorship):

> *I had prostate cancer, and I had an operation. And, if you're looking at an app to try and get patients who've had cancer, y'know, back and fit again, I'm not sure that these exercises [on J&J and Gorilla Workout] were the right ones. I personally felt, that if I were being... had this been about six years ago [around time of treatment], they were too physical. I needed gentler exercises.* [Male, aged 70 years, prostate cancer]

However, there were also several noncancer-specific factors that influenced the perceived relevance of the app to the participants. These factors included the extent to which the app(s) aligned with the participants' PA goals:

> *I suppose it depends what you're trying to get out of it and, for me, it's looking at trying to regain a level of fitness, because I've probably lost it over the last four months or so. And I see the Seven Minute Workout as the one that will specifically do that whereas, [Human] is just monitoring what I will tend to do anyway.* [Male, aged 65 years, prostate cancer]

The extent to which the difficulty level of the app was suitable for the user also affected perceived relevance. This was particularly apparent for the strength- and resistance-based training apps:

> *[Human] was, as I say, very easy. It doesn't cause you any difficulties or problems. So I think anybody can use it. You know, it doesn't really matter how fit you are or how unfit you are, it's not going to be a problem...[with Gorilla Workout] I found, even on the easy level, that some of the exercises were impossible...Level 1 is you can perform 0-10 push-ups, but they still kind of think you're gonna be able to do some. It's, like, I can't do any. And I don't think I'm ever gonna be.* [Female, aged 43 years, breast cancer]

> *Um, and then [J&J] had things like press-ups and the plank. I mean, I just thought it was a joke, to be honest...I had a go on a couple of different days. Um, but it was, it was just much too difficult...I felt quite demoralised when they were so difficult. But yeah, something that's, um, you know, much more gentle to build up from, um, I think is quite a nice idea.* [Female, aged 47 years, breast cancer]

The participants also described that the way they interact with their mobile phone affects the perceived relevance of certain types of PA apps, namely activity trackers that require you to carry the smartphone to measure PA behavior:

> *[Human] assumes your phone is always on you…mine never is, unless I go out. So, it stays on the hall table...So of course, if it's left on the hall table, you're not moving around at all. So it'll say, "You're pretty inactive," y'know, "How about a walk around the block?" and you think, err, I've been doing the housework all morning. I'm exhausted.* [Male, aged 65 years, prostate cancer]

Finally, in terms of the participants' self-identity and their perception of whether the app fits with this identity affected their opinion of its perceived relevance:

> *And it is a man, isn't it, doing the exercises?...[J&J] was quite masculine, I think...I know it's a silly thing but even if it, if there was a choice of having a woman or a man to watch, you know.* [Female, aged 47 years, breast cancer]

> *And of course, umm, on both of them [J&J and Gorilla Workout]...the videos, err, show the sort of slim, fit young, ultra-fit, young men doing it. You think, "Gosh, I...I haven't looked like that for about 40 years."* [Male, aged 69 years, prostate cancer]

Perceived Quality of the App

The participants described several factors that affected their perceived quality of the apps to promote PA. The participants expressed greater engagement with the apps that were perceived to be higher quality, although they did not necessarily agree which those apps were. The factors affecting perceived quality differed between users.

Primarily, the users described the importance of ensuring that an app is easy to use and intuitive to foster engagement from the first usage:

> *...the bottom line is that...[The Walk's] not intuitive...Perhaps I should have looked for a help area, or something, if I wanted to make full use of it, but then I also think, if an app is gonna be good, then it, it needs to lend itself to the user...with Human, again, I didn't look out for any help areas. It's just, you start using it, it tells you what, what's going on, what you've done, and you can interpret it quite easily.* [Male, aged 51 years, prostate cancer]

The participants described the importance of ensuring that an app, which tracks PA behavior, does so accurately:

> *...the main issue I had was that [Human] would record activity, but it would get it wrong. So when I was out on a bike ride, umm, it had me doing a mixture of walking, cycling, umm or running...so I just felt that it didn't really work that well for me.* [Male, aged 68 years, colorectal cancer]

Furthermore, the participants' description of how well produced the app was affected their perceived quality of the apps:

> *I kept getting a bit confused with the voices. They weren't different enough in the story. Mainly, as I say, because, um, it was a bit frenetic and people were noisy and speaking quickly and it was a bit jumpy...and just the production of [The Walk], you know...it was a bit jumbled and thrown together almost.* [Female, aged 65 years, breast cancer]

The J&J app provided an explanation of the scientific evidence-base behind the recommended exercises and workout program, and this was described as increasing the perceived quality and credibility of the app to benefit health:

> *I did like the mass of support documentation you could delve down into to find out why the exercises were what they were, and the, umm, sort of, a bit of medical stuff behind it…I felt [J&J] was more medical-oriented…it was looking at your total body, total welfare – and I thought that it felt very professional…I felt the regime was based on good scientific basis.* [Male, aged 70 years, prostate cancer]

Opinions of Behavior Change Techniques Used to Promote Physical Activity

Opinions of BCTs used to promote PA within the apps were sought during the interviews and grouped into the following categories: "video demonstrations;" "prompts/cues (reminders);" "goal setting, self-monitoring, and feedback on behavior;" and "incentives, rewards, and gamification." Participants' views toward each of these strategies varied considerably, and their opinions on these BCTs determined the extent to which the participants engaged with the apps to which they were allocated.

Video Demonstrations

The use of video demonstrations to illustrate how to perform specific exercises correctly was well received:

> *...the method of presentation, brilliant. [J&J] was very clear...the bloke was there doing it with you...because you can sort of follow along, without just trying to remember how you should be doing it, and you can look at him to see how he's got his legs, straight or bent a bit.* [Male, aged 51 years, colorectal cancer]

Prompts/Cues (Reminders)

There was mixed feedback on the use of push notifications/reminders to prompt users to engage in PA and how effective they were. This depended on the users' opinion on reminders, their tone, and how appropriate they were in terms of the time or context in which they were delivered:

> *...mixed feelings about the sort of constant reminders [Human] gave you...it's quite good in some respects, because it does make you think, "Oh, yeah. Okay. I'll just go and have a quick walk to the end of the road and back." Err...But then when...three or four are coming, you're thinking, "Oh god, would you shut up?"...I didn't mind "Oh, what about a quick walk after lunch?" that sort of thing...they were quite positive.* [Female, aged 65 years, breast cancer]

> *...there was at least one of those prompts on Human, that actually we followed it. It said something like, "Let's go for a walk," and we said, "do you know what? Let's do that"...on other occasions, er, we said, well, actually, it's dark so we're not...you tend to start ignoring it 'cause it might not be appropriate at that time...so it wasn't a bad thing – but it wasn't always the right thing at the right time.* [Male, aged 65 years, prostate cancer]

Goal-Setting, Self-Monitoring, and Feedback on Behavior

These BCTs were grouped as they are frequently used alongside each other to promote PA. For instance, the Human app presents the daily 30-min PA goal, facilitates self-monitoring of progress toward the goal by presenting data collected by the smartphone's activity tracker, and then presents feedback on their behavior to indicate whether that goal was met or not. Therefore, it is difficult to separate out the participants' opinions of each of these BCTs individually; however, the participants generally responded positively to this approach to promote PA:

> *[Human] does show you like summaries and averages. It gives you some interesting information so you can see whether you're doing better or worse than you were doing yesterday and that kind of thing...it's nice to have a target and a challenge to work on.* [Female, aged 43 years, breast cancer]

> *I could see that I was actually walking more than I thought. So it all adds up...I think it is interesting to monitor because you can actually see how much you're doing, and...how quickly actually you reach your target. So you could think, like, "Oh, instead of half hour walking, maybe I could increase it to 45 minutes" or an hour if you want to push yourself. So I think that's definitely a benefit to monitor it...for me, just the data it was interesting and nice to see what I'm actually doing, and be more aware, and in that sense actually that...that already motivated me...to walk a bit extra instead of the bus...so in that sense...I did walk more with the app.* [Female, aged 54 years, breast cancer]

Some participants also discussed their positive experience of these types of BCTs using other digital technologies to support PA before this study:

> *I've just got the [Apple] Health one on my iPhone, which we check the steps every day. So because that's nicely how many steps you've done, how far you've done, and that 10,000 steps...we've both taken that on-board as a very good target...[which is] good because you could have a look and say, "Oh, crikey. I haven't done enough today" or "I haven't done enough this week," or whatever.* [Male, aged 69 years, prostate cancer]

> *I find the, you know, the completion of the steps quite satisfying...if I've got to the evening and I'm on, you know, nine thousand and something, I want to make sure I've got that to 10,000 if I walk up and down the stairs a few times, and then actually when you go over, you know, you do feel quite pleased with yourself...[Fitbit] would plot how many days you'd done, how many steps and what your average was for the week and what your average was for the month and that was quite rewarding, because you do feel like you are achieving something.* [Female, aged 47 years, breast cancer]

Incentives/Rewards and Gamification

There was mixed feedback on the use of incentives/rewards and gamification to increase engagement with the app and PA. This type of BCT was most prevalent in The Walk; however, participants were generally put off using this app by some of the usability issues mentioned above and the extent to which the app was perceived as relevant to them:

> *[The Walk's] trying to show you where, you could possibly take alternate...you could select to do a slightly longer walk, and have the chance of getting more points from other things. Like picking up packages, but I haven't really looked at that.* [Male, aged 60 years, prostate cancer]

Many of the participants said they felt that the gaming aspect to the app was inappropriate for them and they did not find it interesting:

> *I'm not interested in doing that, you know. I mean, even listening to [The Walk], it just got boring...I listened to it as I was walking along and I thought this is not for me really, you know, there was people missing here and people hiding there. I didn't know what it was talking about really. I'm not into that sort of thing.* [Male, aged 71 years, prostate cancer]

Apps That Promote Walking Are Most Appealing for Cancer Survivors

In acknowledging cancer survivors' varying needs (above), and incorporating their personal experience of cancer with their experience of using the apps in this study, the participants generally agreed that a walking-based app would be most appealing for cancer survivors. Walking was perceived to be safe, accessible, and achievable for the vast majority of people regardless of their ability, cancer type, treatment type, side effects, or where they are in their cancer journey. They also said that walking was enjoyable, which increased the likelihood that it would be sustainable and consequently effective:

> *First thing to do when you're coming back from the surgery, or any kind of treatment, I think walking is probably the safest way to introduce yourself back into [an] exercise routine.* [Male, aged 51 years, prostate cancer]

> *I couldn't use my upper body because of the surgery and then I had the chemo and I just couldn't go to the classes, so…but what I did do was walking, because I thought even if I can't do anything else you can always walk...if you really talk about something people can do right after or maybe even during treatment, I think walking is the easiest, the safest and the best way to start.* [Female, aged 54 years, breast cancer]

However, they did acknowledge the need to ensure that participants are engaging in PA that is of high enough intensity to meet the PA recommendations:

> *People might be having a 10 minute dawdle round the garden centre and think that they've done their exercise...I can see the sort of, the, the challenge with getting the balance, um, between the...it being achievable but also being effective isn't it?* [Female, aged 47 years, breast cancer]

Some participants recognized the importance of resistance training:

> *I think walking is very good, but equally I think it's overall, y'know, a balanced body strength and, and flexibility's important. So, I think it's worth persevering with that approach as well.* [Male, aged 68 years, colorectal cancer]

However, others reported that they did not enjoy or want to do these types of exercises:

> *I like the walking better than the exercises...the workouts and that sort of thing...I would hate to get...right into the heavy stuff, er, and tiring myself out, you know, cause we are getting older.* [Male, aged 70 years, prostate cancer]

> *I don't like doing exercise, and yet, as I mean in doing strengthening exercise and that sort of thing to build my muscle up, but I don't mind walking.* [Male, aged 70 years, colorectal cancer]

Physical Activity Apps Should Be Integrated Into Cancer Care

The participants agreed that routinely discussing PA and being directed toward onward support (including apps) within the cancer care pathway would ensure everyone diagnosed with cancer receives support. The participants discussed who would be best placed to direct them toward a PA app and when and how this should be introduced:

Patients Should Be Directed to Physical Activity Apps

Participants said that discussions around PA, including being directed toward resources to support behavior change (apps or otherwise), should be discussed with patients as a routine part of cancer care:

> *I think...there being some sort of formal introduction to the possibility of doing this, then rather it being sort of left for you to find it by yourself...that's what your expert's for.* [Male, aged 69 years, prostate cancer]

> *I don't think a lot of people would bother to go out and look, to see what apps they can find to do exercise. So, I think, if you're gonna do one, I think you've got to encourage somehow, you've got to encourage people to say, or to go, "Oh, that looks good. I'll use that one."* [Male, aged 70 years, prostate cancer]

Health Care Professionals' Recommendations Are Valued

There was a general consensus that the medical team, in particular the Clinical Nurse Specialist (CNS), would be best placed to discuss PA and possible interventions with patients. Participants reported feeling that they had built a relationship with their nurse and medical care team over the course of treatment and that they would trust the advice they provided as safe, accurate, and beneficial for their recovery:

> *The specialist nurses – so you always have a breast care specialist nurse who looks after you and if they started talking about it and telling you it was a good thing to do – I would have, I would have definitely done it...because you develop such a relationship with the specialist nurse who's in charge of your case.* [Female, aged 52 years, breast cancer]

> *The nurses. I was assigned a support nurse...She was very good at giving me advice, and support...If she had said to me, "Look, there's a jolly good app. You will need some ex...you need to get back into fitness again, you've had a big op...have a look at this one." I'd have taken that.* [Male, aged 70 years, prostate cancer]

This was also discussed in the context of the fear and uncertainty that is often raised when trying to increase PA post cancer and the potential for inaccurate and potentially unsafe information but that they would trust the medical team and CNS:

> *I didn't go to any of the support groups although I think they're a good idea, because people do get, you know, a lot from them. I do think it's dodgy if you haven't got a professional person there, because, as I found just sitting in...in the waiting room, um, you know, people have misconceptions...they've got their own ideas about their own treatment and their own health, and um, they start feeding people with, as I say, wrong information and wrong facts...so I was sort of aware that I'd just listen to what [the nurses] told me.* [Female, aged 65 years, breast cancer]

Some participants discussed the impact that receiving PA recommendations and feedback from trusted health professionals had on their subsequent participation in PA:

> *I had one of my check-ups with my consultant, and she said it might be a good time to introduce a tiny little bit of gentle exercise...and so from that point I then got a Fitbit and starting doing 10,000 steps a day, and by the next time I saw her I'd lost a stone and, um, she was very pleased really.* [Female, aged 47 years, breast cancer]

Other participants acknowledged that people seek information from different sources, in different ways, so having the information and direction toward an app available via a range of channels might be beneficial:

> *I think if you want to promote an app like this, it's, er, it's a good idea maybe to go, er, yeah, do it via various channels, so both a Clinical Nurse Specialist, er, the oncology physios, or charities, like, er, like Prostate Cancer UK or Breast Cancer Care.* [Female, 54 years, breast cancer]

Physical Activity Should Be Recommended Before and After Treatment

Participants suggested that PA interventions should be discussed at diagnosis or before treatment as a way to help manage or reduce side effects during treatment and after treatment to promote recovery and self-management:

> *I think if it...if it came as part of the pre-treatment package then I think that would be fantastic, 'cause you're already kind of…yes, you're in a state of shock, but if you're being given stuff to help and start playing with it before you actually start your treatment...because once you're in it, it's quite hard...and then another option, definitely after you finish treatment. Like, if you're feeling fatigued around radiotherapy time or after, definitely then.* [Female, aged 38 years, breast cancer]

> *What I've been trialling out [Human] that should be in your initial pack. So you...once you're diagnosed with the cancer, then you're given the pack and everything else, what to expect and go through, and I think it should be at that stage, as early as possible...that's the time you need that information.* [Male, aged 54 years, colorectal cancer]

Discussion

Principal Findings

The sample of breast, prostate, and colorectal cancer survivors interviewed in this qualitative study was receptive to the idea of apps to increase PA but highlighted that it is important to acknowledge the varying needs and preferences of this heterogeneous group. Participants recognized that the impact of cancer on each individual in terms of cancer type, treatment, prognosis, and experience of side effects can be very different, and successful app-based PA interventions must account for that diversity. The results demonstrate the subjective and dynamic nature of engagement with digital interventions and revealed factors that affected engagement for each individual (eg, their perceptions of the advantages and disadvantages of using apps to promote PA, relevance of the app, the quality of the app, and of the BCTs used to promote PA).

Participants recommended that walking would be the most appealing form of PA to recommend using an app and could be recommended at any stage across the cancer trajectory. This was because it was described as feeling safe, achievable, accessible, and enjoyable, regardless of cancer type, treatments received, or ability and could be used to increase confidence and fitness before incorporating strength/resistance-based training as recovery progresses. In terms of the strength/resistance-based training apps in this study (J&J and Gorilla Workout), there was a perception that even the beginner levels of these apps were too difficult and potentially unsafe, given the age, fitness level of many of the participants, in addition to their experience of side effects and recovery from cancer treatment. However, the participants were receptive to the format of these types of apps, with detailed video demonstrations illustrating how to perform each exercise. Activity tracking/walking-based apps did not provoke the same level of unease and the participants said that they felt that these need not be tailored specifically toward people who have had cancer. Although most participants recognized the benefit of strength- and resistance-based training, there was a consensus that apps that promote this type of PA would need to be tailored more specifically toward specific cancer types (eg, with regard to location of surgery) and for people with a lower starting level of ability, confidence, and familiarity with these types of exercises. Some participants also described strength and resistance training as unenjoyable and that they would be unlikely to adhere to these types of regimes. This illustrates the need to increase awareness about other ways of incorporating the strength and resistance training element of the PA recommendations in a way that is more enjoyable or feasible and might be more appealing to this group (eg, yoga, carrying shopping bags) compared with specific workout routines.

The participants suggested that to effectively direct cancer survivors toward an app-based PA intervention, this should be integrated within the existing cancer care pathway and recommended by their health care professionals, particularly CNSs. They described being directed toward an app within the medical setting as providing an opportunity to increase knowledge about the cancer-specific benefits of PA from a trusted source. The participants recommended that discussing PA/directing to ongoing support would be most beneficial before or after treatment, and particularly if it was highlighted as a way to alleviate side effects and promote recovery. They also felt that recommending walking specifically would be appropriate at any point after diagnosis for the majority of cancer survivors.

There is ongoing debate about the most appropriate, feasible, and effective way to support cancer survivors to increase PA within routine cancer care [45-49]. The results of our study support the use of existing PA apps to support low-risk moderate intensity PA (eg, walking) that could help cancer survivors to achieve the recommended minimum of 150 min of at least moderate-intensity PA per week [18-21]. However, one of the main issues of concern for the participants in this study was the lack of supervision and the potential for harm, particularly regarding the resistance training apps, especially for patients who are unfamiliar with these types of exercises or who might require specialist support. Although patients might receive more appropriate and tailored support if delivered and supervised by appropriate allied health professionals (eg, clinical exercise physiologists and physiotherapists) in specialist facilities [48] where adherence to the regimen can be monitored, there are issues regarding access and uptake [50]. A recent UK study found that despite national guidelines recommending that prostate cancer survivors treated with androgen deprivation therapy should receive 12 weeks of supervised exercise training, only 17% of National Health Service (NHS) trusts are able to provide this [51]. This reflects the lack of availability of these programs and the difficulty of implementation in routine care, particularly if uptake is poor. Future work should aim to better understand the potential for apps to support PA, which is likely to require greater involvement and supervision from exercise oncology specialists (eg, resistance training) and with greater adaptation/tailoring based on the individual's type of cancer, experience of treatment (eg, surgery, hormone therapy, chemotherapy, or radiotherapy) and associated consequences of treatment and side effects (eg, stoma, cachexia, or lymphedema). Greater supervision is also likely to be required for people with advanced/metastatic disease.

However, as highlighted by the participants in this study, there is little debate about the value that patients place on the recommendations provided by their clinical team, particularly the CNS and consultants [51-53]. Despite this, few cancer survivors receive PA recommendations or referrals to exercise programs within routine care [51,54]; health professionals report little discussion about PA with their patients and low awareness of PA recommendations for cancer survivors [52,55-57]. Therefore, it is crucial that oncology staff are supported to have discussions about PA with patients, direct them toward behavioral support to increase PA, and refer to specialist programs, where available. The implementation of recommendations to appropriate PA apps in cancer care requires greater exploration.

Most research in PA and cancer has been overrepresented by female cancer survivors' and primarily by women who have had breast cancer. For instance, in a meta-analysis exploring the effects of PA after cancer conducted by Fong et al [58], 25 of the 39 included studies were conducted exclusively in breast

cancer patients. Although only 6 of the 15 studies included in our review exploring the impact of digital interventions on PA in cancer survivors were conducted exclusively with breast/endometrial cancer survivors, the other 9 studies were all overrepresented by female participants [35]. However, in this study, 69% (22/32) of our sample were male, driven by the 50% of our participants with prostate cancer. It would be interesting to explore the demographic characteristics or particular cancer types for which PA apps are most appealing on a larger scale.

Our approach, enabling participants to experience searching for, downloading, and using selected apps *in the wild* for a period of time, proved to be a time- and resource-efficient method, allowing us to understand how cancer survivors actually experience different types of apps and BCTs. We suggest this provides greater ecological validity than previous studies in the area that have, for instance, sought feedback of hypothetical app features and example text messages from slideshows shown to focus groups of cancer survivors [40]. Digital health research has come to appreciate the importance of usability, design, and tailoring for engagement [38,59]; however, recent reviews have conceptualized engagement with digital health interventions more broadly [42,60]. These reviews have highlighted factors such as personal agency and motivation, personal life and values, the engagement and recruitment approach, and the quality of the digital health intervention [60] and the delivery method (eg, aesthetics/design, ease of use, personalization, and message tone), content (eg, BCTs such as feedback and reminders), the population (eg, demographic characteristics, personal relevance, and self-efficacy), and both the social (eg, norms and social cues) and physical (eg, health care system, location, and time) settings as being important for engagement [42]. Our methodology has allowed us to demonstrate these broader influences on engagement, and we suggest that this methodology could be useful in the development and evaluation of other mobile health (mHealth) products for other health conditions and other health behaviors.

But, how should we respond to the demand from participants for highly tailored interventions that feel relevant to each individual user? Will it be more appropriate to identify/develop a number of PA apps that are suitable for different groups of cancer survivors and from which they could choose the one they think is most suited to them rather than attempting to develop one app that is flexible enough to meet all needs and preferences of a heterogeneous group of individuals? Should we focus on making apps that are cancer specific, or choosing among existing noncancer-specific apps and focusing on how the app is introduced to the individual? In light of this challenge, Short et al [38] have developed a PA app referral scheme to select the most appropriate publicly available, noncancer-specific PA app for a cancer survivor based on a referral matrix, taking into account the participant's fitness level, PA interests, app preferences, and personality characteristics [41]. This novel approach to evaluation of multiple PA apps within a referral scheme takes advantage of the large number of appropriate and relevant publicly available PA interventions, while offering flexibility, choice, and tailoring to the users' needs and preferences.

Limitations

This study should be viewed in light of a number of limitations. The sample was self-selecting. This led to a high proportion of participants who were already physically active and who were interested in technology and their health and recovery. We did not quantify the participants' current level of PA; however, none of the participants reported being completely inactive. Although this study intended to explore initial opinions of the use of PA apps among cancer survivors, we need to understand the views of those who are inactive or engaging in very little PA, who might feel less confident in engaging in PA or using apps, and who might be unaware of the benefits of PA postcancer diagnosis. Our approach to recruitment means we cannot estimate the number of eligible people who saw the advertisements versus those who responded. Although the participants in this study were able to use the selected apps for between 2 and 3 weeks, a more realistic experience than discussing hypothetical app features in a single session, this does not completely reflect *real-life* app usage or engagement. Participants did not choose the apps, and we did not assess experiences in the longer term. This might be amplified by the fact the participants knew they were taking part in a research study and so might have been more inclined to persevere with some of the apps they disliked and may have discontinued using otherwise.

Conclusions

In conclusion, this sample of breast, prostate, and colorectal cancer survivors were receptive to the use of apps to promote PA but felt that for apps to be effective among this group, they must feel relevant to the individual. This includes accounting for the needs of those who have been diagnosed with different types of cancer, experienced different types of treatment and side effects, and have different levels of PA ability. Walking was highlighted as the most appealing type of PA to promote via an app as it is perceived as safe, achievable, accessible, and enjoyable. We suggest it is useful to also consider the impact of the users' perception of the relevance of an app and how an app relates to their self-identity. This can arise from the app features, but might also be affected by how the app is introduced (eg, by a trusted health professional). Digital health research has come to appreciate the importance of usability and its impact on engagement. Our methodology has allowed us to demonstrate the broader and more dynamic influences on engagement with apps, and we believe this work could, therefore, generalize to evaluations of mHealth products for other health conditions and other health behaviors.

Acknowledgments

ALR is funded by a Medical Research Council PhD studentship. AF, HWWP, and LS are funded by the Higher Education Funding Council for England. DAK is funded by the National Institute for Health Research (NIHR) Oxford Biomedical Research Centre. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. The authors have received no funding from any organizations that might have an interest in the presented work.

Conflicts of Interest

The authors have no personal financial interests related to the presented work. AF, HWWP, and LS currently work alongside Six to Start (the developers of the app, The Walk, used in this study) on an unrelated research project. HWWP has a current PhD student (unrelated to this study) who was employed by Johnson & Johnson and whose employment ceased before the beginning of this study. HWWP has received consultancy fees from Crystallise, System Analytic, and The HELP Trust and received funding from myownteam and Shift.ms. These relationships/activities were unrelated to the presented work or analyses. All other authors have no conflicts of interest to declare.

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015 Mar 1;136(5):E359-E386. [doi: 10.1002/ijc.29210] [Medline: 25220842]
2. Weis J. Cancer-related fatigue: prevalence, assessment and treatment strategies. Expert Rev Pharmacoecon Outcomes Res 2011 Aug;11(4):441-446. [doi: 10.1586/erp.11.44] [Medline: 21831025]
3. van den Beuken-van Everdingen MH, De Rijke JM, Kessels AG, Schouten HC, Van Kleef M, Patijn J. Prevalence of pain in patients with cancer: a systematic review of the past 40 years. Ann Oncol 2007 Sep;18(9):1437-1449 [FREE Full text] [doi: 10.1093/annonc/mdm056] [Medline: 17355955]
4. Sharma N, Hansen CH, O'Connor M, Thekkumpurath P, Walker J, Kleiboer A, et al. Sleep problems in cancer patients: prevalence and association with distress and pain. Psychooncology 2012 Sep;21(9):1003-1009. [doi: 10.1002/pon.2004] [Medline: 21726015]
5. Koo H, Seo Y, Cho M, Kim M, Choi H. Weight change and associated factors in long-term breast cancer survivors. PLoS One 2016;11(7):e0159098 [FREE Full text] [doi: 10.1371/journal.pone.0159098] [Medline: 27391162]
6. Smith MR. Changes in fat and lean body mass during androgen-deprivation therapy for prostate cancer. Urology 2004 Apr;63(4):742-745. [doi: 10.1016/j.urology.2003.10.063] [Medline: 15072892]
7. Krebber AM, Buffart LM, Kleijn G, Riepma IC, de Bree R, Leemans CR, et al. Prevalence of depression in cancer patients: a meta-analysis of diagnostic interviews and self-report instruments. Psychooncology 2014 Feb;23(2):121-130 [FREE Full text] [doi: 10.1002/pon.3409] [Medline: 24105788]
8. Mitchell AJ, Ferguson DW, Gill J, Paul J, Symonds P. Depression and anxiety in long-term cancer survivors compared with spouses and healthy controls: a systematic review and meta-analysis. Lancet Oncology 2013 Jul;14(8):721-732. [doi: 10.1016/S1470-2045(13)70244-4]
9. Macmillan Cancer Support. 2015. The burden of cancer and other long-term health conditions URL: http://www.macmillan.org.uk/documents/press/cancerandotherlong-termconditions.pdf [accessed 2016-09-09] [WebCite Cache ID 6kOLbNdhG]
10. Hayes SC, Janda M, Cornish B, Battistutta D, Newman B. Lymphedema after breast cancer: incidence, risk factors, and effect on upper body function. J Clin Oncol 2008 Jul 20;26(21):3536-3542 [FREE Full text] [doi: 10.1200/JCO.2007.14.4899] [Medline: 18640935]
11. Peeters KC, van de Velde CJ, Leer JW, Martijn H, Junggeburt JM, Kranenbarg EK, et al. Late side effects of short-course preoperative radiotherapy combined with total mesorectal excision for rectal cancer: increased bowel dysfunction in irradiated patients--a Dutch colorectal cancer group study. J Clin Oncol 2005 Sep 01;23(25):6199-6206. [doi: 10.1200/JCO.2005.14.779] [Medline: 16135487]
12. Miller DC, Sanda MG, Dunn RL, Montie JE, Pimentel H, Sandler HM, et al. Long-term outcomes among localized prostate cancer survivors: health-related quality-of-life changes after radical prostatectomy, external radiation, and brachytherapy. J Clin Oncol 2005 Apr 20;23(12):2772-2780. [doi: 10.1200/JCO.2005.07.116] [Medline: 15837992]
13. Quinten C, Coens C, Ghislain I, Zikos E, Sprangers MA, Ringash J, PROBE, EORTC Clinical Groups. The effects of age on health-related quality of life in cancer populations: a pooled analysis of randomized controlled trials using the European Organisation for Research and Treatment of Cancer (EORTC) QLQ-C30 involving 6024 cancer patients. Eur J Cancer 2015 Dec;51(18):2808-2819. [doi: 10.1016/j.ejca.2015.08.027] [Medline: 26602015]
14. Lahart IM, Metsios GS, Nevill AM, Carmichael AR. Physical activity, risk of death and recurrence in breast cancer survivors: a systematic review and meta-analysis of epidemiological studies. Acta Oncol 2015 May;54(5):635-654. [doi: 10.3109/0284186X.2014.998275] [Medline: 25752971]

15. Schmid D, Leitzmann MF. Association between physical activity and mortality among breast cancer and colorectal cancer survivors: a systematic review and meta-analysis. Ann Oncol 2014 Jul;25(7):1293-1311. [doi: 10.1093/annonc/mdu012] [Medline: 24644304]
16. Friedenreich CM, Neilson HK, Farris MS, Courneya KS. Physical activity and cancer outcomes: a precision medicine approach. Clin Cancer Res 2016 Oct 01;22(19):4766-4775 [FREE Full text] [doi: 10.1158/1078-0432.CCR-16-0067] [Medline: 27407093]
17. Mishra SI, Scherer RW, Geigle PM, Berlanstein DR, Topaloglu O, Gotay CC, et al. Exercise interventions on health-related quality of life for cancer survivors. Cochrane Database Syst Rev 2012;8:CD007566. [doi: 10.1002/14651858.CD007566.pub2] [Medline: 22895961]
18. Rock CL, Doyle C, Demark-Wahnefried W, Meyerhardt J, Courneya KS, Schwartz AL, et al. Nutrition and physical activity guidelines for cancer survivors. CA Cancer J Clin 2012;62(4):243-274 [FREE Full text] [doi: 10.3322/caac.21142] [Medline: 22539238]
19. US Department of Health and Human Services. 2008. Physical Activity Guidelines for Americans URL: https://health.gov/paguidelines/pdf/paguide.pdf [accessed 2018-04-30] [WebCite Cache ID 6z4M5lhF5]
20. Schmitz KH, Courneya KS, Matthews C, Demark-Wahnefried W, Galvão DA, Pinto BM, American College of Sports Medicine. American College of Sports Medicine roundtable on exercise guidelines for cancer survivors. Med Sci Sports Exerc 2010 Jul;42(7):1409-1426. [doi: 10.1249/MSS.0b013e3181e0c112] [Medline: 20559064]
21. Denlinger CS, Ligibel JA, Are M, Baker KS, Demark-Wahnefried W, Dizon D, National comprehensive cancer network. Survivorship: healthy lifestyles, version 2.2014. J Natl Compr Canc Netw 2014 Sep;12(9):1222-1237 [FREE Full text] [Medline: 25190692]
22. Boyle T, Vallance JK, Ransom EK, Lynch BM. How sedentary and physically active are breast cancer survivors, and which population subgroups have higher or lower levels of these behaviors? Support Care Cancer 2016 May;24(5):2181-2190. [doi: 10.1007/s00520-015-3011-3] [Medline: 26563180]
23. Wang Z, McLoone P, Morrison DS. Diet, exercise, obesity, smoking and alcohol consumption in cancer survivors and the general population: a comparative study of 16 282 individuals. Br J Cancer 2015 Feb 3;112(3):572-575 [FREE Full text] [doi: 10.1038/bjc.2014.598] [Medline: 25429527]
24. Mason C, Alfano CM, Smith AW, Wang C, Neuhouser ML, Duggan C, et al. Long-term physical activity trends in breast cancer survivors. Cancer Epidemiol Biomarkers Prev 2013 Jun;22(6):1153-1161 [FREE Full text] [doi: 10.1158/1055-9965.EPI-13-0141] [Medline: 23576689]
25. Sander AP, Wilson J, Izzo N, Mountford SA, Hayes KW. Factors that affect decisions about physical activity and exercise in survivors of breast cancer: a qualitative study. Phys Ther 2012 Apr;92(4):525-536 [FREE Full text] [doi: 10.2522/ptj.20110115] [Medline: 22156026]
26. Paxton RJ, Anderson A, Sarkar S, Taylor WC. Breaking up sedentary behavior: perceptions from cancer survivors. Cancer Nurs 2016;39(4):272-278 [FREE Full text] [doi: 10.1097/NCC.0000000000000330] [Medline: 26713501]
27. Fisher A, Wardle J, Beeken RJ, Croker H, Williams K, Grimmett C. Perceived barriers and benefits to physical activity in colorectal cancer patients. Support Care Cancer 2016 Feb;24(2):903-910 [FREE Full text] [doi: 10.1007/s00520-015-2860-0] [Medline: 26268781]
28. Wurz A, St-Aubin A, Brunet J. Breast cancer survivors' barriers and motives for participating in a group-based physical activity program offered in the community. Support Care Cancer 2015 Aug;23(8):2407-2416 [FREE Full text] [doi: 10.1007/s00520-014-2596-2] [Medline: 25605568]
29. Koutoukidis DA, Beeken RJ, Manchanda R, Michalopoulou M, Burnell M, Knobf MT, et al. Recruitment, adherence, and retention of endometrial cancer survivors in a behavioural lifestyle programme: the Diet and Exercise in Uterine Cancer Survivors (DEUS) parallel randomised pilot trial. BMJ Open 2017 Oct 08;7(10):e018015 [FREE Full text] [doi: 10.1136/bmjopen-2017-018015] [Medline: 28993394]
30. National Cancer Institute. Age and Cancer Risk URL: https://www.cancer.gov/about-cancer/causes-prevention/risk/age [accessed 2017-11-15] [WebCite Cache ID 6uzpF5HnL]
31. Ofcom. 2015. The Communications Market Report URL: http://stakeholders.ofcom.org.uk/binaries/research/cmr/cmr15/CMR_UK_2015.pdf [accessed 2015-11-19] [WebCite Cache ID 6d9WicRgI]
32. Ofcom. 2017. The Communications Market Report URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0017/105074/cmr-2017-uk.pdf [accessed 2017-11-15] [WebCite Cache ID 6uzkykPGY]
33. Statista. 2018. Share of adults in the United States who owned a smartphone from 2015 to 2018, by age group URL: https://www.statista.com/statistics/489255/percentage-of-us-smartphone-owners-by-age-group/ [accessed 2018-04-25] [WebCite Cache ID 6ywSOdGLO]
34. Jahangiry L, Farhangi MA, Shab-Bidar S, Rezaei F, Pashaei T. Web-based physical activity interventions: a systematic review and meta-analysis of randomized controlled trials. Public Health 2017 Nov;152:36-46. [doi: 10.1016/j.puhe.2017.06.005] [Medline: 28734170]
35. Roberts AL, Fisher A, Smith L, Heinrich M, Potts HW. Digital health behaviour change interventions targeting physical activity and diet in cancer survivors: a systematic review and meta-analysis. J Cancer Surviv 2017 Dec;11(6):704-719 [FREE Full text] [doi: 10.1007/s11764-017-0632-1] [Medline: 28779220]

36. Puszkiewicz P, Roberts AL, Smith L, Wardle J, Fisher A. Assessment of cancer survivors' experiences of using a publicly available physical activity mobile application. JMIR Cancer 2016 May 31;2(1):e7 [FREE Full text] [doi: 10.2196/cancer.5380] [Medline: 28410168]
37. McCarroll ML, Armbruster S, Pohle-Krauza RJ, Lyzen AM, Min S, Nash DW, et al. Feasibility of a lifestyle intervention for overweight/obese endometrial and breast cancer survivors using an interactive mobile application. Gynecol Oncol 2015 Jun;137(3):508-515. [doi: 10.1016/j.ygyno.2014.12.025] [Medline: 25681782]
38. Bradbury K, Watts S, Arden-Close E, Yardley L, Lewith G. Developing digital interventions: a methodological guide. Evid Based Complement Alternat Med 2014;2014:561320 [FREE Full text] [doi: 10.1155/2014/561320] [Medline: 24648848]
39. Yardley L, Morrison L, Bradbury K, Muller I. The person-based approach to intervention development: application to digital health-related behavior change interventions. J Med Internet Res 2015 Jan;17(1):e30 [FREE Full text] [doi: 10.2196/jmir.4055] [Medline: 25639757]
40. Robertson MC, Tsai E, Lyons EJ, Srinivasan S, Swartz MC, Baum ML, et al. Mobile health physical activity intervention preferences in cancer survivors: a qualitative study. JMIR Mhealth Uhealth 2017 Jan 24;5(1):e3 [FREE Full text] [doi: 10.2196/mhealth.6970] [Medline: 28119278]
41. Short CE, Finlay A, Sanders I, Maher C. Development and pilot evaluation of a clinic-based mHealth app referral service to support adult cancer survivors increase their participation in physical activity using publicly available mobile apps. BMC Health Serv Res 2018 Dec 16;18(1):27 [FREE Full text] [doi: 10.1186/s12913-017-2818-7] [Medline: 29338722]
42. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Dec;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
43. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
44. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
45. Hardcastle SJ, Cohen PA. Effective physical activity promotion to survivors of cancer is likely to be home based and to require oncologist participation. J Clin Oncol 2017 Nov 10;35(32):3635-3637. [doi: 10.1200/JCO.2017.74.6032] [Medline: 28915086]
46. Adams SC, Iyengar NM, Scott JM, Jones LW. Exercise implementation in oncology: one size does not fit all. J Clin Oncol 2018 Mar 20;36(9):925-926. [doi: 10.1200/JCO.2017.76.2906] [Medline: 29373093]
47. Lopez C, Jones J, Alibhai SM, Santa MD. What is the “home” in home-based exercise? The need to define independent exercise for survivors of cancer. J Clin Oncol 2018 Mar 20;36(9):926-927. [doi: 10.1200/JCO.2017.76.4365] [Medline: 29373096]
48. Newton RU, Taaffe DR, Chambers SK, Spry N, Galvão DA. Effective exercise interventions for patients and survivors of cancer should be supervised, targeted, and prescribed with referrals from oncologists and general physicians. J Clin Oncol 2018 Mar 20;36(9):927-928. [doi: 10.1200/JCO.2017.76.7400] [Medline: 29373097]
49. Hardcastle SJ, Cohen PA. Reply to S.C. Adams et al, C. Lopez et al, and R.U. Newton et al. J Clin Oncol 2018 Mar 20;36(9):928-930. [doi: 10.1200/JCO.2017.76.8218] [Medline: 29373095]
50. Queen M, Karatzaferi C, Bloxham SR, Panwar U, Drew P, Barton AG, et al. How can physical activity referral rates for breast cancer patients be increased? Front Oncol 2016;6:198 [FREE Full text] [doi: 10.3389/fonc.2016.00198] [Medline: 27672620]
51. Bourke L, Turner R, Greasley R, Sutton E, Steed L, Smith D, STAMINA investigators. A multi-centre investigation of delivering national guidelines on exercise training for men with advanced prostate cancer undergoing androgen deprivation therapy in the UK NHS. PLoS One 2018;13(7):e0197606 [FREE Full text] [doi: 10.1371/journal.pone.0197606] [Medline: 29975707]
52. Koutoukidis DA, Beeken RJ, Lopes S, Knobf MT, Lanceley A. Attitudes, challenges and needs about diet and physical activity in endometrial cancer survivors: a qualitative study. Eur J Cancer Care (Engl) 2017 Nov;26(6). [doi: 10.1111/ecc.12531] [Medline: 27324208]
53. Hardcastle SJ, Glassey R, Salfinger S, Tan J, Cohen P. Factors influencing participation in health behaviors in endometrial cancer survivors. Psychooncology 2017 Aug;26(8):1099-1104. [doi: 10.1002/pon.4288] [Medline: 27665487]
54. Fisher A, Williams K, Beeken R, Wardle J. Recall of physical activity advice was associated with higher levels of physical activity in colorectal cancer patients. BMJ Open 2015 Apr 28;5(4):e006853 [FREE Full text] [Medline: 25922098]
55. Nadler M, Bainbridge D, Tomasone J, Cheifetz O, Juergens RA, Sussman J. Oncology care provider perspectives on exercise promotion in people with cancer: an examination of knowledge, practices, barriers, and facilitators. Support Care Cancer 2017 Dec;25(7):2297-2304. [doi: 10.1007/s00520-017-3640-9] [Medline: 28258503]
56. Williams K, Beeken RJ, Fisher A, Wardle J. Health professionals' provision of lifestyle advice in the oncology context in the United Kingdom. Eur J Cancer Care (Engl) 2015 Jul;24(4):522-530. [doi: 10.1111/ecc.12305] [Medline: 25732397]

57. Daley AJ, Bowden SJ, Rea DW, Billingham L, Carmicheal AR. What advice are oncologists and surgeons in the United Kingdom giving to breast cancer patients about physical activity? Int J Behav Nutr Phys Act 2008;5:46 [FREE Full text] [doi: 10.1186/1479-5868-5-46] [Medline: 18803812]
58. Fong DY, Ho JW, Hui BP, Lee AM, Macfarlane DJ, Leung SS, et al. Physical activity for cancer survivors: meta-analysis of randomised controlled trials. Br Med J 2012;344:e70 [FREE Full text] [Medline: 22294757]
59. Garnett C, Crane D, West R, Brown J, Michie S. Identification of behavior change techniques and engagement strategies to design a smartphone app to reduce alcohol consumption using a formal consensus method. JMIR Mhealth Uhealth 2015;3(2):e73 [FREE Full text] [doi: 10.2196/mhealth.3895] [Medline: 26123578]
60. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Sep 15;16(1):120 [FREE Full text] [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]

Abbreviations

BCT: behavior change technique
CNS: clinical nurse specialist
J&J: the Johnson & Johnson Official 7 Minute Workout
mHealth: mobile health
NHS: National Health Service
NIHR: National Institute for Health Research
PA: physical activity
QoL: quality of life

Edited by G Eysenbach; submitted 01.05.18; peer-reviewed by C Short, F Mckay, C Valle; comments to author 28.06.18; revised version received 30.07.18; accepted 10.09.18; published 04.01.19.

Please cite as:
Roberts AL, Potts HWW, Koutoukidis DA, Smith L, Fisher A
Breast, Prostate, and Colorectal Cancer Survivors' Experiences of Using Publicly Available Physical Activity Mobile Apps: Qualitative Study
JMIR Mhealth Uhealth 2019;7(1):e10918
URL: http://mhealth.jmir.org/2019/1/e10918/
doi:10.2196/10918
PMID:30609982

Original Paper

Health Benefits and Cost-Effectiveness From Promoting Smartphone Apps for Weight Loss: Multistate Life Table Modeling

Christine Cleghorn[1], BSc, MSc, PhD; Nick Wilson[1], DIH, MBChB, MPH; Nisha Nair[1], MD, MPH; Giorgi Kvizhinadze[1], BSc, MSc, PhD; Nhung Nghiem[1], BEcon, MEcon, PhD; Melissa McLeod[1], MBChB, MPH, FNZCPHM; Tony Blakely[1], MBChB, MPH, PhD, FAFPHM

BODE[3] Programme, Department of Public Health, University of Otago, Wellington, New Zealand

Corresponding Author:
Christine Cleghorn, BSc, MSc, PhD
BODE³ Programme
Department of Public Health
University of Otago
PO Box 7343
Wellington South
Wellington, 6242
New Zealand
Phone: 64 4 918 6182 ext 6182
Email: cristina.cleghorn@otago.ac.nz

Abstract

Background: Obesity is an important risk factor for many chronic diseases. Mobile health interventions such as smartphone apps can potentially provide a convenient low-cost addition to other obesity reduction strategies.

Objective: This study aimed to estimate the impacts on quality-adjusted life-years (QALYs) gained and health system costs over the remainder of the life span of the New Zealand population (N=4.4 million) for a smartphone app promotion intervention in 1 calendar year (2011) using currently available apps for weight loss.

Methods: The intervention was a national mass media promotion of selected smartphone apps for weight loss compared with no dedicated promotion. A multistate life table model including 14 body mass index–related diseases was used to estimate QALYs gained and health systems costs. A lifetime horizon, 3% discount rate, and health system perspective were used. The proportion of the target population receiving the intervention (1.36%) was calculated using the best evidence for the proportion who have access to smartphones, are likely to see the mass media campaign promoting the app, are likely to download a weight loss app, and are likely to continue using this app.

Results: In the base-case model, the smartphone app promotion intervention generated 29 QALYs (95% uncertainty interval, UI: 14-52) and cost the health system US $1.6 million (95% UI: 1.1-2.0 million) with the standard download rate. Under plausible assumptions, QALYs increased to 59 (95% UI: 27-107) and costs decreased to US $1.2 million (95% UI: 0.5-1.8) when standard download rates were doubled. Costs per QALY gained were US $53,600 for the standard download rate and US $20,100 when download rates were doubled. On the basis of a threshold of US $30,000 per QALY, this intervention was cost-effective for Māori when the standard download rates were increased by 50% and also for the total population when download rates were doubled.

Conclusions: In this modeling study, the mass media promotion of a smartphone app for weight loss produced relatively small health gains on a population level and was of borderline cost-effectiveness for the total population. Nevertheless, the scope for this type of intervention may expand with increasing smartphone use, more easy-to-use and effective apps becoming available, and with recommendations to use such apps being integrated into dietary counseling by health workers.

(JMIR Mhealth Uhealth 2019;7(1):e11118) doi:10.2196/11118

KEYWORDS

weight loss diet; telemedicine; smartphone; cost-utility analysis; life tables; quality-adjusted life years

Introduction

Obesity is an important risk factor for many chronic diseases that impact people's quality of life and incur substantial health system costs. Obesity is an established risk factor for cardiovascular disease (CVD), diabetes, osteoarthritis, and various cancers [1].

Mobile health (mHealth) has been defined as "the application of mobile technologies, including phones, tablets, telemonitoring, and tracking devices, to support and enhance the performance of health care and public health practice" [2]. In the modeling study presented here, mHealth refers to using smartphone apps to deliver diet, exercise and health information, and behavior change support to participants to help them lose weight. mHealth tools can be accessed at people's convenience, from their homes or using their phones on the go. These interventions, therefore, have the potential to provide more regular information than face-to-face weight loss programs and may therefore be an important low-cost addition to current obesity reduction strategies.

mHealth technologies have the potential to improve public health in the future despite the current absence of strong evidence of effectiveness [3]. The evidence base for the effectiveness of smartphone apps for weight loss is, however, growing. A systematic review including interventions that utilized smartphone apps, text messaging, and Web resources was published in 2014 [4]. It included 12 primary studies of randomized controlled trials (RCTs) investigating mHealth weight loss interventions through diet and physical activity. The meta-analysis of these 12 studies estimated an additional 0.43 kg weight loss (95% CI: 0.25-0.61) for the intervention groups compared with controls. However, evidence on how long and at what magnitude this weight loss persists is poor, an important source of uncertainty we explore in this paper as it will have an impact on future health gains and costs or cost savings.

Another key determinant of overall population impact is the uptake of smartphone apps. Uptake will depend on marketing, placement, and word-of-mouth. Governments or health system providers could also promote use of apps, particularly if such promotion increases health benefits cost-effectively. Therefore, the aim of this study was to estimate the likely future health impacts, costs, and cost-effectiveness of mass media promotion of mHealth programs that use smartphone apps to deliver health information and behavior change support to participants for weight loss, compared with the existing levels of promotion and use of mHealth in a fairly typical developed country setting: New Zealand. A secondary aim was to identify targets for future research that would improve the precision of cost-effectiveness modeling for these types of interventions.

Methods

Overview

Outputs from this modeling include incremental quality-adjusted life-years (QALYs) gained and costs or cost savings in New Zealand dollars (NZ $). Outputs were discounted at 3%, with 0% and 6% used in scenario analyses. A health system perspective was used, and benefits and costs were modeled using a lifetime horizon. The intervention was modeled as a one-off intervention implemented in 1 year in overweight and obese New Zealand adults.

To estimate the difference in QALYs and health system costs between the model's intervention and business-as-usual (BAU) comparator, the entire New Zealand population, alive in 2011, was simulated out until death using a dietary multistate life table (MSLT) model built in Excel. The structure and BAU inputs for this generic model are described in detail in the model's technical report (see Multimedia Appendix 1 [5]). The remainder of this Methods section provides a summary of general structure and BAU inputs and more details on specific intervention parameters for this mHealth intervention.

The BAU comparator was assumed to include the existing level of mHealth promotion—which is negligible in New Zealand (ie, we are aware of no health agencies in the country that promote specific smartphone apps for weight loss). Therefore, we did not strip the baseline component of the model back to a hypothetical "no mHealth" comparator. Briefly, the BAU model uses projected all-cause mortality and morbidity rates by sex and age and separately for Māori (indigenous population) and non-Māori ethnic groups. Running alongside this main life table were 14 body mass index (BMI)–related disease life tables, where proportions of the population simultaneously resided: coronary heart disease (CHD), stroke, type 2 diabetes, osteoarthritis, and multiple cancers (ie, endometrial, kidney, liver, esophageal, pancreatic, thyroid, colorectal, breast, ovarian, and gallbladder). The proportion of the New Zealand population in each disease life table was a function of the disease incidence, case fatality, and remission (the latter in cancers only).

The intervention was modeled as a one-off smartphone app promotion that occurred in year one, 2011. The intervention effect was captured through changes in BMI resulting from the mHealth intervention. The change in BMI was then combined with relative risks for the associations between BMI and diseases through population impact fractions (PIFs; percentage reductions in future BMI-related disease incidence) that alter the inflow to the BMI-related disease life tables. Time lags from change in BMI to change in disease incidence were allowed for by using the average BMI change over a previous window of time of 0 to 5 years for CVD, diabetes, and osteoarthritis and 10 to 30 years for cancers. Probabilistic uncertainty about the boundaries (5, 10, and 30 years) was also specified (see Multimedia Appendix 1 [5]).

Input Parameters

Business-As-Usual Parameters

All input parameters (specified by sex, age, and ethnicity unless stated differently) are shown in Table 1 and described in more detail in Multimedia Appendix 1 [5]. Briefly, each BMI-related disease had incidence, prevalence, and case fatality in 2011. Remission rates were specified for cancers but set to 0 for chronic diseases of CHD, stroke, type 2 diabetes, and osteoarthritis (ie, lifetime diagnoses). These parameters were calculated using DISMOD II (World Health Organization

2001-2009, created by Jan J Barendregt) [6], which is a program used to calculate epidemiologically and mathematically coherent sets of parameters for each disease. Future trends in cancer incidence, case fatality, and remission were specified using regression estimates of trends from historic data. Trends in other diseases were obtained from the New Zealand Burden of Disease Study (NZBDS) [7].

Morbidity was quantified (separately by sex, age, and ethnic groups) for each disease using the years of life lived with disability (YLDs) from the NZBDS, divided by the population count to give prevalent YLDs. Disability weights from the Global Burden of Disease Study 2010 were used to estimate the health status valuation of these YLDs [8].

Health system costs (sex- and age-specific) were calculated in 2011 in NZ $ using individually linked data for publicly funded (and some privately funded) health events occurring in 2006 to 2010, including hospitalizations, inpatient procedures, outpatients, pharmaceuticals, laboratories, and expected primary care usage. Building on an existing framework [9] for calculating the timing of health system costs, the whole cohort was assigned an (sex- and age-specific) annual health system cost of a citizen without a BMI-related disease and not in the last 6 months of their life. Additional disease-specific excess costs were assigned to people (1) in the first year of a BMI-related disease diagnosis, (2) in the last 6 months of life if dying of the given disease, and (3) otherwise prevalent cases of each disease. Costs were modeled over the lifetime of the cohort, including costs both related and unrelated to the BMI-related diseases modeled (meaning increased longevity because of weight loss interventions contributes to increased health system costs for some cohort members). Organisation for Economic Cooperation and Development purchasing power parity for 2011 was used when costs were converted to US dollars (US $1.486 to NZ $1).

Intervention Parameters

In this study, mHealth programs are those that use smartphone apps to deliver health information and behavior change support to participants for weight loss. It was assumed that smartphone apps would be promoted nationally through 1 main medium. First, weblinks to the best 5 iOS and 5 Android weight loss apps (all under NZ $4 to download), as recently identified through a New Zealand study [12], would be displayed on the Ministry of Health (MoH) and other health promotion websites. These apps largely work through their calorie counting and exercise tracking features with extra tips and support features. Promotion of these apps would be through a government-funded mass media campaign.

The proportion of the population that would receive this intervention and how this is calculated is presented in Figure 1. The target population was overweight or obese adults (the target population was those older than 18 years, but relative risks for the association between BMI and disease apply from age 25 years onward) living in New Zealand, who have access to smartphones and who want to try and lose weight. The proportion of New Zealand adults (aged >18 years) who were overweight or obese was taken from the National Nutrition Survey (2008-2009) and was estimated by sex, ethnic, and age groups. The proportion of this population who take up this intervention was calculated as follows.

First, it was estimated that 74.42% (with an uncertainty interval [UI] of 57.49%-88.19%; see Table 2) of the population have access to smartphones apps. A total of 2 estimates of smartphone usage were used: a survey carried out by Research New Zealand [13] reporting 59% smartphone ownership or access by New Zealand adults in 2013 and a forecast of 90% smartphone access by New Zealanders by 2018 [14]. (We use 2011 baseline data but have used more current smartphone usage in New Zealand to give more relevant outputs).

Second, the number of people that would be reached through a mass media campaign was estimated based on the reach of a number of previous national-level Health Promotion Agency (HPA) campaigns listed below. The HPA is a state-funded organization that leads programs to promote health in New Zealand.

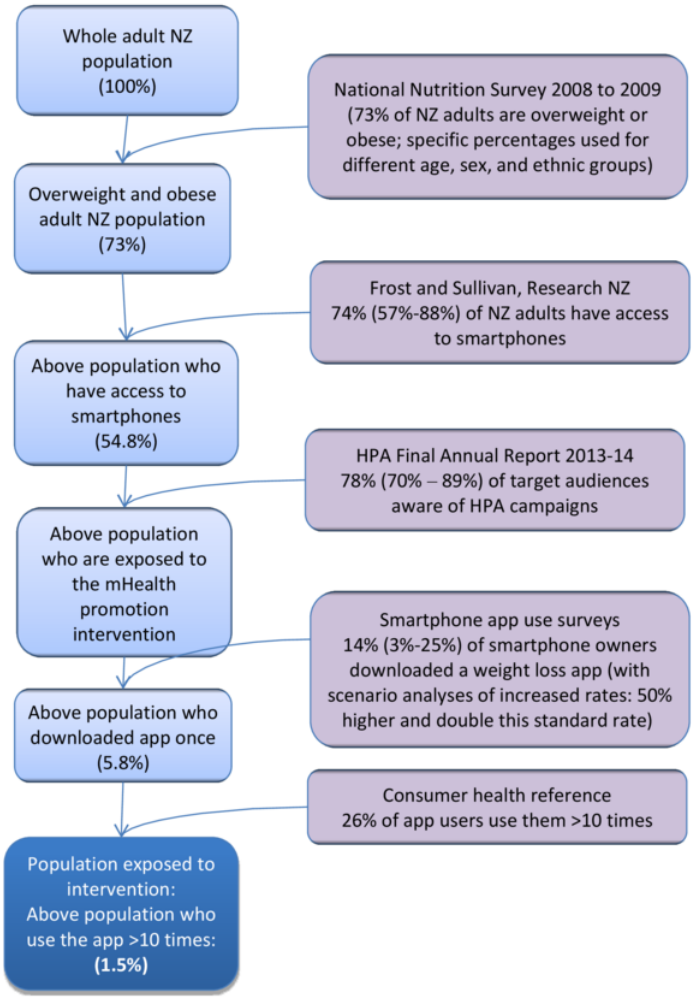
- For the HPA Heart and Diabetes Checks campaign in 2013, 70% of the core audience was reached at least once each month, with 50% of the target audience seeing the commercials 3 times [15].
- A total of 2 alcohol awareness campaigns achieved 89% total awareness of the key marketing messages, and 75% of adults were made aware of key elements of an alcohol law change [15].
- A television campaign on rheumatic fever prevention reached 76% of the target audience (parents and caregivers of at-risk children and young people) [16].

Our estimate of the likely reach of the mass media campaign promoting the smartphone apps for weight loss was based on an average of these mass media campaign figures with uncertainty spanning its range: a central estimate of 77.94% with a UI of 70.00% to 89.00%. Nearly half (46.8%) of all New Zealand adults (78% of the 60% of New Zealanders that have access to smartphones) were assumed to be exposed to the promotion of the intervention and have a smartphone. This is probably a relatively conservative estimate as there may be additional reach through mechanisms we did not model, that is, via Web-based activity as suggested in the HPA healthy eating program (the average total reach for the HPA healthy eating program for all Web-based activity for 2013-14 was 2,272,525 hits per month [15]), through referrals by health professionals, and through word-of-mouth.

Table 1. Baseline input parameters used in modeling the promotion of smartphone apps for weight loss.

Key parameter	Source and application to model	Uncertainty	Distribution and heterogeneity
Baseline population count	Statistics New Zealand (SNZ) population estimates for 2011	Nil uncertainty	Sex; age; ethnicity
All-cause mortality rates	SNZ mortality rates for 2011	Nil uncertainty	Sex; age; ethnicity
Disease-specific incidence, prevalence, case fatality rates, and remission rates	For each disease, coherent sets of incidence rates, prevalence, case fatality rates (CFR), and remission rates (zero for noncancers, the complement of the CFR for cancers to give the expected 5-year relative survival) were estimated using DISMOD II using data from New Zealand Burden of Disease Study (NZBDS), HealthTracker, and the Ministry of Health	Uncertainty: rates all ±5% SD	Log-normal; sex; age; ethnicity
Disease trends	Trends are applied to incidence, case fatality, and remission. These are switched on until 2026 and then kept constant for the remainder of the lifetimes of the modeled population	Uncertainty ±0.5% absolute change; diabetes: uncertainty ±1.5% absolute change	Normal; sex; ethnicity
Total morbidity per capita in 2011	The per capita rate of years of life lived with disability (YLD) from the NZBDS	Uncertainty ±10% SD	Log-normal; sex; age; ethnicity
Disease morbidity rate per capita	2006 NZBDS (projected to 2011); each disease was assigned a disability rate (DR; by sex and age) equal to YLDs for that disease (scaled down to adjust for comorbidities) from the 2006 NZBDS projected forward to 2011, divided by the disease prevalence. This DR was assigned to the proportion of the cohort in each disease state	Uncertainty: ±10% SD	Normal; sex; age
Health system costs	Linked health data (hospitalizations, inpatient procedures, outpatients, pharmaceuticals, laboratories, and expected primary care usage) for each individual in New Zealand for the period 2006 to 2010 had unit costs assigned to each event, and then health system costs (NZ $2011) were estimated	Estimated at SD ±10% of the point estimate	Gamma; sex; age
Time lags for intervention effect	It takes time for a change in body mass index (BMI) to impact on disease incidence. As there are no precise data on just how long these are, we have used wide windows of time lags. For cancers, the time lag is assumed to range between 10 and 30 years. For CHD, stroke, diabetes, and osteoarthritis (the noncancers), the time lag is assumed to be shorter and ranges between 0 and 5 years. Wide uncertainty is included around these estimates	Uncertainty: ±20% SD	Normal
BMI theoretical minimum risk exposure level (TMREL)	TMREL is the level of risk exposure that is theoretically possible and minimizes overall risk and is derived from the latest Global Burden of Disease 2013 study [10]. This allows us to estimate how much of the disease burden could be lowered by shifting the distribution of a risk factor to the level that would lead to the greatest improvement in population health	Uncertainty: uniform distribution between 0 and 1	Uniform
Height of the New Zealand adult population (for BMI calculations)	Mean and SD of height from the New Zealand Adult Nutrition Survey 2008 to 2009 [11]	Uncertainty using reported SD	Normal; sex; ethnicity

Figure 1. Flow diagram illustrating the targeting of the smartphone weight loss app promotion intervention in the New Zealand adult population. HPA: Health Promotion Agency; mHealth: mobile health; NZ: New Zealand.



We used 2 US surveys which reported the percentage of smartphone owners who have downloaded a weight loss app to estimate the proportion of the above population that would download a weight loss app in this study. Fronstin [17] reports 25% of smartphone owners with private health insurance had used a weight management or diet app. In an internet report published by Pew Research Center, 12% of the 19% of smartphone owners that had reported downloading a health app downloaded a weight loss app [18]. This equates to 2.3% of all smartphone owners. This is a wide range of estimates and reflects baseline use of weight loss apps in the US population (not the impact of mass media campaigns) but is the best evidence we could identify as currently available. Wide uncertainty has been incorporated into the estimate to reflect this. The average (13.46% with a UI of 2.50%-25.00%; see Table 2) of these 2 figures was used as an estimate of the overweight and obese smartphone owners who have been reached by the mass media campaign, who are likely to download a weight loss app. Under this standard rate of app downloads, the mass media promotion does not increase the proportion of the target population downloading weight loss apps, but may still change the type of app downloaded to those promoted.

As this figure represents the usual download rate without a mass media campaign, we modeled a range of different but still plausible download rates: standard download rate (the base-case), 50% increase in downloading, and doubling of this download rate, with corresponding proportional UIs.

A Web survey in 2011 by the Consumer Health Information Corporation (N=395) in the United States found that 26% of downloaded smartphone apps are used repeatedly (ie, 10 times or more) [19]. Assuming this applies to overweight or obese people in New Zealand, this gave 1.36% of the target population who are likely to use the app more than 10 times (see Figure 1; note that in this figure, the final percentage of the population that is exposed to the intervention is 1.5% based on the average proportion of New Zealand adults that are overweight or obese, and when this number is calculated in the model using population weightings, it is 1.36%). Uncertainty around previously outlined parameters contributed to uncertainty around this final figure (1.36% [0.39%-2.99%]). See Figure 1 for a pictorial summary of this population selection process. All parameters used to target this intervention were the same for all age, sex, and ethnic groups except for the proportion of the population that is overweight or obese, which differs by age, sex, and ethnicity.

Table 2. Intervention input parameters used in modeling the promotion of smartphone apps for weight loss.

Parameters	Source and application to model	Expected value and uncertainty	Distribution and heterogeneity
Effect size	The meta-analysis generated an effect size of 0.43 kg (95% CI 0.25-0.61) of mobile device interventions compared with control groups [4]; effect size operational only for overweight and obese adults in the model. Effect size in kg was converted to body mass index (BMI) using average heights for the New Zealand population for the 4 demographic groups	0.43 kg (95% CI 0.25-0.61)	Normal
BMI decay	Meta-analysis evidence of weight loss decay [20]; at the end of intervention delivery, the modeled BMI reduction decays back to the preintervention BMI at a rate of 0.03 units per month	Uncertainty±20% SD	Log-normal
Proportion of New Zealanders with smartphones	Frost and Sullivan press release [14], Research New Zealand [13]	74.42% (57.49%-88.19%), CI based on the range of estimates available	Beta
Proportion of the above population who are likely to be exposed to the mobile health (mHealth) promotion intervention	Heath Promotion Agency final annual report 2013-14 [15]	77.94% (70.00%-89.00%), CI based on the range of estimates available	Beta
Above population who are likely to have downloaded a weight loss app once	Smartphone app use surveys [18]	13.46% (2.50%-25.00%), CI based on the range of estimates available	Beta
Above population who use the app >10 times	Consumer health information corporation [19]	26%; uncertainty±SD (SD: 20% of mean)	Beta
Intervention costs	Total intervention costs are NZ $2,883,000	Uncertainty±SD (SD: 20% of mean)	Gamma
Relative risks for BMI and disease incidence	See Multimedia Appendix 1 [5] for disease-specific relative risks		Sex, age

The effect size for reduction in BMI among *successful* app users (ie, 10 or more uses) was taken from a systematic review of RCTs for mobile devices and weight loss in adults [4]. It included 17 RCTs, 12 of which were primary studies and 5 were secondary analyses of primary studies. Of the 12 studies, 8 used a mobile phone as the intervention medium, specifically smartphones. Of these 12 studies, 9 targeted both diet and physical activity to induce weight loss. The remaining 3 studies concentrated primarily on physical activity to induce weight loss, whereas none of the studies targeted just dietary change. Intervention duration for the 12 studies (ie, not including follow-up) ranged from 4 weeks to 2 years, and studies were all carried out in high-income countries: the United Kingdom, United States, Finland, and Australia. All included studies used an intention-to-treat analysis in accordance with the original assignment. Interventions included a variety of approaches including weight, energy intake and energy expenditure goal setting and self-monitoring, text reminders on various topics, meal and physical activity planning, buddy components, trophy rooms for rewards, and even some group sessions and calls from counselors. The meta-analysis of these 12 studies with weight as the outcome estimated an additional 0.43 kg weight loss (95% CI 0.25-0.61) for the intervention groups compared with controls. This weight change was converted to a change in BMI using average height for the 4 demographic groups in the MSLT model (Māori men, Māori women, non-Māori men, non-Māori women).

The effect size of −0.43 kg (−0.61 to −0.25) equated to a reduction in 0.14 to 0.17 BMI units across sex by ethnic groups in those successfully completing the mHealth intervention. Regarding decay of effect, the trials included in the systematic review [4] did not measure maintenance of the weight loss over time. A meta-analysis of face-to-face dietary advice by Dansinger et al [20] found that BMI increased by 0.03 BMI units per month post dietary counseling from an initial BMI decrease of 1.9 units. Evidence on how weight regain differs by type of weight loss intervention and magnitude of initial weight loss is currently limited, so we used this 0.03 BMI units per month as an estimate of how the effect of the mHealth intervention would decay post intervention. With such a small initial effect size, the BMI decrease returned to 0 approximately 5 months post the year of the intervention.

Intervention Costs

For this intervention, it was assumed that already existing smartphone apps (as identified in a recent survey [12]) were promoted, which avoids costs associated with any new software development. Costs of a media campaign by the MoH or the HPA have been estimated from previous health promotion media campaigns (Table 3).

Relative Risks

The change in BMI was then combined with the disease-specific relative risks (see Multimedia Appendix 1 [5]) through PIFs, which altered the incidence of BMI-related diseases.

Table 3. Costs associated with the smartphone app for weight loss promotion intervention.

Cost component[a]	Cost (NZ $)	Details
One-off costs for the promotion of the smartphone apps	$72,000	The cost of promotion on relevant government-funded websites (Ministry of Health, district health board, Health Promotion Agency [HPA]). Estimate based on the HPA Breakfast-eaters campaign (Personal Communication, HPA, October 2015) for Web-based promotions (Google adwords, Facebook adverts, promoting Facebook posts, etc) to drive consumers to the Breakfast-eaters website
Mass media promotion	$2,791,000	Cost of 1 year mass media promotion (assumed to be the same as the 2013-14 Quitline marketing budget; the promotion required for this intervention was assumed to be similar to the level of marketing undertaken by Quitline); $2,887,000) [21]. The cost of Quitline advertising and promotion for the 12-month period in 2013-2014 was NZ $2,165,000, and the staff management costs for "marketing and communications" were NZ $722,000 [21]. These 2014 costs were consumer price index–adjusted to the 2011 base year, giving an annual cost of NZ $2,791,000
Identifying top apps	$20,000	Cost of a one-off upgrade of previous New Zealand work [12] in identifying the top 5 apps for Apple and Android weight loss apps for promotion on the websites (NZ $20,000 contract)
Total intervention costs	$2,883,000	Uncertainty: estimated at SD±20% of the point estimate, gamma distribution. Correlated (0.75) with intervention parameters (access to smartphones, exposure to promotion campaign, and weight loss app downloaded)

[a]Costs to the individual were not included as they were out of scope with the health system perspective used but would include a proportion of the cost of a smartphone and its running costs, the usually trivial cost of the app (though most are free) and any costs (or cost-savings) for dietary changes and increased physical activity.

Modeling and Analysis

Microsoft Excel using an Ersatz add-in (Epigear International, created by Jan J Barendregt) was used to run each of the scenarios presented with uncertainty through the model 2000 times. Each of these simulations involved a random draw from the probability density function about those parameters specified with uncertainty in Tables 1 and 2. The main results produced were incremental QALYs gained and net health system costs accrued. The net health system cost was the sum of the intervention cost and any difference in projected future health system expenditure resulting from changes in disease incidence because of the mHealth intervention (including extra health costs from any increased life span).

Results

The estimated impact of the base-case intervention was a health gain of 29 QALYs (95% UI: 14-52; with 3% discounting) and costs to the health system of NZ $2.3 million (95% UI: NZ $1.6-3.0 million, US $1.6 million [95% UI: 1.1-2.0]) over the lifetime of the modeled population (Table 4). This was assuming the standard rate of app downloading (ie, the mass media promotion does not increase the proportion of the target population downloading weight loss apps but may still change the type of app downloaded to those promoted). QALY gains increased to 45 (95% UI: 21-81) and 59 (95% UI: 27-107) when the proportion of the target population downloading the app was modified in plausible directions, that is, increased by 50% from the standard download rate and doubled, respectively. Costs decreased to NZ $2.0 million (95% UI: NZ $1.1-2.8 million, US $1.4 million [95% UI: 0.7-1.9]) and NZ $1.8 million (95% UI: NZ $0.7-2.6 million, US $1.2 million [95% UI: 0.5-1.8]), respectively.

Costs per QALY gained (or the incremental cost-effectiveness ratio) were NZ $79,700 (US $53,600) for the standard download rate, NZ $45,500 (US $30,600) for the 50% increase in download rate, and NZ $29,900 (US $20,100) for the doubling download rate scenarios. On the basis of a threshold of NZ $45,000 (US $30,000), this intervention would appear to be of borderline cost-effectiveness for the total population and cost-effective for Māori when standard download rates increased by 50%. The intervention was cost-effective when download rates doubled as a result of the mass media campaign.

QALYs and associated costs were similar between men and women. As Māori make up only 15% of the total population in New Zealand, the majority of absolute QALYs gained and costs occurred in the non-Māori population. Health gains for the target population, those that are overweight or obese, were 0.011 QALYs per 1000 people for standard download rates and 0.021 QALYs per 1000 people when download rates were doubled. The age-standardized per capita QALY gains from the intervention for Māori were double of those for non-Māori at 0.010 per 1000 population for Māori and 0.005 for non-Māori in the base-case and were 0.021 in Māori and 0.009 in non-Māori when download rates were doubled. Adjusting for higher background mortality and morbidity rates for Māori, in an equity analysis where non-Māori mortality and morbidity rates were applied to Māori [22], QALYs gained for Māori increased from 5 to 6 total QALYs.

Undiscounted base-case results gave a health gain of 55 QALYs as a result of the intervention (Table 5) and 19 QALYs with 6% discounting. In the hypothetical scenario where weight loss is maintained over time, the total QALYs gained would increase to 2420 over the lifetime of the cohort and provide NZ $44.3 million in cost savings.

Table 4. Health gain (in quality-adjusted life-years) and health system costs saved over the life course from the promotion of smartphone apps for weight loss among the New Zealand population alive in 2011 (population N=4.4 million; 3% discounting; 95% UI in brackets). Results presented for those older than 25 years as relative risks for the associations between risk factors and disease start at age 25 years.

Subpopulation	Non-Māori	Māori	Ethnic groups combined	
	QALYs[a]	QALYs	QALYs	Net costs to the health system (NZ $ million)[b]
Base-case with no increase in standard downloading of apps (13.5% of exposed population download a weight loss app)				
All	24 (10-47)	5 (2-10)	29 (14-52)	2.3 (1.6-3.0)
Men	12	2	14	1.1
Women	12	3	15	1.2
Per capita[c]	0.006 (0.005)	0.007 (0.009)	0.007	0.53
Per capita for those overweight and obese[c]	0.010 (0.005)	0.011 (0.010)	0.011	0.84
Scenario: 50% increase in standard downloading of apps (20.3% of exposed population download a weight loss app)				
All	37 (15-73)	8 (3-15)	45 (21-81)	2.0 (1.1-2.8)
Men	18	4	22	1.0
Women	19	4	23	1.0
Per capita[c]	0.010 (0.008)	0.011 (0.014)	0.010	0.46
Per capita for those overweight and obese[c]	0.016 (0.007)	0.016 (0.016)	0.016	0.73
Scenario: doubling the standard downloading of apps (27.0% of exposed population download a weight loss app)				
All	49 (19-97)	10 (4-20)	59 (27-107)	1.8 (0.7-2.6)
Men	24	5	29	0.9
Women	25	5	30	0.9
Per capita[c]	0.013 (0.010)	0.015 (0.018)	0.013	0.40
Per capita for those overweight and obese[c]	0.021 (0.009)	0.021 (0.021)	0.021	0.63

[a]QALYs: quality-adjusted life-years.

[b]Includes both the cost offsets and intervention cost (see Table 3), distributed pro rata across all people alive in 2011.

[c]All per capita results are QALYs per 1000 adults and NZ $ per adult. Results in brackets for Māori and non-Māori are age-standardized. Results rounded to either 2 or 3 meaningful digits.

Table 5. Scenario analyses about health gain in quality-adjusted life-years and health system costs for the promotion of smartphone apps for weight loss compared with business as usual (expected value analysis; no uncertainty).

Scenario	QALYs[a] gained	Net costs to the health system (NZ $ million)
Base-case model[b]	30	2.3
Discount rate		
0% per annum	55	2.1
6% per annum	19	2.4
No decay in intervention benefit (permanent weight loss)	2420	−44.3 (ie, cost saving)

[a]QALY: quality-adjusted life-years.

[b]Discount rate 3%, standard app download rates, and intervention effect decays at a rate of 0.03 body mass index (BMI) units per month.

Figure 2. Tornado plot indicating which parameters drive uncertainty in the model results for health gain (in quality-adjusted life-years; QALYs) for the population. BMI: body mass index; CF: case fatality; inc: incidence; mHealth: mobile health; NZ: New Zealand; rem: remission; RR: relative risks; TMREL: theoretical minimum risk exposure level.

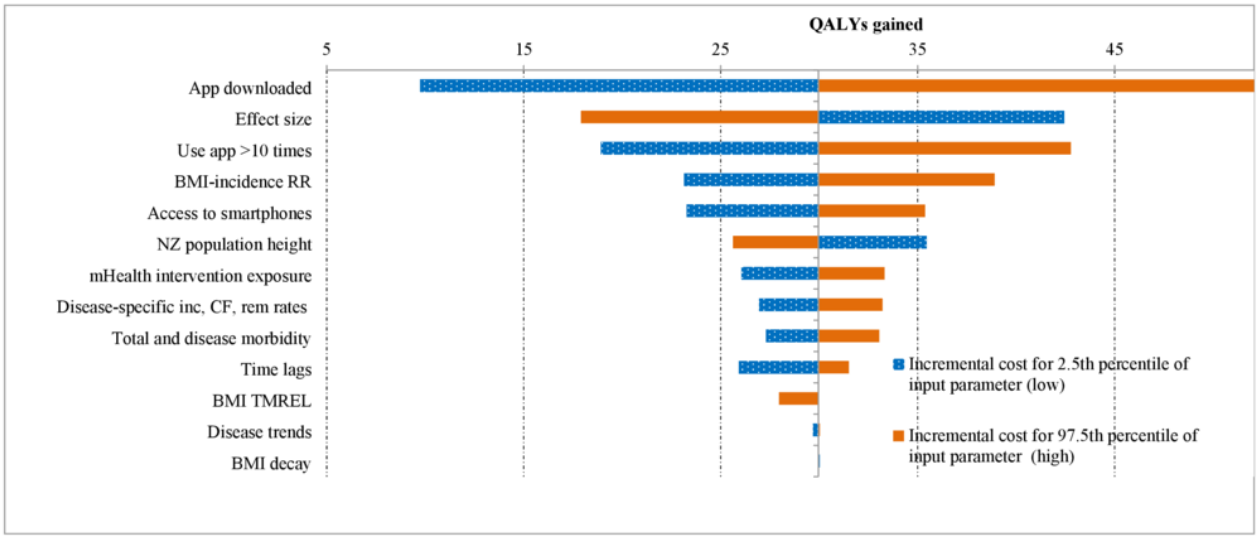
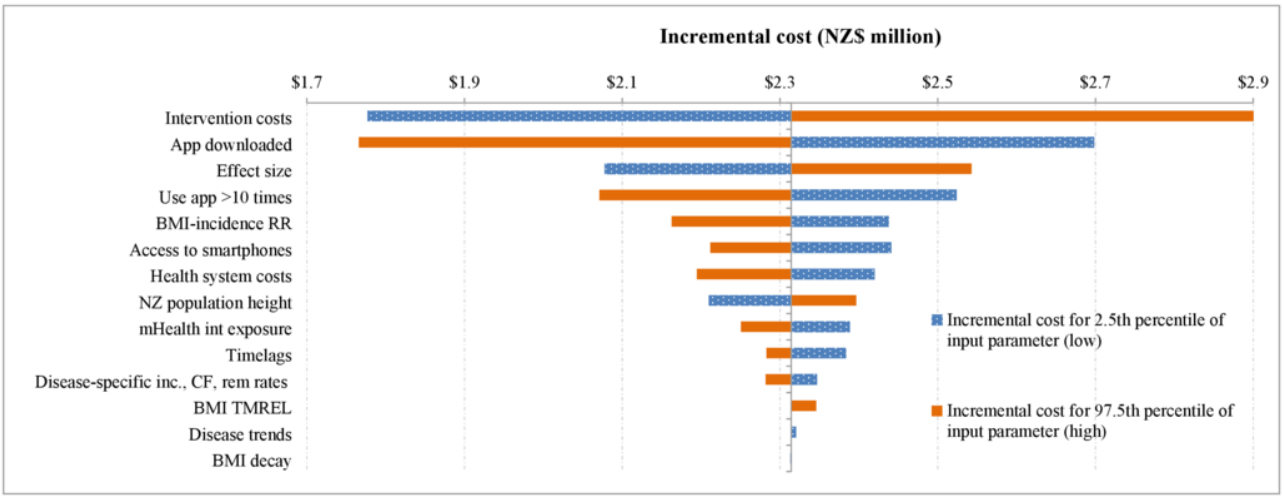


Figure 3. Tornado plot indicating which parameters drive uncertainty in the model results for health system costs for the population. BMI: body mass index; CF: case fatality; inc: incidence; mHealth: mobile health; NZ: New Zealand; rem: remission; RR: relative risks; TMREL: theoretical minimum risk exposure level.



Parameters contributing to the uncertainty in the model are shown in tornado plots in Figures 2 and 3. The parameters contributing the most to the uncertainty around the QALYs are whether the app was downloaded (which was varied in the main results presented in Table 4), the effect size, and the estimate of regular use of the app (defined as 10 or more uses). Other parameters that also contributed to overall uncertainty included uncertainty around the relative risks for the association between BMI and disease incidence; access to smartphones; the height of New Zealanders; estimates of the proportion of the target population being exposed to the mass media campaign; the disease-specific estimates of incidence, case fatality, and remission; and morbidity estimates. Uncertainty around the time lag between the intervention and onset of disease, the theoretical minimum risk exposure level around the association between BMI and disease, disease trends, and the BMI decay contributed the least to overall uncertainty.

Similar results are seen for the uncertainty contributing to the UIs around costs, but the biggest contributor was the intervention costs. The next top 3 parameters are as follows: (1) whether the app was downloaded, (2) the effect size, (3) and regular use of the app. Uncertainty around the disease-specific health system costs also makes a moderate contribution to the overall uncertainty.

Discussion

Principal Findings and Interpretation

The modeled intervention is based on a meta-analysis of mobile device interventions, which reported that those following the mHealth weight loss interventions lost an average of 0.43 kg (95% CI −0.61 to −0.25) more weight than the controls [4]. When modeled through to changes in incidences of BMI-associated diseases and then health gain in QALYs and health system costs, this intervention was found to be of

borderline cost-effectiveness for the total population when download rates increased by 50% but cost-effective for Māori, based on threshold of NZ $45,000 per QALY gained. However, it would be cost-effective for the total population if download rates doubled (NZ $29,900 or US $20,100 per QALY) as a result of the mass media campaign. Further research to improve the estimate of download rates resulting from this type of campaign would be useful. The total impact on health was small, with total QALYs gained ranging from 29 to 59 over the remaining life span of 4.4 million people entering the simulation. Costs to the health system ranged from NZ $1.8 to NZ $2.3 million.

Even with higher download rates (and therefore higher health gain), the total health gain in these scenarios was still small from a population per capita perspective. For example, health gains seen with annual tobacco tax increases gave a total of 60,400 QALYs in the New Zealand population [23]. Reducing dietary salt intake by 35% (through mandatory maximum salt levels in packaged foods and reductions in salt through fast foods and restaurant food and discretionary intake) generated 235,000 QALYs [24] over the lifetime of the New Zealand population.

Assuming that the intervention itself is equally effective for Māori and non-Māori, age-standardized population per capita QALYs for Māori were double of those for non-Māori, reflecting higher rates of overweight and obesity in Māori compared with non-Māori. Even greater benefit for Māori could potentially be achieved from an app and promotional campaign designed specifically for this population group.

Study Strengths and Limitations

The effect size used for modeling is an important parameter that drives the health and cost outputs presented. This was taken from a meta-analysis, the best quality evidence available, but the effect seen in the New Zealand context may vary from the meta-analysis effect size. The mHealth apps in the meta-analysis are different from the apps that would be promoted through this intervention if it was implemented in New Zealand. The meta-analysis effect size also included other mobile devices other than smartphones and other elements of mHealth. These differences between the modeled effect and the likely real effect are important to consider along with the fact that the uncertainty around the effect size was the second largest contributor to the overall uncertainty of the results. The lack of evidence on the effectiveness of a mass media campaign on download rates is another limitation of this work. This parameter contributes the most uncertainty to the health outcomes, and variation in this parameter changes the intervention from being cost-ineffective to being cost-effective. Therefore, future research regarding this parameter would better inform understanding of the cost-effectiveness of promoting smartphone apps for weight loss.

The modeled health benefits may actually be underestimated for a number of reasons. This study only models the effect of the intervention in those that actually take up the intervention, but there may be additional spill-over benefits to other members of the household through health-promoting changes in household meals or additional family physical activity. Furthermore, the impact of the intervention on physical activity itself was not modeled through to disease incidence. The intervention was also modeled as a one-off intervention when in reality the intervention could be ongoing, having the potential to both recruit more people over a number of years and/or sustain behavior change among initial participants. It is also likely that smartphone app usage and the quality of the apps available will increase over time, with effectiveness potentially increasing the effect size relative to that from the meta-analysis published in 2014. For example, higher-quality apps can integrate the collection of data on dietary energy intake with automatic estimates of energy expenditure based on the pedometer built into the smartphone. Additionally, the scope for this type of intervention may expand if smartphone weight loss apps were integrated with dietary counseling in primary care. These types of apps might become more integrated into daily routines and so weight loss achievements might be sustained for longer periods into the future.

The evidence for the rate at which the weight loss attenuates back to baseline, the BMI decay, is sourced from a meta-analysis [20] and is based on weight loss dietary counseling interventions. However, weight loss decay may differ for mHealth interventions where individuals can continue to access the app in the future, unlike with face-to-face dietary counseling, which is time limited. The scenario analysis where weight loss is maintained over the life course shows much greater health gains (2420 QALYs) and produces substantial cost savings (NZ $44.3 million). It is likely that the truth lies somewhere between the base-case and this scenario. Furthermore, as discussed above, app design and also changes to the obesogenic environment may impact the intervention decay rate. These factors may warrant additional research to improve estimation of health gains.

Finally, this study takes a health system perspective, but this intervention might result in wider societal benefits, for example, modifications made to people's diets for weight loss could result in lower consumption of energy-dense dairy products and meat products, therefore reducing greenhouse gas emissions [25] and other livestock-related environmental damage (water use, water pollution, erosion, and reduced biodiversity). Improved health from a lower BMI and increased physical activity could also result in higher productivity in the workplace (eg, from reduced illness-related absenteeism, early retirement because of illness, and premature death before retirement age).

Potential Implications for Research

As the field of mHealth develops, further research into the proportion of overweight and obese people who would regularly use an mHealth weight loss intervention, how this would be influenced by mass media campaigns, their subsequent weight loss, and how long their weight loss is maintained would all be useful. Consideration could also be given to determining app usage in the context of smartphone-based digital assistants, which can access apps (eg, Google's "Google Assistant" and Apple's "Siri"), or the provision of weight loss support from home-based digital assistants (eg, Amazon's "Alexa").

Potential Implications for Health Agencies

The results of modeling this mHealth intervention suggest it is likely to have relatively small absolute health gains at a population level (given the current levels of app use and current app design). As such, smartphone weight loss apps should not be a priority for inclusion in current obesity reduction strategies. Resources should instead be prioritized toward cost-effective or cost-saving interventions likely to have greater health impacts, such as food or beverage taxes and subsidies [26], restrictions on marketing of unhealthy foods [27], and improved nutrition labeling [28]. The scope for smartphone weight loss apps may expand with increasing smartphone use, more easy-to-use and effective apps becoming available, and integration of app promotion with dietary counseling by health workers. mHealth for weight loss may therefore become a more viable component of obesity prevention strategies in the future.

Acknowledgments

This research was funded by the Ministry of Business Innovation & Employment (project number UOOX1406). The authors are supported by the Burden of Disease Epidemiology, Equity and Cost Effectiveness (BODE^3) Programme, which is studying the effectiveness and cost-effectiveness of various preventative interventions and has received and receives funding support from the Health Research Council of New Zealand (Project numbers 10/248 and 16/443). The funding bodies had no role in the design of the study; the collection, analysis, and interpretation of data; or in writing the manuscript.

Authors' Contributions

CC transformed the existing tobacco MSLT to model dietary risk factors, helped conceptualize the intervention, modeled the intervention, and wrote the first draft of the paper. NW conceived the idea, helped conceptualize the intervention, and contributed to drafts of the paper. N Nair, helped conceptualize the intervention and contributed to drafts of the paper. MM contributed to the equity analyses and contributed to drafts of the paper. GK and N Nghiem contributed to the model build and drafts of the paper. TB contributed to the model build, helped conceptualize the intervention, and contributed to drafts of the paper.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Technical Report for BODE^3 Diet Intervention and Multistate Lifetable Models Version 1.

[PDF File (Adobe PDF File), 5MB - mhealth_v7i1e11118_app1.pdf]

References

1. Lim SS, Vos T, Flaxman AD, Danaei G, Shibuya K, Adair-Rohani H, et al. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2224-2260 [FREE Full text] [doi: 10.1016/S0140-6736(12)61766-8] [Medline: 23245609]
2. Hall AK, Cole-Lewis H, Bernhardt JM. Mobile text messaging for health: a systematic review of reviews. Annu Rev Public Health 2015 Mar 18;36:393-415 [FREE Full text] [doi: 10.1146/annurev-publhealth-031914-122855] [Medline: 25785892]
3. Grady A, Yoong S, Sutherland R, Lee H, Nathan N, Wolfenden L. Improving the public health impact of eHealth and mHealth interventions. Aust N Z J Public Health 2018 Apr;42(2):118-119. [doi: 10.1111/1753-6405.12771] [Medline: 29384248]
4. Lyzwinski LN. A systematic review and meta-analysis of mobile devices and weight loss with an intervention content analysis. J Pers Med 2014;4(3):311-385 [FREE Full text] [doi: 10.3390/jpm4030311] [Medline: 25563356]
5. Cleghorn C, Blakely T, Nghiem N, Mizdrak A. University of Otago. 2017. Technical report for BODE³ intervention and DIET MSLT models, Version 1 URL: http://www.otago.ac.nz/wellington/otago670797.pdf [accessed 2018-01-16] [WebCite Cache ID 6wWShODcI]
6. Barendregt JJ, Van Oortmarssen GJ, Vos T, Murray CJ. A generic model for the assessment of disease epidemiology: the computational basis of DisMod II. Popul Health Metr 2003 Apr 14;1(1):4 [FREE Full text] [Medline: 12773212]
7. Ministry of Health. Superu. Wellington: Ministry of Health; 2013. Health Loss in New Zealand: A report from the New Zealand Burden of Diseases, Injuries and Risk Factors Study URL: http://thehub.superu.govt.nz/sites/default/files/41073_health-loss-in-new-zealand-final_0.pdf [accessed 2018-01-17] [WebCite Cache ID 6wWSCiC9h]
8. Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A, et al. Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. Lancet 2012 Dec 15;380(9859):2129-2143. [doi: 10.1016/S0140-6736(12)61680-8] [Medline: 23245605]
9. van Baal PH, Feenstra TL, Polder JJ, Hoogenveen RT, Brouwer WB. Economic evaluation and the postponement of health care costs. Health Econ 2011 Apr;20(4):432-445. [doi: 10.1002/hec.1599] [Medline: 21210494]

10. GBD 2013 Risk Factors Collaborators, Forouzanfar MH, Alexander L, Anderson HR, Bachman VF, Biryukov S, et al. . [doi: 10.1016/S0140-6736(15)00128-2] [Medline: 26364544]
11. University of Otago, Ministry of Health. Ministry of Health. 2011. A Focus on Nutrition: Key findings of the 2008/09 New Zealand Adult Nutrition Survey URL: https://www.health.govt.nz/system/files/documents/publications/a-focus-on-nutrition-v2.pdf [accessed 2018-05-15] [WebCite Cache ID 6zQEebczH]
12. Patel R, Sulzberger L, Li G, Mair J, Morley H, Shing MN, et al. Smartphone apps for weight loss and smoking cessation: quality ranking of 120 apps. N Z Med J 2015 Sep 04;128(1421):73-76. [Medline: 26370762]
13. Research New Zealand. Research New Zealand. 2014. A Report on a Survey of New Zealanders? Use of Mobile Electronic Devices URL: http://www.researchnz.com/pdf/special%20reports/research%20new%20zealand%20special%20report%20-%20use%20of%20smartphones.pdf]Accessed: [accessed 2018-01-17] [WebCite Cache ID 6wWT7CYx0]
14. SCOOP. 2013 Dec 11. NZ will have 90% smartphone and 78% tablet ownership by 2018 URL: http://www.scoop.co.nz/stories/BU1312/S00422/nz-will-have-90-smartphone-and-78-tablet-ownership-by-2018.htm [accessed 2017-10-30] [WebCite Cache ID 6uc9NXIZs]
15. Health Promotion Agency. Health Promotion Agency. 2014 Oct. Annual Report For the Year Ended 30 June 2014 URL: https://www.hpa.org.nz/sites/default/files/Final%20annual%20report%202013-14_0_0.pdf [accessed 2018-01-17] [WebCite Cache ID 6wWVuIZAm]
16. TNS New Zealand Limited. Health Promotion Agency. Wellington: Health Promotion Agency; 2015. Rheumatic fever campaign evaluation URL: https://www.hpa.org.nz/sites/default/files/RF%202014%20campaign%20evaluation.pdf [accessed 2018-01-16] [WebCite Cache ID 6wWWG0fSw]
17. Fronstin P. Findings from the 2012 EBRI/MGA consumer engagement in health care survey. EBRI Issue Brief 2012 Dec(379):1-27. [Medline: 23390734]
18. Fox S, Duggan M. Pew Research Centre. 2012 Nov 08. Mobile Health 2012 URL: http://www.pewinternet.org/2012/11/08/mobile-health-2012/ [accessed 2017-10-30] [WebCite Cache ID 6ucA0C4C3]
19. Smith D. Consumer Health Information Corporation. Motivating patients to use smartphone health apps URL: http://www.consumer-health.com/motivating-patients-to-use-smartphone-health-apps/ [accessed 2017-10-31] [WebCite Cache ID 6ucAA6rPN]
20. Dansinger ML, Tatsioni A, Wong JB, Chung M, Balk EM. Meta-analysis: the effect of dietary counseling for weight loss. Ann Intern Med 2007 Jul 3;147(1):41-50. [Medline: 17606960]
21. Quitline. Wellington: Quitline; 2014. Quitline Me Mutu Annual Review 2013/2014 URL: https://quit.org.nz/-/media/Images/Quitline/PDFs-and-Docs/The-Quit-Group/qtl7826_annual_review_web_aw.pdf?la=en [accessed 2018-01-17] [WebCite Cache ID 6wWWhJnJJ]
22. McLeod M, Blakely T, Kvizhinadze G, Harris R. Why equal treatment is not always equitable: the impact of existing ethnic health inequalities in cost-effectiveness modeling. Popul Health Metr 2014;12:15 [FREE Full text] [doi: 10.1186/1478-7954-12-15] [Medline: 24910540]
23. Blakely T, Cobiac LJ, Cleghorn CL, Pearson AL, van der Deen FS, Kvizhinadze G, et al. Health, health inequality, and cost impacts of annual increases in tobacco tax: multistate life table modeling in New Zealand. PLoS Med 2015 Jul;12(7):e1001856 [FREE Full text] [doi: 10.1371/journal.pmed.1001856] [Medline: 26218517]
24. Wilson N, Nghiem N, Eyles H, Mhurchu CN, Shields E, Cobiac LJ, et al. Modeling health gains and cost savings for ten dietary salt reduction targets. Nutr J 2016 Apr 26;15:44 [FREE Full text] [doi: 10.1186/s12937-016-0161-1] [Medline: 27118548]
25. Friel S, Barosh LJ, Lawrence M. Towards healthy and sustainable food consumption: an Australian case study. Public Health Nutr 2014 May;17(5):1156-1166. [doi: 10.1017/S1368980013001523] [Medline: 23759140]
26. Cobiac LJ, Tam K, Veerman L, Blakely T. Taxes and subsidies for improving diet and population health in Australia: a cost-effectiveness modelling study. PLoS Med 2017 Feb;14(2):e1002232 [FREE Full text] [doi: 10.1371/journal.pmed.1002232] [Medline: 28196089]
27. Cairns G, Angus K, Hastings G, Caraher M. Systematic reviews of the evidence on the nature, extent and effects of food marketing to children. A retrospective summary. Appetite 2013 Mar;62:209-215. [doi: 10.1016/j.appet.2012.04.017] [Medline: 22561190]
28. Cecchini M, Warin L. Impact of food labelling systems on food choices and eating behaviours: a systematic review and meta-analysis of randomized studies. Obes Rev 2016 Mar;17(3):201-210. [doi: 10.1111/obr.12364] [Medline: 26693944]

Abbreviations

BAU: business-as-usual
BMI: body mass index
CHD: coronary heart disease
CPI: consumer price index
CVD: cardiovascular disease

DHB: district health board
HPA: Health Promotion Agency
mHealth: mobile health
MoH: Ministry of Health
MSLT: multistate life table
NZBDS: New Zealand Burden of Disease Study
PIF: population impact fraction
QALY: quality-adjusted life-year
RCT: randomized controlled trial
UI: uncertainty interval
YLD: years of life lived with disability

Edited by G Eysenbach; submitted 24.05.18; peer-reviewed by Y Yang, P Reeves, A Mohammadabadi; comments to author 21.09.18; revised version received 04.10.18; accepted 04.10.18; published 15.01.19.

Please cite as:
Cleghorn C, Wilson N, Nair N, Kvizhinadze G, Nghiem N, McLeod M, Blakely T
Health Benefits and Cost-Effectiveness From Promoting Smartphone Apps for Weight Loss: Multistate Life Table Modeling
JMIR Mhealth Uhealth 2019;7(1):e11118
URL: https://mhealth.jmir.org/2019/1/e11118/
doi:10.2196/11118
PMID:30664471

Original Paper

Comparison of mHealth and Face-to-Face Interventions for Smoking Cessation Among People Living With HIV: Meta-Analysis

Olalekan A Uthman[1], MD, MPH, PhD; Chidozie U Nduka[2], MD, MPH, PhD; Mustapha Abba[2], BSC, MSc; Rocio Enriquez[3], BA, MPH; Helena Nordenstedt[3], MD, MPH, PhD; Fred Nalugoda[4], MHSc, PhD; Andre P Kengne[5], MD, PhD; Anna M Ekström[3], MD, MPH, PhD

[1]Warwick-Centre for Applied Health Research and Delivery, Division of Health Sciences, University Warwick, Coventry, United Kingdom

[2]Division of Health Sciences, University of Warwick, Coventry, United Kingdom

[3]Global and Sexual Health, Department of Public Health, Karolinska Institutet, Stockholm, United Kingdom

[4]Uganda Virus Research Institute, Rakai Health Sciences Program, Rakai, Uganda

[5]Non-Communicable Diseases Research Unit, South African Medical Research Council, Cape Town, South Africa

Corresponding Author:
Olalekan A Uthman, MD, MPH, PhD
Warwick-Centre for Applied Health Research and Delivery
Division of Health Sciences
University Warwick
Gibbet Hill campus
Coventry, CV4 7AL
United Kingdom
Phone: 44 2476573163
Email: olalekan.uthman@warwick.ac.uk

Abstract

Background: The prevalence of smoking among people living with HIV (PLHIV) is higher than that reported in the general population, and it is a significant risk factor for noncommunicable diseases in this group. Mobile phone interventions to promote healthier behaviors (mobile health, mHealth) have the potential to reach a large number of people at a low cost. It has been hypothesized that mHealth interventions may not be as effective as face-to-face strategies in achieving smoking cessation, but there is no systematic evidence to support this, especially among PLHIV.

Objective: This study aimed to compare two modes of intervention delivery (mHealth vs face-to-face) for smoking cessation among PLHIV.

Methods: Literature on randomized controlled trials (RCTs) investigating effects of mHealth or face-to-face intervention strategies on short-term (4 weeks to <6 months) and long-term (≥6 months) smoking abstinence among PLHIV was sought. We systematically reviewed relevant RCTs and conducted pairwise meta-analyses to estimate relative treatment effects of mHealth and face-to-face interventions using standard care as comparison. Given the absence of head-to-head trials comparing mHealth with face-to-face interventions, we performed adjusted indirect comparison meta-analyses to compare these interventions.

Results: A total of 10 studies involving 1772 PLHIV met the inclusion criteria. The average age of the study population was 45 years, and women comprised about 37%. In the short term, mHealth-delivered interventions were significantly more efficacious in increasing smoking cessation than no intervention control (risk ratio, RR, 2.81, 95% CI 1.44-5.49; n=726) and face-to-face interventions (RR 2.31, 95% CI 1.13-4.72; n=726). In the short term, face-to-face interventions were no more effective than no intervention in increasing smoking cessation (RR 1.22, 95% CI 0.94-1.58; n=1144). In terms of achieving long-term results among PLHIV, there was no significant difference in the rates of smoking cessation between those who received mHealth-delivered interventions, face-to-face interventions, or no intervention. Trial sequential analysis showed that only 15.16% (726/1304) and 5.56% (632/11,364) of the required information sizes were accrued to accept or reject a 25% relative risk reduction for short- and long-term smoking cessation treatment effects. In addition, sequential monitoring boundaries were not crossed, indicating that the cumulative evidence may be unreliable and inconclusive.

Conclusions: Compared with face-to-face interventions, mHealth-delivered interventions can better increase smoking cessation rate in the short term. The evidence that mHealth increases smoking cessation rate in the short term is encouraging but not

sufficient to allow a definitive conclusion presently. Future research should focus on strategies for sustaining smoking cessation treatment effects among PLHIV in the long term.

(JMIR Mhealth Uhealth 2019;7(1):e203) doi:10.2196/mhealth.9329

KEYWORDS

HIV; mHealth; smoking cessation

Introduction

The introduction of effective antiretroviral therapy has resulted in a marked reduction in AIDS-related mortality worldwide. Patterns of morbidity and mortality have shifted from AIDS-related opportunistic infections to age-related comorbidities; moreover, it is now recognized that people living with HIV (PLHIV) are at increased risk of developing cardiovascular disease [1,2]. This increased risk is likely to be of multifactorial origin [3]: the disease and its treatment. Moreover, PLHIV are predisposed to engage in unhealthy behaviors [1,2].

Tobacco use is the single most common cause of preventable death worldwide and an important modifiable risk factor for several chronic conditions, including coronary heart diseases, chronic obstructive pulmonary diseases, and certain cancers [4]. About 15% of the world's population smoked tobacco in 2015 [5]. However, prevalence estimates of tobacco use in vulnerable populations are much higher: 32% among people with mental health disorders [6], 73% among homeless people [7], 77% among substance abusers [8], and 84% among prisoners [9]. Of note, the prevalence of smoking in PLHIV ranges between 50% and 70% [10], and like other vulnerable groups, success rates of quitting attempts and sustained abstinence are much lower than in the general population [11]. Smoking for stress relief, inadequate support from health service providers, and high smoking acceptance rates among communities of PLHIV are among the perceived barriers to abstinence in this high-risk group, and these considerably differ from self-reported barriers in apparently healthy populations without known chronic conditions [11]. For these reasons, intervention strategies for smoking cessation in the general population may not be as effective in HIV-positive populations. Although there have been several reports on the effectiveness of smoking cessation interventions among PLHIV [12], there are still questions left unanswered, notably about the mode of intervention delivery and its impact on smoking cessation. Short message service (SMS) text messages and other phone-based strategies have the potential to be more cost effective in service delivery than face-to-face contact, but it has been hypothesized that such mobile health (mHealth) strategies might be less effective or no different in terms of achieving smoking abstinence [12]. mHealth services provide unique opportunities for delivering smoking cessation interventions to large number of people at a low cost. However, there is no systematic evidence to substantiate this hypothesis. Therefore, we first sought to review all existing literature investigating mHealth and face-to-face interventions for smoking cessation among PLHIV. Second, we examined whether the required amount of information has been reached to confidently conclude that mHealth is more effective than no mHealth and that future trials need not examine this question any longer.

Methods

Information Sources and Search Strategy

We conducted searches on the following major databases: Embase, Medical Literature Analysis and Retrieval System Online, the Cochrane Library Central Register of Controlled Trials, the ClinicalTrials.gov registry, and cross-references of relevant articles for randomized controlled trials (RCTs) investigating the effectiveness of smoking cessation interventions among HIV-positive smokers and published until up to May 2018. Multimedia Appendix 1 shows the database search strategy, including the search term combinations used.

Selection Criteria

We evaluated each identified study against the following predetermined selection criteria: *Types of population* (PLHIV); *types of intervention and comparator* (face-to-face counseling, mHealth-delivered intervention, or no intervention country group); *types of outcome* (smoking abstinence); and *study design* (RCTs).

Selection of Studies

Two authors (CUN and MA) screened the titles and abstracts of all the potential studies we identified as a result of the search and coded them as "retrieve" (eligible or potentially eligible or unclear) or as "do not retrieve." Any disagreements were arbitrated by a third author (OAU). Subsequently, CUN and MA assessed the full-text study reports to confirm their eligibility for inclusion while noting the reasons for excluding studies considered ineligible for the meta-analysis. Again, any disagreements were resolved following discussions with OAU.

Data Extraction

CUN and MA extracted demographic and clinical data from the included studies where available. Data on trial design, sample size, mean age, proportion of women, average daily number of cigarettes, interventions, outcomes, and follow-up durations were extracted. Any disagreements were resolved following discussions with OAU.

Outcome Measures

The main outcome was short-term smoking abstinence, which has been defined as abstinence of at least 4 weeks duration, but less than 6 months after the intervention was initiated [12]. The secondary outcome was smoking abstinence of at least 6 months duration (long-term abstinence) [13].

Risk of Bias Assessment

CUN and MA judged the risk of bias in each included study using the Cochrane risk of bias assessment tool, which includes the following domains: randomization sequence generation (selection bias); allocation concealment (selection bias); blinding of participants, providers, (performance bias) and outcome assessors (detection bias); completeness of outcome data (attrition bias); and selective outcome reporting (reporting bias). Each RCT was classified as having "high," "low," or "unclear" risk of bias in each domain [14]. OAU resolved any differences in the assessments.

Statistical Analysis

We adopted an adjusted indirect comparison meta-analysis [15-18], a logical extension of standard meta-analysis to infer relative effectiveness of mHealth-delivered versus face-to-face interventions when the direct "head-to-head" evidence is lacking, that is, not directly addressed within any of the included trials. To illustrate this, in a situation where we have 3 treatments A (mHealth), B (face-to-face), and C (no intervention control), A and C have been compared in RCTs; B and C have been compared in other RCTs; and A and B had not been directly compared. The approach enabled the indirect comparisons (eg, A vs B) constructed from 2 trials that have one treatment in common to be incorporated (eg, A vs C and B vs C; Figure 1). Using Bucher adjusted indirect comparison method, the treatment effect for T_{AB} can be calculated by using the following equation:

$$T_{AB}=T_{AC}-T_{BC}$$

where T represents the treatment effect (eg, log risk ratio, RR) between the 2 interventions. SE is calculated as follows:

$$SE(T_{AB})=\sqrt{(SE(T_{AC})^2-SE(T_{BC})^2)}$$

All data were analyzed using R package "stats" (version 3.2.2). As part of the primary analysis, subgroup analysis was conducted based on the intensity and duration of follow-up period. Analysis was performed separately for short-term (4 weeks to <6 months) and long-term (≥6 months) smoking abstinence. We also quantified heterogeneity by computing the I^2 statistic; a value greater than 50% implied that the treatment estimates were considerably heterogeneous across the included studies. The pooled treatment estimates were reported using RRs and 95% CIs.

We examined the reliability and conclusiveness of the available evidence using trial sequential analyses (TSA) [19-21]. The sample size required for a reliable and conclusive meta-analysis is at least as large as that of a single optimally powered RCT. Therefore, we calculated the sample size (ie, the heterogeneity-corrected optimal information size) required to detect or reject a minimal 25% relative risk reduction intervention effect. We then used the heterogeneity-corrected optimal information size to help construct Lan-DeMets sequential monitoring boundaries for our cumulative meta-analyses [22], analogous to interim monitoring in an RCT, to determine when sufficient evidence had been accrued (Figure 2): Significant (P<.05) meta-analysis included potentially spurious evidence of effect, that is, the cumulative Z-curve did not cross the monitoring boundaries (curve A), or firm evidence of effect, that is, the cumulative Z-curve crossed the monitoring boundaries (curve B). Nonsignificant (P ≥.05) meta-analysis included absence of evidence, that is, the meta-analysis included less patients than the required information size (curve C), or lack of effect, that is, the meta-analysis included more patients than the required information size (curve D). We conducted TSA using TSA version 0.937 with an intention to maintain an overall 5% risk of a type I error and 20% risk of a type II error (power of 80%).

Figure 1. Adjusted indirect comparison network meta-analysis framework. A: mhealth delivered; B: face to face; C: standard of care.

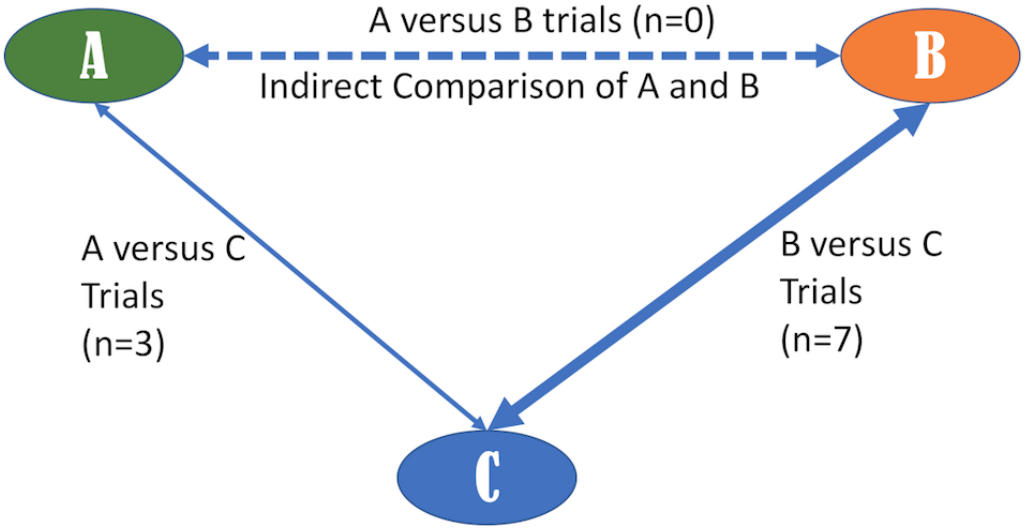
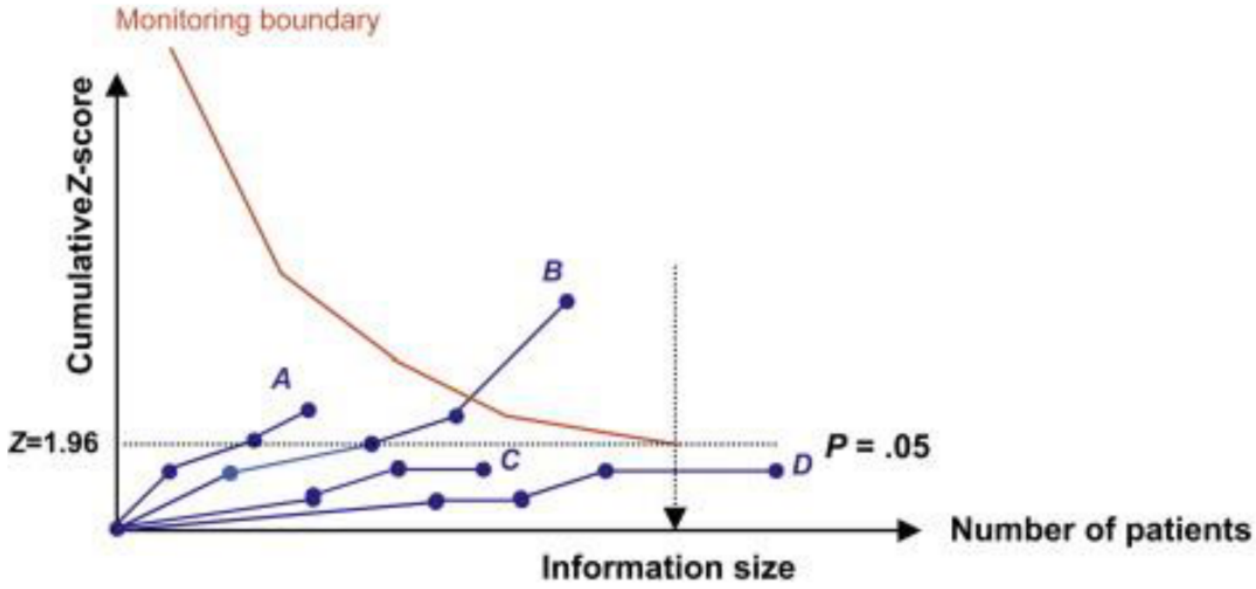


Figure 2. Trial sequential analysis.



Results

Study Selection and Characteristics of Included Studies

The study selection process is illustrated in Figure 3. The search strategy yielded 308 records, and upon screening abstracts and duplicate records, we identified 23 potentially eligible studies. We excluded 13 additional studies for the following reasons: the outcome was not abstinence (4 studies), no control group existed (4 studies), follow-up was less than 4 weeks (1 study), the study was quasi-experimental (1 study), face-to-face or mHealth interventions were not specified (1 study), a zero smoking abstinence rate was found in both the intervention and control groups, a computational error was identified (1 study), and participants in the intervention arm received either mHealth or face-to-face interventions but not both (1 study). Hence, 10 RCTs met the inclusion criteria and were included in the analysis [23-32].

The 10 included studies comprised a total of 1772 participants, all of whom were current smokers living with HIV. All studies were from a high-income country, that is, the United States (Table 1). On average, each participant at baseline smoked about 16 cigarettes daily (range 11-20). The average age of the study population was 45 years (range 42-50), and women comprised about 37% (range 8%-100%). Smoking cessation strategies were administered face-to-face in 7 studies [23-27,29,30], and sustained smoking abstinence estimates were reported in 4 studies [23,30-32]. The intensity and maximum follow-up period ranged from 4 weeks to as much 52 weeks. The number of counseling also varied across the studies, from 1 session to 11 sessions.

Risk of Bias in Included Studies

Among the 10 included studies, 6 reported the use of computer-generated lists of random numbers for randomization, whereas the other 4 studies did not describe the random sequence generation process. Allocation concealment was not described in any study; therefore, the risks of selection bias were unclear. Blinding of the participants and investigators was not described in 9 of the 10 studies, leaving 1 study in which investigators facilitated counseling sessions, which was thus judged to have a high risk of performance bias. Outcomes assessors were masked to the intervention in 2 studies, whereas the other 8 studies were assessed to have unclear risks of detection bias. Attrition bias was low in 7 studies, and reporting bias high or unclear in 5 studies (Figure 4).

Short-Term Smoking Cessation (4 Weeks to <6 Months)

Figure 5 displays a caterpillar plot of the relative RRs and 95% CIs of efficacy for all possible pairwise comparisons of the different treatment strategies. For short-term smoking cessation (ie, ≥4 weeks of smoking abstinence within 6 months of the intervention), 7 trials compared face-to-face intervention versus no intervention control group, and 4 compared mHealth-delivered interventions versus no intervention control group (n=1870). Participants randomized to SMS-delivered interventions were 2.81 times more likely to have stopped smoking compared to those who received standard care (RR 2.81, 95% CI 1.44-5.49). In addition, PLHIV who received mHealth-delivered interventions were twice as likely to have stopped smoking compared to those who received face-to-face interventions (RR 2.31, 95% 1.13-4.72). On average, face-to-face interventions were no more effective than no intervention in increasing short-term smoking cessation (RR 1.22, 95% CI 0.94-1.58). The measure of inconsistency between studies (I^2) was 6.3%, suggesting that the included studies were not statistically heterogeneous.

Figure 3. Study selection flow diagram. SMS: short message service.

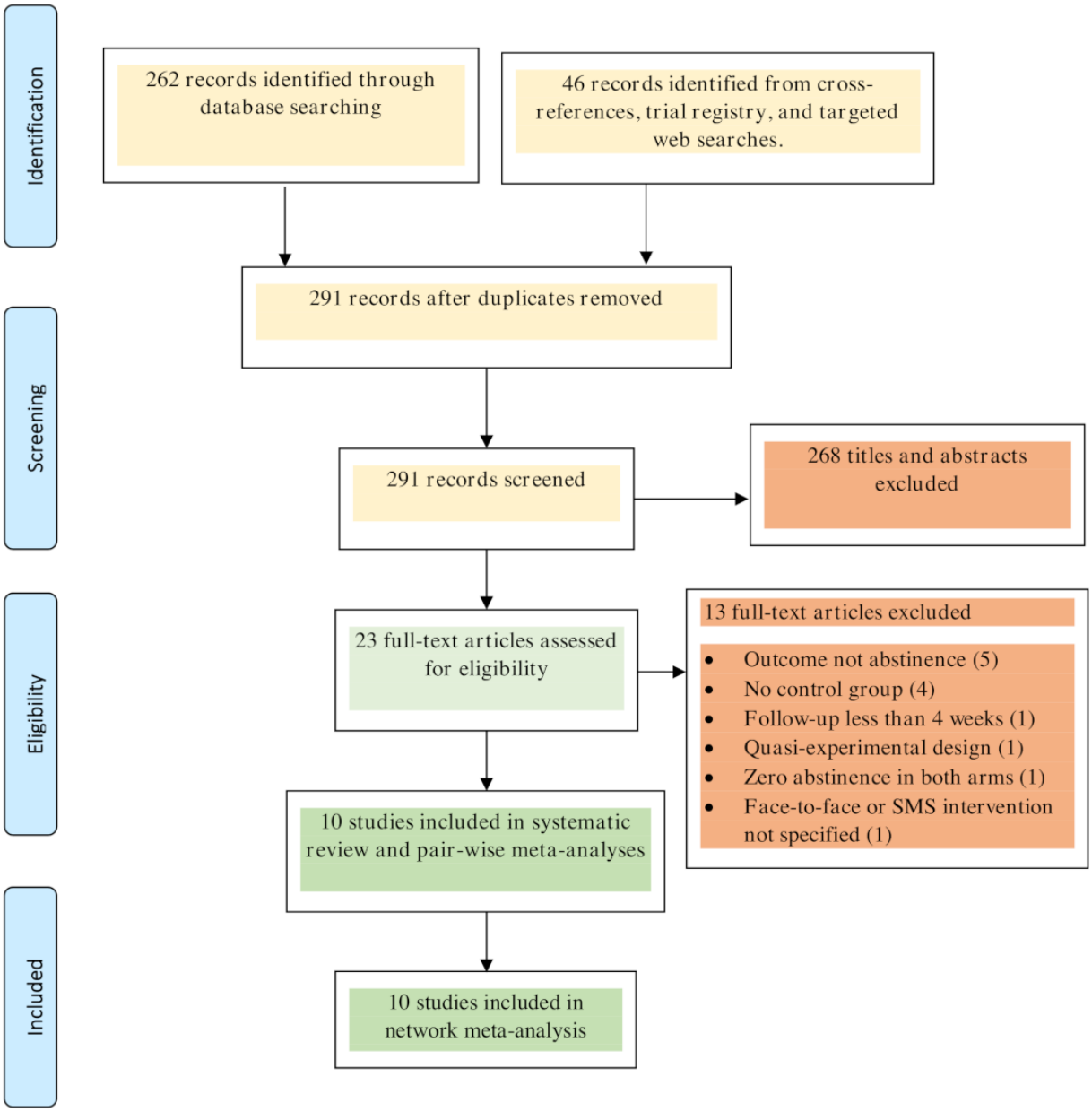


Table 1. Characteristics of the included studies.

Study	Sample size	Average age, years	Women (%)	Average daily number of cigarettes	Intervention	Outcome	Follow-up durations
Humfleet 2015 [23]	209	45	18	19.8	1 face-to-face CBT[b] session and 6 face-to-face CBT sessions	Sustained smoking abstinence	12, 24, 26, and 52 weeks
Ingersoll 2009 [24]	40	42	45	17.3	1 face-to-face counseling session	7-day point prevalence smoking abstinence	4 and 12 weeks
Lloyd-Richardson 2009 [25]	444	42	37	18.2	4 face-to-face counseling sessions	7-day point prevalence smoking abstinence	2, 4, and 6 months
Manuel 2013 [26]	30	49	100	16.1	1 face-to-face counseling session	7-day point prevalence smoking abstinence	4 weeks
Moadel 2012 [27]	145	49	51	12.0	8 face-to-face counseling sessions	7-day point prevalence smoking abstinence	42 and 132 days
Shelley 2015 [28]	158	50	16	15.0	Twice daily short message service text messages and twice daily short message service text messages+7 phone counseling sessions	7-day point prevalence smoking abstinence	1, 4, 8, 12, and 24 weeks
Shuter 2014 [29]	138	46	40	10.9	8 face-to-face CBT sessions	7-day point prevalence smoking abstinence	6 weeks, 3 months
Tucker 2017 [30]	40	42.9	8	13.0	Face-to-face counseling	Sustained smoking abstinence	3 months
Vidrine 2006 [31]	94	43	22	20.1	8 phone counseling sessions	Sustained smoking abstinence	3 months
Vidrine 2012 [32]	474	45	30	19.2	11 phone counseling sessions	Sustained smoking abstinence	3, 6, and 12 months

[a]RCT: randomized controlled trial.

[b]CBT: cognitive behavioral therapy.

Figure 4. Risk of bias in included studies.

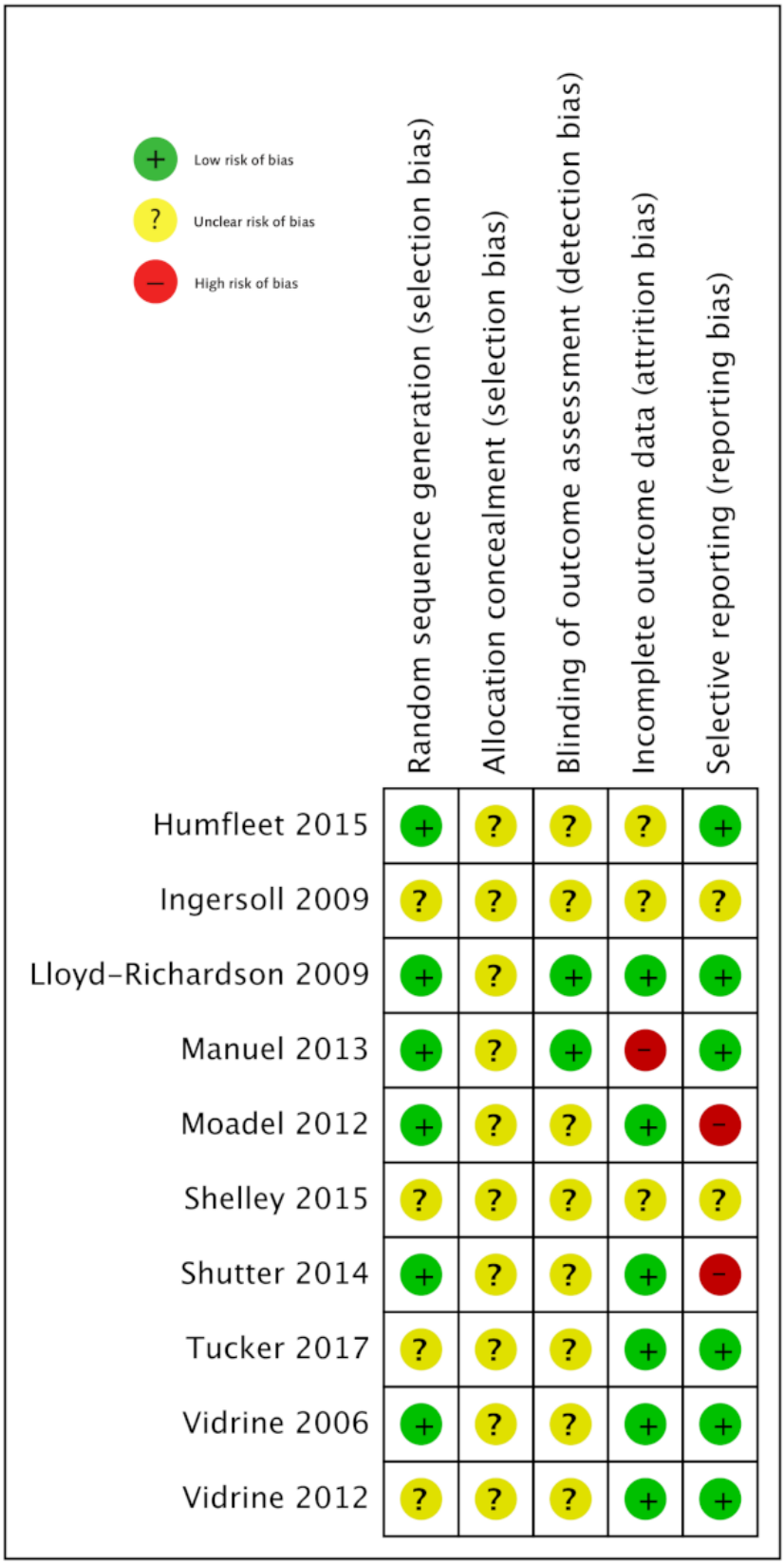


Figure 5. Pairwise comparisons of all interventions, short-term effect. RR: risk ratio; SMS: short message service; SoC: standard of care, no intervention control.

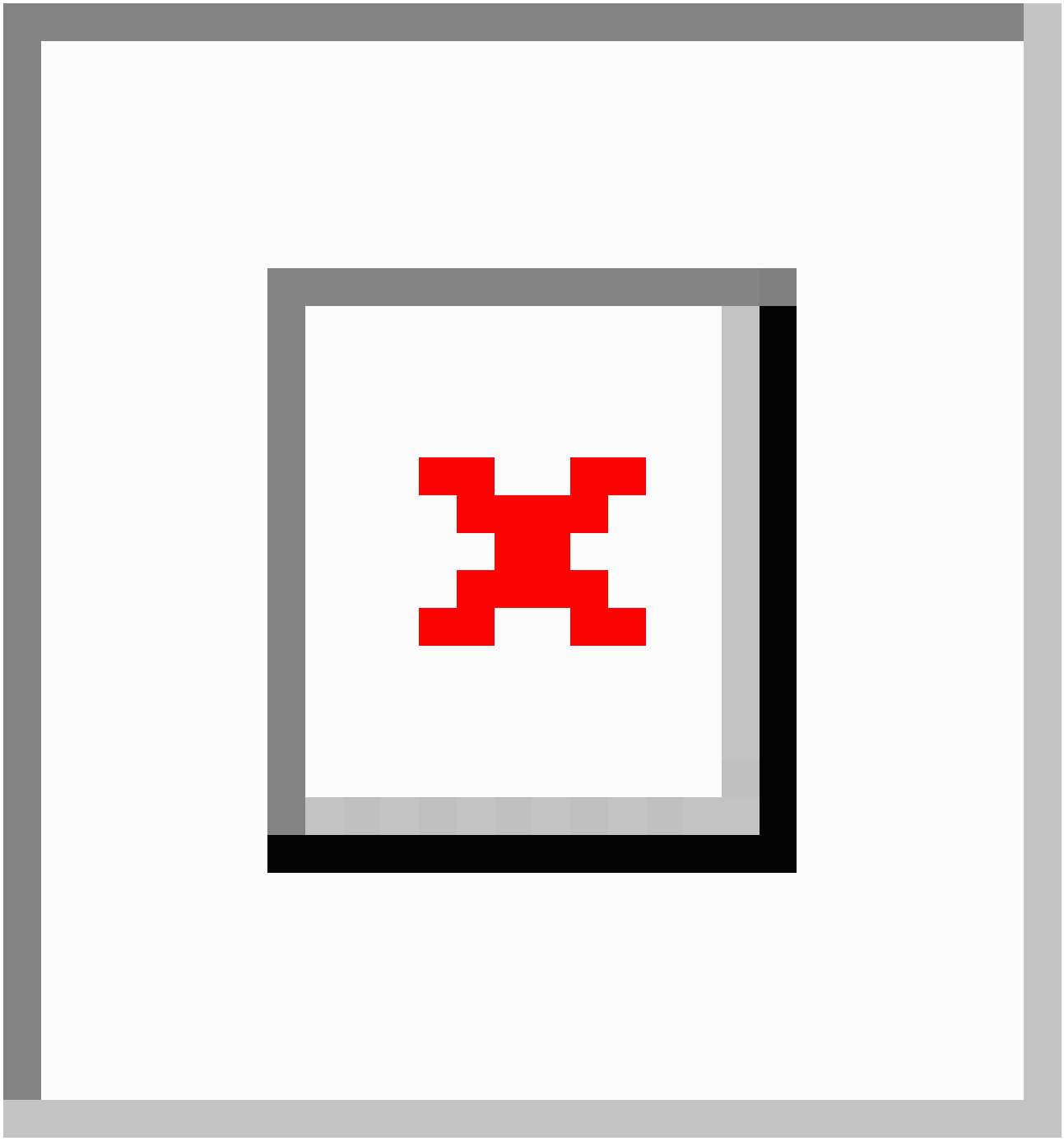
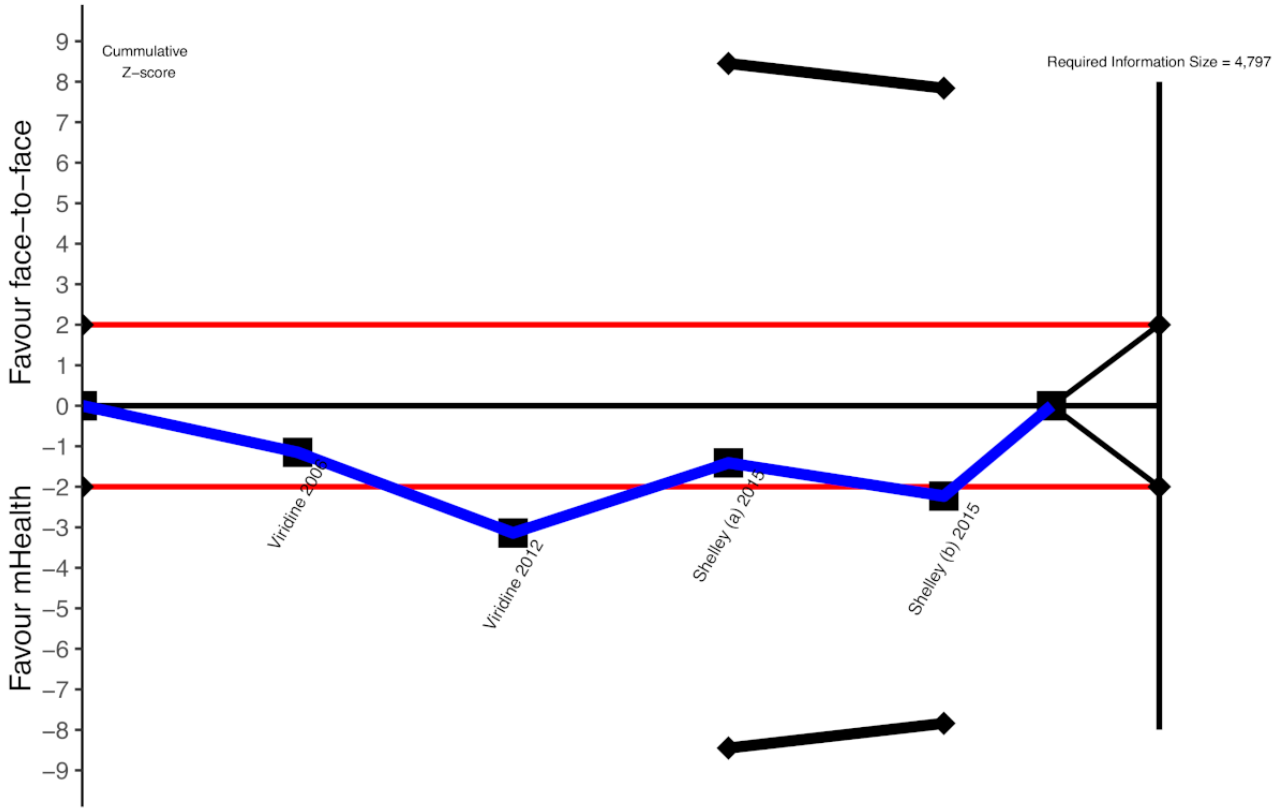


Figure 6. Trial sequential analysis for mHealth for smoking cessation in people living with HIV, short-term effect. mHealth: mobile health.



Our calculations indicated that the optimal information size needed to reliably detect a plausible treatment effect for short-term smoking cessation is 4797 participants (Figure 6). However, only 15.16% (726/4787) of the participants of the required information size were accrued. More so, the sequential monitoring boundary has not been crossed, indicating that the cumulative evidence is unreliable and inconclusive (Figure 6).

Long-Term Smoking Cessation (≥6 Months)

For long-term smoking cessation, that is, abstinence ≥6 months, 3 trials compared face-to-face intervention versus no intervention control group and 2 compared mHealth-delivered interventions versus no intervention control group (n=1546). There was no significant difference in smoking cessation rates between PLHIV randomized to mHealth-delivered interventions and those in the no intervention control group (RR 0.67, 95% CI 0.27-1.67). Similarly, there was no significant difference in smoking cessation rates between face-to-face interventions and no intervention control groups (RR 1.02, 95% CI 0.68-1.53). In addition, adjusted indirect treatment comparison between face-to-face and mHealth interventions revealed no significant difference (RR 0.65, 95% CI 0.29-1.47; Figure 7). The measure of inconsistency between the included studies (I^2) was 0%, suggesting no evidence that the included studies were statistically heterogeneous.

Figure 7. Pairwise comparisons of all interventions, long-term effect. RR: risk ratio; SMS: short message service; SoC: standard of care, no intervention control.

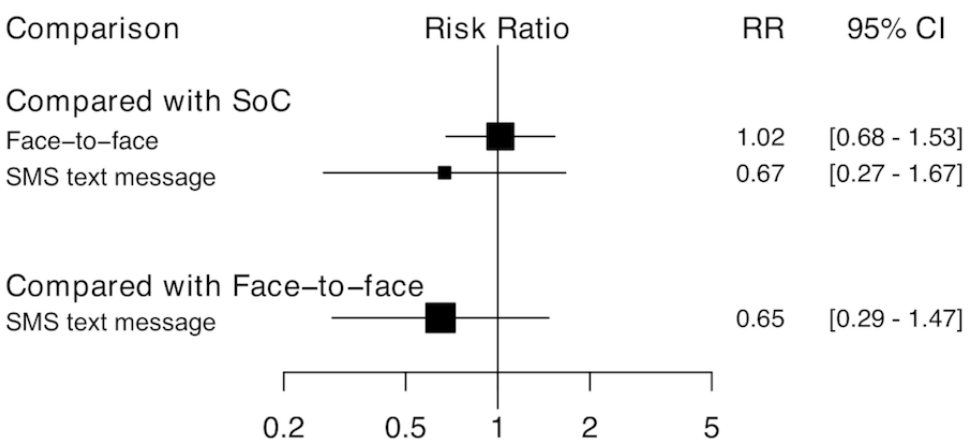
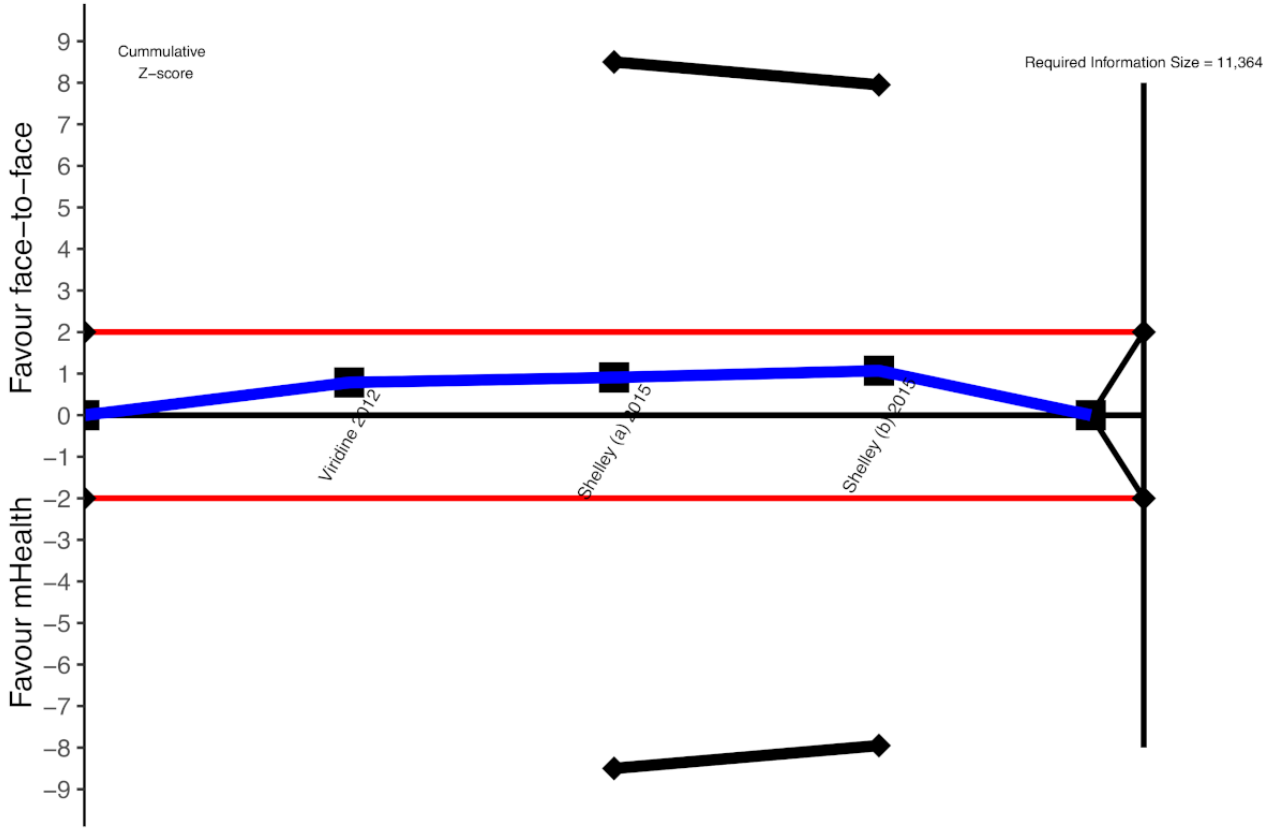


Figure 8. Trial sequential analysis for mHealth for smoking cessation in people living with HIV, long-term effect. mHealth: mobile health.



Our calculations indicated that the optimal information size needed to reliably detect a plausible treatment effect for long-term smoking cessation was 11,364 participants (Figure 8). However, only 5.56% (632/11,364) participants of the required information size were accrued in the pooled analysis. The sequential monitoring boundary has also not been crossed, indicating that the cumulative evidence is unreliable and inconclusive (Figure 8).

In Figure 8, dashed blue cumulative Z curves do not cross solid black trial sequential monitoring boundaries for benefit and horizontal red lines illustrate the traditional level of statistical significance (P=.05).

Discussion

Principal Findings

The results suggest that mHealth interventions for smoking cessation in PLHIV leads to better short-term improvement in smoking cessation rates than face-to-face interventions. However, from 6 months after the intervention and onward, there is no evidence of any effect regardless of the mode of intervention delivery. Our findings are broadly consistent with a previous meta-analysis of studies conducted in the general population, which reported a higher pooled smoking abstinence rate associated with SMS text messaging for 3 months compared to SMS text messaging for 6 months [33]. It is important to note that the absence of evidence is not evidence of absence of long-term effects of mHealth-delivered interventions. The lack of significant differences in long-term abstinence, however, may be due to the small number of studies contributing to this indirect evidence network, and as such, the evidence is inconclusive. Accurate understanding of the strength of the evidence for mHealth requires a systematic, comprehensive, and unbiased accumulation of the available evidence and methods adopted from formal interim monitoring boundaries applied to cumulative meta-analysis. The results of our TSA showed that the evidence that mHealth increases smoking cessation rates in the short term is encouraging but may be unreliable to make conclusive inferences.

Given the crucial need for the prevention of cardiovascular disease risk in PLHIV, there is a need for future pragmatic trials comparing mHealth and face-to-face intervention, especially in resource-limited settings that bear the highest burden of HIV and where smoking is now a bigger problem [2,34]. Furthermore, low-income settings are now experiencing an epidemiological transition from infectious diseases to chronic diseases [2] as a result of dramatic changes in diet and lifestyle. The epidemiological transition in resource-limited settings is happening over a shorter time frame than that experienced historically by high-income countries [34]. In addition, there is a need to identify mHealth-delivered interventions that are most beneficial for PLHIV. We should also investigate innovative specific features of mHealth interventions that can achieve long-term effects, for example, by varying the mode of delivery (weekly SMS text messaging) or by personalized and more tailored SMS text messages.

Limitations

The limitations in our study warrant consideration. First, the included studies were conducted in a high-income setting, which

potentially limits generalization of the results to low- and middle-income settings in which the burden of HIV and tobacco-related illnesses and deaths are currently most severe [5]. Nonetheless, our findings may be generalizable to other vulnerable groups in high-income countries. Second, the intervention arms in the included studies all comprised multicomponent strategies, which may have influenced our results; however, tests for heterogeneity revealed that the studies included in our analyses were in fact homogeneous. Third, we could not compare mHealth or face-to-face interventions with strategies that entailed a combination of both interventions because none of the included studies allowed this dual treatment in the intervention arm. Wewers et al [35] examined mHealth and face-to-face interventions in their study; however, we considered this study ineligible because the participants received either mHealth or face-to-face interventions and not both and because the numbers were not specified. Furthermore, with only 10 studies considered eligible for our review, we could not perform meta-regression analyses to explore potential effect-modifiers such as age, sex, coexisting substance abuse, and average number of cigarettes smoked daily at baseline. In spite of these limitations, we present novel systematic evidence evaluating the preferred mode of contact to be employed for improving smoking abstinence among PLHIV.

Conclusion

Compared to face-to-face interventions, mHealth-delivered interventions can boost smoking cessation rates, at least in the short term, among PLHIV with higher smoking prevalence rates than the general population. However, it remains unclear how long the effects of such interventions last. Future research should focus on strategies for sustaining the treatment effect in the long term and move beyond high-income settings.

Conflicts of Interest

OAU is supported by the National Institute of Health Research using Official Development Assistance funding and Wellcome Trust strategic award. The views expressed in this publication are those of the author(s) and not necessarily those of the NHS, the National Institute for Health Research, the Department of Health, or Wellcome Trust.

Multimedia Appendix 1

Medline search strategy.

[PNG File, 590KB - mhealth_v7i1e203_fig.ocx]

References

1. Dalal S, Beunza JJ, Volmink J, Adebamowo C, Bajunirwe F, Njelekela M, et al. Non-communicable diseases in sub-Saharan Africa: what we know now. Int J Epidemiol 2011 Aug;40(4):885-901. [doi: 10.1093/ije/dyr050] [Medline: 21527446]
2. Gaziano TA, Bitton A, Anand S, Abrahams-Gessel S, Murphy A. Growing epidemic of coronary heart disease in low- and middle-income countries. Curr Probl Cardiol 2010 Feb;35(2):72-115 [FREE Full text] [doi: 10.1016/j.cpcardiol.2009.10.002] [Medline: 20109979]
3. Petersen M, Yiannoutsos CT, Justice A, Egger M. Observational research on NCDs in HIV-positive populations: conceptual and methodological considerations. J Acquir Immune Defic Syndr 2014 Sep;67 Suppl 1:S8-16 [FREE Full text] [doi: 10.1097/QAI.0000000000000253] [Medline: 25117964]
4. WHO. Geneva, Switzerland: World Health Organization; 2008. WHO REPORT on the global TOBACCO epidemic, 2008 URL: http://www.who.int/sorry/ [accessed 2018-11-15] [WebCite Cache ID 73wWFLoO2]
5. WHO. Geneva, Switzerland: World Health Organization; 2017. WHO global report on trends in prevalence of tobacco smoking 2015 URL: http://www.who.int/sorry/ [accessed 2018-11-15] [WebCite Cache ID 73wXrrbPe]
6. Lawrence D, Mitrou F, Zubrick SR. Smoking and mental illness: results from population surveys in Australia and the United States. BMC Public Health 2009 Aug;9:285 [FREE Full text] [doi: 10.1186/1471-2458-9-285] [Medline: 19664203]
7. Baggett TP, Rigotti NA. Cigarette smoking and advice to quit in a national sample of homeless adults. Am J Prev Med 2010 Aug;39(2):164-172. [doi: 10.1016/j.amepre.2010.03.024] [Medline: 20621264]
8. Kelly PJ, Baker AL, Deane FP, Kay-Lambkin FJ, Bonevski B, Tregarthen J. Prevalence of smoking and other health risk factors in people attending residential substance abuse treatment. Drug Alcohol Rev 2012 Jul;31(5):638-644. [doi: 10.1111/j.1465-3362.2012.00465.x] [Medline: 22642401]
9. Cropsey K, Eldridge G, Weaver M, Villalobos G, Stitzer M, Best A. Smoking cessation intervention for female prisoners: addressing an urgent public health need. Am J Public Health 2008 Oct;98(10):1894-1901. [doi: 10.2105/AJPH.2007.128207] [Medline: 18703440]
10. Shuter J, Bernstein S. Cigarette smoking is an independent predictor of nonadherence in HIV-infected individuals receiving highly active antiretroviral therapy. Nicotine Tob Res 2008 Apr;10(4):731-736. [doi: 10.1080/14622200801908190] [Medline: 18418794]
11. Twyman L, Bonevski B, Paul C, Bryant J. Perceived barriers to smoking cessation in selected vulnerable groups: a systematic review of the qualitative and quantitative literature. BMJ Open 2014 Dec;4(12):e006414 [FREE Full text] [doi: 10.1136/bmjopen-2014-006414] [Medline: 25534212]

12. Pool ERM, Dogar O, Lindsay RP, Weatherburn P, Siddiqi K. Interventions for tobacco use cessation in people living with HIV and AIDS. Cochrane Database Syst Rev 2016 Jun(6):CD011120. [doi: 10.1002/14651858.CD011120.pub2] [Medline: 27292836]
13. West R, Hajek P, Stead L, Stapleton J. Outcome criteria in smoking cessation trials: proposal for a common standard. Addiction 2005 Mar;100(3):299-303. [doi: 10.1111/j.1360-0443.2004.00995.x] [Medline: 15733243]
14. Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, Cochrane Bias Methods Group, Cochrane Statistical Methods Group. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ 2011 Oct;343:d5928 [FREE Full text] [Medline: 22008217]
15. Bucher HC, Guyatt GH, Griffith LE, Walter SD. The results of direct and indirect treatment comparisons in meta-analysis of randomized controlled trials. J Clin Epidemiol 1997 Jun;50(6):683-691. [Medline: 9250266]
16. Song F, Altman DG, Glenny A, Deeks JJ. Validity of indirect comparison for estimating efficacy of competing interventions: empirical evidence from published meta-analyses. BMJ 2003 Mar 01;326(7387):472 [FREE Full text] [doi: 10.1136/bmj.326.7387.472] [Medline: 12609941]
17. Glenny AM, Altman DG, Song F, Sakarovitch C, Deeks JJ, D'Amico R, International Stroke Trial Collaborative Group. Indirect comparisons of competing interventions. Health Technol Assess 2005 Jul;9(26):1-134, iii [FREE Full text] [Medline: 16014203]
18. Gartlehner G, Moore CG. Direct versus indirect comparisons: a summary of the evidence. Int J Technol Assess Health Care 2008;24(2):170-177. [doi: 10.1017/S0266462308080240] [Medline: 18400120]
19. Brok J, Thorlund K, Gluud C, Wetterslev J. Trial sequential analysis reveals insufficient information size and potentially false positive results in many meta-analyses. J Clin Epidemiol 2008 Aug;61(8):763-769. [doi: 10.1016/j.jclinepi.2007.10.007] [Medline: 18411040]
20. Higgins JPT, Whitehead A, Simmonds M. Sequential methods for random-effects meta-analysis. Stat Med 2011 Apr 30;30(9):903-921 [FREE Full text] [doi: 10.1002/sim.4088] [Medline: 21472757]
21. Wetterslev J, Thorlund K, Brok J, Gluud C. Estimating required information size by quantifying diversity in random-effects model meta-analyses. BMC Med Res Methodol 2009 Dec 30;9:86 [FREE Full text] [doi: 10.1186/1471-2288-9-86] [Medline: 20042080]
22. Pogue JM, Yusuf S. Cumulating evidence from randomized trials: utilizing sequential monitoring boundaries for cumulative meta-analysis. Control Clin Trials 1997 Dec;18(6):580-93; discussion 661. [Medline: 9408720]
23. Humfleet GL, Hall SM, Delucchi KL, Dilley JW. A randomized clinical trial of smoking cessation treatments provided in HIV clinical care settings. Nicotine Tob Res 2013 Aug;15(8):1436-1445 [FREE Full text] [doi: 10.1093/ntr/ntt005] [Medline: 23430708]
24. Ingersoll KS, Cropsey KL, Heckman CJ. A test of motivational plus nicotine replacement interventions for HIV positive smokers. AIDS Behav 2009 Jun;13(3):545-554 [FREE Full text] [doi: 10.1007/s10461-007-9334-4] [Medline: 18066659]
25. Lloyd-Richardson EE, Stanton CA, Papandonatos GD, Shadel WG, Stein M, Tashima K, et al. Motivation and patch treatment for HIV+ smokers: a randomized controlled trial. Addiction 2009 Nov;104(11):1891-1900 [FREE Full text] [doi: 10.1111/j.1360-0443.2009.02623.x] [Medline: 19719796]
26. Manuel JK, Lum PJ, Hengl NS, Sorensen JL. Smoking cessation interventions with female smokers living with HIV/AIDS: a randomized pilot study of motivational interviewing. AIDS Care 2013;25(7):820-827 [FREE Full text] [doi: 10.1080/09540121.2012.733331] [Medline: 23116051]
27. Moadel AB, Bernstein SL, Mermelstein RJ, Arnsten JH, Dolce EH, Shuter J. A randomized controlled trial of a tailored group smoking cessation intervention for HIV-infected smokers. J Acquir Immune Defic Syndr 2012 Oct 01;61(2):208-215 [FREE Full text] [doi: 10.1097/QAI.0b013e3182645679] [Medline: 22732470]
28. Shelley D, Tseng T, Gonzalez M, Krebs P, Wong S, Furberg R, et al. Correlates of Adherence to Varenicline Among HIV+ Smokers. Nicotine Tob Res 2015 Aug;17(8):968-974 [FREE Full text] [doi: 10.1093/ntr/ntv068] [Medline: 26180221]
29. Shuter J, Morales DA, Considine-Dunn SE, An LC, Stanton CA. Feasibility and preliminary efficacy of a web-based smoking cessation intervention for HIV-infected smokers: a randomized controlled trial. J Acquir Immune Defic Syndr 2014 Sep 01;67(1):59-66 [FREE Full text] [doi: 10.1097/QAI.0000000000000226] [Medline: 25118794]
30. Tucker JS, Shadel WG, Galvan FH, Naranjo D, Lopez C, Setodji C. Pilot evaluation of a brief intervention to improve nicotine patch adherence among smokers living with HIV/AIDS. Psychol Addict Behav 2017 Dec;31(2):148-153 [FREE Full text] [doi: 10.1037/adb0000221] [Medline: 27736144]
31. Vidrine DJ, Arduino RC, Lazev AB, Gritz ER. A randomized trial of a proactive cellular telephone intervention for smokers living with HIV/AIDS. AIDS 2006 Jan 09;20(2):253-260. [doi: 10.1097/01.aids.0000198094.23691.58] [Medline: 16511419]
32. Vidrine DJ, Marks RM, Arduino RC, Gritz ER. Efficacy of cell phone-delivered smoking cessation counseling for persons living with HIV/AIDS: 3-month outcomes. Nicotine Tob Res 2012 Jan;14(1):106-110 [FREE Full text] [doi: 10.1093/ntr/ntr121] [Medline: 21669958]
33. Spohr SA, Nandy R, Gandhiraj D, Vemulapalli A, Anne S, Walters ST. Efficacy of SMS Text Message Interventions for Smoking Cessation: A Meta-Analysis. J Subst Abuse Treat 2015 Sep;56:1-10. [doi: 10.1016/j.jsat.2015.01.011] [Medline: 25720333]

34. Miranda JJ, Kinra S, Casas JP, Davey SG, Ebrahim S. Non-communicable diseases in low- and middle-income countries: context, determinants and health policy. Trop Med Int Health 2008 Oct;13(10):1225-1234 [FREE Full text] [doi: 10.1111/j.1365-3156.2008.02116.x] [Medline: 18937743]
35. Wewers ME, Neidig JL, Kihm KE. The feasibility of a nurse-managed, peer-led tobacco cessation intervention among HIV-positive smokers. J Assoc Nurses AIDS Care 2000;11(6):37-44. [doi: 10.1016/S1055-3290(06)60353-1] [Medline: 11082801]

Abbreviations

mHealth: mobile health
PLHIV: people living with HIV
RCT: randomized controlled trial
RR: risk ratio
SMS: short message service
TSA: trial sequential analyses

Edited by G Eysenbach; submitted 16.11.17; peer-reviewed by P Krebs, X Ling; comments to author 15.03.18; revised version received 25.05.18; accepted 18.06.18; published 07.01.19.

Please cite as:
Uthman OA, Nduka CU, Abba M, Enriquez R, Nordenstedt H, Nalugoda F, Kengne AP, Ekström AM
Comparison of mHealth and Face-to-Face Interventions for Smoking Cessation Among People Living With HIV: Meta-Analysis
JMIR Mhealth Uhealth 2019;7(1):e203
URL: https://mhealth.jmir.org/2019/1/e203/
doi:10.2196/mhealth.9329
PMID:30617044

Original Paper

A Mobile Phone-Based Program to Promote Healthy Behaviors Among Adults With Prediabetes Who Declined Participation in Free Diabetes Prevention Programs: Mixed-Methods Pilot Randomized Controlled Trial

Dina Griauzde[1,2,3], MD, MS; Jeffrey T Kullgren[4], MD, MPH; Brad Liestenfeltz[5], BS, RN; Tahoora Ansari[6], MPH; Emily H Johnson[2], BS; Allison Fedewa[4], MSc; Laura R Saslow[5], PhD; Caroline Richardson[2,3], MD; Michele Heisler[2,3,4,6], MD, MPH

[1]Veterans Affairs Center for Clinical Management Research, Veterans Affairs Ann Arbor Healthcare System, Ann Arbor, MI, United States

[2]University of Michigan Medical School, Ann Arbor, MI, United States

[3]Institute for Healthcare Policy and Innovation, University of Michigan, Ann Arbor, MI, United States

[4]University of Michigan School of Dentistry, Ann Arbor, MI, United States

[5]University of Michigan School of Nursing, Ann Arbor, MI, United States

[6]University of Michigan School of Public Health, Ann Arbor, MI, United States

Corresponding Author:
Michele Heisler, MD, MPH
University of Michigan School of Dentistry
1011 North University Avenue
Ann Arbor, MI, 48109
United States
Phone: 1 734 647 4844
Fax: 1 734 647 3301
Email: mheisler@umich.edu

Abstract

Background: Despite evidence that Diabetes Prevention Programs (DPPs) can delay or prevent progression to type 2 diabetes mellitus (T2DM), few individuals with prediabetes enroll in offered programs. This may be in part because many individuals with prediabetes have low levels of autonomous motivation (ie, motivation that arises from internal sources) to prevent T2DM.

Objective: This study aims to examine the feasibility and acceptability of a mobile health (mHealth) intervention designed to increase autonomous motivation and healthy behaviors among adults with prediabetes who previously declined participation free DPPs. In addition, the study aims to examine changes in autonomous motivation among adults offered 2 versions of the mHealth program compared with an information-only control group.

Methods: In this 12-week, parallel, 3-arm, mixed-methods pilot randomized controlled trial, participants were randomized to (1) a group that received information about prediabetes and strategies to prevent T2DM (control); (2) a group that received a mHealth app that aims to increase autonomous motivation among users (app-only); or (3) a group that received the app plus a physical activity tracker and wireless-enabled digital scale for self-monitoring (app-plus). Primary outcome measures included rates of intervention uptake (number of individuals enrolled/number of individuals assessed for eligibility), retention (number of 12-week survey completers/number of participants), and adherence (number of device-usage days). The secondary outcome measure was change in autonomous motivation (measured using the Treatment Self-Regulation Questionnaire), which was examined using difference-in-difference analysis. Furthermore, we conducted postintervention qualitative interviews with participants.

Results: Overall, 28% (69/244) of eligible individuals were randomized; of these, 80% (55/69) completed the 12-week survey. Retention rates were significantly higher among app-plus participants than participants in the other 2 study arms combined (P=.004, χ^2). No significant differences were observed in adherence rates between app-only and app-plus participants (43 days vs 37 days; P=.34). Among all participants, mean autonomous motivation measures were relatively high at baseline (6.0 of 7.0 scale), with no statistically significant within- or between-group differences in follow-up scores. In qualitative interviews (n=15), participants identified reasons that they enjoyed using the app (eg, encouraged self-reflection), reasons that they did not enjoy

using the app (eg, did not consider personal circumstances), and strategies to improve the intervention (eg, increased interpersonal contact).

Conclusions: Among individuals with prediabetes who did not engage in free DPPs, this mHealth intervention was feasible and acceptable. Future work should (1) examine the effectiveness of a refined intervention on clinically relevant outcomes (eg, weight loss) among a larger population of DPP nonenrollees with low baseline autonomous motivation and (2) identify other factors associated with DPP nonenrollment, which may serve as additional potential targets for interventions.

Trial Registration: ClinicalTrials.gov NCT03025607; https://clinicaltrials.gov/ct2/show/NCT03025607 (Archived by WebCite at http://www.webcitation.org/73cvaSAie)

(JMIR Mhealth Uhealth 2019;7(1):e11267) doi:10.2196/11267

KEYWORDS

autonomous motivation; behavioral change; mHealth; mobile phone; prediabetes; prevention; type 2 diabetes mellitus

Introduction

Type 2 diabetes mellitus (T2DM) is a key driver of death, disability, and health care spending in the United States [1,2]. In 2015, >30 million US adults had T2DM, while 84 million more were estimated to have prediabetes, a condition associated with an increased risk of developing T2DM [1]. Diabetes Prevention Programs (DPPs) can help individuals with prediabetes to achieve modest weight loss through diet and physical activity changes that reduce the 3-year risk of developing T2DM by >50% [3,4]. Accordingly, DPPs are now offered throughout the United States, and a growing number of health plans [5], including Medicare [6], offer DPPs to eligible plan members at no out-of-pocket cost.

Despite the widespread availability of DPPs and public health efforts that aim to increase DPP engagement, rates of program uptake remain extremely low [7,8]. To date, strategies to increase the DPP uptake have targeted extrinsic barriers to participation (eg, lack of time and cost) through the provision of Web-based DPPs [9] and insurance coverage with limited success [5,6]. In contrast, to our knowledge, no current strategies address intrinsic barriers to participation, such as low levels of motivation to prevent T2DM, yet prior literature suggests that a lack of motivation may be a key barrier to DPP engagement [10]. Accordingly, it is necessary to develop and test scalable approaches to help increase the motivation of millions of Americans who have prediabetes but are not yet taking actions to reduce their risk of progression to T2DM. Such strategies may be most effective if they draw on the principles of self-determination theory to increase autonomous motivation (ie, motivation that arises from internal sources and aligns with personal interests and values) [11,12]. Greater levels of autonomous motivation correlate positively with dietary adherence [13], weight loss [14,15], physical activity [16,17], DPP participation [10], and maintenance of healthy behaviors over time [18,19].

Mobile health (mHealth) apps that are easy to use and do not require a significant time commitment may be effective and highly scalable approaches to increase autonomous motivation to prevent T2DM among those with prediabetes [20,21]. One mHealth app under development, for example, promotes personal well-being by helping users to (1) identify their core values (eg, to be a good parent); (2) reflect on their adherence to these values; and (3) develop the energy and willpower to live in accordance with their core values by improving key health behaviors (eg, sleep, physical activity, and diet). The mHealth app integrates user-entered health information with contextual data (eg, local weather and day of the week) and then delivers brief tailored messages and health tips to help individuals gain awareness of and control over the factors in real-time that influence their ability to engage in self-care behaviors. In this way, the app helps users connect their daily habits and routines with personal interests and values, thereby strengthening autonomous motivation to engage in healthy behaviors. Yet, it is not known whether adults who have already declined participation in offered DPPs are willing to participate in and then engage in offered mHealth programs.

Accordingly, in this 3-arm, mixed-methods pilot randomized controlled trial, we tested the feasibility of recruiting DPP nonenrollees into an mHealth intervention and the acceptability of the mhealth program—used alone and also in conjunction with Fitbit devices (eg, activity tracker and wireless internet-enabled scale) to encourage self-monitoring—among individuals with prediabetes who had declined participation in Web-based or face-to-face DPPs offered at no out-of-pocket expense by their health plans. As we hypothesized that autonomous motivation would be a key proximal mediator of behavioral changes among those who did engage with the intervention, we also estimated the change in study participants' autonomous motivation during the 12-week intervention period. In addition, as Fitbit devices can enhance motivation and self-efficacy through self-determination theory principles [11,22] and self-monitoring techniques [23], we further hypothesized that autonomous motivation to prevent T2DM would increase to a greater degree among individuals who used the app in conjunction with Fitbit devices compared with individuals who used the app alone or who were assigned to the control arm.

Methods

Design

We conducted a 12-week, parallel, 3-arm, mixed-methods pilot randomized controlled trial between May 2017 and February 2018 (NCT03025607). Overall, 69 participants were randomized to 1 of 3 arms (Figure 1) as follows: (1) a group that received information about prediabetes and evidence-based ways to decrease the progression to T2DM, as well as a list of resources

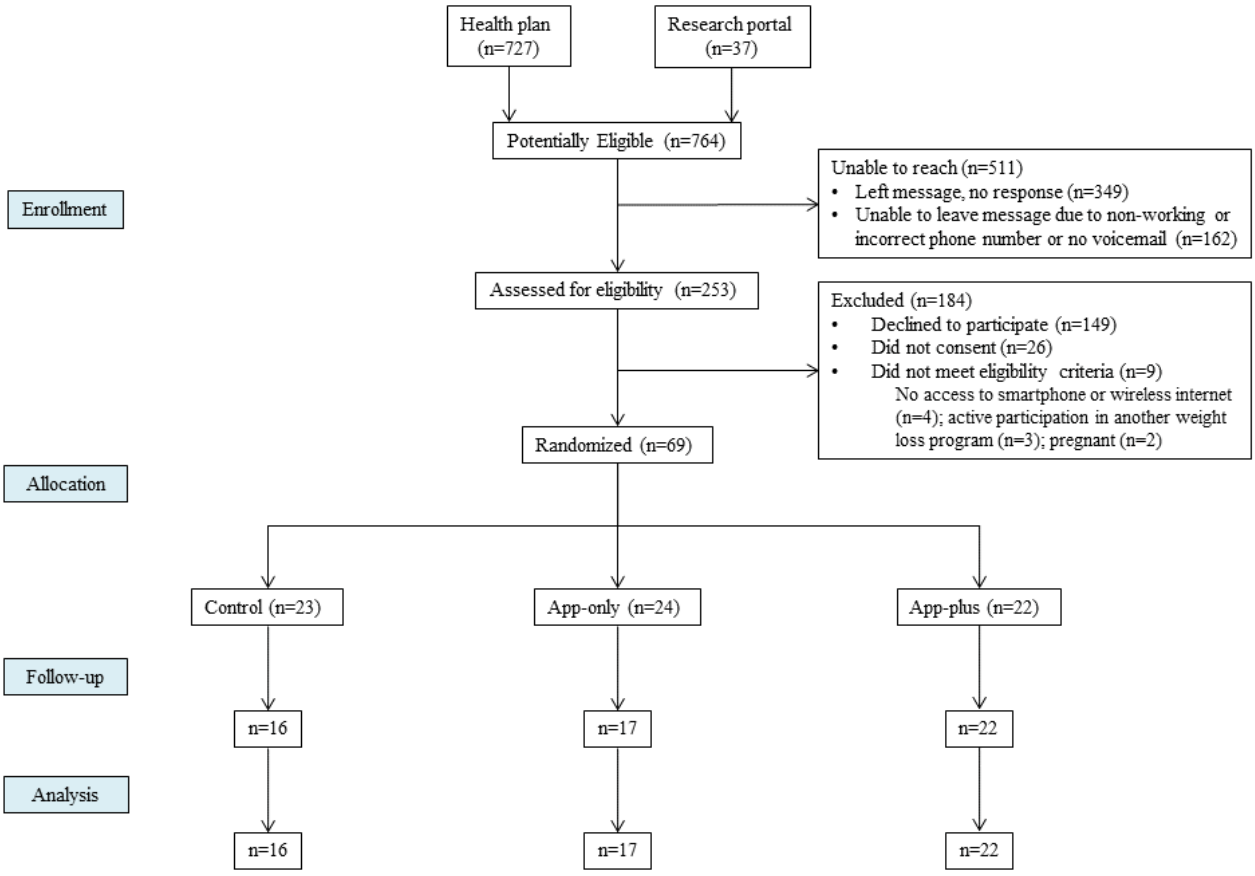
for mHealth tools for monitoring diet, physical activity, and weight (control group); (2) a group that received the same information as the control group and the mobile smartphone app (app-only); and (3) a group that received the same information as the control group, as well as the mobile smartphone app and Fitbit devices (eg, activity tracker and wireless internet-enabled scale) whose results were automatically synced with the mobile app and informed the app's tailored messaging (app-plus). This commercially available app is hosted on Amazon Web Services, with all data encrypted at rest, in transit, and when backed up. We used a mixed-methods sequential explanatory design [24]; quantitative and qualitative data were collected in 2 consecutive phases during the study and then integrated into the final stage of data analysis. This approach enabled us to interpret our quantitative data in the context of qualitative participant experiences. The protocol was approved by the University of Michigan Institutional Review Board (HUM00111389).

Setting and Participants

The intervention was delivered remotely. The inclusion criteria were as follows: (1) nonenrollment in a DPP at least 6 months after invitation from one's health plan to participate at no out-of-pocket cost (ie, DPP nonenrollee); (2) prediabetes based on American Diabetes Association criteria of a hemoglobin A_{1c} (HbA_{1c}) level between 5.7% and 6.4%; (3) access to a personal smartphone; and (4) access to home wireless internet. We excluded women who were pregnant or intended to become pregnant during the intervention period.

We had a unique opportunity to recruit locally, as our institution's self-funded health insurers recently began to offer face-to-face and Web-based DPP options to health plan members (ie, employees, retirees, and students of the University of Michigan or their dependents) with prediabetes at no out-of-pocket cost, yet only 6% of program invitees enrolled in a DPP within the first 6 months (September 2015-February 2016) of the program (unpublished communication). For this pilot study, the University's health plans provided the study team with a random 18.5% (727/3926) sample of DPP nonenrollees. In addition, we posted study recruitment information on the University's health research website to allow interested and potentially eligible individuals to contact our team directly [25]. We attempted to contact all individuals by telephone to invite them to participate in this study. Three attempts were made to contact each individual; a voicemail with the study team's contact information was left after the second attempt. Individuals interested in study participation were screened by telephone to ensure they met the study eligibility criteria, and informed consent was obtained electronically using the RedCap survey platform [26].

Figure 1. Study flow diagram.



Allocation

Individuals who met the study inclusion criteria, provided written informed consent, and completed a baseline questionnaire were assigned to the 3 study groups using 1:1:1 central computerized randomization. The allocation sequence was generated using Stata 14. A Web-based tool, the University of Michigan computerized randomization system (Treatment Assignment Tool-UM, TATUM), was used to allow for blinded treatment allocation. We used stratified randomization with variable block lengths to ensure a balance of age and gender between groups. Owing to the nature of the intervention, it was not possible to blind participants; those performing the analyses, however, were blinded to treatment assignment arms.

Intervention

All participants received the Center for Disease Control and Prevention's 2-page educational handout on prediabetes and evidence-based strategies to prevent the progression to T2DM, as well as a list of free mHealth resources for monitoring weight and physical activity. In addition, app-only and app-plus participants received emailed instructions for setting up the app. App-plus participants received their Fitbit devices and set-up instructions through postal mail. A study team member was available by telephone and email to answer study-related questions and troubleshoot technical issues. Three app-only participants (12.5%) contacted the study team to request assistance with the app set-up. Six app-plus participants (27.2%) contacted the study team to request assistance with setting up the app or Fitbit devices. Once participants began using the app and/or Fitbit devices, there was no further contact with the study team for technology support, and there was no additional planned contact between participants and study team members during the study period.

App-only and app-plus participants were asked to use the smartphone app daily to chart the following health-related habits and behaviors: (1) Sleep; (2) Presence; (3) Activity; (4) Creativity; and (5) Eating (S.P.A.C.E). In addition to charting S.P.A.C.E. on a daily basis, users were asked to reflect on and chart their alignment with personal core values (ie, life purpose); these charted data then informed tailored messages and health tips, as well as predictions of an individual's energy and willpower for the coming day. These predictions are intended to help individuals gain awareness of and control over the factors that influence their health behaviors. Furthermore, app-plus participants were asked to use the Fitbit scale and activity tracker daily to self-monitor weight and physical activity, respectively. These devices interfaced with the app platform such that the Fitbit data informed delivered tailored messages and health tips.

Within the app, users were asked if they wished to receive a daily reminder to chart their day. Users who desired a daily reminder received a push notification at a self-selected time, which reminded them to chart their day. Users who did not desire a daily reminder received no other reminders to use the study-specific device(s).

Primary Quantitative Measures: Feasibility and Acceptability

We evaluated the intervention's feasibility (uptake and retention rates) and acceptability (adherence and qualitative experience). The program feasibility was determined by calculating the intervention uptake rate, defined as the number of participants recruited to the intervention divided by the total number of potentially eligible participants. Furthermore, we calculated the rate of intervention uptake among only those who were reached by telephone. To determine the study retention rate, we calculated the rate of completion of the 12-week survey among all individuals enrolled in the study.

Among app-only and app-plus participants, we measured adherence to the app, defined as the number of days that users entered data into the app during the 12-week intervention period. Among app-plus participants, we measured participant adherence to the Fitbit activity tracker and scale, defined as the number of total days that each of these devices were used during the intervention period.

Secondary Quantitative Measures: Web-based Surveys

Prior to randomization, individuals who consented to study participation were asked to complete a Web-based survey via RedCap, a secure Web app [26]; this first survey was used to collect demographic and socioeconomic information, including age, gender, race, ethnicity, education, and household income. We used the 7-item, validated Treatment Self-Regulation Questionnaire (TSRQ) to measure autonomous motivation to prevent T2DM [27]. Following the 12-week intervention period, participants were emailed a link to the second survey. This survey asked participants to complete the same validated instrument that was collected at baseline. Participants were provided with a US $10 gift card following the completion of each survey (ie, baseline and 12-weeks).

Qualitative Measures: Semistructured Telephone Interviews

Following the 12-week intervention period, we invited all individuals in the app-only and app-plus groups to participate in a semistructured telephone interview. We planned to conduct a minimum of 20 interviews with additional interviews to be conducted only if thematic saturation was not achieved at this point [28]. During the interviews, we explored participants' experiences with the app and Fitbit devices, if applicable. In addition, participants discussed health behavioral changes that occurred as a result of program participation and suggested potential strategies to strengthen and refine the intervention. Of note, interview participants received a US $20 gift card as compensation for their time.

Sample Size

Based on prior studies of autonomous motivation among University of Michigan employees [10], we anticipated that the baseline level of autonomous motivation to prevent T2DM among those who declined DPP participation after invitation by their health plan to be 5.7 (measured on a 1-7 scale with 1 being the lowest and 7 being the highest). During the 12-week intervention period, we anticipated that autonomous motivation

would increase by 0.6 points in the app-only arm and by 0.8 points in the app-plus arm. Assuming an SD of 1.0 for change in autonomous motivation in both arms, we required 29 participants in each arm to provide 80% power to detect these changes in autonomous motivation in the intervention arms compared with the control arm. Prior research demonstrates that a 0.5-point increase in autonomous motivation is associated with markedly higher weight loss and increased physical activity compared with individuals who did not achieve this increase in autonomous motivation [18]. To account for the possibility that some participants may be lost to follow-up during our 12-week intervention, we conservatively inflated our sample size by 20% to enroll 35 participants in each arm.

Owing to administrative changes within the health plan and competing research interests within our institution, the plan provided us with a limited list (727/3926, 18.5% sample) of individuals who were potentially eligible for our study. As such, we were unable to meet our recruitment target. Using our realized sample size (n=69), we conducted a post-hoc power analysis, which showed that we had 80% power to detect a mean difference of ≥0.38 in the intervention arms compared with the control arm.

Statistical Analysis

Quantitative Data Analysis

We used logistic regression to compare differences in rates of engagement between the 2 intervention arms. We used linear regression to compare differences in adherence (ie, app-usage days) between the intervention arms. In addition, we compared changes in autonomous motivation among app-only and app-plus participants versus control participants using a difference-in-differences analytic approach. For continuous outcome measures, we modeled the effect using linear regression, and for dichotomous outcomes, we modeled the effect using logistic regression. The difference-in-difference is an interaction term between a categorical variable indicating the study group (ie, control vs app-only vs app-plus) and a categorical variable indicating the data collection time-point (ie, baseline vs 12-week follow-up). The difference-in-differences design accounts for the possibility that temporal trends unrelated to the intervention may have influenced the study outcome. All analyses were conducted using Stata 14 (StataCorp LP).

Qualitative Data Analysis

Semistructured interviews were recorded, transcribed verbatim, and imported into qualitative analysis software, Dedoose (SocioCultural Research Consultants, Los Angeles, CA, USA). Two investigators independently read and coded transcribed interviews. Interviews were then coded jointly using consensus conferences and analyzed using directed content analysis [29]. Although we planned to conduct a minimum of 20 interviews, no new themes emerged after coding 8 transcripts. Given that thematic saturation was achieved earlier than anticipated, we conducted only 15 interviews.

Results

Intervention Uptake

Figure 1 shows the flow of participants through the study. Contact information for a total of 740 individuals identified as potentially eligible for study participation was provided to us by their health plan, and 37 individuals identified as potentially eligible by self-report through a health research portal. We were unable to reach the majority of potentially eligible individuals (527/777, 68%). Among 253 individuals assessed for eligibility, 244 were eligible to participate, and 28% (69/244) of these eligible individuals consented to study participation and were randomized to 1 of 3 study arms.

Baseline Characteristics

Demographic and socioeconomic characteristics were assessed at baseline (Table 1). Most participants were females (64%), white (65%), and educated, with 91% attaining education beyond high school. The mean age was 51.7 years (11.2). At baseline, mean autonomous motivation score was 6.0 (SD 1.0) among control group participants, 5.8 (SD 1.0) among app-only participants, and 6.0 (SD 1.0) among app-plus participants.

Quantitative Analyses

Retention

Among those randomized (n=69), 55 (80%) completed the 12-week survey. Rates of survey completion varied across study arms. Among participants in control, app-only, and app-plus groups, completion rates were 70% (16/23), 71% (17/24), and 100% (22/22), respectively. Retention differed significantly between app-plus participants and participants in the other 2 study arms combined (P=.004, χ^2).

Adherence

During the 12-week (84-day) intervention period, app-only participants used the app for a mean of 43 days (SD 26.6; 51% of study days), while app-plus participants used the app for a mean of 37 days (SD 26.2; 44% of study days); P value (.34).

Among app-plus participants (n=22), 73% (16/22) used the Fitbit activity tracker for a mean of 32 days (SD 12.0), and 59% (13/22) used the Fitbit scale for a mean of 15.9 days (SD 15.4). Of note, 3 app-only participants paired their personal Fitbits with the app, although they were not instructed to do so as part of the study; these individuals used the Fitbit for a mean of 21 (SD 8) days.

Exploratory Quantitative Outcomes

Table 2 shows the changes in autonomous motivation scores across the study groups. The scores were measured on a scale of 1-7 using the TSRQ; higher scores indicate greater levels. No statistically significant within- or between-group differences were observed in self-reported autonomous motivation.

Table 1. Baseline characteristics of study participants.

Characteristics	Control (n=23)	App-only (n=24)	App-plus (n=22)
Demographics			
Mean age (years), mean (SD)	51.3 (11.0)	52.1 (12.0)	51.6 (11.1)
Female, n (%)	15 (65.2)	15 (62.5)	14 (63.6)
Body mass index in kg/m^2, mean (SD)	33.0 (10.4)	30.7 (9.3)	33.4 (7.8)
Minority race[a], n (%)	6 (28.6)	11 (45.8)	7 (31.8)
Education, n (%)			
High school graduate	3 (13.0)	1 (4.2)	1 (4.6)
More than high school	20 (87.0)	22 (91.7)	21 (95.5)
Household income (in US $), n (%)			
<50,000	7 (31.8)	6 (27.3)	6 (28.6)
50,000-100,000	8 (36.4)	12 (54.6)	6 (28.6)
>100,000	7 (31.8)	4 (18.2)	9 (42.9)
Autonomous motivation to prevent type 2 diabetes mellitus[b], mean (SD)	6.01 (1.0)	5.80 (1.0)	5.96 (1.0)

[a]Defined as any race other than white.

[b]Measured on a scale of 1-7 using the Treatment Self-Regulation Questionnaire. Higher scores indicate greater levels.

Table 2. Difference-in-difference analysis for autonomous motivation scores at 12 weeks compared with baseline.

Study groups	Baseline mean (SE)[a]	12-week mean (SE)	*P* value	
			Within-group difference at 12 weeks)	Difference-in-difference from baseline to 12 weeks
Control (n=16)	6.01 (0.21)	5.87 (0.25)	–0.14 (.57)	Not applicable
App-only (n=17)	5.80 (0.21)	5.88 (0.25)	0.08 (.73)	0.22 (.51)
App-plus (n=22)	5.96 (0.21)	5.90 (0.21)	–0.06 (.72)	0.08 (.77)

[a]All values in this table are predicted from the model.

Participant Experiences With the Intervention

Among 24 app-only participants invited to participate in an interview, 5 individuals (20%) agreed to take part. Among 22 app-plus participants invited to participate in an interview, 10 individuals (45%) agreed to take part. During these interviews, key themes emerged regarding participants' perceptions of the app, capturing those aspects of the app that they liked or disliked (Table 3).

Among 13 interviewees who identified components of the app that they enjoyed, the majority (n=8) appreciated the app's support for self-reflection. For example, one app-only participant commented, "I liked how you had to rank how [you were] feeling [each day]...I thought [that was] an interesting way just to take a step back, just sort of a self-assessment." Others (n=5) noted that the app supported adherence to healthy behaviors over time through daily charting of health habits (eg, diet, physical activity, and sleep), light-touch health tips, and educational videos. As noted by one app-only participant, "[the app] was a good reminder...to help push [me] to keep moving...doing more and more. "

Among 11 participants who identified components of the app that they did not enjoy, almost half (n=5) commented that daily use of the app felt burdensome as a result of the minimal day-to-day variation in individual health behaviors, redundancy of educational content, and perceived arbitrariness of future predictions. One app-plus participant commented that he was initially motivated to chart daily; however, he also said:

> *...after a while...I lost interest in trying to understand what it was doing for me other than just keeping track and telling me that tomorrow it's supposed to rain. You might have a bad day.*

Table 3. Participants' perceptions of the mHealth app and representative quotes.

Participant perceptions	Representative quotes
Encouraged reflection on factors that influence health	• "[The App] helps me think about how I can use [my family] to support me...even though they live far away, I can just have a conversation with them and try to use them as part of my support, as well as my community, which are my friends, my church, my school parents, things like that. 'Cause I realize that these are actually part of the environment that could help me be a healthier person." *(App-only)* • "It makes me decompress from my day and just think, "How could I have made my day better? What did I do? What didn't I do?" *(App-plus)*
Supported healthy behaviors	• "I was more conscious of what I ate. I started...drinking more water, less caffeinated beverages, less carbonated beverages...I wasn't as tired. I set a goal where I was going to bed by a certain time." *(App-only)* • "When I go see my doctor, it's kind of like, '...you need to exercise more...you need to change your diet'. But the nice thing about [the app] was [it] broke it down into these things that you could learn about that allowed you to have a better understanding [of] your health condition...and also how you can sort of prevent certain health risks from happening." *(App-only)*
Daily use was burdensome	• "There [were] a lot of questions about how I feel today...it just seemed to be a little bit of the same old same old every day or every time I looked at it." *(App-only)* • "[The App] just got too time consuming and I just lost interest in keeping track of all that data. It just became too overwhelming, I was doing other things." *(App-plus)*
Failed to consider personal circumstances	• "I [have] paroxysmal afib, which means some days...I didn't feel very energetic...[But] there was no way to [tell the app], 'this day is different for completely non-purpose related reasons'." *(App-plus)* • "Sometimes [things] go completely awry and just change what's gonna happen, my plan for the day. So outside factors...absolutely [have] an impact on your day. So you can still be positive, you can still have a plan for exercise. But sometimes, there's things that come up..." *(App-plus)*

Four individuals voiced frustration with certain health tips delivered by the app, as these failed to recognize personal or environmental circumstances that transiently influenced one's health habits, energy, or willpower. For example, an app-plus participant noted:

> *...time I was on vacation, and I have to work really hard to get the vacation. And I had a drink every single day, not a lot, just maybe one, and there was a thing that came up about sleeping and limiting your alcohol intake, and I'm going, "Oh, for God's sakes. I shouldn't even put any of that down."*

Among app-plus interviewees, all (n=10) used the Fitbit activity tracker, and most (n=6) noted that it facilitated engagement in routine physical activity. For example, one participant said:

> *I live about two miles away from our office. I ended up much more in the mode of, "I'm gonna walk if it's all possible."*

Several participants specifically appreciated the activity tracker's concrete step count goal, and one noted:

> *...looking at [activity] from a more lucid mathematical standpoint was very helpful. It made me more active without having to engage in an abrupt behavior or thought change.*

Among app-plus interviewees, 8 used the scale and appreciated the ease with which the data synced with the Fitbit app. One participant commented "I thought it was wonderful...[you just] step on this little device and magically it goes into your statistics, and I get a running account of if my weight's going up or down or whatever." Similarly, another noted, "I just step on the scale and it's recorded in the Fitbit app, and that was handy 'cause it keeps a record."

Thirteen interview participants identified specific health behavioral changes that resulted from participation in this intervention. These included increased physical activity (n=9), improved dietary habits (n=8), increased awareness of other factors that influence health and well-being such as social connectedness and adequate sleep (n=6).

Thirteen interview participants suggested strategies to enhance the intervention. Five participants recommended adding some level of "human contact" to support behavioral change better. An app-plus participant commented, "I would have enjoyed talking with an actual person…to get more advice." Three participants thought that more concrete goal-setting could better help participants achieve health goals. For example, an app-plus participant noted:

> *[The app] didn't seem to offer...concrete things to do. It just sort of asking me to reflect on how I did in sort of pretty unstructured ways. [I wanted to] be able to set concrete things to do...Instead of just asking me how active I was, ask if I [met my goal of] walking at least four miles a day...*

Another suggested the addition of concrete nutritional advice so that participants may know:

> *...what not to eat, what to eat, and what are the nutritional values of different things, and how you can manage your day based on your work schedule, when you should be eating, what you should be eating, how much you should be eating and you could still feel hungry.*

Discussion

Principal Findings

To the best of our knowledge, this is the first study to test an intervention to support healthy behaviors among individuals with prediabetes who had recently declined participation in Web-based or face-to-face DPPs offered at no cost. Our findings demonstrate that it is indeed feasible to recruit DPP nonenrollees to an mHealth intervention. Nearly one-third of eligible individuals enrolled in this intervention despite previously declining to participate in free Web-based and group-based DPPs offered by our University's self-funded insurers. Furthermore, the app—used alone and also in conjunction with Fitbit devices—was acceptable among intervention group participants, as indicated by high levels of adherence and positive qualitative experiences.

Retention differed markedly between app-plus participants and participants in the other 2 study arms. One explanation for between-arm differences in retention is that the Fitbit devices enhanced the intervention's acceptability and perceived value to participants. Fitbit devices incorporate established behavioral change techniques (eg, self-monitoring, feedback, and goal-setting) [30], and our qualitative data suggest that participants' enjoyment of these features may have motivated study retention. Alternatively, because app-plus participants received a more robust intervention, they may have felt a greater sense of obligation to the study, making them more likely to complete the 12-week survey. Adherence to the app did not differ markedly between intervention groups. In qualitative interviews, participants indicated that they discontinued the app daily use owing to the perceived burden of data entry and lack of personal relevance. These reasons for the discontinued app use are consistent with those previously described in the literature [31].

We examined the intervention's preliminary efficacy on autonomous motivation to prevent T2DM, which we hypothesized to be a key proximal mediator of behavioral change. Our analyses did not demonstrate statistically significant differences in levels of autonomous motivation between intervention arms. It is plausible that we were unable to discern changes in autonomous motivation owing to higher-than-predicted baseline levels of autonomous motivation and resultant ceiling effect of the TSRQ. While high baseline levels of autonomous motivation may have occurred by random chance, it is also possible that these high levels identify a nonrandom subset of DPP nonenrollees who are motivated to prevent T2DM, yet face other barriers to DPP enrollment (eg, lack of time). Accordingly, high levels of postintervention autonomous motivation across arms may reflect intrinsic characteristics of our study participants rather than the intervention's effect. Given the importance of autonomous motivation for initiating and sustaining healthy behaviors, it is critically important to characterize autonomous motivation levels among the broader population of DPP nonenrollees and conduct a larger-scale effectiveness trial to examine changes in autonomous motivation specifically among individuals with lower baseline levels. Another possibility is that 12 weeks was a too short period to observe marked improvements in autonomous motivation; prior studies have examined changes over longer time periods. In addition, future research should explore factors other than low levels of autonomous motivation that may deter the DPP uptake to inform additional targeted interventions to address these barriers specifically.

Mobile smartphone apps and other mHealth technologies are increasingly used as tools to promote lifestyle changes [32], and technology-assisted translations of the DPP have been used to improve program reach [33]. While such programs may be cost-effective and convenient, their effectiveness is variable, and little is known about the populations most likely to engage in or benefit from mHealth programs [33,34]. Without such knowledge, these programs cannot be adequately tailored or disseminated to those most likely to benefit from them. In this study, we specifically recruited individuals who declined participation in free DPPs, and, through qualitative interviews, we gained insight into key opportunities to augment the effectiveness of this low-intensity mHealth program. Notably, several participants expressed a desire for enhanced interpersonal contact during the study period. In addition to fostering a sense of personal connection, such contact may facilitate concrete goal-setting and follow-up, thereby optimizing behavioral change outcomes; prior mHealth interventions for weight loss, for example, have proven most effective when combined with health coaching [35-37]. Furthermore, we demonstrated that some interpersonal contact is necessary for program on-boarding.

Limitations

First, we aimed to enroll 35 individuals in each study arm, but we were unable to meet this recruitment target owing to administrative changes within the health plan and competing research interests within our institution. Thus, we were not powered to detect our hypothesized changes in autonomous motivation, and baseline autonomous motivation scores were higher than expected among our study participants. Second, we recruited individuals from a single regional health plan, and our results may not be generalizable to other populations; our study participants were highly educated with access to personal smartphones and home wireless internet. As such, they may have been more willing and able to engage in a mHealth intervention for diabetes prevention than less educated or resourced individuals [38,39]. Future work could aim to engage a broader cohort of DPP nonparticipants with lower levels of baseline autonomous motivation and more diverse sociodemographic characteristics. Finally, because this was a pilot study designed to assess the feasibility and acceptability, we were not powered to examine changes in clinically relevant behaviors for T2DM prevention (eg, weight loss and increased physical activity); these outcomes warrant investigation in larger-scale trials.

Conclusions

National initiative [40,41] and policies [42] promote DPPs as the dominant diabetes prevention strategy, yet the ability of DPPs to improve population health is compromised by the low program uptake. Alternative strategies are urgently needed to help the large majority of individuals with prediabetes prevent

T2DM and T2DM-related complications. In this pilot study, we demonstrate the feasibility and acceptability of a low-intensity mHealth program among some individuals with prediabetes who do not desire participation in formal DPPs. However, additional strategies are also needed to engage those DPP nonparticipants who also decline mHealth programs. In future work, we will refine the existing intervention by incorporating participant-identified preferences for increased interpersonal contact and concrete goal-setting. We will then conduct a larger-scale effectiveness trial to examine changes in key proximal mediators of behavioral change (eg, autonomous motivation and self-efficacy), as well as changes in clinically relevant outcomes (eg, weight, HbA_{1c}, and physical activity). Furthermore, we will explore needs and preferences for lifestyle change approaches among a broad population of DPP nonparticipants, and these data will be used to develop additional tailored interventions for T2DM prevention.

Acknowledgments

DG acknowledges support from the Veterans Health Administration in her role as a Health Services Research Fellow. The authors acknowledge funding support from the Blue Cross Blue Shield Foundation of Michigan, the National Med-Peds Residency Association, the Robert Wood Johnson Clinical Scholars Program, and the National Institute of Diabetes and Digestive and Kidney Diseases (P30DK092926). JTK is a Veterans Affairs (VA) Health Services Research and Development Service Career Development awardee at the Ann Arbor VA. JTK also received grants from the Robert Wood Johnson Foundation, from the US Department of Veterans Affairs, and from the Donaghue Foundation. LRS is a K01 awardee from the National Institute of Diabetes and Digestive and Kidney Diseases (DK107456).

Conflicts of Interest

DG, BL, TA, EHJ, AF, CR, and MH declare that they have no conflicts of interest. JTK has received consulting fees from SeeChange Health and HealthMine, and a speaking honorarium from AbilTo, Inc.

Multimedia Appendix 1

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - mhealth_v7i1e11267_fig.pdf]

References

1. Centers for Disease Control and Prevention. National Diabetes Statistics Report 2017. Atlanta, GA; 2017. URL: https://www.cdc.gov/diabetes/data/statistics/statistics-report.html [accessed 2018-04-04] [WebCite Cache ID 73ktt1YUX]
2. American Diabetes Association. Economic costs of diabetes in the U.S. in 2012. Diabetes Care 2013 Apr;36(4):1033-1046 [FREE Full text] [doi: 10.2337/dc12-2625] [Medline: 23468086]
3. Knowler WC, Barrett-Connor E, Fowler SE, Hamman RF, Lachin JM, Walker EA, Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. N Engl J Med 2002 Feb 07;346(6):393-403 [FREE Full text] [doi: 10.1056/NEJMoa012512] [Medline: 11832527]
4. Dunkley AJ, Bodicoat DH, Greaves CJ, Russell C, Yates T, Davies MJ, et al. Diabetes prevention in the real world: effectiveness of pragmatic lifestyle interventions for the prevention of type 2 diabetes and of the impact of adherence to guideline recommendations: a systematic review and meta-analysis. Diabetes Care 2014 Apr;37(4):922-933 [FREE Full text] [doi: 10.2337/dc13-2195] [Medline: 24652723]
5. AHIP. Diabetes. 2016. URL: https://www.ahip.org/diabetes/ [accessed 2018-11-07] [WebCite Cache ID 73kt4GRiX]
6. Centers for Medicare and Medicaid Services. Medicare Diabetes Prevention Program (MDPP) Expanded Model. 2017 Nov 02. URL: https://innovation.cms.gov/initiatives/medicare-diabetes-prevention-program/ [accessed 2018-04-04] [WebCite Cache ID 73ktURHo1]
7. Wareham N. Mind the gap: efficacy versus effectiveness of lifestyle interventions to prevent diabetes. The Lancet Diabetes & Endocrinology 2015 Mar;3(3):160-161 [FREE Full text] [doi: 10.1016/S2213-8587(15)70015-X]
8. Jackson SL, Long Q, Rhee MK, Olson DE, Tomolo AM, Cunningham SA, et al. Weight loss and incidence of diabetes with the Veterans Health Administration MOVE! lifestyle change programme: an observational study. The Lancet Diabetes & Endocrinology 2015 Mar;3(3):173-180. [doi: 10.1016/S2213-8587(14)70267-0]
9. Sepah SC, Jiang L, Peters AL. Long-term outcomes of a Web-based diabetes prevention program: 2-year results of a single-arm longitudinal study. J Med Internet Res 2015;17(4):e92 [FREE Full text] [doi: 10.2196/jmir.4052] [Medline: 25863515]
10. Kullgren JT, Knaus M, Jenkins KR, Heisler M. Mixed methods study of engagement in behaviors to prevent type 2 diabetes among employees with pre-diabetes. BMJ Open Diabetes Res Care 2016;4(1):e000212 [FREE Full text] [doi: 10.1136/bmjdrc-2016-000212] [Medline: 27738513]
11. Donnachie C, Wyke S, Mutrie N, Hunt K. 'It's like a personal motivator that you carried around wi' you': utilising self-determination theory to understand men's experiences of using pedometers to increase physical activity in a weight

management programme. Int J Behav Nutr Phys Act 2017 Dec 05;14(1):61 [FREE Full text] [doi: 10.1186/s12966-017-0505-z] [Medline: 28476118]
12. Deci E, Ryan R. Self-determination theory in health care and its relations to motivational interviewing: a few comments. Int J Behav Nutr Phys Act 2012 Mar 02;9(1):24 [FREE Full text] [doi: 10.1186/1479-5868-9-24] [Medline: 22385839]
13. Shigaki C, Kruse RL, Mehr D, Sheldon KM, Bin G, Moore C, et al. Motivation and diabetes self-management. Chronic Illn 2010 Sep;6(3):202-214. [doi: 10.1177/1742395310375630] [Medline: 20675362]
14. Williams GC, Grow VM, Freedman ZR, Ryan RM, Deci EL. Motivational predictors of weight loss and weight-loss maintenance. J Pers Soc Psychol 1996 Jan;70(1):115-126. [Medline: 8558405]
15. Teixeira PJ, Carraça EV, Marques MM, Rutter H, Oppert J, De BI, et al. Successful behavior change in obesity interventions in adults: a systematic review of self-regulation mediators. BMC Med 2015 Apr 16;13:84 [FREE Full text] [doi: 10.1186/s12916-015-0323-6] [Medline: 25907778]
16. Hurkmans EJ, Maes S, de Gucht V, Knittle K, Peeters AJ, Ronday HK, et al. Motivation as a determinant of physical activity in patients with rheumatoid arthritis. Arthritis Care Res (Hoboken) 2010 Mar;62(3):371-377 [FREE Full text] [doi: 10.1002/acr.20106] [Medline: 20391483]
17. Koponen AM, Simonsen N, Suominen S. Determinants of physical activity among patients with type 2 diabetes: the role of perceived autonomy support, autonomous motivation and self-care competence. Psychol Health Med 2017 Dec;22(3):332-344. [doi: 10.1080/13548506.2016.1154179] [Medline: 26952696]
18. Silva MN, Markland D, Carraça EV, Vieira PN, Coutinho SR, Minderico CS, et al. Exercise autonomous motivation predicts 3-yr weight loss in women. Med Sci Sports Exerc 2011 Apr;43(4):728-737. [doi: 10.1249/MSS.0b013e3181f3818f] [Medline: 20689448]
19. Teixeira PJ, Silva MN, Mata J, Palmeira AL, Markland D. Motivation, self-determination, and long-term weight control. Int J Behav Nutr Phys Act 2012 Mar 02;9:22 [FREE Full text] [doi: 10.1186/1479-5868-9-22] [Medline: 22385818]
20. Azar KMJ, Lesser LI, Laing BY, Stephens J, Aurora MS, Burke LE, et al. Mobile applications for weight management: theory-based content analysis. Am J Prev Med 2013 Nov;45(5):583-589. [doi: 10.1016/j.amepre.2013.07.005] [Medline: 24139771]
21. JOOL Health. The JOOL Platform. 2018. URL: https://www.joolhealth.com/the-jool-platform/mobile-and-web-app/ [accessed 2017-06-14] [WebCite Cache ID 73kuuq4Be]
22. Asimakopoulos S, Asimakopoulos G, Spillers F. Motivation and User Engagement in Fitness Tracking: Heuristics for Mobile Healthcare Wearables. Informatics 2017 Jan 22;4(1):5. [doi: 10.3390/informatics4010005]
23. Lutes LD, Steinbaugh EK. Theoretical models for pedometer use in physical activity interventions. Physical Therapy Reviews 2013 Jul 19;15(3):143-153. [doi: 10.1179/1743288X10Y.0000000002]
24. Ivankova NV. Using Mixed-Methods Sequential Explanatory Design: From Theory to Practice. Field Methods 2006 Feb 01;18(1):3-20. [doi: 10.1177/1525822X05282260]
25. UMHealthResearch. 2018 May 08. URL: https://umhealthresearch.org/ [accessed 2018-05-08] [WebCite Cache ID 73kv7IW6Q]
26. REDCap. Accessed June 19, 2017 URL: https://www.project-redcap.org/ [accessed 2018-11-07] [WebCite Cache ID 73kzFNtii]
27. Levesque CS, Williams GC, Elliot D, Pickering MA, Bodenhamer B, Finley PJ. Validating the theoretical structure of the Treatment Self-Regulation Questionnaire (TSRQ) across three different health behaviors. Health Educ Res 2007 Oct;22(5):691-702. [doi: 10.1093/her/cyl148] [Medline: 17138613]
28. Francis JJ, Johnston M, Robertson C, Glidewell L, Entwistle V, Eccles MP, et al. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. Psychol Health 2010 Dec;25(10):1229-1245. [doi: 10.1080/08870440903194015] [Medline: 20204937]
29. Hsieh H, Shannon SE. Three approaches to qualitative content analysis. Qual Health Res 2005 Nov;15(9):1277-1288. [doi: 10.1177/1049732305276687] [Medline: 16204405]
30. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res 2014 Aug;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
31. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR Mhealth Uhealth 2015 Nov;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
32. Zhao J, Freeman B, Li M. Can Mobile Phone Apps Influence People's Health Behavior Change? An Evidence Review. J Med Internet Res 2016 Oct 31;18(11):e287 [FREE Full text] [doi: 10.2196/jmir.5692] [Medline: 27806926]
33. Grock S, Ku J, Kim J, Moin T. A Review of Technology-Assisted Interventions for Diabetes Prevention. Curr Diab Rep 2017 Sep 23;17(11):107. [doi: 10.1007/s11892-017-0948-2] [Medline: 28942537]
34. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
35. Partridge SR, McGeechan K, Hebden L, Balestracci K, Wong AT, Denney-Wilson E, et al. Effectiveness of a mHealth Lifestyle Program With Telephone Support (TXT2BFiT) to Prevent Unhealthy Weight Gain in Young Adults: Randomized

Controlled Trial. JMIR Mhealth Uhealth 2015 Jun 15;3(2):e66 [FREE Full text] [doi: 10.2196/mhealth.4530] [Medline: 26076688]
36. Mao AY, Chen C, Magana C, Caballero BK, Olayiwola JN. A Mobile Phone-Based Health Coaching Intervention for Weight Loss and Blood Pressure Reduction in a National Payer Population: A Retrospective Study. JMIR Mhealth Uhealth 2017 Jun 08;5(6):e80 [FREE Full text] [doi: 10.2196/mhealth.7591] [Medline: 28596147]
37. Spring B, Pellegrini C, McFadden HG, Pfammatter AF, Stump TK, Siddique J, et al. Multicomponent mHealth Intervention for Large, Sustained Change in Multiple Diet and Activity Risk Behaviors: The Make Better Choices 2 Randomized Controlled Trial. J Med Internet Res 2018 Jun 19;20(6):e10528 [FREE Full text] [doi: 10.2196/10528] [Medline: 29921561]
38. Margolis R. Educational differences in healthy behavior changes and adherence among middle-aged Americans. J Health Soc Behav 2013 Aug;54(3):353-368 [FREE Full text] [doi: 10.1177/0022146513489312] [Medline: 23988727]
39. Joiner KL, Nam S, Whittemore R. Lifestyle interventions based on the diabetes prevention program delivered via eHealth: A systematic review and meta-analysis. Prev Med 2017 Jul;100:194-207 [FREE Full text] [doi: 10.1016/j.ypmed.2017.04.033] [Medline: 28456513]
40. American Medical Association. Prevent Diabetes STAT. URL: https://preventdiabetesstat.org/ [accessed 2018-11-07] [WebCite Cache ID 73kyTbbYB]
41. Ad Council. 1 in 3 American adults has prediabetes. Do you?. 2018. URL: https://doihaveprediabetes.org/ [accessed 2018-11-07] [WebCite Cache ID 73kz9Io6x]
42. Centers for Medicare and Medicaid Services. Webinar: Medicare Diabetes Prevention Program - Overview. 2018. URL: https://innovation.cms.gov/resources/dpp-overview.html [accessed 2018-11-07] [WebCite Cache ID 73kz9Io6x]

Abbreviations

DPP: Diabetes Prevention Program
HbA $_{1c}$: hemoglobin A_{1c}
mHealth: mobile health
T2DM: type 2 diabetes mellitus
TSRQ: Treatment Self-Regulation Questionnaire

Edited by G Eysenbach; submitted 09.06.18; peer-reviewed by L Mantoani, J Claes, D Nault; comments to author 26.07.18; revised version received 23.08.18; accepted 10.09.18; published 09.01.19.

Please cite as:
Griauzde D, Kullgren JT, Liestenfeltz B, Ansari T, Johnson EH, Fedewa A, Saslow LR, Richardson C, Heisler M
A Mobile Phone-Based Program to Promote Healthy Behaviors Among Adults With Prediabetes Who Declined Participation in Free Diabetes Prevention Programs: Mixed-Methods Pilot Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(1):e11267
URL: http://mhealth.jmir.org/2019/1/e11267/
doi:10.2196/11267
PMID:30626566

Original Paper

The QuitIT Coping Skills Game for Promoting Tobacco Cessation Among Smokers Diagnosed With Cancer: Pilot Randomized Controlled Trial

Paul Krebs[1], PhD; Jack Burkhalter[2], PhD; Jeffrey Fiske[3], BS; Herbert Snow[3], BFA; Elizabeth Schofield[2], MPH; Michelle Iocolano[2], MA; Sarah Borderud[2], MA; Jamie S Ostroff[2], PhD

[1]Department of Population Health, New York University School of Medicine, New York, NY, United States

[2]Department of Psychiatry and Behavioral Sciences, Memorial Sloan-Kettering Cancer Center, New York, NY, United States

[3]Muzzy Lane Software, Newburyport, MA, United States

Corresponding Author:
Paul Krebs, PhD
Department of Population Health
New York University School of Medicine
227 East 30th Street
New York, NY, 10016
United States
Phone: 1 6465012637
Email: Paul.Krebs@nyumc.org

Abstract

Background: Although smoking cessation apps have become popular, few have been tested in randomized clinical trials or undergone formative evaluation with target users.

Objective: We developed a cessation app targeting tobacco-dependent cancer patients. Game design and behavioral rehearsal principles were incorporated to help smokers identify, model, and practice coping strategies to avoid relapse to smoking. In this randomized pilot trial, we examined feasibility (recruitment and retention rates), acceptability (patient satisfaction), quitting self-confidence, and other cessation-related indices to guide the development of a larger trial.

Methods: We randomized 42 English-speaking cancer patients scheduled for surgical treatment to either the Standard Care (SC; telecounseling and cessation pharmacotherapies) or the experimental QuitIT study arm (SC and QuitIT game). Gameplay parameters were captured in-game; satisfaction with the game was assessed at 1-month follow-up. We report study screening, exclusion, and refusal reasons; compare refusal and attrition by key demographic and clinical variables; and report tobacco-related outcomes.

Results: Follow-up data were collected from 65% (13/20) patients in the QuitIT and 61% (11/18) in SC arms. Study enrollees were 71% (27/38) females, 92% (35/38) white people, and 95% (36/38) non-Hispanic people. Most had either lung (12/38, 32%) or gastrointestinal (9/38, 24%) cancer. Those dropping out were less likely than completers to have used a tablet (P<.01) and have played the game at all (P=.02) and more likely to be older (P=.05). Of 20 patients in the QuitIT arm, 40% (8/20) played the game (system data). There were no differences between those who played and did not play by demographic, clinical, technology use, and tobacco-related variables. Users completed an average of 2.5 (SD 4.0) episodes out of 10. A nonsignificant trend was found for increased confidence to quit in the QuitIT arm (d=0.25, 95% CI −0.56 to 1.06), and more participants were abstinent in the QuitIT group than in the SC arm (4/13, 30%, vs 2/11, 18%). Satisfaction with gameplay was largely positive, with most respondents enjoying use, relating to the characters, and endorsing that gameplay helped them cope with actual smoking urges.

Conclusions: Recruitment and retention difficulties suggest that the perihospitalization period may be a less than ideal time for delivering a smoking cessation app intervention. Framing of the app as a “game” may have decreased receptivity as participants may have been preoccupied with hospitalization demands and illness concerns. Less tablet experience and older age were associated with participant dropout. Although satisfaction with the gameplay was high, 60% (12/20) of QuitIT participants did not play the game. Paying more attention to patient engagement, changing the intervention delivery period, providing additional reward and support for use, and improving cessation app training may bolster feasibility for a larger trial.

Trial Registration: ClinicalTrials.gov NCT01915836; https://clinicaltrials.gov/ct2/show/NCT01915836 (Archived by WebCite at http://www.webcitation.org/73vGsjG0Y)

(JMIR Mhealth Uhealth 2019;7(1):e10071) doi:10.2196/10071

KEYWORDS

tobacco; cancer; mHealth; app; mobile phone

Introduction

Smartphones and tablet computers are approaching universal use and have opened up new possibilities for the delivery of smoking cessation intervention apps [1,2]. As of 2015, 441 English-language smoking cessation apps were available in app stores [3]. Significant possibilities exist for the development of evidence-based smoking cessation apps targeted to meeting the needs of specific populations of smokers, including young adults, persons with chronic health conditions, etc.

While hundreds of smoking cessation apps are available, few have been subjected to clinical trials or formative evaluation with target users [3,4]. A review of the scientific literature through 2015 found only 3 cessation apps that had been tested in a randomized trial design [3]. This lack of research is concerning because 66% of users of mobile health (mHealth) apps report logging in at least once a day, and 55% users spend at least 11 minutes interacting with them [5]. In addition, most (97%) existing tobacco-related apps have not incorporated principles of effective behavior change [6]. The largest category of smoking-related apps are "calculators," which track money saved and other quitting metrics. Studies have demonstrated, however, that this simple approach does not help smokers quit [4,7,8].

Cessation apps also often fail to take advantage of features to improve user engagement [9]. In a national survey of health app developers, about half of respondents (46%) had stopped using some health apps, primarily due to high data entry burden, loss of interest, and hidden costs [5]. Gamification or "the application of game design principles in order to change behaviors in nongame situations" [10] offers potential to engage mHealth users and promote behavior change [11-14]. Gamification of smoking cessation interventions has led to higher reported levels of engagement in a comparison of mobile apps that featured education and progress tracking [15]. For instance, the Super Smoky game, which focused on youth smokers, was found to increase motivation to quit smoking [16]. However, depth of gameplay was still limited and consisted of having users turn their avatar away from cigarettes and smoking opportunities and did not contain comprehensive information on quitting.

Our research group has developed a smoking cessation intervention based on game design principles and behavioral rehearsal therapy to help smokers identify, model, and practice coping strategies to avoid relapse to smoking [17-19]. Our development process included expert focus groups, prototyping with game developers, and think-aloud testing with a sample of 20 smokers with a history of cancer [18]. We developed the game for use with hospitalized smokers due to our observation of high relapse following hospital discharge. Effective smoking cessation and relapse prevention interventions require participant engagement in a range of complex challenges and strategies such as identifying tobacco use triggers, engaging in alternative coping behaviors, seeking social support from family and friends about tobacco use, modifying one's internal dialogue, and dealing with inevitable slips to prevent relapse [20]. Translating these evidence-based strategies into a game required an immersive app in which users can learn and practice these techniques in a realistic context. Through repeated exposure to conditioned cues to smoke, (eg, socializing with friends who smoke), such a game environment may be able to help smokers virtually practice and master coping skills and build crucial self-efficacy for managing smoking urges. Our premise is that an intervention that combines virtually presented smoking cues with engaging narrative and personally relevant coping skills practice may help smokers overcome barriers to quitting and maintaining tobacco abstinence. The goal of the project was to develop a cessation treatment app using an immersive laptop or tablet-based game environment to help smokers cope with smoking urges and prevent smoking lapses. After developing the game through a formative evaluation process, we conducted a randomized pilot study to examine the feasibility of conducting a subsequent clinical trial (NCT01915836), its acceptability, and preliminary data regarding the game's effects on tobacco cessation coping and use outcomes.

Methods

Participants and Procedures

Memorial Sloan-Kettering Cancer Center (MSKCC) patients who reported being current smokers at their initial medical appointments were advised to quit by their attending physician and referred to the MSKCC Tobacco Treatment Program. Research staff contacted potentially eligible patients via telephone to describe study procedures and screen for additional eligibility criteria. Eligible participants were English-speaking patients with a recent (within the past 6 months) cancer diagnosis or mass suspicious of cancer, those scheduled for surgery to remove a localized tumor, those who reported smoking cigarettes within the past 30 days, and those who had sufficient sensory acuity and manual dexterity to use a computer game. Those with a distant metastatic disease, major psychiatric illness, cognitive impairment, and inability to comply with study procedures or provide informed consent were excluded from the study. The research assistant (RA) met participants in the hospital a day or two following their surgery to complete informed consent and baseline assessments and to provide training on the use of the intervention.

Randomization or Design

Participants were stratified by age (<65, ≥65 years) and randomly assigned with a 1:1 ratio to two study arms. We stratified by age to control for the potential relationship between age and prior mobile devices experience. Participants completed a baseline survey at enrollment and a follow-up survey via phone or mail 1 month following hospital discharge.

Intervention Conditions

Standard Care

Participants were offered at no cost 4 telephone or bedside counseling sessions and our in-house print cessation educational materials [21]. Trained oncology nurses certified as tobacco treatment specialists who follow evidence-based behavioral and pharmacological best practices conducted the counseling sessions. The initial session focused on motivation building, choosing a quit date, if applicable, reviewing the print educational materials, and providing information about and arranging the provision of cessation pharmacotherapies. The second counseling session focused on coping with smoking urges and preventing smoking relapse. The third and fourth counseling sessions focused on relapse prevention or recycling to a repeat quit attempt for those who had resumed smoking. At the end of each session, the tobacco treatment specialists completed a checklist outlining the goals of each session to track patient adherence and treatment fidelity.

Standard Care + Smoking Cues Coping Skills Game (QuitIT)

Patients assigned to the QuitIT condition were offered Standard Care (SC) in addition to having the QuitIT game installed on an iPad device. The RA trained participants during their hospital stay on how to use the game. Training sessions included a verbal overview of the game, its rules, and its objectives. Then, participants were asked to watch a brief tutorial video, play a practice game episode, and afterwards, the RA evaluated patients' comprehension of gameplay. Patients were encouraged to play 3-4 game episodes per week for a 1-month period posthospitalization. Participants were loaned an iPad for 1 month and instructed to contact the RA if they encountered technical difficulties.

The intervention used gaming techniques to exemplify key behavioral strategies based on the social cognitive theory [18]. The game was conceptualized in a narrative structure meant to engage users in each of 10 episodes featuring different characters across 9 situations, all depicting common smoking-related triggers. These included getting ready in the morning, coming home from work, driving in a car, having a frustrating phone conversation, and being offered a cigarette while drinking with friends. The goal of each episode was for users to guide the character through a series of tempting situations and thoughts without resorting to smoking. Users clicked on response choices to guide the character and direct the story line. The screen presented an "urge to smoke" meter, which helped prompt users to monitor and engage the character in appropriate strategies. If a player did not assist the character, the character would slip and smoke, presenting the user with feedback and the opportunity to try again. After users successfully negotiated an episode, the game moved on to the next one. Users received points and badges for avoiding smoking and completing episodes. Multimedia Appendix 1 shows 3 screenshots from the intervention. Screenshot 1 shows the main game screen; here one can see progress across the 9 scenarios and tutorial as well as the badges that users have earned (on the right column). New scenarios get unlocked when one is completed, and users can replay previous scenarios again to earn more points and badges. Badges represent the various coping strategies users employ in the scenarios, with the idea that they should diversify their coping strategies to earn different badges. Screenshot 2 shows a prototypical challenge scenario, in this case, getting the character Ann out of the apartment without smoking. Across the top of the screen are meters showing that Ann's triggers in this scene are nicotine withdrawal and pain. In addition, we see her smoking urge meter. Users click on the cards at the bottom to choose her next action. The cards represent various coping strategies, including cognitive self-talk, relaxation strategies, and using cessation medications, as well as a distractor card that contains a nonhelpful choice that would increase her smoking urge. In screenshot 3, the challenge is for Liz to not smoke during a night out. Here the challenges at the top are wanting to celebrate and feel social; she has chosen a behavioral strategy, switching to drinking ginger ale to stay in control of the situation and avoid smoking. To reinforce coping strategies outside of gameplay, we provided participants with a set of real-life "coping cards," which resembled playing cards and outlined the primary coping strategies in the game and featured scenes from the game.

Measures

Demographics and Medical Status

At baseline, participants reported demographics including age, sex, ethnicity, education, occupation, comorbid medical conditions, and smoking and quitting history. Medical charts were reviewed, and data on cancer diagnosis and treatment were extracted.

Tobacco Use

At baseline, tobacco use and quitting history were assessed with standard measures adapted from the National Adult Tobacco Survey [22]. One month after study entry (1 month), smoking abstinence ("have you smoked combustible cigarettes, even a puff, in the last 7 days?"), relapse ("Have you smoked cigarettes, even a puff, since you were discharged from the hospital?"), quitting attempts ("Have you tried to quit smoking since you entered the study?"), and use of cessation medications and other interventions were assessed via self-report. Abstinence was biochemically verified with salivary cotinine assays. Saliva samples were collected and analyzed for cotinine concentrations using gas chromatography, consistent with standardized methods. Active, passive, and no smoking exposure were defined as cotinine concentrations of ≥31.5 ng/mL, 0.5-31.4 ng/mL and <0.5 ng/mL, respectively [23]. For participants reporting follow-up use of nicotine replacement therapies or electronic cigarettes, breath samples were conducted in person to test for levels of expired carbon monoxide (CO). Tobacco abstinence was confirmed by <10 ppm CO in the expired air.

Tobacco-Related Variables

At both baseline and 1-month follow-up, the following measures were assessed: (1) the Fagerstrom Test of Nicotine Dependence, a 6-item self-report scale with a summary score between 0 (very low dependence) and 10 (heavy dependence) [24], was used to assess dependence; (2) the 10-item Questionnaire of Smoking Urges-brief form [25] measured the intensity of smoking urges; and (3) the Confidence Questionnaire assessed situational self-efficacy in being able to resist urges to smoke across 16 everyday situations [26].

Gameplay

The game database tracked the number of log-ins, unique sessions, length of play, and episodes completed.

Satisfaction

Questions regarding participants' experience, satisfaction, and perceived game helpfulness were implemented in-game after 30 minutes of gameplay and at 1-month follow-up for participants to evaluate their experiences using the QuitIT game. We used 10 items developed by our research team to evaluate general experiences with the game and perceived helpfulness to improve the subjective experience of the game for future iterations.

Analysis

The primary goal of this pilot study was to examine feasibility (recruitment and retention rates), acceptability (patient satisfaction), and cessation-related trends associated with the intervention to guide the development of a larger trial. We, therefore, detail screening, exclusion, and refusal reasons and compare refusal and attrition by key demographic and clinical variables. For tobacco-related outcomes, we report means and SDs at baseline and 1-month follow-up. Between-group differences were not analyzed on tobacco-related outcomes as the pilot was not powered to detect statistically significant differences. Analyses were conducted using SAS version 9.3 (SAS Institute, Inc) and the MBESS package in R (version 3.3.3) for CIs of effect sizes.

Results

Participants

A total of 525 patients were screened for participation. Of those, 388 were determined to be ineligible, primarily because of having been diagnosed >6 months ago, not having smoked in the past 30 days, and having metastatic disease. Of 137 eligible patients, 71% (98/137) refused, reporting that they were not interested in the study (n=42), preferred to quit on their own (n=34), did not want to quit at all (n=11). A total of 38 patients were randomized. At 1-month follow-up, data were collected from 13 patients in the QuitIT arm and 11 in the SC arm (see Figure 1 for full study flow details).

Table 1 shows that 40% (15/38) of enrolled participants were between the age of 50 and 59 years, 71% (27/38) were females, 92% (35/38) were white people, and 95% (36/38) were non-Hispanic people. The most common cancer diagnoses were lung (12/38, 32%) and gastrointestinal (9/38, 24%) cancers. Most patients were diagnosed with disease stages I (14/38, 37%) or II (7/38, 18%).

Feasibility and Use Metrics

Refusal and attrition rates were examined to assess the feasibility of conducting a future intervention trial. Those who enrolled versus refused were more likely to be female (P=.003) but did not differ by other demographic or clinical characteristics (see Table 1). Across both QuitIT and SC arms, a total of 24 participants completed the 1-month follow-up. As shown in Table 2, those who less frequently used a tablet computer at baseline were more likely to drop out (P<.01), less likely to have used the game at all (P=.02), and more likely to be older (P=.05). Of 20 in the QuitIT arm, 40% (8/20) played the game (as determined by system data). No statistically significant differences in demographic, clinical, technology use, and tobacco-related variables were found between those who played and did not play (Multimedia Appendix 2). Users completed an average of 2.5 (SD 4.0) episodes, with a range of 0-10 episodes completed.

Tobacco-Related Outcomes

At 1-month assessment, data were available from 24 participants. A trend was found for the primary tobacco-related outcome, increased situational self-efficacy (confidence to quit), in the QuitIT arm (d=0.25, 95% CI –0.56 to 1.06). In addition, the QuitIT participants reported higher intention to stay quit (d=1.03, 95% CI 0.14-1.89; Table 3). Confirmed abstinence was higher in the QuitIT arm, with 30% (4/13) of the sample reporting abstinence versus 18% (2/11) in the SC arm. Nicotine replacement therapy and other cessation medications were used by a minority of participants in each arm with 5/11 (46%) using them in the SC and 4/13 (31%) in the QuitIT arm.

Satisfaction and Participant Feedback

At 1-month follow-up, 8 participants who used the game completed survey items related to satisfaction and helpfulness (Table 4). Most respondents thought the game kept their attention and 63% (5/8) thought it was fun to use. Of them, 88% (7/8) said they could relate to the characters and 63% (5/8) indicated that they got interested in their stories. All participants said they learned at least something about coping with smoking urges. A little more than half (63%, 5/8) thought playing helped them cope with urges to smoke; 75% (6/8) would apply what they learned in real life and learned at least a moderate amount from the coping cards and 75% (6/8) thought the game was the right length.

Figure 1. Study flow. Asterisk indicates that five participants were deemed ineligible after randomization due to a change in prognosis; their data were not analyzed. Cessation tx: cessation treatment; cog imp: cognitive impairment; psych illness: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition psychological diagnosis.

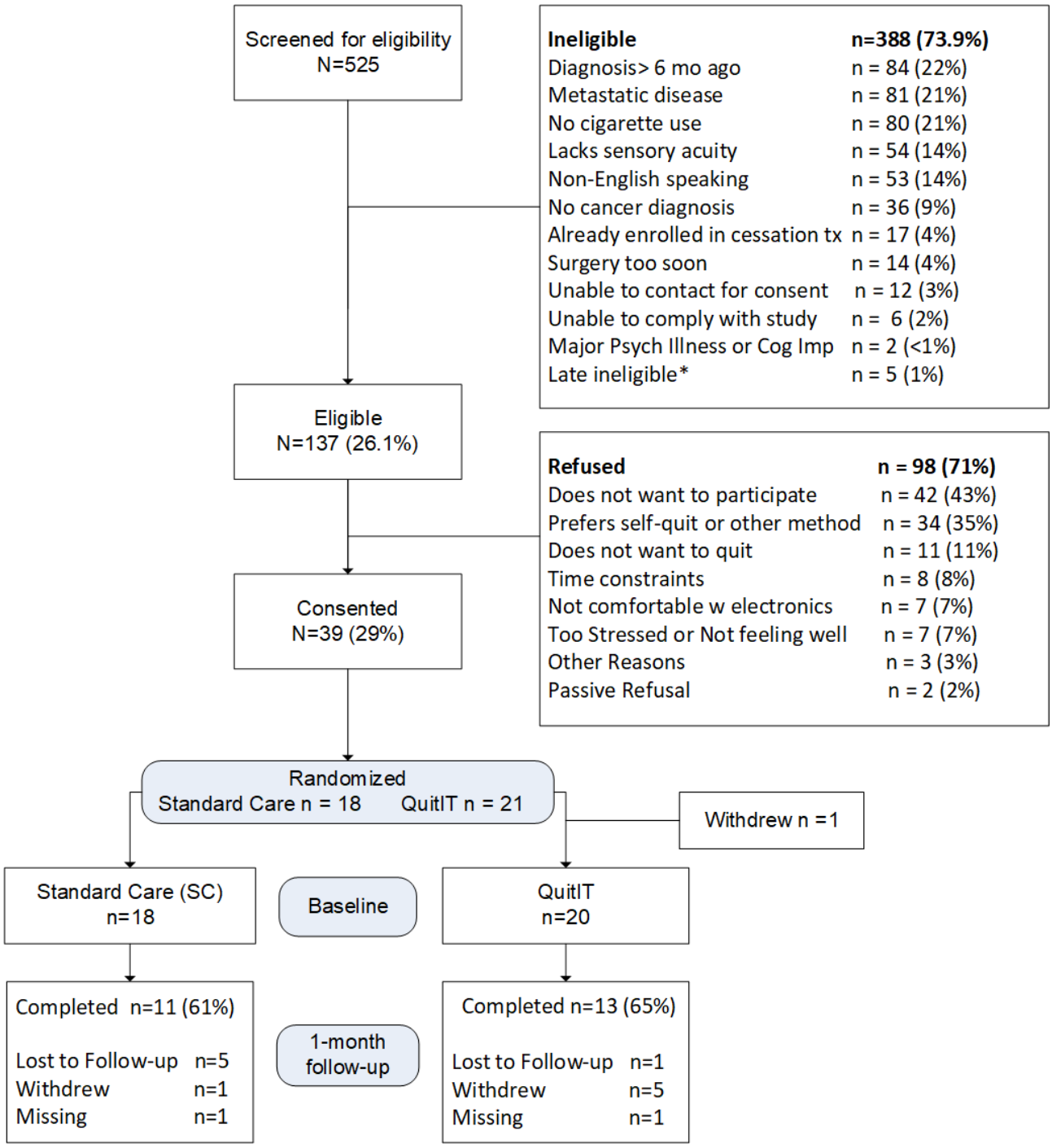


Table 1. Comparison of participant characteristics by enrollment status.

Characteristic	All (n=136), n (%)	Enrolled (n=38), n (%)	Refused (n=98), n (%)	*P* value[a]
Age (years)				.50
<40	7 (5.2)	3 (7.9)	4 (4.1)	
40-49	15 (11.0)	4 (10.5)	11 (11.2)	
50-59	52 (38.2)	15 (39.5)	37 (37.8)	
60-69	45 (33.1)	13 (34.2)	32 (32.7)	
70+	17 (12.5)	3 (7.9)	14 (14.3)	
Sex				.003
Female	69 (50.7)	27 (71.1)	42 (42.9)	
Male	67 (49.3)	11 (29.0)	56 (57.1)	
Race				.18
White	114 (83.8)	35 (92.1)	79 (80.6)	
Black	7 (5.2)	1 (2.6)	6 (6.1)	
Asian	4 (2.9)	2 (5.3)	2 (2.0)	
Other	1 (0.7)	0 (0)	1 (1.0)	
Refused	10 (7.4)	0 (0)	10 (10.2)	
Ethnicity				.54
Non-Hispanic	131 (96.3)	36 (94.7)	95 (96.9)	
Hispanic	5 (3.7)	2 (5.3)	3 (3.1)	
Cancer site				.08
Gastrointestinal	40 (29.4)	9 (23.7)	31 (31.6)	
Lung	32 (23.5)	12 (31.6)	20 (20.4)	
Urologic	18 (13.2)	2 (5.3)	16 (16.3)	
Colorectal	14 (10.3)	5 (13.2)	9 (9.2)	
Gynecologic	11 (8.1)	6 (15.8)	5 (5.1)	
Other	21 (15.4)	4 (10.5)	17 (17.4)	
Cancer stage				.72
0	7 (5.2)	3 (7.9)	4 (4.1)	
I	54 (39.7)	14 (36.8)	40 (40.8)	
II	27 (19.9)	7 (18.4)	20 (20.4)	
III	26 (19.1)	6 (15.8)	20 (20.4)	
IV	7 (5.2)	4 (10.5)	3 (3.1)	
Missing (or 88 or 99)	15 (11.0)	4 (10.5)	11 (11.2)	

[a]*P* value is based on *t* test for age, Mantel-Haenszel chi-square for clinical stage, and Pearson's chi-square for all other variables.

Table 2. Attrition by participant characteristics (n=38).

Characteristic	All (n=38)	Completers (n=24)	Lost to follow-up (n=14)	*P* value
Sex, n (%)				.48
Female	27 (71)	18 (67)	9 (33)	
Male	11 (29)	6 (55)	5 (46)	
Race, n (%)				.34
White	35 (92)	21 (60)	14 (40)	
Asian	2 (5)	2 (100)	0 (0)	
Black	1 (3)	1 (100)	0 (0)	
Marital status, n (%)				.28
Married	21 (55)	11 (52)	10 (48)	
Divorced or widowed	11 (29)	8 (73)	3 (27)	
Single	6 (16)	5 (83)	1 (17)	
Education, n (%)				.42
≤High school	9 (24)	4 (44)	5 (56)	
Some college	12 (32)	9 (75)	3 (25)	
College grad	17 (45)	11 (65)	6 (35)	
Employment, n (%)				.20
Employed	14 (37)	9 (64)	5 (36)	
Retired	13 (34)	6 (46)	7 (54)	
Unemployed or on leave	11 (29)	9 (82)	2 (18)	
Income, n (%)				.66
<US $10k	4 (11)	4 (100)	0 (0)	
US $10k-$30k	4 (11)	2 (50)	2 (50)	
US $30k-$70k	8 (21)	4 (50)	4 (50)	
>US $70k	20 (53)	13 (65)	7 (35)	
Missing	2 (5)	1 (50)	1 (50)	
Baseline tablet use, n (%)				<.01
Never or rarely or monthly	14 (37)	4 (29)	10 (71)	
Occasionally or more	24 (63)	20 (83)	4 (17)	
Baseline gameplay history, n (%)				.85
Never or rarely or monthly	6 (16)	4 (67)	2 (33)	
Occasionally or more	32 (84)	20 (63)	12 (38)	
Smoking since diagnosis, n (%)				.97
Maintained or increased	8 (21)	5 (63)	3 (38)	
Decreased	30 (79)	19 (63)	11 (37)	
Quit attempts of >24 hours in past year, n (%)				.33
No	9 (24)	4 (44)	5 (56)	
Yes, once	9 (24)	7 (78)	2 (22)	
Yes, more than once	20 (53)	13 (65)	7 (35)	
Cancer site, n (%)				.62

Characteristic	All (n=38)	Completers (n=24)	Lost to follow-up (n=14)	*P* value
Lung	12 (32)	7 (58)	5 (42)	
Gastrointestinal	9 (24)	4 (44)	5 (56)	
Gynecologic	6 (16)	5 (83)	1 (17)	
Colorectal	5 (13)	4 (80)	1 (20)	
Urologic	2 (5)	1 (50)	1 (50)	
Other	4 (11)	3 (75)	1 (25)	
Clinical stage, n (%)				.46
0	3 (8)	1 (33)	2 (66.7)	
I	14 (37)	8 (57)	6 (42.9)	
II	7 (18)	5 (71)	2 (28.6)	
III	6 (16)	3 (50)	3 (50.0)	
IV	4 (11)	3 (75)	1 (25.0)	
Missing (or 88 or 99)	4 (11)	4 (100)	0 (0)	
Arm, n (%)				.80
Standard Care	18 (47)	11 (61)	7 (39)	
Quit It	20 (53)	13 (65)	7 (35)	
Gameplay (n=20), n (%)				.02
Any	8 (21)	8 (100)	0 (0)	
None	30 (79)	16 (53)	14 (47)	
Age (current; years), mean (SD)	57.11 (9.6)	54.79 (9.1)	61.07 (9.5)	.05
Baseline Fagerstrom score[a], mean (SD)	3.68 (2.2)	3.41 (2.1)	4.17 (2.5)	.35
Years smoking (n=40), mean (SD)	36.7 (12.6)	34.43 (14.1)	40.50 (8.8)	.16
Baseline cigarettes per day (n=41), mean (SD)	12.34 (14.7)	11.10 (7.1)	14.46 (22.8)	.50
Baseline intention to abstain for 30 days (n=40), mean (SD)	2.73 (2.1)	3.04 (2.3)	2.15 (1.8)	.23
Baseline coping strategies, mean (SD)				.57
Number used (of 13)	8.47 (3.8)	8.75 (4.2)	8.00 (3.2)	
Baseline situational self-efficacy, mean (SD)				.76
Mean (of 16 items), range 0-100	57.05 (22.6)	56.20 (24.2)	58.52 (20.2)	

[a]0-2: Very low; 3-4: Low; 5: Medium; 6-7: High; 8-10: Very high or heavy.

Table 3. Tobacco-related outcomes at 1-month follow-up.

Characteristic	Standard care (n=11), mean (SD)	QuitIT (n=13), mean (SD)	Effect size *d* (95% CI)
Situational self-efficacy (range 0-100)	80.01 (19.3)	84.43 (15.9)	0.25 (−0.56 to 1.06)
Days abstinent postdischarge	25.64 (11.8)	23.85 (23.3)	−0.09 (−0.90 to 0.71)
Length of admission (days)	8.64 (5.9)	6.69 (5.0)	0.36 (−0.45 to 1.17)
Intend to abstain for next 30 days (range 1-5)	1.55 (0.9)	3.75 (2.8)	1.03 (0.14 to 1.89)
Importance of quitting (range 1-10)	10 (0)	9.46 (1.9)	−0.38 (−1.18 to 0.44)

Table 4. Satisfaction with the game at 1-month follow-up (n=8).

Item and response	Value, n (%)
The game kept my attention.	
Strongly agree or agree	6 (75)
Neither agree nor disagree	0 (0)
Disagree or strongly disagree	2 (25)
The game has been fun to use.	
Strongly agree or agree	5 (63)
Neither agree nor disagree	2 (25)
Disagree or strongly disagree	1 (13)
I could relate to the characters as they dealt with smoking temptations.	
Strongly agree or agree	7 (88)
Neither agree nor disagree	0 (0)
Disagree or strongly disagree	1 (13)
I got interested in the characters' stories.	
Strongly agree	5 (63)
Neither agree nor disagree	1 (13)
Disagree or strongly disagree	2 (25)
How much did you learn from the game about ways to help you cope with smoking urges?	
I learned a lot	4 (50)
I learned a moderate amount	2 (25)
I learned a little bit	2 (25)
I didn't learn much at all	0 (0)
Playing the game helped me cope with urges to smoke.	
Strongly agree	5 (63)
Neither agree nor disagree	0 (0)
Disagree or strongly disagree	3 (38)
How likely are you to apply what you learned in the game to real like smoking temptations?	
Extremely likely or likely	4 (50)
Likely	2 (25)
Neither likely nor unlikely	2 (25)
Unlikely or extremely unlikely	0 (0)
How would you rate your experience with this game session?	
Extremely satisfied	2 (25)
Satisfied	3 (38)
Neither satisfied nor dissatisfied	2 (25)
Dissatisfied or extremely dissatisfied	1 (13)
In terms of length, the game sessions:	
Took too long	2 (25)
Were just about right	6 (75)
Were too short	0 (0)
How useful were the game cards in helping you cope with smoking urges?	

Item and response	Value, n (%)
Extremely helpful	2 (25)
Very helpful	2 (25)
Moderately helpful	1 (13)
A little bit helpful	1 (13)
Not at all helpful	2 (25)
In the past month, how many days did you look at the deck of QuitIT cards?	
0	3 (38)
1-4	2 (25)
5-19	0 (0)
20+	2 (25)
Missing	1 (13)

Discussion

This study presents findings from a pilot randomized controlled trial evaluating acceptability, use, and preliminary outcomes from an interactive tablet-based game to promote abstinence from tobacco use following cancer-related hospitalization. Results should be interpreted with regard to the feasibility of pursing a larger powered trial. The following criteria were assessed as indicators of whether to pursue follow-up work: recruitment of the target sample size in the allotted timeframe, acceptance rate of at least 50%, retention of at least 80%, at least small effect sizes (d=0.2) on primary outcomes, minimal adverse events, and patients' reports of interest in and acceptability of the intervention [27]. In light of these criteria, our goal of 190 patients was not met primarily due to insufficient number of eligible patients (N=137) despite screening all Tobacco Treatment Program-referred patients for 2 years. Among those eligible, 71% (98/137) declined participation citing lack of interest in the study (42/98, 43%) or preferring to quit on their own (34/98, 35%) as the most common reasons. The study retention was 61% (11/18) and 65% (13/20) in the SC and QuitIT conditions, respectively, which is lower than the 80% feasibility criterion.

Recruiting study participants in the context of recent cancer diagnosis and treatment is complicated by patient anxiety, disruptions in daily patterns due to multiple medical appointments, and worry about treatment outcomes. Nonetheless, the recruitment rate is markedly lower than observed in other cessation studies we have conducted with patients recently diagnosed with cancer. For example, we found a 30% recruitment rate in a previous trial testing a handheld computer that guided smoking reduction with presurgical, tobacco-dependent, cancer patients [28]. Other studies [7,15,16,29] that have examined cessation apps were not clinical trials or used Web-based volunteer recruitment, precluding ascertainment of recruitment rates from a specified cohort of patients. Nevertheless, the low recruitment rate warrants consideration for future app studies. It is possible that describing the app as a "game" may have appeared as inappropriately frivolous in the context of a cancer diagnosis and surgery, contributing to the low rate of study participation. Using a more serious term such as "mobile app or guide for smoking cessation," "video simulation," or "games for behavior change" may be more appropriate to the cancer context. We primarily used a telephone recruitment approach, which presents difficulties for explaining such a novel intervention to potential participants. Recruiting participants in person, or having an interactive ad on the hospital website, during which the intervention could be demonstrated may increase interest and willingness to participate. In addition, 35% (34/98) noted they wanted to quit on their own and refused all cessation services, suggesting optimistic bias and, perhaps, low awareness of the effectiveness of cessation interventions [30]. Further work, including qualitative interviews with this population, may assist in framing app-based interventions for future trials.

Although the study was not powered to detect a difference in quit rates, 6 people had confirmed smoking abstinence. Despite this, more participants in the QuitIT group were abstinent than the SC arm. In addition, while not statistically significant, there was a trend for increased confidence in quitting for the QuitIT arm (d=0.25), as well as for intention to abstain (d=1.03). These moderate effects suggest some potential promise for the intervention and meet the specified criteria for feasibility [27]. Given that only 8 of 21 people in the QuitIT arm actually played the game and only for an average of 2.5 sessions, it is likely that these treatment effects might be more pronounced with greater adherence to gameplay or app use. This suggests that the game likely has potential, but that increased attention should be paid to addressing barriers for use. Gameplay during perihospitalization and recovery may be particularly challenging due to the healing process and presence of physical symptoms adversely affecting energy and quality of life. Introducing the game at another time and a longer intervention period with more frequent prompts to play the game may ameliorate these challenges.

In terms of app use and acceptability, satisfaction data from the 1-month follow-up indicate that users found that the game kept their attention and was fun to use, but that it could be improved in terms of relating to the characters and helping cope with urges. In addition, there were no study-related adverse events. It may be necessary to provide an additional reward for use and to support use with counselor check-ins, or monetary incentives

[31], as well as improved training, as noted previously. We found that prior tablet use and younger age were found to be significant predictors of app use; this is not surprising given that younger age has been associated with the health app use [5]. Persons who already owned a tablet would be more familiar and likely to integrate it into their daily activities, using it for reading, email, internet, etc, whereas, persons who received it only for the study would not otherwise interact with the tablet (study tablets were locked down for other uses). Older persons with greater income and education are just as likely as younger persons to use smartphones and tablets [32], but also less likely to be smokers [33]. Our sample of cancer patients skews toward older persons, who would likely require increased coaching about tablet use as well as technical support. One training session was likely insufficient to help nonusers become comfortable with using a tablet. Furthermore, allowing participants to use the tablet for other activities would likely increase instances of app use.

This pilot clinical trial is, nevertheless, an important advance in establishing an evidence base for health-related apps, and tobacco cessation apps, in particular. The great majority of tobacco cessation-related apps are of poor quality [4]. In an analysis of apps being guided by behavior change theory, Choi et al found that only 10.3% (18/175) apps examined used 3 key theoretical domains [34]. Of apps that have been developed by smoking behavior change experts, *QuitSTART* from Smokefree.gov and the Truth campaign's *This is Quitting* app have not yet been evaluated in a clinical trial. *SmartQuit*, an app based on the Acceptance and Commitment therapy, was found to have superior engagement compared with the *QuitSTART* app [7,35]. Greater evidence for apps is necessary to assess this treatment modality and promote greater use. Only 20% of health app users have had a doctor recommend an app [5]. This is not surprising as a recent survey of 264 health care providers found that although most (203/264, 76.9%) believed that apps had potential to change smoking behavior, fewer (112/264, 42.4%) believed that the currently available apps were useful in treatment [9]. Our findings suggest that better patient engagement and greater participant training are essential for conducting trials of apps in clinical populations and that greater adherence is needed to properly assess intervention effects. Next steps should involve another round of formative qualitative interviews with potential users to identify how to improve descriptions of the app to improve recruitment, increase adherence to game use, and meet expectations for help to improve skills for coping with smoking urges.

Acknowledgments

This work was supported by R42 DA029353 (QuitIT: Development of a Web-based, 3D Coping Skills Game to Increase Self-Efficacy for Maintaining Smoking Abstinence).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots from game.

[PDF File (Adobe PDF File), 477KB - mhealth_v7i1e10071_app1.pdf]

Multimedia Appendix 2

Baseline characteristic differences by actual gameplay (n=20 in QuitIT arm).

[PDF File (Adobe PDF File), 43KB - mhealth_v7i1e10071_app2.pdf]

Multimedia Appendix 3

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - mhealth_v7i1e10071_app3.pdf]

References

1. Mawston N. Strategy Analytics Wireless Smartphone Strategies (WSS) service. 2012. Global Smartphone Installed Base by Operating System for 88 Countries: 2007 to 2012 URL: https://tinyurl.com/yc28xogt [accessed 2017-06-10] [WebCite Cache ID 73pFdVAjm]
2. Valdivieso-López E, Flores-Mateo G, Molina-Gómez JD, Rey-Reñones C, Barrera Uriarte ML, Duch J, et al. Efficacy of a mobile application for smoking cessation in young people: study protocol for a clustered, randomized trial. BMC Public Health 2013 Aug 01;13:704 [FREE Full text] [doi: 10.1186/1471-2458-13-704] [Medline: 23915067]
3. Haskins BL, Lesperance D, Gibbons P, Boudreaux ED. A systematic review of smartphone applications for smoking cessation. Transl Behav Med 2017 Jun;7(2):292-299 [FREE Full text] [doi: 10.1007/s13142-017-0492-2] [Medline: 28527027]

4. Abroms LC, Lee WJ, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. Am J Prev Med 2013 Dec;45(6):732-736 [FREE Full text] [doi: 10.1016/j.amepre.2013.07.008] [Medline: 24237915]
5. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR Mhealth Uhealth 2015;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
6. Choi J, Noh G, Park D. Smoking cessation apps for smartphones: content analysis with the self-determination theory. J Med Internet Res 2014 Feb;16(2):e44 [FREE Full text] [doi: 10.2196/jmir.3061] [Medline: 24521881]
7. Bricker JB, Mull KE, Kientz JA, Vilardaga R, Mercer LD, Akioka KJ, et al. Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. Drug Alcohol Depend 2014 Oct 1;143:87-94. [doi: 10.1016/j.drugalcdep.2014.07.006] [Medline: 25085225]
8. Ubhi HK, Kotz D, Michie S, van Schayck OC, Sheard D, Selladurai A, et al. Comparative analysis of smoking cessation smartphone applications available in 2012 versus 2014. Addict Behav 2016 Jul;58:175-181 [FREE Full text] [doi: 10.1016/j.addbeh.2016.02.026] [Medline: 26950256]
9. McClure J, Hartzler A, Catz S. Design Considerations for Smoking Cessation Apps: Feedback From Nicotine Dependence Treatment Providers and Smokers. JMIR Mhealth Uhealth 2016 Feb 12;4(1):e17 [FREE Full text] [doi: 10.2196/mhealth.5181] [Medline: 26872940]
10. Robson K, Plangger K, Kietzmann J, McCarthy I, Pitt L. Is it all a game? Understanding the principles of gamification. Business Horizons 2015 Jul;58(4):411-420. [doi: 10.1016/j.bushor.2015.03.006]
11. Primack B, Carroll M, McNamara M, Klem M, King B, Rich M, et al. Role of video games in improving health-related outcomes: a systematic review. Am J Prev Med 2012 Jun;42(6):630-638 [FREE Full text] [doi: 10.1016/j.amepre.2012.02.023] [Medline: 22608382]
12. Dickey M. Engaging by design: How engagement strategies in popular computer and video games can inform instructional design. ETR&D 2005 Jun;53(2):67-83. [doi: 10.1007/BF02504866]
13. Okazaki S, Skapa R, Grande I. Capturing Global Youth: Mobile Gaming in the U.S., Spain, and the Czech Republic. Journal of Computer-Mediated Communication 2008;13(4):827-855 [FREE Full text] [doi: 10.1111/j.1083-6101.2008.00421.x]
14. Hamari J, Koivisto J, Sarsa H. Does gamification work? A literature review of empirical studies on gamification. 2014 Jan 6 Presented at: 2014 47th Hawaii International Conference on System Sciences; 2014; Hilton Waikoloa, Big Island.
15. El-Hilly AA, Iqbal SS, Ahmed M, Sherwani Y, Muntasir M, Siddiqui S, et al. Game On? Smoking Cessation Through the Gamification of mHealth: A Longitudinal Qualitative Study. JMIR Serious Games 2016 Oct 24;4(2):e18 [FREE Full text] [doi: 10.2196/games.5678] [Medline: 27777216]
16. Song H, Kim J, Kwon R, Jung Y. Anti-smoking educational game using avatars as visualized possible selves. Computers in Human Behavior 2013 Sep;29(5):2029-2036 [FREE Full text] [doi: 10.1016/j.chb.2013.04.008]
17. Krebs P, Burkhalter J, Lewis S, Hendrickson T, Chiu O, Fearn P, Memorial Sloan-Kettering Cancer Center. Development of a Virtual Reality Coping Skills Game to Prevent Post-Hospitalization Smoking Relapse in Tobacco Dependent Cancer Patients. J Virtual Worlds Res 2009 Aug;2(2) [FREE Full text] [Medline: 28736598]
18. Krebs P, Burkhalter JE, Snow B, Fiske J, Ostroff JS. Development and Alpha Testing of QuitIT: An Interactive Video Game to Enhance Skills for Coping With Smoking Urges. JMIR Res Protoc 2013;2(2):e35 [FREE Full text] [doi: 10.2196/resprot.2416] [Medline: 24025236]
19. Baranowski MT, Bower PK, Krebs P, Lamoth CJ, Lyons EJ. Effective Feedback Procedures in Games for Health. Games Health J 2013 Dec;2(6):320-326. [doi: 10.1089/g4h.2013.1328] [Medline: 26197072]
20. Marlatt G, Donovan D. Relapse Prevention: Maintenance Strategies in the Treatment of Addictive Behaviors. New York: Guilford Publications, Inc; 2005.
21. Memorial Sloan-Kettering Cancer Center. 2015. Your Guide to Quitting Smoking URL: https://www.mskcc.org/cancer-care/patient-education/your-guide-quitting-smoking [accessed 2018-11-10] [WebCite Cache ID 73pGgiAoK]
22. CDC. US Department of Health & Human Services. 2016. National Adult Tobacco Survey URL: https://www.cdc.gov/tobacco/data_statistics/surveys/nats/index.htm [accessed 2018-11-10] [WebCite Cache ID 73pGxUNGc]
23. SRNT Subcommittee on Biochemical Verification. Biochemical verification of tobacco use and cessation. Nicotine Tob Res 2002 May;4(2):149-159. [doi: 10.1080/14622200210123581] [Medline: 12028847]
24. Heatherton TF, Kozlowski LT, Frecker RC, Rickert W, Robinson J. Measuring the heaviness of smoking: using self-reported time to the first cigarette of the day and number of cigarettes smoked per day. Br J Addict 1989 Jul;84(7):791-799. [Medline: 2758152]
25. Cox LS, Tiffany ST, Christen AG. Evaluation of the brief questionnaire of smoking urges (QSU-brief) in laboratory and clinical settings. Nicotine Tob Res 2001 Feb;3(1):7-16. [doi: 10.1080/14622200020032051] [Medline: 11260806]
26. Baer JS, Lichtenstein E. Classification and prediction of smoking relapse episodes: an exploration of individual differences. J Consult Clin Psychol 1988 Feb;56(1):104-110. [Medline: 3346434]
27. NHLBI. 2014. Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies URL: https://www.nhlbi.nih.gov/health-pro/guidelines/in-develop/cardiovascular-risk-reduction/tools/cohort [accessed 2017-02-01] [WebCite Cache ID 6nxeJyirH]

28. Ostroff JS, Burkhalter JE, Cinciripini PM, Li Y, Shiyko MP, Lam CY, et al. Randomized trial of a presurgical scheduled reduced smoking intervention for patients newly diagnosed with cancer. Health Psychol 2014 Jul;33(7):737-747 [FREE Full text] [doi: 10.1037/a0033186] [Medline: 23895203]
29. DeLaughter KL, Sadasivam RS, Kamberi A, English TM, Seward GL, Chan SW, et al. Crave-Out: A Distraction/Motivation Mobile Game to Assist in Smoking Cessation. JMIR Serious Games 2016 May 26;4(1):e3 [FREE Full text] [doi: 10.2196/games.4566] [Medline: 27229772]
30. Burkhalter J, Lord-Bessen J, Li Y, Holland S, Hay J, Ostroff J. How do Self-Efficacy Evaluations Perfrom Longitudinally in a Smoking Cessation Trial? 2014 Presented at: Society of Nicotine and Tobacco Research Annual Conference; 2014; Seattle, WA.
31. Halpern SD, French B, Small DS, Saulsgiver K, Harhay MO, Audrain-McGovern J, et al. Randomized Trial of Four Financial-Incentive Programs for Smoking Cessation. N Engl J Med 2015 May 28;372(22):2108-2117. [doi: 10.1056/NEJMoa1414293]
32. Anderson M, Perrin A. Pew Research Center. 2017. Technology use among seniors URL: http://www.pewinternet.org/2017/05/17/technology-use-among-seniors/ [accessed 2017-10-19] [WebCite Cache ID 6uKse8X5t]
33. Centers for Disease Control and Prevention. US Department of Health & Human Services. 2015. Current Cigarette Smoking Among Adults in the United States URL: https://www.cdc.gov/tobacco/data_statistics/fact_sheets/adult_data/cig_smoking/index.htm [accessed 2018-11-10] [WebCite Cache ID 73pHI77lo]
34. Choi J, Noh G, Park D. Smoking cessation apps for smartphones: content analysis with the self-determination theory. J Med Internet Res 2014;16(2):e44 [FREE Full text] [doi: 10.2196/jmir.3061] [Medline: 24521881]
35. Bricker J. Fred Hutchinson Cancer Center. 2017. Randomized Trial of an Innovative Smartphone Intervention for Smoking Cessation URL: https://clinicaltrials.gov/ct2/show/NCT02724462 [accessed 2018-11-10] [WebCite Cache ID 73pHUd9gK]

Abbreviations

CO: carbon monoxide
MSKCC: Memorial Sloan-Kettering Cancer Center
RA: research assistant
SC: standard care

Edited by M Bars; submitted 09.02.18; peer-reviewed by R Windsor, S Theis, A Eisingerich, J Volkman, E Laird; comments to author 27.03.18; revised version received 27.04.18; accepted 06.06.18; published 10.01.19.

Please cite as:
Krebs P, Burkhalter J, Fiske J, Snow H, Schofield E, Iocolano M, Borderud S, Ostroff JS
The QuitIT Coping Skills Game for Promoting Tobacco Cessation Among Smokers Diagnosed With Cancer: Pilot Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(1):e10071
URL: http://mhealth.jmir.org/2019/1/e10071/
doi:10.2196/10071
PMID:30632971

Review

The Efficacy of Mobile Phone Apps for Lifestyle Modification in Diabetes: Systematic Review and Meta-Analysis

Xinghan Wu[1], MSc; Xitong Guo[1], PhD; Zhiwei Zhang[2], PhD

[1]eHealth Research Institute, School of Management, Harbin Institute of Technology, Harbin, China

[2]Department of Statistics, University of California at Riverside, Riverside, CA, United States

Corresponding Author:
Xitong Guo, PhD
eHealth Research Institute
School of Management
Harbin Institute of Technology
92 West Dazhi Street, Nangang District
Harbin,
China
Phone: 86 451 86414022
Email: xitongguo@gmail.com

Abstract

Background: Diabetes and related complications are estimated to cost US $727 billion worldwide annually. Type 1 diabetes, type 2 diabetes, and gestational diabetes are three subtypes of diabetes that share the same behavioral risk factors. Efforts in lifestyle modification, such as daily physical activity and healthy diets, can reduce the risk of prediabetes, improve the health levels of people with diabetes, and prevent complications. Lifestyle modification is commonly performed in a face-to-face interaction, which can prove costly. Mobile phone apps provide a more accessible platform for lifestyle modification in diabetes.

Objective: This review aimed to summarize and synthesize the clinical evidence of the efficacy of mobile phone apps for lifestyle modification in different subtypes of diabetes.

Methods: In June 2018, we conducted a literature search in 5 databases (Cochrane Central Register of Controlled Trials, MEDLINE, Embase, CINAHL, and PsycINFO). We evaluated the studies that passed screening using The Cochrane Collaboration's risk of bias tool. We conducted a meta-analysis for each subtype on the mean difference (between intervention and control groups) at the posttreatment glycated hemoglobin (HbA_{1c}) level. Where possible, we analyzed subgroups for short-term (3-6 months) and long-term (9-12 months) studies. Heterogeneity was assessed using the I^2 statistic.

Results: We identified total of 2669 articles through database searching. After the screening, we included 26 articles (23 studies) in the systematic review, of which 18 studies (5 type 1 diabetes, 11 type 2 diabetes, and 2 prediabetes studies) were eligible for meta-analysis. For type 1 diabetes, the overall effect on HbA_{1c} was statistically insignificant (P=.46) with acceptable heterogeneity (I^2=39%) in the short-term subgroup (4 studies) and significant heterogeneity between the short-term and long-term subgroups (I^2=64%). Regarding type 2 diabetes, the overall effect on HbA_{1c} was statistically significant (P<.01) in both subgroups, and when the 2 subgroups were combined, there was virtually no heterogeneity within and between the subgroups (I^2 range 0%-2%). The effect remained statistically significant (P<.01) after adjusting for publication bias using the trim and fill method. For the prediabetes condition, the overall effect on HbA_{1c} was statistically insignificant (P=.67) with a large heterogeneity (I^2=65%) between the 2 studies.

Conclusions: There is strong evidence for the efficacy of mobile phone apps for lifestyle modification in type 2 diabetes. The evidence is inconclusive for the other diabetes subtypes.

(JMIR Mhealth Uhealth 2019;7(1):e12297) doi:10.2196/12297

KEYWORDS

smartphone; mobile applications; diabetes mellitus; lifestyle; physical activity; diet; behavior therapy

Introduction

Background

Diabetes mellitus is a chronic disease that has a negative effect on people's quality of life and results in a series of unfavorable outcomes [1]. Diabetes mellitus can be divided into three subtypes (type 1 diabetes mellitus [T1DM], type 2 diabetes mellitus [T2DM], and gestational diabetes mellitus [GDM]), which share the same behavioral risk factors, such as inactivity and unhealthy diets [2]. A guideline released by the International Diabetes Federation in 2017 estimated a yearly cost of US $727 billion globally due to diabetes and related complications [2,3]. Efforts in lifestyle modifications, such as daily physical activity and healthy diets, can reduce the risk of prediabetes, improve the health level of people with diabetes, and prevent complications [4-6]. Lifestyle modification is commonly performed through face-to-face consultations at medical institutions, periodic monitoring by rehabilitation specialists, or both. Other, and more personalized, types of lifestyle modifications, such as personal coaching for physical activity interventions and personal consultations for healthy diet interventions, are not widely available [4].

Mobile phone apps are widely used in both developed and developing countries and have shown great potential to deliver personalized medical advice. In prior studies, apps were demonstrated to facilitate patients' health promotion by improving their self-management awareness and compliance [7-9]. In practice, apps have been used to help people living with various health conditions and problems, including mental health [10,11], heart failure [12], and smoking cessation [13,14]. In addition, more than 120 apps are available in iTunes and Google Play for diabetes management [15].

Moreover, apps for diabetes management have shown great promise toward improving mental and physical health. Research has shown that the use of apps has statistically significant effects in improving self-efficacy, increasing disease knowledge, enhancing physician-patient communication, and lowering diabetes incidence through delivering information, education, self-management, therapeutic advice, and drug guidance [16].

Objective

Despite growing interest in the efficacy of apps for lifestyle modification in diabetes management, it is unclear what evidence is available and what this evidence suggests. This lack of knowledge hampers the development of practical guidelines on the use of apps for lifestyle modification in the specific types of diabetes. This review aimed to summarize and synthesize the clinical evidence about the efficacy of mobile phone apps for lifestyle modification in the different subtypes of diabetes.

Methods

Data Sources and Search Strategy

We conducted a systematic review and report the results according to the guidance of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [17]. To identify relevant studies, we systematically searched 5 bibliographic databases: the Cochrane Central Register of Controlled Trials in the Cochrane Library, MEDLINE (via the Web of Science), Embase, CINAHL (via EBSCOhost), and PsycINFO (via EBSCOhost). Multimedia Appendix 1 presents the search strategy for each database based on Boolean operators. The scope of the search was defined by publication dates between January 1, 2006 and May 14, 2018 with no restrictions on the languages used.

Inclusion and Exclusion Criteria

We included studies if they met all of the following conditions: (1) participants had T1DM, T2DM, or GDM or were prediabetes patients; (2) participants were 18 years of age or older; (3) the study included interventions that used apps as a major component; (4) lifestyle modification (eg, physical activity and healthy diets) was provided via apps; (5) the study measured the participants' glycated hemoglobin (HbA_{1c}), weight, fasting blood sugar, body mass index, or other health-related outcomes; (6) the study was a randomized controlled trial; and (7) the full text of article was available. If 1 or more of the inclusion criteria were not met, we excluded such studies.

Reviewers (XW and XG) searched the extant literature and assessed the studies independently. Any disputes were discussed with a third reviewer (ZZ) to reach a consensus.

Study Selection

We selected studies in 3 phases: identification, preliminary screening, and full-text screening. In the identification phase, the first author conducted electronic searches. The titles and abstracts of all articles identified were collated into 1 database. In the preliminary screening phase, duplicated records were removed. Two authors (XW and XG) independently screened the titles and abstracts of the identified articles according to the inclusion criteria. Studies that did not meet all inclusion criteria were excluded. When disputes arose, a third author (ZZ) was asked to arbitrate. Relevant reviews were retained in the preliminary screening phase. In the full-text screening phase, the same 2 authors (XW and XG) independently screened the full text in accordance with the stated inclusion criteria. Thereafter, XW and XG hand searched the reference lists of all relevant studies for additional relevant ones. In cases of disagreement, ZZ participated to achieve consensus.

Data Extraction

XW and XG independently extracted data from the acceptable studies (which had passed the full-text screening) and incorporated them into a spreadsheet. The data extracted included details of the studies such as author, year, country, sample size, study design, diabetes subtype, details of intervention and control, outcomes of interest, and the key results. As in previous reviews and meta-analyses, if a study had multiple intervention arms, we limited data extraction to the most active intervention arm based on the use of apps (ie, the intervention arm that provided the largest collection of interventions based on apps). If the desired data had not been reported in an article, we contacted the first author of the article to retrieve any missing information. In some cases, we back-calculated unreported standard deviations from reported data such as confidence intervals [18].

Quality Assessment

Two reviewers (XW and XG) independently assessed the risk of bias in the eligible studies in accordance with The Cochrane Collaboration's risk of bias tool [19]. The instrument has been widely used in evidence-based medical research to evaluate the risk of bias in 6 different aspects (selection, performance, detection, attrition, reporting, and others). In the case of a dispute, we invited another one of the authors to participate in the discussion to help resolve this dispute. we then exported the results of the risk of bias assessment to the software Review Manager (RevMan) version 5.3 (The Nordic Cochrane Centre, The Cochrane Collaboration) to create a visual representation for publication. We handled difficulties in scoring some of the studies by reading the protocols if available (either published in journals or at clinicaltrials.gov, or obtained directly from the authors of each study).

Data Synthesis and Statistical Analysis

The first author extracted all the data from the appropriate studies. All authors evaluated the preliminary results of the reviews. We conducted a meta-analysis on the mean posttreatment HbA_{1c} values for the intervention and control groups with standard deviations. For studies that only reported mean changes from the baseline, we used this in addition to standard deviations for both groups. If adjusted and unadjusted estimates of treatment effects were both presented, we chose to use the adjusted estimate as reported in the article. When an intention-to-treat and per-protocol analysis were both presented, we chose to use the intention-to-treat results for better internal validity.

We assessed heterogeneity using the I^2 statistic. We used a fixed effects model when I^2 was less than 40%; otherwise, we used a random effects model. We conducted subgroup analyses on the long-term (9-12 months) and short-term (3-6 months) effects for each type of diabetes (to the extent possible). All analyses were performed in R version 3.3.3 (R Foundation).

Outcome Measures

The primary outcome of interest was HbA_{1c}. Secondary outcomes included body mass index, weight loss, change in waist circumference, and behavioral changes in physical activity and healthy diets. Physical activity could be measured by step counts, walking activity, and gait performance. Healthy diets were measured by diet balance, food intake, nutrition consumption, and changes in intestinal microflora. Both physical activity and healthy diets were measured using standardized questionnaires.

Results

Identified and Included Studies

The PRISMA diagram in Figure 1 shows our search process and results. We identified a total of 2669 articles through database searching. After we eliminated duplicates, 2232 articles were left for the preliminary screening. In the preliminary screening, we excluded 2093 articles for not meeting all inclusion criteria, leaving 139 articles for full-text screening. Following the full-text screening, we included 17 articles in this review. Separately, we included 9 articles through hand-searched reference lists. Finally, we included 26 articles (based on 23 different studies) in the systematic review, and 18 studies were eligible for meta-analysis. Of the 18 studies eligible for meta-analysis, 5 [20-24] examined T1DM, 11 [25-35] examined T2DM, and 2 [36,37] examined prediabetes. We excluded 1 study [38], which included both T1DM and T2DM patients, from the meta-analysis because the article did not stratify participants according to disease type in reporting efficacy data.

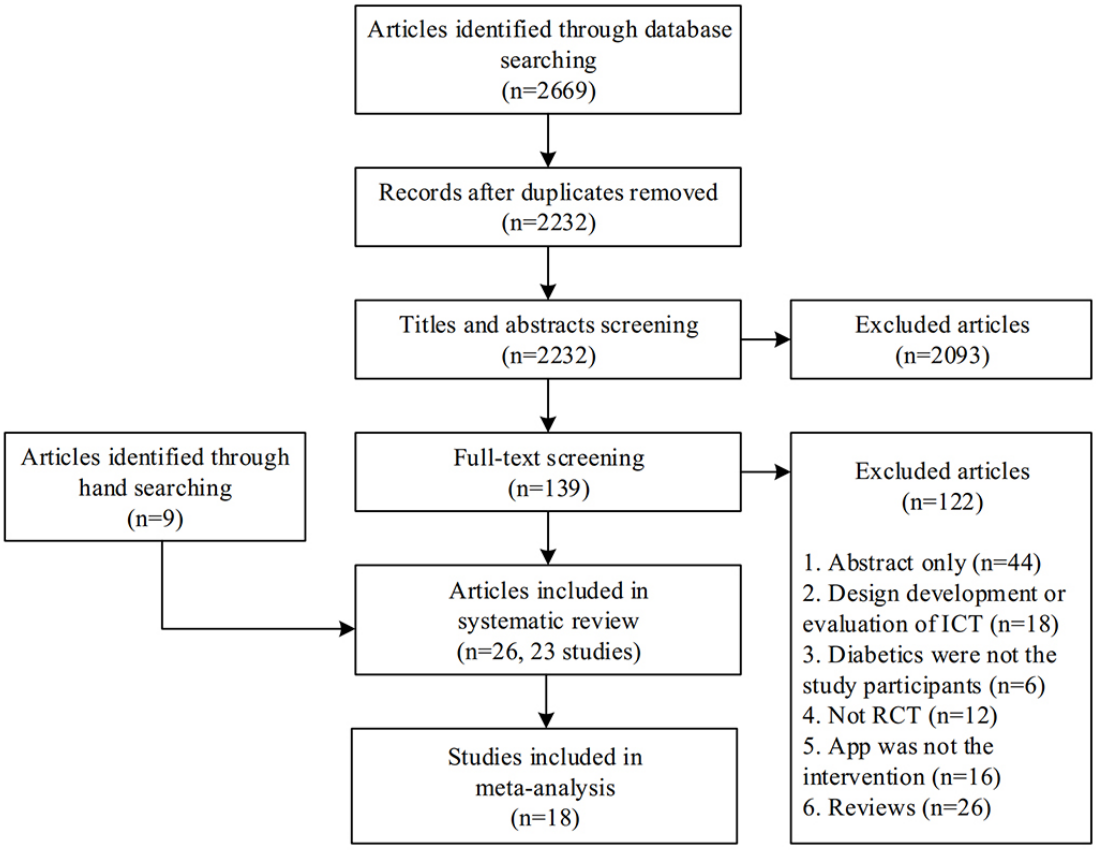
Study Characteristics

Multimedia Appendix 2 shows the characteristics of all 23 studies. All the studies were randomized controlled trials with apps as the main component of the intervention. In our review, a total of 2526 participants were enrolled. (We included only 1 intervention arm when a study had multiple intervention arms.) In the T1DM studies [20-24], the mean age of participants ranged from 34.9 (SD 13.1) to 39.7 (SD 10.8) years. Of these studies, 4 were undertaken in Europe [20-23] and 1 in Australia [24]. In the T2DM studies [25-35,39,40], the mean age of participants was much higher, ranging from 44.7 (SD 14.0) to 66.3 (SD 8.6) years. Of these studies, 5 were undertaken in Europe [26-28,32,39], 4 in North America [25,29,30,33], 3 in Asia [31,34,35], and 1 in Australia [40]. The 3 prediabetic studies [36,37,41] were undertaken in the United States, and the mean ages of the participants were 40.3 (SD 10.8), 55.2 (SD 9.0), and 55.0 (SD 8.9) years, respectively. Only 1 study focused on pregnant women at high risk of GDM [42], and this study was conducted in the United States, with participants having a mean age of 32.4 (SD 4.4) years. One study [38] enrolled both T1DM and T2DM patients in China with a mean age of 54.3 (SD 12.7) years [38]. The 23 studies ranged from 3 months to 1 year in terms of participant follow-up and investigated the efficacy of apps with respect to physical activity, healthy diets, physiological measures, physical measures, and quality of life. A total of 19 studies measured HbA_{1c} outcomes at the baseline, at posttreatment, or both time points in both the intervention and control groups; Multimedia Appendix 3 presents the relevant summary statistics. Of the 5 studies [22,25,27,39,41] that had multiple intervention arms based on apps, we included the most active arm in the meta-analysis.

App Characteristics

Multimedia Appendix 4 identifies and summarizes the apps used in the 23 included studies. We describe these below in greater detail.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses flowchart of included studies. ICT: information and communication technology; RCT: randomized controlled trial.



Type 1 Diabetes

All 5 apps in the T1DM studies had lifestyle modification as a major component and included the self-monitoring of participants' physical activity and healthy diets. Blood glucose was the only clinical measurement that had to be monitored and uploaded by patients. Of these apps, 4 [20-22,24] required professional input on patient health conditions from health care providers (HCPs) via text messages and telephone calls. The frequency of required HCP feedback ranged from once every week to once every 3 weeks. Other feedback modes, including graphical feedback and automated feedback, were employed to promote holistic awareness and to set personal health goals [20-24]. Because of T1DM's pathological characteristics, all 5 apps included an insulin bolus calculator or a medication adjustment supporter for glycemic control, or both [20-24]. In addition, 1 study that introduced the use of apps did not place an extra time cost on patients' self-management processes [22].

Type 2 Diabetes

The T2DM studies used 13 apps, 12 of which [25-35,39] were designed to modify patient self-management behavior through at least one type of feedback; 4 of these apps [26,31,33,34] provided HCP feedback when necessary. The clinical measurements logged into the apps were blood glucose [25-35], blood pressure [26,28,31,34], body weight [26,28,30,31,34], and mood [33]. Only 3 of these apps provided adjustment support for medications [25,29,32]. None of these apps had an insulin bolus calculator function. Physical activity monitoring was provided by 10 apps [26-28,30,31,33-35,39,40], and healthy diet monitoring was provided by 6 apps [25,27,29,33-35]. One app integrated context exercises into the physical activity function component and aimed to increase motivation and promote positive physical activity behavior [40]. This app, however, did not support assistance from other personnel or any form of feedback.

Gestational Diabetes

Kennelly and colleagues [42] studied an app for pregnant women that provided educational sessions on targeted nutrition and physical activity advice. The research team sent emails every other week to address specific problems. Throughout the study, follow-up hospital visits were carried out to ensure proper delivery of the intervention.

Prediabetes

The prediabetes studies used 3 apps [36,37,41] designed to help patients with personal weight management. All 3 apps monitored physical activity and healthy diet behavior. Body weight was the only clinical measurement tracked and recorded by patients themselves. Medication adjustment support and insulin bolus calculation were not specified in the articles. Only 1 app provided HCP feedback on a weekly to monthly basis via personalized messages and phone calls [41].

Risk of Bias Within Studies

We assessed the risk of bias in the 23 included studies using The Cochrane Collaboration's risk of bias tool. In some cases

[23-25,27,28,33,36,39-42], we found relevant details in study protocols. Figure 2 (biases by study) and Figure 3 (by overall percentage of each risk) show detailed results. Blinding of participants and personnel is the only area where the risk of bias was high; however, this high risk of performance bias was unlikely to have a large impact on HbA_{1c}, whose measurement is fairly objective. In all other areas, the risk of bias was low for most of the studies.

Figure 2. Risk of bias in each study. Green: low risk of bias; yellow: unclear risk of bias; red: high risk of bias.

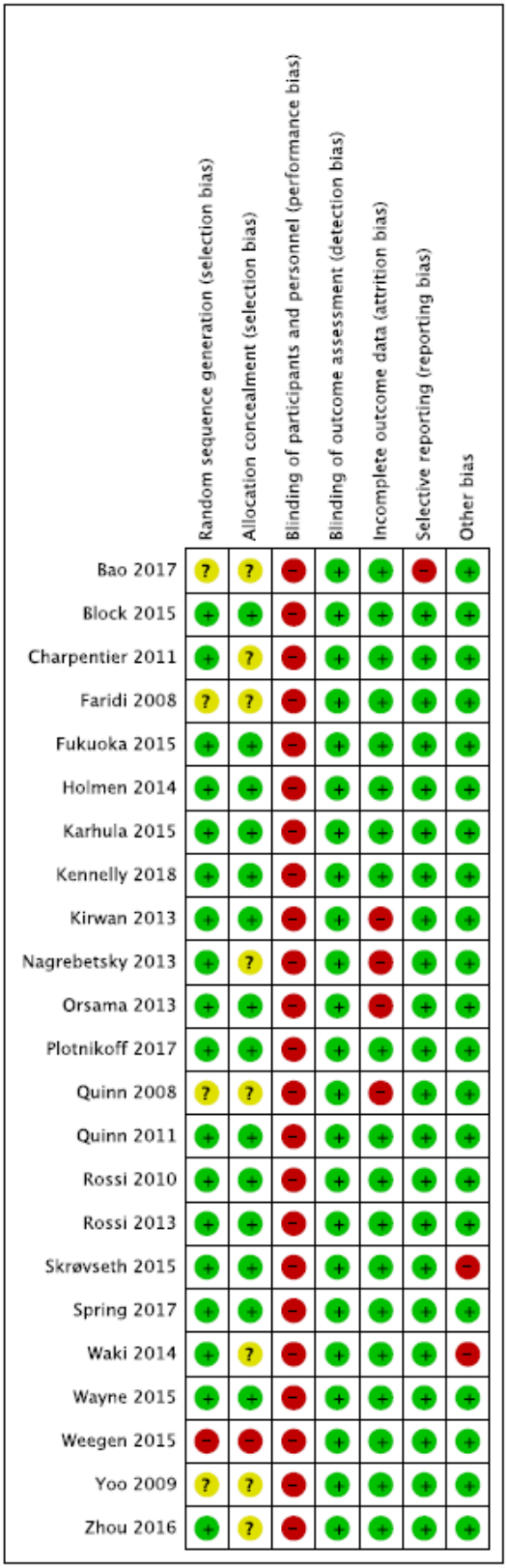
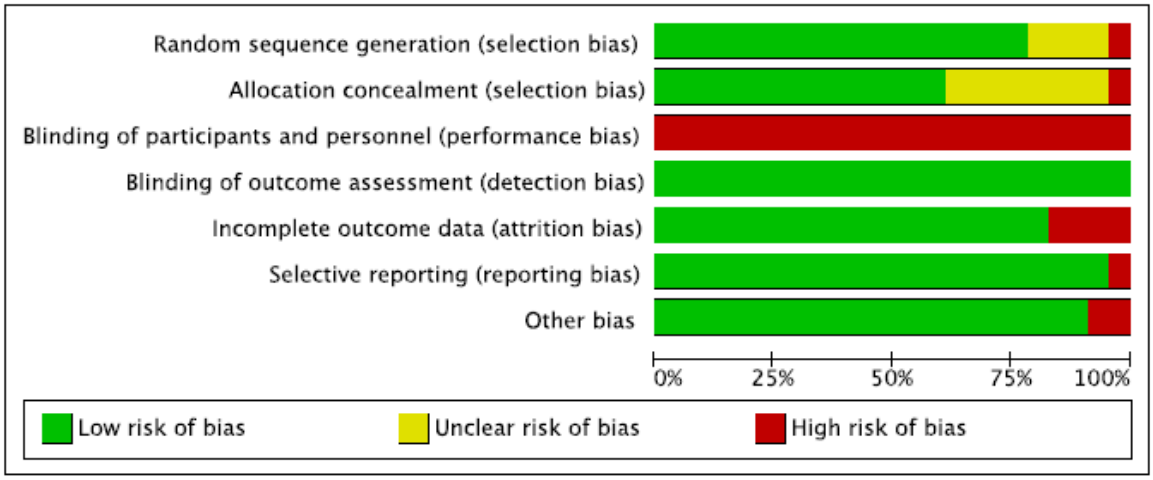


Figure 3. Overall risk of each type of bias.



Efficacy of Apps in Diabetes HbA $_{1c}$ Control

Type 1 Diabetes Apps

We included 5 T1DM studies in the quantitative synthesis, with 4 studies looking at the short-term effect and 1 looking at the long-term effect (Figure 4). In the short-term effect subgroup, the degree of heterogeneity was acceptable (I^2=39%), and we estimated the overall difference in HbA_{1c} between the app intervention and control groups to be –0.09 (95% CI –0.34 to 0.15), which was not significantly different from 0 (*P*=.18). There was significant heterogeneity in pooling the only long-term effect study with the short-term effect subgroup (I^2=64%). After pooling, the overall mean difference was statistically insignificant at –0.21 (95% CI –0.52 to 0.09; *P*=.17).

Type 2 Diabetes Apps

We included 11 T2DM studies in the quantitative synthesis, with 7 studies looking at the short-term effect and 4 looking at the long-term effect (Figure 5). In the short-term effect subgroup, there was virtually no heterogeneity (I^2=0%), and we estimated the overall difference in HbA_{1c} between the app intervention and control groups to be –0.48 (95% CI –0.69 to –0.28), which was significantly different from 0 (*P*<.01). In the long-term effect subgroup, the degree of heterogeneity was acceptable with I^2=2%, and we estimated the overall difference in HbA_{1c} between the app and control groups to be –0.25 (95% CI –0.43 to –0.07), which was significantly different from 0 (*P*<.01). There was virtually no heterogeneity in pooling the 2 subgroups together (I^2=0%), and the pooled difference in mean was statistically significant at –0.35 (95% CI –0.48 to –0.21; *P*<.01).

Figure 4. Forest plot of short- and long-term effects of apps for type 1 diabetes mellitus. IV: inverse variance.

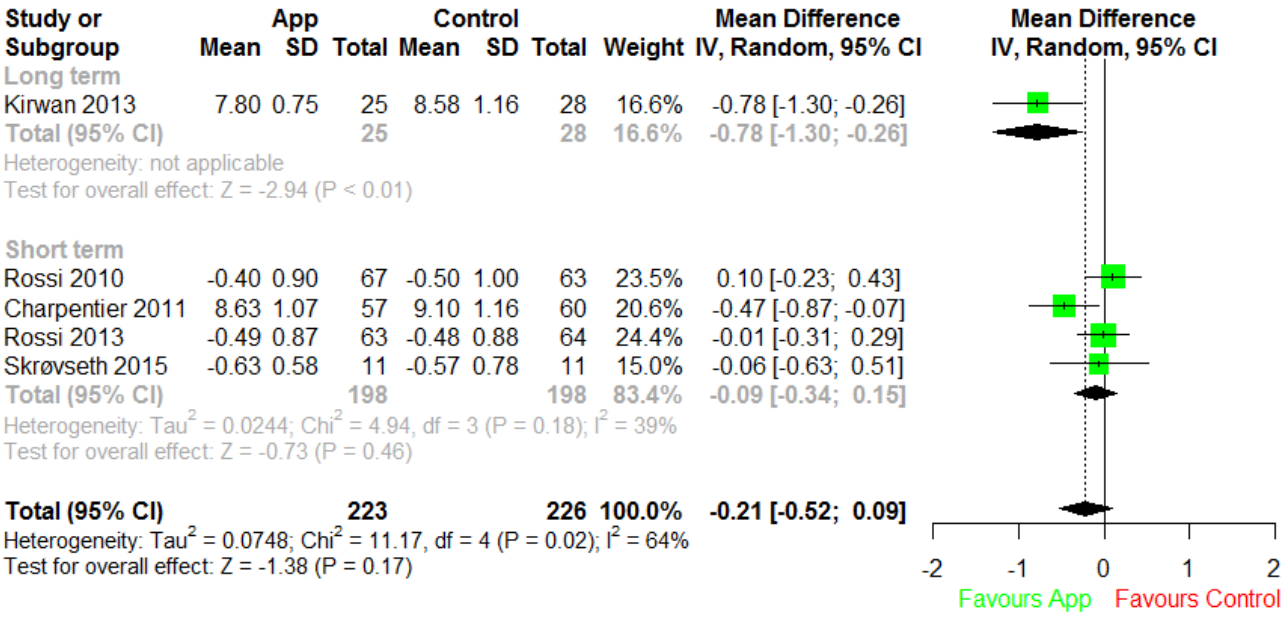
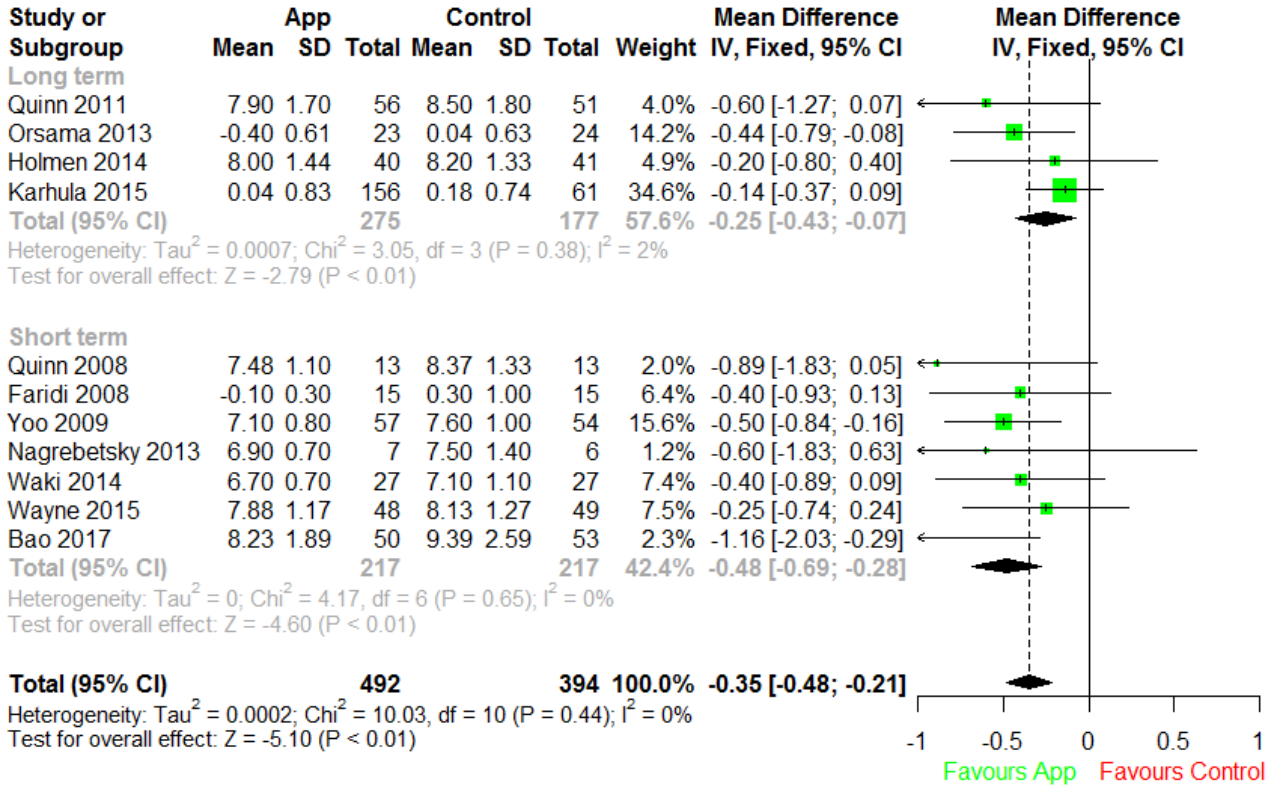


Figure 5. Forest plot of short- and long-term effects of apps for type 2 diabetes mellitus. IV: inverse variance.



We used the funnel plot method [43] to examine the publication bias of all 11 T2DM studies. Visual inspection of the funnel plot (Figure 6) revealed a fair amount of asymmetry. Furthermore, Egger and colleagues' linear regression test [44] was statistically significant (*P*=.02), indicating that the funnel plot was significantly asymmetric, possibly due to publication bias. Therefore, we undertook a sensitivity analysis using the trim and fill method [45] to adjust for possible publication bias in estimating the overall effect size. The trim and fill method estimated that there were 4 unpublished studies with negative findings, shown as open circles in Figure 7. After imputing the 4 unpublished studies, the funnel plot became symmetrical, and the pooled difference in mean HbA_{1c} remained statistically significant at –0.300 (95% CI –0.43 to –0.17; *P*<.001).

Figure 6. Funnel plot of publication bias. HbA_{1c}: glycated hemoglobin.

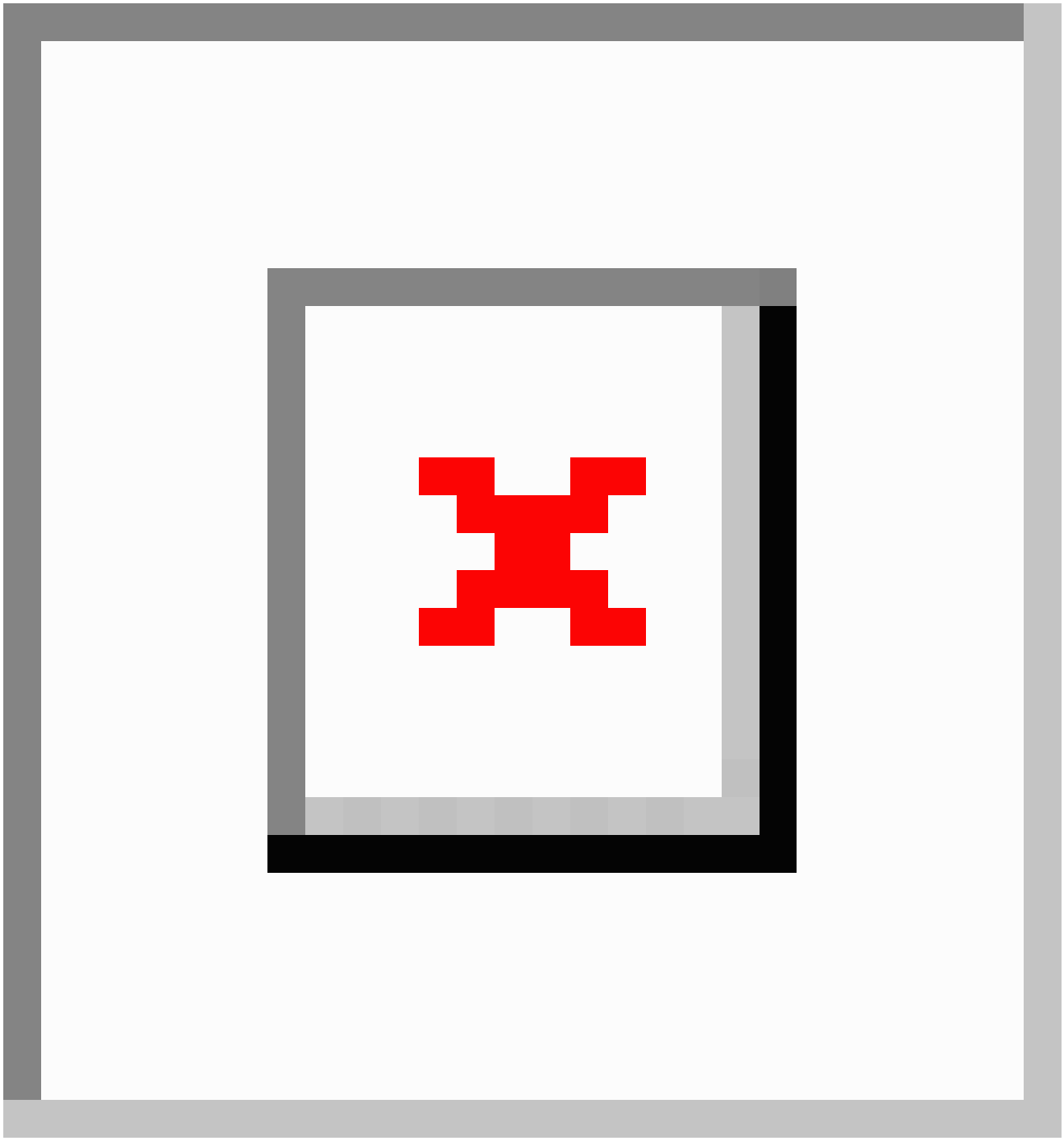


Figure 7. Trim and fill plot of publication bias. HbA_{1c}: glycated hemoglobin; open circles: estimated unpublished studies with negative findings.

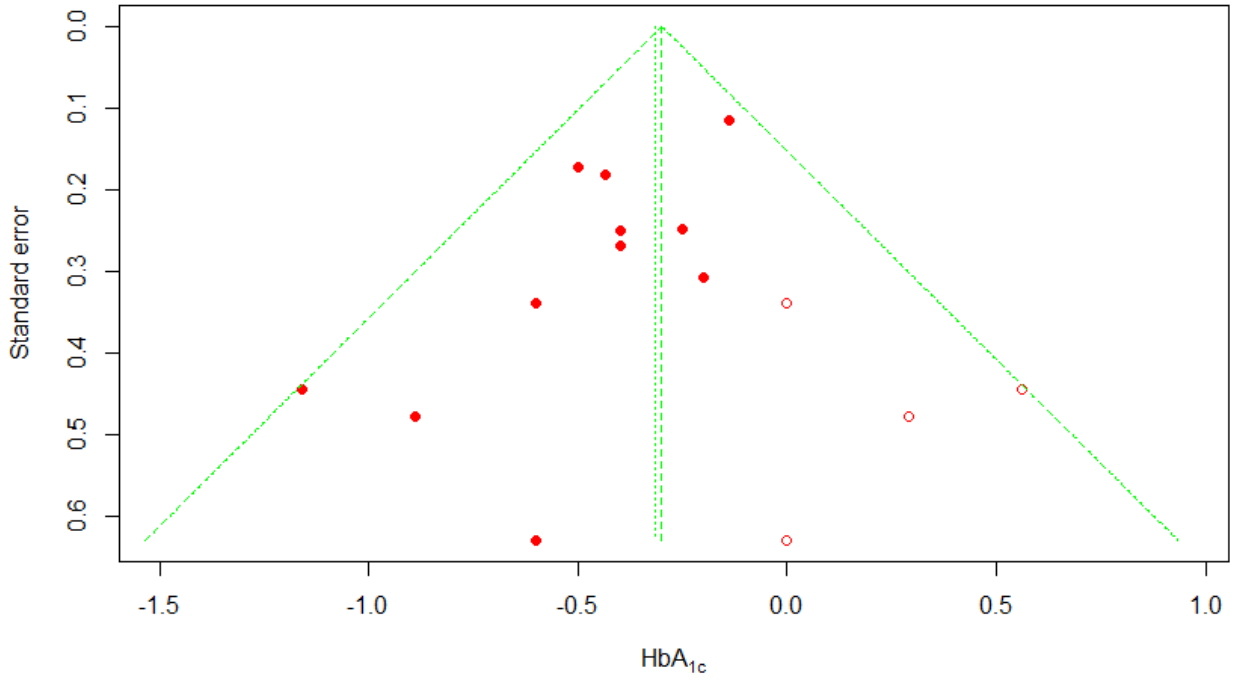
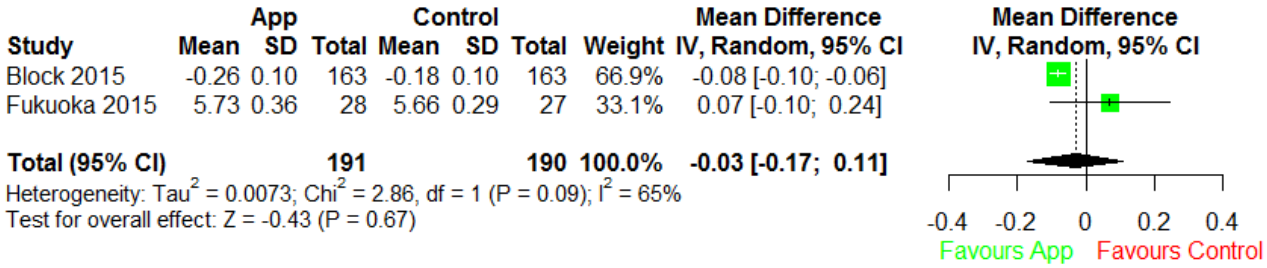


Figure 8. Forest plot the effect of prediabetes apps. IV: inverse variance.



Prediabetes Apps

We included 2 prediabetes studies in the quantitative synthesis, and both were short-term studies (Figure 8). Heterogeneity between the studies was significant (I^2=65%). After pooling, the mean reduction was statistically insignificant at –0.03 (95% CI –0.17 to 0.11; *P*=.67).

Discussion

Principal Findings

We included a total of 18 studies with HbA_{1c} outcomes in the quantitative meta-analysis. Among these, there were 5 T1DM studies (4 short-term studies and 1 long-term study). The short-term T1DM studies indicated an insignificant reduction in the HbA_{1c} level with acceptable heterogeneity. The only long-term T1DM study reported a significant reduction in the mean HbA_{1c} level. For T2DM, the meta-analysis included 11 studies, which together showed a significant reduction in the mean HbA_{1c} level, presumably owing to the persuasiveness of the apps for lifestyle modification. In subgroup analyses based on study duration, the long-term and short-term effects were both significant. The long-term effect estimate was slightly smaller than the short-term effect estimate, but the difference was not significant. The meta-analysis also included 2 studies of the prediabetes condition, both of which were of a short-term duration. There was significant heterogeneity between the 2 studies of the prediabetes condition, and the overall difference in the mean HbA_{1c} level between the app and control groups was not statistically significant.

Comparison With Prior Studies

To the best of our knowledge, this is the first systematic review and meta-analysis with subgroup analyses of apps based on diabetes subtypes and study duration. Previous reviews [46-51] did not include subgroup analyses based on diabetes subtypes and study duration. Some of them [46-48] were limited to one or two specific technologies (eg, pedometer, short message service); others [49,50] addressed different interventions (eg, nonapps and computer-based interventions), and 1 of them [51] focused on usability.

For T1DM, we explored 1 systematic review and meta-analysis published in 2016 [16]. That review found a nonsignificant reduction in the mean HbA_{1c} level of –0.10 (95% CI –0.41 to

0.21) with a high degree of heterogeneity (I^2=59.15%). Our review included 5 T1DM studies, 4 of which were of short-term duration. Our meta-analysis produced a similar overall effect estimate with less heterogeneity and better precision.

For T2DM, a systematic review and meta-analyses was previously published on the effectiveness of apps for noncommunicable diseases [52]. The 7 studies included in that meta-analysis were T2DM studies of major concern. We also included these studies in our meta-analysis [25-28,33,34]. In the previous reviews, the overall difference in the mean HbA_{1c} level between the app and control groups was estimated to be –0.50 (95% CI –0.91 to –0.08; I^2=41%) for short-term effects and –0.24 (95% CI –0.43 to –0.06; I^2=0%) for long-term effects. For the short-term effect, the results are similar to our results in this review. For the long-term effect, our meta-analysis produced a similar effect estimate with less heterogeneity and better precision. An important factor in the decreased heterogeneity and improved precision was the removal of the 1 study that involved both T1DM and T2DM patients [38].

We are not aware of a previous meta-analysis of the overall effect of apps for lifestyle modification in prediabetic patients. However, self-management and continuous care are pivotal issues in prediabetes care for clinical prognosis. Our meta-analysis based on 2 prediabetic studies produced an overall effect estimate of –0.03 (95% CI –0.17 to 0.11). The high degree of heterogeneity between our 2 studies (I^2=65%) suggests that important effect modifiers may exist for the effects of apps in this context. Searching for effect modifiers should be an important objective for future studies in this area.

As mobile platforms, apps can incorporate different function modules, such as lifestyle modification monitoring, health education, medication or insulin adjustment, logging of clinical measurements, and health management feedback. Feedback, a major behavior change technique, may be implemented as graphical, automated, or HCP feedback. Graphical feedback is probably the most elementary of the three forms; it is frequently used to visualize patient health data [23,24,27,28,30-32,34,36,38,39,41]. Automated feedback is usually provided in a personalized manner. In our review, many studies provided automated feedback based on algorithms or theories [20-22,25-31,34,36-39]. HCP feedback is provided by health care professionals, either in person or remotely [20-22,24-29,31-35,38,39,41,42].

Limitations

A major limitation of this research was the lack of outcome data beyond 12 months due to the limited duration of the studies reviewed. Diabetes is a chronic condition requiring sustained lifestyle modification, and it is important to understand the longer-term (beyond 12 months) efficacy and safety of apps in our elderly patient population. This gap may be closed by future studies with longer follow-up periods and more complete collection of outcome data (eg, mortality and adverse events). Another limitation is the reliance of some apps on self-reported food intake, which may be unreliable. The effectiveness of such apps in the real world will depend on the quality of user input.

Conclusion

The results of our review indicate that there is strong evidence for the efficacy of apps for lifestyle modification in T2DM, and that additional evidence is needed for the other subtypes of diabetes. The ambiguous results for T1DM may be related to the pathogenesis of the disease. The efficacy of T1DM self-management is heavily dependent on the administration of glucose with insulin and medication in the short term, which makes it difficult to demonstrate the efficacy of apps. Prediabetes conditions and GDM may be considered transition stages of diabetes, in which the continuum of care can affect the clinical prognosis directly. The different subtypes of diabetes clearly entail different considerations in designing and developing future apps for lifestyle modification in people with diabetes.

Acknowledgments

This study was partially funded by the National Natural Science of China (grants 71622002, 71531007, 71471048, 71871074, 71871073, 71801062 and 71771059).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[PDF File (Adobe PDF File), 121KB - mhealth_v7i1e12297_app1.pdf]

Multimedia Appendix 2

Characteristics of all included studies.

[PDF File (Adobe PDF File), 75KB - mhealth_v7i1e12297_app2.pdf]

Multimedia Appendix 3

Summary of glycated hemoglobin (HbA_{1c}) data by study, treatment arm, and time.

[PDF File (Adobe PDF File), 21KB - mhealth_v7i1e12297_app3.pdf]

Multimedia Appendix 4

Overview of key functions of the included apps.

[PDF File (Adobe PDF File), 41KB - mhealth_v7i1e12297_app4.pdf]

References

1. American Diabetes Association. 3. Comprehensive Medical Evaluation and Assessment of Comorbidities: Standards of Medical Care in Diabetes-2018. Diabetes Care 2018 Dec;41(Suppl 1):S28-S37. [doi: 10.2337/dc18-S003] [Medline: 29222374]
2. IDF Diabetes Atlas, 8th edition. Brussels, Belgium: International Diabetes Federation; 2017.
3. Ogurtsova K, da Rocha Fernandes JD, Huang Y, Linnenkamp U, Guariguata L, Cho NH, et al. IDF Diabetes Atlas: global estimates for the prevalence of diabetes for 2015 and 2040. Diabetes Res Clin Pract 2017 Jun;128:40-50. [doi: 10.1016/j.diabres.2017.03.024] [Medline: 28437734]
4. Schellenberg ES, Dryden DM, Vandermeer B, Ha C, Korownyk C. Lifestyle interventions for patients with and at risk for type 2 diabetes: a systematic review and meta-analysis. Ann Intern Med 2013 Oct 15;159(8):543-551. [doi: 10.7326/0003-4819-159-8-201310150-00007] [Medline: 24126648]
5. Kebede MM, Liedtke TP, Möllers T, Pischke CR. Characterizing active ingredients of eHealth interventions targeting persons with poorly controlled type 2 diabetes mellitus using the behavior change techniques taxonomy: scoping review. J Med Internet Res 2017 Oct 12;19(10):e348 [FREE Full text] [doi: 10.2196/jmir.7135] [Medline: 29025693]
6. Hemmingsen B, Gimenez-Perez G, Mauricio D, Roqué IFM, Metzendorf M, Richter B. Diet, physical activity or both for prevention or delay of type 2 diabetes mellitus and its associated complications in people at increased risk of developing type 2 diabetes mellitus. Cochrane Database Syst Rev 2017 Dec 04;12:CD003054. [doi: 10.1002/14651858.CD003054.pub4] [Medline: 29205264]
7. American Diabetes Association. 1. Promoting Health and Reducing Disparities in Populations. Diabetes Care 2017 Dec;40(Suppl 1):S6-S10. [doi: 10.2337/dc17-S004] [Medline: 27979888]
8. 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]
9. Waite M, Martin C, Franklin R, Duce D, Harrison R. Human factors and data logging processes with the use of advanced technology for adults with type 1 diabetes: systematic integrative review. JMIR Hum Factors 2018 Mar 15;5(1):e11 [FREE Full text] [doi: 10.2196/humanfactors.9049] [Medline: 29535079]
10. Van Ameringen M, Turna J, Khalesi Z, Pullia K, Patterson B. There is an app for that! The current state of mobile applications (apps) for DSM-5 obsessive-compulsive disorder, posttraumatic stress disorder, anxiety and mood disorders. Depress Anxiety 2017 Jun;34(6):526-539. [doi: 10.1002/da.22657] [Medline: 28569409]
11. Mantani A, Kato T, Furukawa TA, Horikoshi M, Imai H, Hiroe T, et al. Smartphone cognitive behavioral therapy as an adjunct to pharmacotherapy for refractory depression: randomized controlled trial. J Med Internet Res 2017 Nov 03;19(11):e373 [FREE Full text] [doi: 10.2196/jmir.8602] [Medline: 29101095]
12. Masterson Creber RM, Maurer MS, Reading M, Hiraldo G, Hickey KT, Iribarren S. Review and analysis of existing mobile phone apps to support heart failure symptom monitoring and self-care management using the mobile application rating scale (MARS). JMIR Mhealth Uhealth 2016 Jun 14;4(2):e74. [doi: 10.2196/mhealth.5882] [Medline: 27302310]
13. Brinker TJ, Brieske CM, Esser S, Klode J, Mons U, Batra A, et al. A face-aging app for smoking cessation in a waiting room setting: pilot study in an HIV outpatient clinic. J Med Internet Res 2018 Aug 15;20(8):e10976 [FREE Full text] [doi: 10.2196/10976] [Medline: 30111525]
14. BinDhim NF, McGeechan K, Trevena L. Smartphone Smoking Cessation Application (SSC App) trial: a multicountry double-blind automated randomised controlled trial of a smoking cessation decision-aid 'app'. BMJ Open 2018 Dec 21;8(1):e017105 [FREE Full text] [doi: 10.1136/bmjopen-2017-017105] [Medline: 29358418]
15. Chavez S, Fedele D, Guo Y, Bernier A, Smith M, Warnick J, et al. Mobile apps for the management of diabetes. Diabetes Care 2017 Oct;40(10):e145-e146. [doi: 10.2337/dc17-0853] [Medline: 28774944]
16. Hou C, Carter B, Hewitt J, Francisa T, Mayor S. Do mobile phone applications improve glycemic control (HbA1c) in the self-management of diabetes? A systematic review, meta-analysis, and GRADE of 14 randomized trials. Diabetes Care 2016 Nov;39(11):2089-2095. [doi: 10.2337/dc16-0346] [Medline: 27926892]
17. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. BMJ 2009;339:b2535 [FREE Full text] [Medline: 19622551]

18. Weir CJ, Butcher I, Assi V, Lewis SC, Murray GD, Langhorne P, et al. Dealing with missing standard deviation and mean values in meta-analysis of continuous outcomes: a systematic review. BMC Med Res Methodol 2018 Mar 07;18(1):25 [FREE Full text] [doi: 10.1186/s12874-018-0483-0] [Medline: 29514597]
19. Higgins JPT, Altman DG, Sterne JAC. Assessing risk of bias in included studies. In: Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0. Oxford, UK: The Cochrane Collaboration; 2011.
20. Rossi MCE, Nicolucci A, Di Bartolo P, Bruttomesso D, Girelli A, Ampudia FJ, et al. Diabetes Interactive Diary: a new telemedicine system enabling flexible diet and insulin therapy while improving quality of life: an open-label, international, multicenter, randomized study. Diabetes Care 2010 Jan;33(1):109-115 [FREE Full text] [doi: 10.2337/dc09-1327] [Medline: 19808926]
21. Rossi MC, Nicolucci A, Lucisano G, Pellegrini F, Di Bartolo P, Miselli V, et al. Impact of the "Diabetes Interactive Diary" telemedicine system on metabolic control, risk of hypoglycemia, and quality of life: a randomized clinical trial in type 1 diabetes. Diabetes Technol Ther 2013 Aug;15(8):670-679. [doi: 10.1089/dia.2013.0021] [Medline: 23844569]
22. Charpentier G, Benhamou P, Dardari D, Clergeot A, Franc S, Schaepelynck-Belicar P, et al. The Diabeo software enabling individualized insulin dose adjustments combined with telemedicine support improves HbA1c in poorly controlled type 1 diabetic patients: a 6-month, randomized, open-label, parallel-group, multicenter trial (TeleDiab 1 Study). Diabetes Care 2011 Mar;34(3):533-539 [FREE Full text] [doi: 10.2337/dc10-1259] [Medline: 21266648]
23. Skrøvseth SO, Årsand E, Godtliebsen F, Joakimsen RM. Data-driven personalized feedback to patients with type 1 diabetes: a randomized trial. Diabetes Technol Ther 2015 Jul;17(7):482-489 [FREE Full text] [doi: 10.1089/dia.2014.0276] [Medline: 25751133]
24. Kirwan M, Vandelanotte C, Fenning A, Duncan MJ. Diabetes self-management smartphone application for adults with type 1 diabetes: randomized controlled trial. J Med Internet Res 2013;15(11):e235 [FREE Full text] [doi: 10.2196/jmir.2588] [Medline: 24225149]
25. Quinn CC, Shardell MD, Terrin ML, Barr EA, Ballew SH, Gruber-Baldini AL. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. Diabetes Care 2011 Sep;34(9):1934-1942 [FREE Full text] [doi: 10.2337/dc11-0366] [Medline: 21788632]
26. Orsama A, Lähteenmäki J, Harno K, Kulju M, Wintergerst E, Schachner H, et al. Active assistance technology reduces glycosylated hemoglobin and weight in individuals with type 2 diabetes: results of a theory-based randomized trial. Diabetes Technol Ther 2013 Aug;15(8):662-669. [doi: 10.1089/dia.2013.0056] [Medline: 23844570]
27. Holmen H, Torbjørnsen A, Wahl AK, Jenum AK, Småstuen MC, Arsand E, et al. A mobile health intervention for self-management and lifestyle change for persons with type 2 diabetes, part 2: one-year results from the Norwegian randomized controlled trial RENEWING HEALTH. JMIR Mhealth Uhealth 2014;2(4):e57 [FREE Full text] [doi: 10.2196/mhealth.3882] [Medline: 25499872]
28. Karhula T, Vuorinen A, Rääpysjärvi K, Pakanen M, Itkonen P, Tepponen M, et al. Telemonitoring and mobile phone-based health coaching among Finnish diabetic and heart disease patients: randomized controlled trial. J Med Internet Res 2015;17(6):e153 [FREE Full text] [doi: 10.2196/jmir.4059] [Medline: 26084979]
29. Quinn CC, Clough SS, Minor JM, Lender D, Okafor MC, Gruber-Baldini A. WellDoc mobile diabetes management randomized controlled trial: change in clinical and behavioral outcomes and patient and physician satisfaction. Diabetes Technol Ther 2008 Jun;10(3):160-168. [doi: 10.1089/dia.2008.0283] [Medline: 18473689]
30. Faridi Z, Liberti L, Shuval K, Northrup V, Ali A, Katz DL. Evaluating the impact of mobile telephone technology on type 2 diabetic patients' self-management: the NICHE pilot study. J Eval Clin Pract 2008 Jun;14(3):465-469. [doi: 10.1111/j.1365-2753.2007.00881.x] [Medline: 18373577]
31. Yoo HJ, Park MS, Kim TN, Yang SJ, Cho GJ, Hwang TG, et al. A ubiquitous chronic disease care system using cellular phones and the internet. Diabet Med 2009 Jun;26(6):628-635. [doi: 10.1111/j.1464-5491.2009.02732.x] [Medline: 19538239]
32. Nagrebetsky A, Larsen M, Craven A, Turner J, McRobert N, Murray E, et al. Stepwise self-titration of oral glucose-lowering medication using a mobile telephone-based telehealth platform in type 2 diabetes: a feasibility trial in primary care. J Diabetes Sci Technol 2013;7(1):123-134 [FREE Full text] [Medline: 23439168]
33. Wayne N, Perez DF, Kaplan DM, Ritvo P. Health coaching reduces HbA1c in type 2 diabetic patients from a lower-socioeconomic status community: a randomized controlled trial. J Med Internet Res 2015;17(10):e224 [FREE Full text] [doi: 10.2196/jmir.4871] [Medline: 26441467]
34. Waki K, Fujita H, Uchimura Y, Omae K, Aramaki E, Kato S, et al. DialBetics: a novel smartphone-based self-management support system for type 2 diabetes patients. J Diabetes Sci Technol 2014 Mar 13;8(2):209-215 [FREE Full text] [doi: 10.1177/1932296814526495] [Medline: 24876569]
35. Bao S, Jiang H, Luo Y, Zhang D. Application of diabetes phone recipe software in diet intervention for patients with type 2 diabetes. Chinese Nurs Res 2017;31(11):1407-1408. [doi: 10.3969/j.issn.1009-6493.2017.11.042]
36. Block G, Azar KM, Romanelli RJ, Block TJ, Hopkins D, Carpenter HA, et al. Diabetes prevention and weight loss with a fully automated behavioral intervention by email, web, and mobile phone: a randomized controlled trial among persons with prediabetes. J Med Internet Res 2015 Oct 23;17(10):e240 [FREE Full text] [doi: 10.2196/jmir.4897] [Medline: 26499966]

37. Fukuoka Y, Gay CL, Joiner KL, Vittinghoff E. A novel diabetes prevention intervention using a mobile app: a randomized controlled trial with overweight adults at risk. Am J Prev Med 2015 Aug;49(2):223-237 [FREE Full text] [doi: 10.1016/j.amepre.2015.01.003] [Medline: 26033349]
38. Zhou W, Chen M, Yuan J, Sun Y. Welltang - a smart phone-based diabetes management application - improves blood glucose control in Chinese people with diabetes. Diabetes Res Clin Pract 2016 Jun;116:105-110. [doi: 10.1016/j.diabres.2016.03.018] [Medline: 27321324]
39. van der Weegen S, Verwey R, Spreeuwenberg M, Tange H, van der Weijden T, de Witte L. It's LiFe! Mobile and web-based monitoring and feedback tool embedded in primary care increases physical activity: a cluster randomized controlled trial. J Med Internet Res 2015 Jul 24;17(7):e184 [FREE Full text] [doi: 10.2196/jmir.4579] [Medline: 26209025]
40. Plotnikoff RC, Wilczynska M, Cohen KE, Smith JJ, Lubans DR. Integrating smartphone technology, social support and the outdoor physical environment to improve fitness among adults at risk of, or diagnosed with, type 2 diabetes: findings from the 'eCoFit' randomized controlled trial. Prev Med 2017 Dec;105:404-411. [doi: 10.1016/j.ypmed.2017.08.027] [Medline: 28887192]
41. Spring B, Pellegrini CA, Pfammatter A, Duncan JM, Pictor A, McFadden HG, et al. Effects of an abbreviated obesity intervention supported by mobile technology: the ENGAGED randomized clinical trial. Obesity (Silver Spring) 2017 Jul;25(7):1191-1198 [FREE Full text] [doi: 10.1002/oby.21842] [Medline: 28494136]
42. Kennelly MA, Ainscough K, Lindsay KL, O Sullivan E, Gibney ER, McCarthy M, et al. Pregnancy exercise and nutrition with smartphone application support: a randomized controlled trial. Obstet Gynecol 2018 May;131(5):818-826. [doi: 10.1097/AOG.0000000000002582] [Medline: 29630009]
43. Sterne JAC, Sutton AJ, Ioannidis JPA, Terrin N, Jones DR, Lau J, et al. Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. BMJ 2011;343:d4002. [Medline: 21784880]
44. Egger M, Davey SG, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ 1997 Sep 13;315(7109):629-634 [FREE Full text] [Medline: 9310563]
45. Duval S, Tweedie R. Trim and fill: a simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics 2000 Jun;56(2):455-463. [Medline: 10877304]
46. Baron J, McBain H, Newman S. The impact of mobile monitoring technologies on glycosylated hemoglobin in diabetes: a systematic review. J Diabetes Sci Technol 2012 Sep;6(5):1185-1196 [FREE Full text] [Medline: 23063046]
47. Buhi ER, Trudnak TE, Martinasek MP, Oberne AB, Fuhrmann HJ, McDermott RJ. Mobile phone-based behavioural interventions for health: a systematic review. Health Educ J 2012 Jul 10;72(5):564-583. [doi: 10.1177/0017896912452071]
48. Russell-Minda E, Jutai J, Speechley M, Bradley K, Chudyk A, Petrella R. Health technologies for monitoring and managing diabetes: a systematic review. J Diabetes Sci Technol 2009 Nov 01;3(6):1460-1471 [FREE Full text] [doi: 10.1177/193229680900300628] [Medline: 20144402]
49. Liang X, Wang Q, Yang X, Cao J, Chen J, Mo X, et al. Effect of mobile phone intervention for diabetes on glycaemic control: a meta-analysis. Diabet Med 2011 Apr;28(4):455-463. [doi: 10.1111/j.1464-5491.2010.03180.x] [Medline: 21392066]
50. Pal K, Eastwood SV, Michie S, Farmer A, Barnard ML, Peacock R, et al. Computer-based interventions to improve self-management in adults with type 2 diabetes: a systematic review and meta-analysis. Diabetes Care 2014 Jun;37(6):1759-1766. [doi: 10.2337/dc13-1386] [Medline: 24855158]
51. Arnhold M, Quade M, Kirch W. Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. J Med Internet Res 2014;16(4):e104 [FREE Full text] [doi: 10.2196/jmir.2968] [Medline: 24718852]
52. Lunde P, Nilsson BB, Bergland A, Kværner KJ, Bye A. The effectiveness of smartphone apps for lifestyle improvement in noncommunicable diseases: systematic review and meta-analyses. J Med Internet Res 2018 May 04;20(5):e162 [FREE Full text] [doi: 10.2196/jmir.9751] [Medline: 29728346]

Abbreviations

GDM: gestational diabetes mellitus
HbA_{1c}: glycated hemoglobin
HCP: health care provider
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
T1DM: type 1 diabetes mellitus
T2DM: type 2 diabetes mellitus

Edited by G Eysenbach; submitted 24.09.18; peer-reviewed by Y Cai, W Min, C Bakker, V Osmani; comments to author 01.11.18; revised version received 21.11.18; accepted 29.12.18; published 15.01.19.

Please cite as:
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;7(1):e12297
URL: http://mhealth.jmir.org/2019/1/e12297/
doi:10.2196/12297
PMID:30664494

Original Paper

Efficacy and Outcomes of a Music-Based Emotion Regulation Mobile App in Distressed Young People: Randomized Controlled Trial

Leanne Hides[1,2,3], PhD (Clin), BBehSc (Hons); Genevieve Dingle[2], PhD (Clin); Catherine Quinn[2], PhD (Clin); Stoyan R Stoyanov[1,2,3], MRes (Psych), BPsych (Hons); Oksana Zelenko[4], PhD; Dian Tjondronegoro[5], PhD; Daniel Johnson[5], PhD; Wendell Cockshaw[2,3], PhD; David J Kavanagh[2,3], PhD

[1]School of Psychology, The University of Queensland, Brisbane, Australia

[2]School of Psychology & Counselling, Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Australia

[3]Centre for Children's Health Research, Queensland University of Technology, Brisbane, Australia

[4]Creative Industries Faculty, Queensland University of Technology, Brisbane, Australia

[5]School of Business and Tourism, Southern Cross University, Gold Coast, Australia

Corresponding Author:
Leanne Hides, PhD (Clin), BBehSc (Hons)
School of Psychology
The University of Queensland
McElwain Building, St Lucia
Brisbane, 4072
Australia
Phone: 1 406185750
Email: l.hides@uq.edu.au

Abstract

Background: Emotion dysregulation increases the risk of depression, anxiety, and substance use disorders. Music can help regulate emotions, and mobile phones provide constant access to it. The *Music eScape* mobile app teaches young people how to identify and manage emotions using music.

Objective: This study aimed to examine the effects of using *Music eScape* on emotion regulation, distress, and well-being at 1, 2, 3, and 6 months. Moderators of outcomes and user ratings of app quality were also examined.

Methods: A randomized controlled trial compared immediate versus 1-month delayed access to *Music eScape* in 169 young people (aged 16 to 25 years) with at least mild levels of mental distress (Kessler 10 score>17).

Results: No significant differences between immediate and delayed groups on emotion regulation, distress, or well-being were found at 1 month. Both groups achieved significant improvements in 5 of the 6 emotion regulation skills, mental distress, and well-being at 2, 3, and 6 months. Unhealthy music use moderated improvements on 3 emotion regulation skills. Users gave the app a high mean quality rating (mean 3.8 [SD 0.6]) out of 5.

Conclusions: Music eScape has the potential to provide a highly accessible way of improving young people's emotion regulation skills, but further testing is required to determine its efficacy. Targeting unhealthy music use in distressed young people may improve their emotion regulation skills.

Trial Registration: Australian New Zealand Clinical Trials Registry ACTRN12615000051549; https://www.anzctr.org.au/Trial/Registration/TrialReview.aspx?id=365974

(JMIR Mhealth Uhealth 2019;7(1):e11482) doi:10.2196/11482

KEYWORDS

depression; anxiety; emotion regulation; music; telehealth; mHealth

Introduction

Young People and Emotion Regulation

Mental health and substance use disorders are at their peak and the leading causes of disability and death worldwide in young people (aged 15-25 years) [1]. Anxiety, depression, and substance use disorders are the most common in this age group. Deficits in emotion regulation or the ability to identify, evaluate, express, and modify emotions are important risk and maintaining factors for these disorders [2-5]. Emotion regulation requires the development and integration of several emotion skills such as an awareness of one's emotional states, directing one's attention away from the cause of a negative emotion; cognitively reappraising the cause of the negative emotion; and accepting, discharging, or suppressing a negative emotional state [6,7]. Young people commonly experience intense emotional states during puberty and the transition from childhood to adulthood [8]. At the same time, their capacity to fully regulate emotional states is still developing [9,10]. Effective emotion regulation skills may help decrease the intensity and duration of dysphoric states, whereas ineffective emotion regulation may increase them [3].

Only a third of young people with mental disorders seek help [11] and many fail to engage in or complete psychological treatment. Around 60% of young people with anxiety disorders respond to cognitive behavioral therapy, with response rates for unipolar major depression ranging from 48% to 81% over 36 weeks [12-15]. Nevertheless, around 40% of young people with anxiety and 20% to 25% with depressive disorders do not respond to treatment [13,15-17], and 47% of responders to depression treatment relapse within 5 years [18].

Music and Emotion

Novel interventions targeting young people's emotion regulation skills could reduce the risk of anxiety or depression disorders and improve help seeking, treatment engagement, and outcomes among those with these disorders. Music listening is one of the favorite leisure activities of young people [19] and one of their most commonly used emotion regulation strategies [20-27], being recently ranked the top stress management strategy in young Australians [28].

Music is commonly used by young people and others to induce, enhance, maintain, or manage moods [29-31]. The impact of music on mood varies according to the goal of the user [32], their pre-existing mood [33], and the type of emotion regulation strategy being used [34]. For example, a correlational study found substantial proportions of adolescents reported improvements in mood when listening to music; these effects were most pronounced when they were happy or bored, rather than angry or sad [33]. Positive effects of music on negative moods appeared constrained by some adolescents preferring angry rather than happy music when in a negative mood [33]. A Web-based experiment found listening to self-selected sad music increased depressive moods and listening to happy music reduced them [34]. However, a partial replication of that study found both participant- and experimenter-selected sad music reduced a depressive mood if a negative mood was induced (via a video clip) before music listening [35]. Finally, music use among people who use cognitive reappraisal as an emotion regulation strategy has been found to enhance well-being, whereas the use of expressive suppression reduced well-being [36]. In summary, music appears to have potential to be used as an effective emotion regulation strategy for improving mood and well-being, although the relationship is complex and controlled studies of adequate power are needed [37].

Programs Using Music to Target Emotion

The relationship between music and emotions has been harnessed in programs aimed at teaching emotion regulation skills to individuals with mental health problems in clinical and community settings, including eating disorders [38], anxiety disorders [39], substance misuse [40], and schizophrenia [41]. One example is the *Tuned In* group program, which uses hypothetical scenarios and participant-selected music to evoke emotions in sessions to increase emotional awareness and emotion regulation skills. This program demonstrated significant improvements in emotional awareness and regulation post treatment among 41 at-risk adolescents attending an education re-engagement program and 216 adolescents attending an independent mainstream secondary school [42]. A 4-session version of the program among dysphoric first-year university students (n=51; aged 18-25 years) found greater emotional awareness and regulation post treatment compared with a 4-week waitlist control group [43]. These findings provide preliminary evidence that music programs such as *Tuned In* result in positive emotion regulation outcomes.

Media has the potential to enhance the emotion regulation skills and mood of young people in their everyday lives. For instance, a survey study of 229 people found that mood-specific media use might be captured by 3 factors: turning to media in a positive mood, in a negative mood, or in a bored mood [10]. Various forms of difficulty regulating emotion (eg, feeling out of control when upset) predicted media use in negative or bored moods only. More specific analyses show that music use in negative moods is predicted by both positive indices (eg, reflection tendencies) and negative indices of emotion regulation (eg, rumination tendencies), whereas television use in negative moods is only predicted by negative indices of emotion regulation [10]. A systemic review of 23 studies on the use of video games for emotion regulation reported that frequent (but not excessive) video game play, including serious games, may enhance emotion regulation, but commercial gaming offered more opportunities for emotion regulation improvement than limited-time (bespoke) games [44].

Music also has the potential to enhance the emotion regulation skills and mood of young people in their everyday lives. Mobile phones that contain digital music players, personal music libraries, and access to digital radio provide a platform for achieving this. Targeted music apps, therefore, provide an anonymous and highly accessible way of providing young people with the skills to identify, express, and manage emotions in their natural environment [19,45,46]. A recent meta-analysis of 21 studies of electronic health (eHealth) interventions for youth concluded that such apps could result in population-level benefits, even with small effects (d=0.13; 95% CI 0.02-0.25) [47].

A growing number of mobile phone apps targeting emotions through music are becoming available. Several approaches are used, including streaming mood-related playlists, providing mood-tagging options for users' own libraries, and playing sounds and soundscapes to promote relaxation [31]. However, it is unclear how the music mood ratings contained in these are derived, and the majority are not specifically designed to help young people regulate emotions.

Objectives

This study aimed to evaluate a new app called *Music eScape*, developed to assist young people with identifying, expressing, and managing emotions using music from their own music library. This study reports the 1-month efficacy and 2-, 3-, and 6-month outcomes of the *Music eScape* app in a sample of young people with at least mild mental distress. Potential moderators of app outcomes, including the amount of music use and healthy or unhealthy music use, were examined. In addition, user ratings of the app's quality were obtained after a month of its use.

Methods

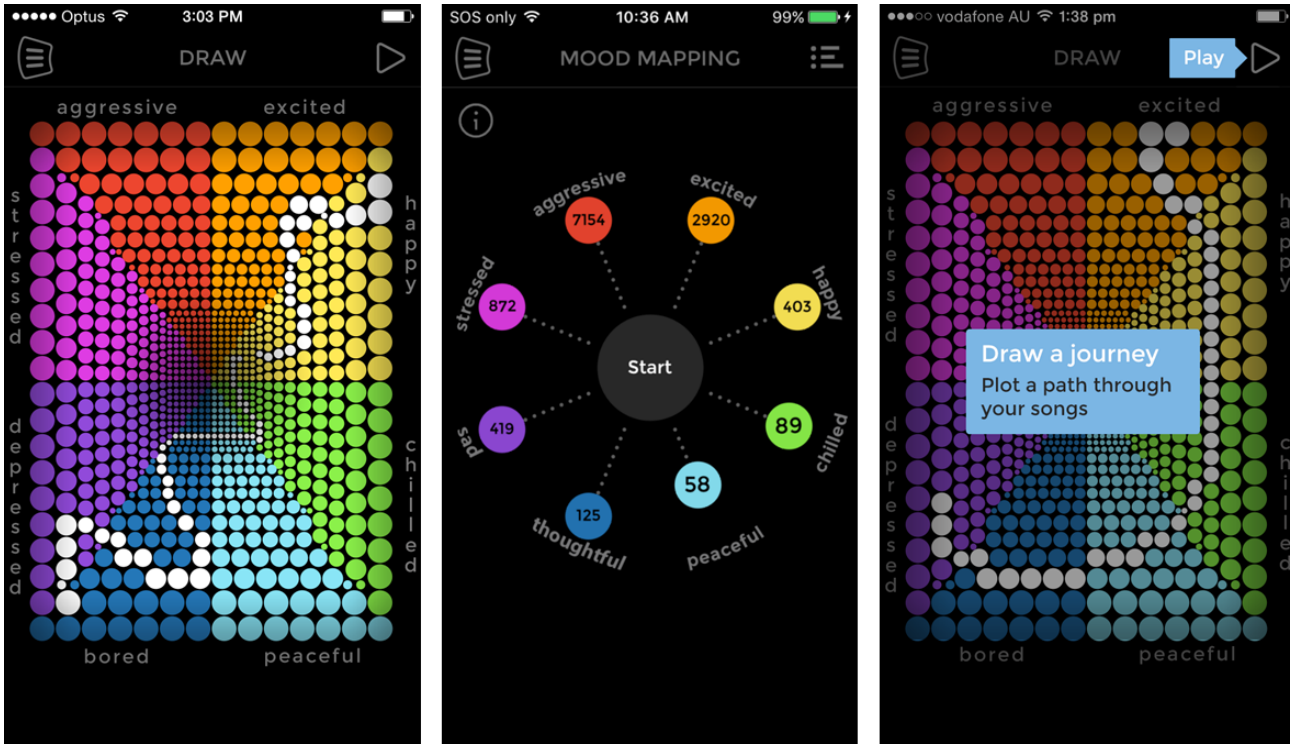
Music eScape App

The *Music eScape* app was co-designed by young people and a multidisciplinary research team using a series of participatory design workshops [48]. App design was informed by the dynamic information-motivation-behavioral skills health behavior model [49,50], and agile development processes were used. The *Music eScape* app analyzes each song in the users' music library according to its level of valence (pleasant to unpleasant) and arousal (very low to very high) using The Echo Nest music data program [51]. The songs are then located in a two-dimensional space consistent with Russell's circumplex model of emotion [52], labeled around the borders with 8 emotions (see Figure 1, left screenshot): aggressive, excited, happy, chilled, peaceful, bored, depressed, and stressed. Once the music scanning is complete, the user is presented with a mood map of their music (the eScape) to help them identify the prevalent moods of their library (see Figure 1, middle screenshot). Before creating a playlist, the app prompts users to reflect on their current and desired mood and then encourages them to plot a mood journey the playlist will support them to create (see Figure 1, right screenshot). This journey comprises a unique trajectory using their own music (eg, starting in the bored segment and ending in the happy one; see Figure 1, right screenshot). Users can save and label their musical mood journeys (eg, *Chill out*) and can also specify the duration of their playlist, from 15 min to 60 min. They can also select preset mood journeys. After completing their mood journey, users are asked to reflect on their current mood and rate the effectiveness of the playlist they just experienced.

Participants and Recruitment

Participants were Australian residents aged 16 to 25 years, who reported at least mild distress in the past month on the Kessler 10 Psychological Distress scale (K10>17) and had an iPhone. Recruitment was via student emails and posters in 2 large universities and snowballing techniques. The advertisements invited young people (aged 16-25 years) who owned an iPhone and *felt stressed* to participate in a study testing a new mood management app. They did not include any mention about music in an attempt to avoid recruiting a selective sample of participants with a high affinity to music. The purpose of the study was also concealed during the consent process, such that participants were not aware of the fact that the mood management app used music until they received access to it.

Figure 1. Music eScape app: (left) home screen; (middle) music mood map; and (right) mood journey instructions.



Measures

Emotion Regulation

The 18-item short-form of the Difficulties in Emotion Regulation Scale (DERS-SF) [53] was used to assesses emotion dysregulation on a scale from 1 (almost never) to 5 (almost always). The DERS-SF has excellent reliability and validity and a similar factor structure to the original 36-item scale in adolescents and adults [54]. There are 6 subscales: lack of emotional awareness (alpha=.80), lack of emotional clarity (alpha=.83), difficulties engaging in goal-directed behaviors (alpha=.88), impulse control difficulties (alpha=.91), nonacceptance of emotional responses (alpha=.85), and limited access to emotion regulation strategies (alpha=.85). Participants were also asked to rate their perceived level of success with using music as an emotion regulation strategy on a Likert scale (1=not at all successful to 9=extremely successful; item derived from Thayer et al's study) [55].

Mental Distress and Well-Being

The K10 scale [56] assessed the frequency of psychological distress in the past month, using items rated from 1 (none of the time) to 5 (all of the time). The K10 is a widely used screening tool developed using item response theory to determine the probable presence or absence of a diagnosable anxiety or depressive disorder. Normative data indicate that a cut-off of ≥17 is indicative of at least mild mental distress and is at the seventy-fifth percentile among Australian youth (aged 16-24 years) [57]. Internal consistency was high in this study sample (alpha=.87).

Mental well-being was measured with the Mental Health Continuum-Short Form (MHC-SF) [58,59]. This 14-item scale measures how frequently respondents experienced emotional, psychological, and social well-being in the past month on a 6-point response scale (1=never to 6=every day). The MHC-SF has high levels of reliability and discriminant, convergent, and cross-cultural validity. Internal consistency was .94.

Music Measures

A total of 10 items designed specifically for this study explored the level of music education and involvement of participants. Example items include the following: "Do you currently play a musical instrument and/or sing in a group or choir?" (yes or no) and "Do you attend concerts or live music on a regular basis (ie, at least once a month)?" (yes or no).

The Healthy-Unhealthy Music Scale [60] assesses healthy (5 items) and unhealthy (8 items) uses of music, with items rated from 1 (never) to 5 (always). Healthy and unhealthy music use refers to protective (eg, "Music gives me the energy to get going") versus risky (eg, "When I listen to music I get stuck in bad memories") forms of music engagement [60]. The healthy subscale has demonstrated concurrent validity with well-being, happiness, and school satisfaction, and its unhealthy subscale is associated with depression, rumination, and stress. Internal consistency in this sample was healthy: alpha=.76 and unhealthy: alpha=.85. A median split was calculated for each of these variables to identify participants scoring either low or high on healthy and low or high on unhealthy music use.

App Use and Quality

Backend data on the date, time, frequency, and length of app use were collected. App engagement was defined as the total number of playlists created per participant. App quality was assessed by the Mobile App Rating Scale-User version (uMARS) [61]. This 20-item scale assesses perceived *objective* app quality on 4 subscales (engagement, functionality, aesthetics, and information) rated on a 5-point scale (1=very poor and 5=excellent). Mean subscale scores and a mean objective quality score were derived. Subjective app quality was assessed using 4 questions: "Would you recommend the app?" (1, not to anyone; to 5, everyone); "How many times would you use it?" (1, 0 times; to 5, >50 times); "Would you pay for this app?" (1, no; 2, maybe; and 3, yes); and overall star-rating (1 to 5 stars).

Procedure

Ethical approval was granted by the relevant university human research ethics committees and the trial was registered with the Australian New Zealand Clinical Trials Registry (ACTRN12615000051549). Informed consent was obtained online before participants completed the baseline online survey. Those meeting study inclusion criteria were automatically identified and randomized via a computerized trial management system to the immediate- or delayed-access groups, with stratification by age group (aged 16-20 years and 21-25 years) and gender.

Those assigned to the immediate-access group were emailed a link to the app. This required users to first download the *TestFlight* app, a beta app distribution platform, which enabled download of the *Music eScape* app before its release in the Apple App Store. Short message service (SMS) text message reminders to access the app were sent at 7-day intervals in the first month.

To minimize attrition, the delayed-access group received 2 SMS text messages during the 1 month wait for access to the app. All baseline and follow-up surveys were completed online. Participants were automatically sent email links to each survey 3 days before, on the day of, and at 3 and 7 days after a follow up was due. Reminder SMS text messages were sent to those who had not completed a follow-up, 8 and 10 days after they were due. Participants were reimbursed Aus $20 for completing each survey.

Statistical Analyses

The immediate- and delayed-access groups were compared on baseline demographic, mental distress and well-being, emotion regulation, and music variables using logistic regressions, with treatment group allocation as the outcome variable. Data screening indicated all outcomes (ie, emotion regulation, mental distress, and well-being) had acceptable skew and kurtosis. Linear mixed models in SPSS version 25 (IBM Corp, Armonk, NY, USA) were used to conduct intent-to-treat analyses, without prediction of missing data, on the primary outcome variable of difficulties in emotion regulation and secondary outcomes of mental distress and well-being. For all outcomes, time and group main effects and time by group interaction from baseline to 1 month were conducted, followed by analyses examining the

impact of the app over time from baseline to the 2-, 3-, and 6-month follow-ups. Gender, baseline duration of music use (hours per week of music listening, with median split into high vs low), and use of music (healthy or unhealthy) were included as control variables and potential moderators of outcomes because of the potential impact of these variables on mood, music, and app use [60]. Two analyses entering app access (yes or no) and app use as additional control variables were also conducted to determine if this varied results. An autoregressive covariance structure (Toeplitz) was specified to account for correlated outcome variables assessed at close time points. Significant effects were probed using pairwise comparisons, and Cohen *d* effect sizes were calculated using SDs pooled across groups and times.

Results

Recruitment and Sample Characteristics

Figure 2 displays the consort diagram. A total of 209 young people responded to recruitment advertisements and completed the online survey. Of those, 80.9% (169/209) met full study inclusion criteria and were allocated to immediate (n=85) or delayed (n=84) app access. Follow-up rates were high (93.5% at 1 month, 87.6% at 2 months, 88.2% at 3 months, and 84.0% at 6 months). There was no significant difference in key demographic factors (eg, age, gender, work status, and education) between those who completed all follow-up surveys and those who missed 1 or more postbaseline assessments.

Demographic characteristics of the sample are displayed in Table 1, and descriptive statistics for the primary and secondary outcome variables are provided in Table 2. There were no significant differences between immediate and delayed groups on any baseline demographic, music, or primary or secondary outcome variables.

Figure 2. Consort diagram.

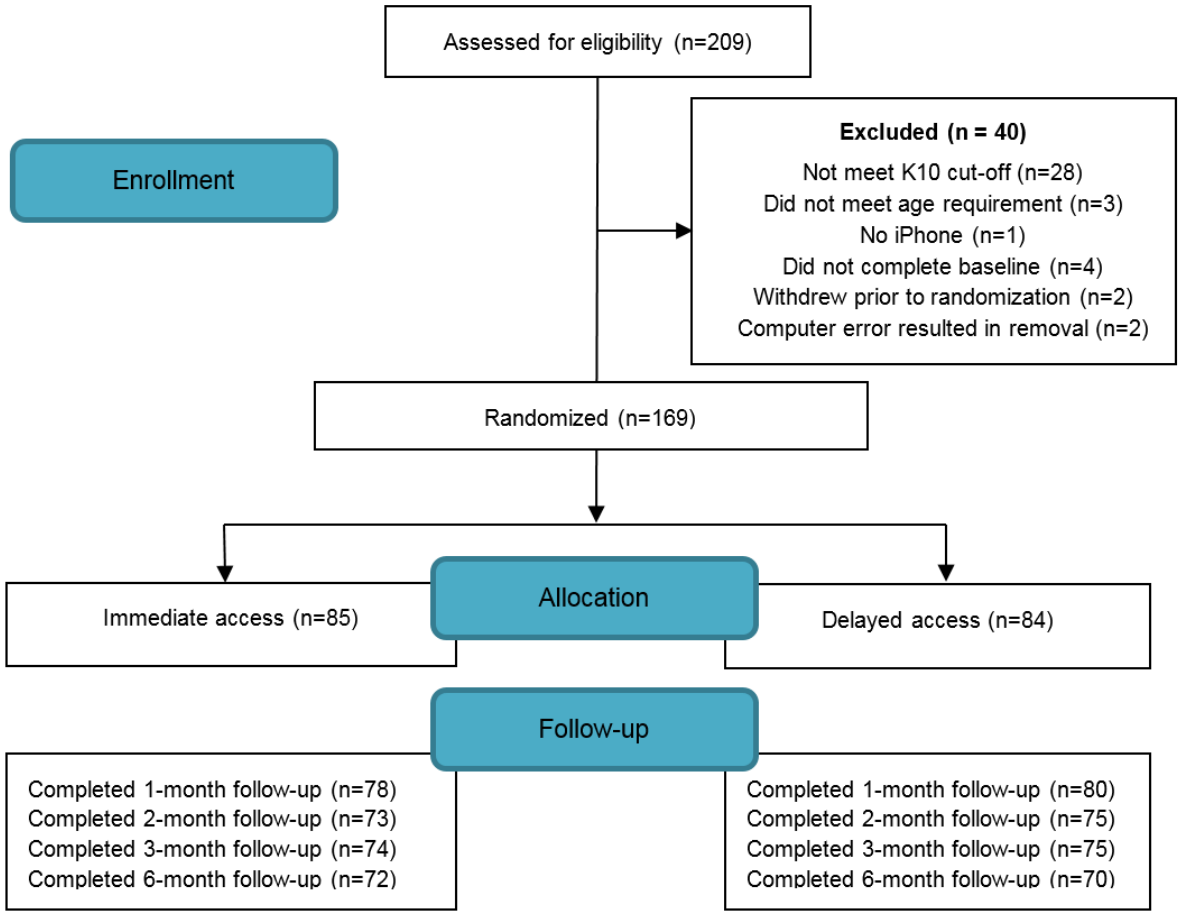


Table 1. Demographic and baseline music characteristics.

Characteristics	Total (N=169)	Immediate (N=85)	Delayed (N=84)	*P* value
Demographics				
Age, mean (SD)	19.9 (2.5)	20.0 (2.6)	19.9 (2.4)	.9
Gender (females), n (%)	134 (79.3)	67 (78.8)	67 (79.8)	.82
English language fluency, n (%)	139 (82.2)	68 (80.0)	71 (84.5)	.44
Aboriginal or Torres Strait Islander, n (%)	2 (1.2)	1 (1.2)	1 (1.2)	.99
Education, n (%)				.29
Degree, postgraduate study	40 (23.7)	17 (19.0)	23 (27.4)	
Certificate or diploma	25 (14.8)	13 (15.3)	12 (14.3)	
High school	104 (61.6)	55 (64.7)	49 (58.4)	
Work status, n (%)				.39
Full-time	16 (9.5)	6 (7.1)	10 (11.9)	
Part-time or casual	43 (25.4)	28 (32.9)	15 (17.9)	
Unemployed or disability allowance	10 (5.9)	4 (4.7)	6 (7.1)	
Home duties	1 (0.6)	0 (0.0)	1 (1.2)	
Volunteer work	2 (1.2)	1 (1.2)	1 (1.2)	
Full-time or part-time student	97 (57.4)	46 (54.1)	51 (60.7)	
Relationship status, n (%)				.06
Married	3 (1.8)	2 (1.8)	1 (1.2)	
In a relationship	85 (50.3)	50 (40.6)	35 (41.7)	
Single	81 (47.9)	33 (38.8)	48 (57.1)	
Current psychological treatment, n (%)	38 (22.5)	17 (20.0)	21 (25.0)	.44
Use smartphone daily, n (%)	155 (91.7)	76 (89.4)	79 (94.0)	.94
Music, n (%)				
Accessed music online in past month	164 (97.0)	82 (96.5)	82 (97.6)	.66
Play musical instrument or sing in choir				
Past	139 (82.2)	71 (83.5)	68 (81.0)	.9
Current	47 (27.8)	24 (28.2)	23 (27.4)	.66
Compose music	52 (30.8)	27 (31.8)	25 (29.8)	.78
Attend concerts or live music at least monthly	50 (29.6)	29 (34.1)	21 (25.0)	.2
Music listening each week, mean (SD)				
Average number of hours per day	2.6 (2.3)	2.8 (2.7)	2.5 (1.9)	.31
Average number of days	6.2 (1.6)	6.3 (1.6)	6.1 (1.5)	.6
Median split (days×hours)—High	98 (46.9)	43 (51.2)	48 (57.1)	.39
HUMS[a]				
High healthy music use	128 (61.2)	61 (71.8)	53 (63.1)	.23
High unhealthy music use	97 (46.4)	41 (48.2)	49 (58.3)	.19
At least moderate success at listening to music to change mood, n (%)	167 (87.9)	75 (88.2)	73 (86.9)	.66

[a]HUMS: Healthy Unhealthy Music scale.

Table 2. Means and SDs for all primary and secondary outcome variables.

Measure	Group	Baseline, mean (SD)	1 month, mean (SD)	2 month, mean (SD)	3 month, mean (SD)	6 month, mean (SD)
DERS[a] aware	Immediate	7.82 (2.66)	8.01 (2.67)	8.10 (2.68)	8.07 (2.74)	7.75 (2.73)
	Delayed	7.49 (2.21)	8.03 (2.65)	8.05 (2.66)	7.77 (2.79)	7.49 (2.73)
	Total	7.66 (2.45)	8.02 (2.65)	8.07 (2.66)	7.92 (2.76)	7.62 (2.72)
Clarity	Immediate	7.51 (2.68)	6.88 (2.40)	6.86 (2.39)	6.80 (2.50)	6.49 (2.68)
	Delayed	7.95 (2.64)	6.98 (2.62)	7.23 (2.64)	6.51 (2.43)	7.20 (2.61)
	Total	7.73 (2.66)	6.93 (2.51)	7.05 (2.52)	6.65 (2.47)	6.84 (2.66)
Goals	Immediate	10.62 (2.99)	9.65 (2.74)	8.84 (3.05)	8.69 (3.23)	8.99 (3.05)
	Delayed	11.17 (2.96)	10.03 (3.06)	9.58 (2.87)	9.49 (3.06)	9.21 (2.67)
	Total	10.89 (2.98)	9.84 (2.91)	9.21 (2.97)	9.09 (3.16)	9.10 (2.86)
Impulse	Immediate	6.69 (3.07)	6.11 (2.69)	5.81 (2.60)	5.53 (2.48)	5.72 (2.41)
	Delayed	7.13 (3.38)	6.36 (2.80)	6.01 (2.68)	6.03 (2.82)	6.10 (2.51)
	Total	6.91 (3.23)	6.24 (2.74)	5.91 (2.64)	5.78 (2.66)	5.91 (2.46)
Nonacceptance	Immediate	8.24 (3.19)	7.32 (2.83)	7.11 (2.48)	7.05 (2.74)	7.18 (3.01
	Delayed	8.35 (3.24)	7.36 (2.83)	7.66 (3.33)	7.56 (3.01)	7.60 (3.08)
	Total	8.29 (3.21)	7.34 (2.82)	7.39 (2.94)	7.31 (2.88)	7.39 (3.04)
Strategies	Immediate	7.47 (3.28)	7.06 (2.84)	6.33 (2.63)	6.36 (2.71)	6.64 (2.71)
	Delayed	7.89 (3.40)	7.00 (2.89)	6.89 (2.77)	6.75 (3.18)	6.33 (2.67)
	Total	7.68 (3.34)	7.03 (2.86)	7.92 (2.71)	6.55 (2.95)	6.49 (2.69)
Kessler 10	Immediate	27.52 (6.91)	23.00 (6.47)	22.51 (6.92)	22.83 (7.55)	22.68 (7.97)
	Delayed	28.33 (6.59)	24.19 (6.94)	22.41 (6.50)	21.96 (6.93)	22.46 (7.55)
	Total	27.92 (6.74)	23.60 (6.72)	22.36 (6.69)	22.39 (7.24)	22.57 (7.74)
MHC-SF[b]	Immediate	52.53 (12.72)	55.85 (12.76)	55.25 (13.81)	57.61 (14.80)	59.88 (13.87)
	Delayed	50.64 (14.88)	52.34 (14.89)	53.08 (15.02)	54.97 (15.32)	54.14 (14.62)
	Total	51.59 (13.82)	54.07 (13.94)	54.16 (14.42)	56.29 (15.07)	57.05 (14.49)

[a]DERS: Difficulties in Emotion Regulation scale.

[b]MHC-SF: Mental Health Continuum-Short Form.

App Use and Quality

Backend data indicated that 12 participants did not download the app, a further 34 downloaded but never used the app, 31 downloaded it but experienced technical flaws, and 7 were allocated to the immediate condition but did not download the app until a month after allocation (included in this group for intent-to-treat purposes). Of those who downloaded the app, the total number of generated playlists ranged from 0 to 71, (median=2). No playlists were generated by 41%, and only 7.5% of the sample generated more than 15 playlists. The number of generated playlists did not vary significantly between immediate- and delayed-access groups or by gender. The duration of app music use variable was considered unreliable as it was not possible to gauge the extent to which participants were listening to the music (vs leaving the app open with music playing).

On the uMARS, the app had a high level of objective app quality ($mean_{overall}$ = 3.8 [SD 0.50]), with good engagement (mean 3.67 [SD 0.61]), aesthetics (mean 4.10 [SD 0.63]), and information (mean 4.05 [SD 0.61]), and acceptable functionality (mean 3.47 [SD 0.66]). Participants reported they would use the app between 10 and 50 times (mean 4.09 [SD 1.04]), and although they were unlikely to pay for the app (mean 2.43 [SD 1.23]), they gave it a 3.6 out of 5-star rating (SD 0.65).

Emotion Regulation Outcomes

The linear mixed model revealed no time by group interaction for any of the 6 difficulties in emotion regulation subscales of the DERS (see Table 3). Time effects were found on 5 of the 6 DERS subscales (clarity, goals, nonacceptance, strategies, and impulse) when comparing baseline both with the 1-month follow-up and with the 2-, 3-, and 6-month follow-ups (see Table 3). These effects did not vary when controlling for whether participants used the app (yes or no) or the level of app use.

To better understand these changes over time, moderating effects of gender, duration of music use, and healthy or unhealthy music use were assessed across all time points. For difficulties engaging in goal-directed behavior when distressed

(DERS-Goal) and nonacceptance of emotional responses (DERS-nonacceptance), no significant moderating effects were found for gender ($F_{4,328}$=2.12; P=.07; $F_{4,351}$=1.16; P=.33), duration of music use ($F_{4,361}$=1.00; P=.40; $F_{4,366}$=0.32; P=.87), unhealthy use of music ($F_{4,369}$=0.53; P=.72; $F_{4,373}$=1.98; P=.09), or healthy use of music ($F_{4,363}$=1.63; P=.17; $F_{4,355}$=0.20; P=.94). When exploring the time main effects, difficulties engaging in goal-directed behavior decreased from baseline to 1 month ($mean_{difference}$=−1.04, 95% CI −1.44 to −0.65; t_{383}=5.24; P<.001; d=0.36) and from 1 to 2 months ($mean_{difference}$=−0.59, 95% CI −1.03 to −0.14; t_{383}=2.59; P=.01; d=0.22), before maintaining stability at 3 and 6 months (P=.72; P=.80; see Figure 3). Nonacceptance of emotional responses decreased from baseline to 1 month ($mean_{difference}$=−0.59, 95% CI −1.41 to −0.55; t_{376}=4.49; P<.001; d=0.32) and maintained stability thereafter (P=.46; P=.81; P=.82).

Table 3. Emotion regulation, mental distress, and well-being outcomes.

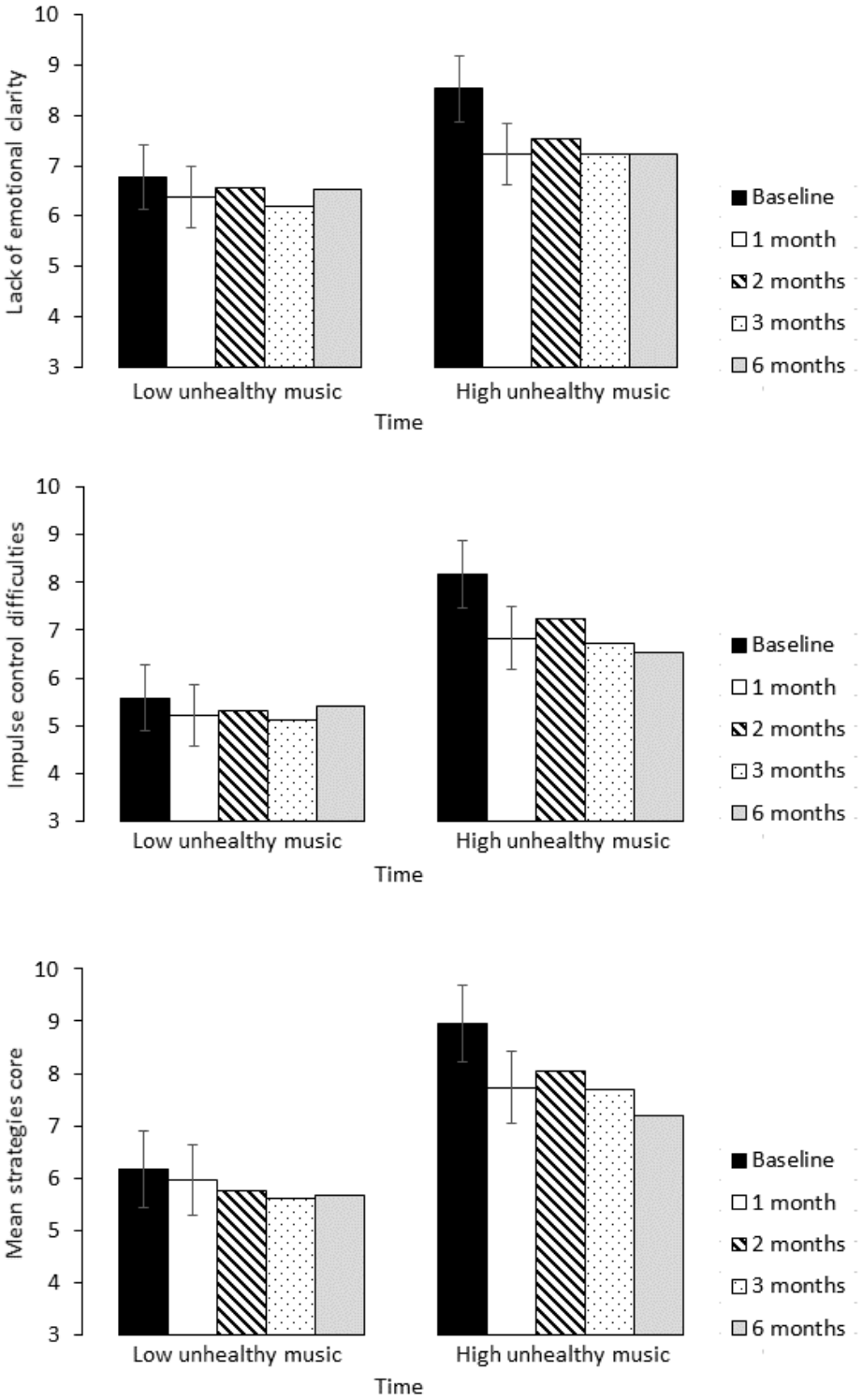
Measure	Comparison	Main effects	*F* test *(df)*	*P* value
DERS[a] awareness	Baseline vs 1 month	Time	3.62 (1,157)	.06
		Time×group	0.66 (1,157)	.42
	Baseline vs 2, 3, and 6 months	Time	1.52 (3,252)	.22
		Time×group	0.18 (3,252)	.91
Clarity	Baseline vs 1 month	Time	19.70 (1,158)	<.001
		Time×group	1.34 (1,158)	.25
	Baseline vs 2, 3, and 6 months	Time	11.66 (3,300)	<.001
		Time×group	1.67 (3,300)	.17
Goals	Baseline vs 1 month	Time	25.73 (1,157)	<.001
		Time×group	0.46 (1,157)	.5
	Baseline vs 2, 3, and 6 months	Time	31.76 (3,325)	<.001
		Time×group	1.21 (3,325)	.31
Impulse	Baseline vs 1 month	Time	9.90 (1,154)	.002
		Time×group	0.73 (1,154)	.39
	Baseline vs 2, 3, and 6 months	Time	10.56 (3,307)	<.001
		Time×group	0.20 (3,307)	.89
Nonacceptance	Baseline vs 1 month	Time	19.99 (1,156)	<.001
		Time×group	0.00 (1,156)	.96
	Baseline vs 2, 3, and 6 months	Time	8.18 (3,366)	<.001
		Time×group	0.67 (3,366)	.57
Strategies	Baseline vs 1 month	Time	11.50 (1,155)	<.001
		Time×group	1.85 (1,155)	.18
	Baseline vs 2, 3, and 6 months	Time	12.53 (3,297)	<.001
		Time×group	1.92 (3,297)	.13
K10[b]	Baseline vs 1 month	Time	74.77 (1,159)	<.001
		Time×group	0.02 (1,159)	.9
	Baseline vs 2, 3, and 6 months	Time	48.11 (3,248)	<.001
		Time×group	0.77 (3,248)	.51
MHC-SF[c]	Baseline vs 1 month	Time	9.60 (3,155)	.002
		Time×group	0.70 (1,155)	.4
	Baseline vs 2, 3, and 6 months	Time	12.75 (3,260)	<.001
		Time×group	1.43 (3,260)	.24

[a]DERS: Difficulties in Emotion Regulation Scale.

[b]K10: Kessler 10.

[c]MHC-SF: Mental Health Continuum-Short Form.

Figure 3. Adjusted mean emotion regulation (difficulties in emotional regulation scale) scores and unhealthy use of music.



Moderating effects were found for unhealthy use of music, for lack of emotional clarity ($F_{4,394}$=2.43; P=.05; Figure 3), impulse control difficulties ($F_{4,316}$=3.18; P=.01; Figure 3), and limited access to emotion regulation strategies ($F_{4,443}$=2.78; P=.03; Figure 3), but not for healthy use of music ($F_{4,387}$=2.06; P=.09; $F_{4,310}$=2.36; P=.05; $F_{4,444}$=0.35; P=.84). No moderating effects were found for gender ($F_{4,158}$=0.02; P=.98; $F_{4,295}$=1.89; P=.11; $F_{4,425}$=1.51; P=.20) or duration of music use ($F_{4,387}$=1.86; P=.12; $F_{4,315}$=0.41; P=.80; $F_{4,441}$=0.52; P=.72).

Post hoc pairwise comparisons with a Bonferroni adjusted *P* value <.01 revealed an adjusted mean decrease in difficulties with emotional clarity from baseline to 1 month follow-up for those who reported high use of unhealthy music ($mean_{difference}$=−1.29, 95% CI −1.83 to −0.75; t_{387}=4.72; *P*<.001; *d*=0.43) but not low ($mean_{difference}$=−0.40, 95% CI −0.98 to 0.18; t_{387}=1.35; *P*=.18; *d*=0.16; Figure 3). Similar results were found for impulse control difficulties (high unhealthy music: $mean_{difference}$=−1.34, 95% CI −1.94 to −0.75; t_{310}=4.45; *P*<.001; *d*=0.34 and low unhealthy music: $mean_{difference}$=−0.09, 95% CI −0.81 to 0.63; t_{310}=0.23; *P*=.81; *d*=0.07; Figure 3) and for limited access to emotion regulation strategies (high unhealthy music: $mean_{difference}$=−1.23, 95% CI −1.84 to −0.63; t_{443}=4.01; *P*<.001; *d*=0.35 and low unhealthy music: $mean_{difference}$=0.20, 95% CI −0.52 to 0.91; t_{443}=0.54; *P*=.59; *d*=0.03; Figure 3). For all 3 emotion regulation variables, reductions were stable at all subsequent time points (Figure 3).

Mental Distress and Well-Being Outcomes

The linear mixed model revealed time main effects but no time by group interaction for mental distress (K10) or well-being (MHC-SF; Table 3). These effects did not vary when controlling for app access (yes or no) or use.

To better understand these changes over time, moderating effects of gender, duration of music use, and unhealthy or healthy music use were assessed across all time points. For mental distress, moderating effects were found for gender ($F_{4,254}$=3.09; *P*=.02) but not for duration of music use ($F_{4,272}$=0.74; *P*=.57) or for healthy ($F_{4,266}$=1.70; *P*=.15) or unhealthy ($F_{4,272}$=0.77; *P*=.55) use of music. Post hoc pairwise comparisons, with a Bonferroni adjusted *P* value of .003, revealed an adjusted mean decrease in mental distress from baseline to 1 month for females ($mean_{difference}$=−4.50, 95% CI −5.66 to −3.34; t_{272}=7.68; *P*<.001; *d*=0.37) but not for males ($mean_{difference}$=−0.12, 95% CI −2.07 to 2.31; t_{272}=0.11; *P*=.92; *d*=0.71).

For well-being, no significant moderating effects were found for gender ($F_{4,283}$=1.55; *P*=.19), duration of music use ($F_{4,308}$=2.41; *P*=.05), unhealthy use of music ($F_{4,310}$=0.13; *P*=.97), or healthy use of music ($F_{4,301}$=0.42; *P*=.79). When exploring the time main effects, there was no change in well-being scores when comparing baseline with 1-month (*P*=.10) or 2-month assessments (*P*=.40). However, there was a significant increase in well-being from baseline to 3 months ($mean_{difference}$=3.09, 95% CI 0.88-5.29; t_{278}=2.76; *P*=.006; *d*=0.33), which was then maintained at the 6 months (3 vs 6 month: *P*=.58).

Discussion

Principal Findings

This study examined the 1-month efficacy and 2-, 3-, and 6-month outcomes of the *Music eScape* app in 169 young people with at least mild levels of mental distress. The trial found no differential improvements from app access at 1 month in emotion regulation, mental distress, or well-being. Nevertheless, improvements on 5 out of the 6 emotion regulation strategies, mental distress, and well-being were evident in both groups over the 6-month trial.

The lack of significant differences between the immediate versus 1-month delayed-access groups indicates that the *Music eScape* app was ineffective in achieving change in emotion regulation, mental distress, or well-being, beyond the impact of research assessments alone. Gender, duration of music use, unhealthy and healthy music use, and app use did not impact these results. Although the use of a 1-month delayed-access control may have limited our ability to find effects, waitlist control conditions are commonly used in mobile health (mHealth) research. For example, 2 recent meta-analyses of smartphone mental health apps reporting 8 out of 18 studies on depression and 4 out of 9 studies on anxiety used waitlist controls [62,63]. The duration of the delay ranged from 4 to 16 weeks, depending on the length of the mHealth intervention [62,63]. Although the meta-analyses found mHealth apps had small to moderate effects on both depression and anxiety outcomes, 2 out of the 3 included studies that used a 1-month waitlist control found no effects [57,58]. Thus, the 1-month delay used in this trial might have been insufficient for participants to receive an adequate dose of the *Music eScape* app. Baseline data also indicated that participants had high levels of music use (2.6 hours per day) and emotional awareness (DERS subscale), and 88% participants reported at least moderate levels of success using music to change their mood, suggesting a ceiling effect may have been present on these variables. Nevertheless, improvements in emotion regulation were found on the 5 DERS subscales across the whole sample, suggesting that this study had the ability to detect a change in emotion regulation across time.

Both groups had access to the *Music eScape* app after the first month. Although improvements in mental distress, well-being, and emotion regulation were found over the 6 months, it is not possible to attribute these results to the app. These improvements may have been because of regression to the mean or assessment effects, particularly given that the 5 assessments were completed over a 6-month period. The amount of app use did not affect any outcomes.

To better understand the changes in emotion regulation strategies, over time, moderating effects of gender, duration of music use, and healthy or unhealthy music use were explored. Results indicated that improvements in emotional clarity, impulse control, and limited access to emotion regulation strategies were only found in distressed young people who engaged in high levels of unhealthy music use at baseline. This finding highlights the potential importance of targeting unhealthy music use to improve the emotion regulation skills of distressed young people. Currently, the app allows users to maintain or intensify their current mood by choosing mood-congruent music. The moderation of outcomes by the degree of unhealthy use of music suggests that some may have used the app to stay in a negative mood, consistent with observations in previous research [30-32]. According to meta-emotion theories, some people are drawn toward emotional experiences, whereas others are motivated to avoid and control emotional experiences. Bartsch et al [44] draw links between these individual meta-emotional tendencies and selective of

media use such as watching melodramatic or horror movies (which some people enjoy, whereas others prefer to avoid). Similarly, some people enjoy listening to sad and angry music [64], and there is some evidence that this music-induced emotional exposure is related to better emotional processing and well-being [65-67]. Other research suggests that immersion in sad and angry music might be unhealthy for some young people who are prone to depression or other mental health problems [37]. Further research is required to determine whether the *Music eScape* app can improve emotion regulation skills through reductions in unhealthy music use, particularly if coaching is provided on its use for that purpose. Prospective research is also required to determine whether unhealthy music use is a correlate or risk factor for depression in young people and whether it moderates emotion regulation skills in young people without depression.

No moderators of well-being outcomes across the 6 months were found, and only female gender, but not the amount or type (healthy or unhealthy) of music use, moderated improvements in mental distress. The reduction in distress for female but not male participants may be partly because of increased power in detecting changes for females, given that 79% of the sample was female. However, no gender differences in app use were found, and gender has not been found to moderate depression or anxiety outcomes in systematic reviews of psychological treatment trials [68,69]. Further research is required to determine if gender moderates eHealth treatment outcomes.

On average, young people gave *Music eScape* a 4 out of 5 rating for overall objective app quality and for information and aesthetics scales and 3 out of 5 for engagement and functionality. These uMARS scores were higher than recent MARS expert ratings on 50 other eHealth apps [61]. Young people reported they would use the app 10 to 50 times and generated a median of 2 playlists, which is sufficient for users to learn how to identify and manage their mood using music. However, 7.1% (12/169) participants did not download the app, 20.1% (34/169) participants downloaded but did not use the app, and a further 18.3% (31/169) participants experienced technical difficulties using the app. Although the level of uptake or usage of music intervention apps is unknown, 2018 Localytics data on 37,000 apps indicate 21% apps are used only once, with use of 62% of apps discontinued before the eleventh use [70]. In comparison, the current data suggest a high level of maintained use.

Strengths and Limitations

A large community sample of 169 young people with at least mild distress (K10 scores of >17) participated in this trial. However, the volunteer sampling method used to recruit participants limits the generalizability of results. Despite efforts to avoid participants with an affinity for music during recruitment, the sample used music for substantial average durations at baseline (2+ hours per 6 days a week). Many also used music for emotion regulation strategy, which may have created ceiling effects on key outcome variables. We were also only able to use the number of app playlists generated as an indicator of app use rather than the duration of app music use.

Strengths of this study include its high participation (80%), app uptake (91.5%), and retention rates in research follow-ups (87%-96%) and the inclusion of a range of potential control variables and moderators of outcomes. However, there might have been other unmeasured moderators that future research may identify. Assessment reactivity could be reduced by minimizing the length of research assessments and masking participants from the research hypotheses (eg, comparing music apps with other health apps that collect information on emotions) and assessment process (eg, including questions on emotions as part of a general health survey).

The music available through the current version of *Music eScape* is limited to the users' own music library. However, 97% of the sample reported accessing music online in the past week. Future versions of the app, which interface with music streaming services that give users access to a much wider repertoire, may enhance its effects by giving users a choice of preferred music for each mood journey step or ensuring that different (or more current) music is offered each time they use the app.

Further testing is required to demonstrate whether the app has effects on emotion regulation, mental health, and well-being over a longer delayed-access period, and if so, whether it has superior effects compared with placebo control apps or other emotion regulation apps or interventions. Additional benefits from adding the app to other interventions for emotion regulation in young people could also be tested.

Conclusions

Mobile app use is increasingly prevalent worldwide, particularly among young people who are also the biggest consumers of music. The *Music eScape* app is freely available to young people, parents, and practitioners via the Google Play store. Unfortunately, since this study has completed, iOS updates have led to the app being currently unavailable via the Apple App Store. Although further testing is required to demonstrate efficacy, the results of this study highlight the potential of music intervention apps such as *Music eScape* to deliver engaging and highly accessible emotion regulation skills training to young people in real time in their natural environment, which in principle could result in population-wide benefits in mental distress and well-being.

Acknowledgments

The authors acknowledge Jessica Lewis for her assistance with data collection. Professor Leanne Hides was supported by an Australian Research Council Future Fellowship (FT120100780; 2012-2016) and is currently supported by a National Health and Medical Research Centre Senior Research Fellowship (APP1119098; 2017-2021). The Young and Well Cooperative Research Centre (Young and Well CRC; 2011-2016) funded this project. The Young and Well CRC was an Australian-based international research center established under the Australian Government's Cooperative Research Centres Program. These funders had no

role in the study design; collection, analysis, or interpretation of data; writing the manuscript; and the decision to submit the manuscript for publication.

Conflicts of Interest

None declared.

Multimedia Appendix 1

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - mhealth_v7i1e11482_app1.pdf]

References

1. Institute for Health Metrics and Evaluation. Institute for Health Metrics and Evaluation. Global Burden of Disease Compare Data Visualization URL: http://www.healthdata.org/data-visualization/gbd-compare [accessed 2018-09-19] [WebCite Cache ID 72XPBEdR0]
2. Aldao A, Nolen-Hoeksema S, Schweizer S. Emotion-regulation strategies across psychopathology: a meta-analytic review. Clin Psychol Rev 2010 Mar;30(2):217-237. [doi: 10.1016/j.cpr.2009.11.004] [Medline: 20015584]
3. Berking M, Wupperman P. Emotion regulation and mental health: recent findings, current challenges, and future directions. Curr Opin Psychiatry 2012 Mar;25(2):128-134. [doi: 10.1097/YCO.0b013e3283503669] [Medline: 22262030]
4. Mennin DS, McLaughlin KA, Flanagan TJ. Emotion regulation deficits in generalized anxiety disorder, social anxiety disorder, and their co-occurrence. J Anxiety Disord 2009 Oct;23(7):866-871 [FREE Full text] [doi: 10.1016/j.janxdis.2009.04.006] [Medline: 19464142]
5. Dingle GA, Neves DD, Alhadad SS, Hides L. Individual and interpersonal emotion regulation among adults with substance use disorders and matched controls. Br J Clin Psychol 2018 Jun;57(2):186-202. [doi: 10.1111/bjc.12168] [Medline: 29277906]
6. Gross JJ. Emotion regulation: affective, cognitive, and social consequences. Psychophysiology 2002 May;39(3):281-291. [Medline: 12212647]
7. Webb TL, Miles E, Sheeran P. Dealing with feeling: a meta-analysis of the effectiveness of strategies derived from the process model of emotion regulation. Psychol Bull 2012 Jul;138(4):775-808. [doi: 10.1037/a0027600] [Medline: 22582737]
8. Casey B, Caudle K. The teenage brain: self control. Curr Dir Psychol Sci 2013 Apr 01;22(2):82-87 [FREE Full text] [doi: 10.1177/0963721413480170] [Medline: 25284961]
9. Beauchaine TP. Future directions in emotion dysregulation and youth psychopathology. J Clin Child Adolesc Psychol 2015;44(5):875-896. [doi: 10.1080/15374416.2015.1038827] [Medline: 26016727]
10. Hannesdottir DK, Ollendick TH. The role of emotion regulation in the treatment of child anxiety disorders. Clin Child Fam Psychol Rev 2007 Sep;10(3):275-293. [doi: 10.1007/s10567-007-0024-6] [Medline: 17705098]
11. Australian Bureau of Statistics (ABS). Australian Bureau of Statistics. Canberra: Australian Bureau of Statistics; 2007. National Survey of Mental Health and Wellbeing: Summary of Results URL: http://www.abs.gov.au/ausstats/abs@.nsf/mf/4326.0 [accessed 2018-11-20] [WebCite Cache ID 744LIJg1d]
12. March JS, Silva S, Petrycki S, Curry J, Wells K, Fairbank J, et al. The Treatment for Adolescents With Depression Study (TADS): long-term effectiveness and safety outcomes. Arch Gen Psychiatry 2007 Oct;64(10):1132-1143. [doi: 10.1001/archpsyc.64.10.1132] [Medline: 17909125]
13. March J, Silva S, Petrycki S, Curry J, Wells K, Fairbank J, Treatment for Adolescents With Depression Study (TADS) Team. Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: Treatment for Adolescents With Depression Study (TADS) randomized controlled trial. J Am Med Assoc 2004 Aug 18;292(7):807-820. [doi: 10.1001/jama.292.7.807] [Medline: 15315995]
14. Seligman LD, Ollendick TH. Cognitive-behavioral therapy for anxiety disorders in youth. Child Adolesc Psychiatr Clin N Am 2011 Apr;20(2):217-238 [FREE Full text] [doi: 10.1016/j.chc.2011.01.003] [Medline: 21440852]
15. James AC, James G, Cowdrey FA, Soler A, Choke A. Cognitive behavioural therapy for anxiety disorders in children and adolescents. Cochrane Database Syst Rev 2013;6:CD004690. [doi: 10.1002/14651858.CD004690.pub3] [Medline: 23733328]
16. Goodyer IM, Dubicka B, Wilkinson P, Kelvin R, Roberts C, Byford S, et al. A randomised controlled trial of cognitive behaviour therapy in adolescents with major depression treated by selective serotonin reuptake inhibitors. The ADAPT trial. Health Technol Assess 2008 May;12(14):iii-iv, ix [FREE Full text] [Medline: 18462573]
17. Vitiello B, Emslie G, Clarke G, Wagner KD, Asarnow JR, Keller MB, et al. Long-term outcome of adolescent depression initially resistant to selective serotonin reuptake inhibitor treatment: a follow-up study of the TORDIA sample. J Clin Psychiatry 2011 Mar;72(3):388-396 [FREE Full text] [doi: 10.4088/JCP.09m05885blu] [Medline: 21208583]
18. Curry J, Silva S, Rohde P, Ginsburg G, Kratochvil C, Simons A, et al. Recovery and recurrence following treatment for adolescent major depression. Arch Gen Psychiatry 2011 Mar;68(3):263-269 [FREE Full text] [doi: 10.1001/archgenpsychiatry.2010.150] [Medline: 21041606]

19. Lonsdale AJ, North AC. Why do we listen to music? A uses and gratifications analysis. Br J Psychol 2011 Feb;102(1):108-134. [doi: 10.1348/000712610X506831] [Medline: 21241288]
20. Juslin PN, Sloboda J, editors. Handbook of music and emotion. Theory, research, applications. New York: Oxford University Press; 2011.
21. Koelsch S. A neuroscientific perspective on music therapy. Ann N Y Acad Sci 2009 Jul;1169:374-384. [doi: 10.1111/j.1749-6632.2009.04592.x] [Medline: 19673812]
22. Koelsch S. Towards a neural basis of processing musical semantics. Phys Life Rev 2011 Jun;8(2):89-105. [doi: 10.1016/j.plrev.2011.04.004] [Medline: 21601541]
23. Koelsch S. Brain and Music. Chichester, West Sussex, UK: John Wiley & Sons Inc; 2012.
24. Levitin DJ, Tirovolas AK. Current advances in the cognitive neuroscience of music. Ann N Y Acad Sci 2009 Mar;1156:211-231. [doi: 10.1111/j.1749-6632.2009.04417.x] [Medline: 19338510]
25. MacDonald R, Kreutz G, Mitchell L, editors. Music, health, and wellbeing. New York: Oxford University Press; 2012.
26. Wheeler BL, Sokhadze E, Baruth J, Behrens GA, Quinn CF. Musically induced emotions: subjective measures of arousal and valence. Music and Medicine 2011 May 13;3(4):224-233. [doi: 10.1177/1943862111407513]
27. Sharman L, Dingle GA. Extreme metal music and anger processing. Front Hum Neurosci 2015;9:272 [FREE Full text] [doi: 10.3389/fnhum.2015.00272] [Medline: 26052277]
28. Australian Psychological Society. Australian Psychological Society. Stress & wellbeing: How Australians are coping with life URL: https://www.headsup.org.au/docs/default-source/default-document-library/stress-and-wellbeing-in-australia-report.pdf?sfvrsn=7f08274d_4 [accessed 2018-11-20] [WebCite Cache ID 744Lqpogm]
29. Thoma MV, Scholz U, Ehlert U, Nater UM. Listening to music and physiological and psychological functioning: the mediating role of emotion regulation and stress reactivity. Psychol Health 2012;27(2):227-241. [doi: 10.1080/08870446.2011.575225] [Medline: 21678187]
30. van Goethem A, Sloboda J. The functions of music for affect regulation. Musicae Scientiae 2011 Jul 13;15(2):208-228. [doi: 10.1177/1029864911401174]
31. Papinczak ZE, Dingle GA, Stoyanov SR, Hides L, Zelenko O. Young people's uses of music for well-being. J Youth Stud 2015 Mar 13;18(9):1119-1134. [doi: 10.1080/13676261.2015.1020935]
32. Baltazar M, Saarikallio S. Toward a better understanding and conceptualization of affect self-regulation through music: a critical, integrative literature review. Psychol Music 2016 Sep 21;44(6):1500-1521. [doi: 10.1177/0305735616663313]
33. McFerran KS, Garrido S, O'Grady L, Grocke D, Sawyer SM. Examining the relationship between self-reported mood management and music preferences of Australian teenagers. Nord J Music Ther 2014 Jun 09;24(3):187-203. [doi: 10.1080/08098131.2014.908942]
34. Garrido S, Schubert E. Moody melodies: Do they cheer us up? A study of the effect of sad music on mood. Psychol Music 2013 Oct 15;43(2):244-261. [doi: 10.1177/0305735613501938]
35. Larwood J, Dingle G. PsyArXiv Preprints. 2018. I get knocked down but I get up again: The effects of listening to music when sad in young adults URL: https://psyarxiv.com/hkqcs/ [accessed 2018-11-16] [WebCite Cache ID 73z3ZIgrm]
36. Chin T, Rickard NS. Emotion regulation strategy mediates both positive and negative relationships between music uses and well-being. Psychol Music 2013 Aug 14;42(5):692-713. [doi: 10.1177/0305735613489916]
37. McFerran KS, Garrido S, Saarikallio S. A critical interpretive synthesis of the literature linking music and adolescent mental health. Youth Soc 2013 Dec 02;48(4):521-538. [doi: 10.1177/0044118X13501343]
38. Hilliard RE. The use of cognitive-behavioral music therapy in the treatment of women with eating disorders. Music Ther Perspect 2001 Jan 01;19(2):109-113. [doi: 10.1093/mtp/19.2.109]
39. Carr C, d'Ardenne P, Sloboda A, Scott C, Wang D, Priebe S. Group music therapy for patients with persistent post-traumatic stress disorder--an exploratory randomized controlled trial with mixed methods evaluation. Psychol Psychother 2012 Jun;85(2):179-202. [doi: 10.1111/j.2044-8341.2011.02026.x] [Medline: 22903909]
40. Dingle GA, Carter NA. Smoke into Sound: a pilot randomised controlled trial of a music cravings management program for chronic smokers attempting to quit. Musicae Scientiae 2017 May 23;21(2):151-177. [doi: 10.1177/1029864916682822]
41. Dingle GA, Williams E, Jetten J, Welch J. Choir singing and creative writing enhance emotion regulation in adults with chronic mental health conditions. Br J Clin Psychol 2017 Nov;56(4):443-457. [doi: 10.1111/bjc.12149] [Medline: 28722166]
42. Dingle GA, Hodges J, Kunde A. Tuned in emotion regulation program using music listening: effectiveness for adolescents in educational settings. Front Psychol 2016;7:859 [FREE Full text] [doi: 10.3389/fpsyg.2016.00859] [Medline: 27375537]
43. Dingle GA, Fay C. Tuned In: the effectiveness for young adults of a group emotion regulation program using music listening. Psychol Music 2016 Oct 19;45(4):513-529. [doi: 10.1177/0305735616668586]
44. Villani D, Carissoli C, Triberti S, Marchetti A, Gilli G, Riva G. Videogames for emotion regulation: a systematic review. Games Health J 2018 Apr;7(2):85-99. [doi: 10.1089/g4h.2017.0108] [Medline: 29424555]
45. Grant S, Pedersen ER, Osilla KC, Kulesza M, D'Amico EJ. Reviewing and interpreting the effects of brief alcohol interventions: comment on a Cochrane review about motivational interviewing for young adults. Addiction 2016 Sep;111(9):1521-1527 [FREE Full text] [doi: 10.1111/add.13136] [Medline: 26508301]
46. Bull M. No dead air! The iPod and the culture of mobile listening. Leis Stud 2005 Jan;24(4):343-355. [doi: 10.1080/0261436052000330447]

47. Fedele DA, Cushing CC, Fritz A, Amaro CM, Ortega A. Mobile health interventions for improving health outcomes in youth: a meta-analysis. JAMA Pediatr 2017 May 01;171(5):461-469. [doi: 10.1001/jamapediatrics.2017.0042] [Medline: 28319239]
48. Spinuzzi C. The methodology of participatory design. Technical Communication 2005;52(2):163-174.
49. Cooperman NA, Richter KP, Bernstein SL, Steinberg ML, Williams JM. Determining smoking cessation related information, motivation and behavioral sills among opiate dependent smokers in methadone treatment. Subst Use Misuse 2015 Apr;50(5):566-581 [FREE Full text] [doi: 10.3109/10826084.2014.991405] [Medline: 25559697]
50. Ingersoll KS, Farrell-Carnahan L, Cohen-Filipic J, Heckman CJ, Ceperich SD, Hettema J, et al. A pilot randomized clinical trial of two medication adherence and drug use interventions for HIV+ crack cocaine users. Drug Alcohol Depend 2011 Jul 01;116(1-3):177-187 [FREE Full text] [doi: 10.1016/j.drugalcdep.2010.12.016] [Medline: 21306837]
51. The Echo Nest. The Echo Nest - Music Data Processing URL: http://the.echonest.com/solutions/musicdata/ [accessed 2018-11-15] [WebCite Cache ID 73xdoF0PG]
52. Russell JA. A circumplex model of affect. J Pers Soc Psychol 1980;39(6):1161-1178. [doi: 10.1037/h0077714]
53. Gratz KL, Roemer L. Multidimensional assessment of emotion regulation and dysregulation: development, factor structure, and initial validation of the difficulties in emotion regulation scale. J Psychopathol Behav Assess 2004 Mar;26(1):41-54. [doi: 10.1023/B:JOBA.0000007455.08539.94]
54. Kaufman EA, Xia M, Fosco G, Yaptangco M, Skidmore CR, Crowell SE. The difficulties in emotion regulation scale short form (DERS-SF): validation and replication in adolescent and adult samples. J Psychopathol Behav Assess 2015 Nov 23;38(3):443-455. [doi: 10.1007/s10862-015-9529-3]
55. Thayer RE, Newman JR, McClain TM. Self-regulation of mood: strategies for changing a bad mood, raising energy, and reducing tension. J Pers Soc Psychol 1994 Nov;67(5):910-925. [Medline: 7983582]
56. Kessler RC, Andrews G, Colpe LJ, Hiripi E, Mroczek DK, Normand SL, et al. Short screening scales to monitor population prevalences and trends in non-specific psychological distress. Psychol Med 2002 Aug;32(6):959-976. [Medline: 12214795]
57. Slade T, Grove R, Burgess P. Kessler Psychological Distress Scale: normative data from the 2007 Australian National Survey of Mental Health and Wellbeing. Aust N Z J Psychiatry 2011 Apr;45(4):308-316. [doi: 10.3109/00048674.2010.543653] [Medline: 21332432]
58. Keyes CL. The mental health continuum: from languishing to flourishing in life. J Health Soc Behav 2002 Jun;43(2):207-222. [Medline: 12096700]
59. Keyes C. The subjective well-being of America's youth: toward a comprehensive assessment. Adolesc Fam Health 2006;4(1):3-11.
60. Saarikallio S, Gold C, McFerran K. Development and validation of the Healthy-Unhealthy Music Scale. Child Adolesc Ment Health 2015 Nov;20(4):210-217 [FREE Full text] [doi: 10.1111/camh.12109] [Medline: 26726295]
61. 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;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
62. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
63. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]
64. Dingle G, Sharman L, Larwood J. Young people's uses of music for emotional immersion. In: McFerran K, Derrington P, Saarikallio S, editors. Handbook of Music and Adolescence. Oxford: Oxford University Press; 2018.
65. Sharman L, Dingle GA. Extreme metal music and anger processing. Front Hum Neurosci 2015;9:272 [FREE Full text] [doi: 10.3389/fnhum.2015.00272] [Medline: 26052277]
66. Van den Tol AJ, Edwards J. Listening to sad music in adverse situations: how music selection strategies relate to self-regulatory goals, listening effects, and mood enhancement. Psychol Music 2014 Jan 29;43(4):473-494. [doi: 10.1177/0305735613517410]
67. Baker FA, Gleadhill LM, Dingle GA. Music therapy and emotional exploration: exposing substance abuse clients to the experiences of non-drug-induced emotions. Arts Psychother 2007 Jan;34(4):321-330. [doi: 10.1016/j.aip.2007.04.005]
68. Cuijpers P, Weitz E, Twisk J, Kuehner C, Cristea I, David D, et al. Gender as predictor and moderator of outcome in cognitive behavior therapy and pharmacotherapy for adult depression: an “individual patient data” meta-analysis. Depress Anxiety 2014 Nov;31(11):941-951. [doi: 10.1002/da.22328] [Medline: 25407584]
69. Nilsen TS, Eisemann M, Kvernmo S. Predictors and moderators of outcome in child and adolescent anxiety and depression: a systematic review of psychological treatment studies. Eur Child Adolesc Psychiatry 2013 Feb;22(2):69-87. [doi: 10.1007/s00787-012-0316-3] [Medline: 22923065]
70. Rodde T. Localytics. 2018. 21% of Users Abandon an App After One Use URL: http://info.localytics.com/blog/21-percent-of-users-abandon-apps-after-one-use [accessed 2018-09-19] [WebCite Cache ID 72XOEpyJQ]

Abbreviations

CRC: Cooperative Research Centre
DERS: Difficulties in Emotion Regulation Scale
DERS-SF: Difficulties in Emotion Regulation Scale-Short Form
eHealth: electronic health
K10: Kessler 10 Psychological Distress scale
MHC-SF: Mental Health Continuum-Short Form
mHealth: mobile health
SMS: short message service
uMARS: Mobile App Rating Scale-User Version

Edited by G Eysenbach; submitted 04.07.18; peer-reviewed by H de Sola, G Pravettoni; comments to author 17.08.18; revised version received 20.09.18; accepted 29.09.18; published 16.01.19.

Please cite as:
Hides L, Dingle G, Quinn C, Stoyanov SR, Zelenko O, Tjondronegoro D, Johnson D, Cockshaw W, Kavanagh DJ
Efficacy and Outcomes of a Music-Based Emotion Regulation Mobile App in Distressed Young People: Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(1):e11482
URL: https://mhealth.jmir.org/2019/1/e11482/
doi:10.2196/11482
PMID:30664457

Original Paper

Health Care Provider Perceptions of Consumer-Grade Devices and Apps for Tracking Health: A Pilot Study

Bree Holtz[1], MSc, PhD; Kerri Vasold[2], BS; Shelia Cotten[3], PhD; Michael Mackert[4], PhD; Mi Zhang[5], PhD

[1]Department of Advertising and Public Relations, College of Communication Arts & Sciences, Michigan State University, East Lansing, MI, United States

[2]Department of Kinesiology, Michigan State University, East Lansing, MI, United States

[3]Department of Media & Information, College of Communication Arts & Sciences, Michigan State University, East Lansing, MI, United States

[4]Center for Health Communication, Stan Richards School of Advertising & Public Relations, University of Texas, Austin, Austin, TX, United States

[5]Department of Electrical and Computer Engineering, College of Engineering, Michigan State University, East Lansing, MI, United States

Corresponding Author:
Bree Holtz, MSc, PhD
Department of Advertising and Public Relations
College of Communication Arts & Sciences
Michigan State University
404 Wilson Road
Room 309
East Lansing, MI, 48824
United States
Phone: 1 517 884 4537
Email: bholtz@msu.edu

Abstract

Background: The use of Web- or mobile phone–based apps for tracking health indicators has increased greatly. However, provider perceptions of consumer-grade devices have not been widely explored.

Objective: The purpose of this study was to determine primary care physicians' and advanced practice registered nurses' perceptions of consumer-grade sensor devices and Web- or mobile phone–based apps that allow patients to track physical activity, diet, and sleep.

Methods: We conducted a cross-sectional mailed survey with a random sample of 300 primary care physicians and 300 advanced practice registered nurses from Michigan, USA. Providers' use and recommendation of these types of technologies, and their perceptions of the benefits of and barriers to patients' use of the technologies for physical activity, diet, and sleep tracking were key outcomes assessed.

Results: Most of the respondents (189/562, 33.6% response rate) were advanced practice registered nurses (107/189, 56.6%). Almost half of the sample (93/189, 49.2%) owned or used behavioral tracking technologies. Providers found these technologies to be helpful in clinical encounters, trusted the data, perceived their patients to be interested in them, and did not have concerns over the privacy of the data. However, the providers did perceive patient barriers to using these technologies. Additionally, those who owned or used these technologies were up to 6.5 times more likely to recommend them to their patients.

Conclusions: Our study demonstrated that many providers perceived benefits for their patients to use these technologies, including improved communication. Providers' concerns included their patients' access and the usability of these technologies. Providers who encountered data from these technologies during patient visits generally perceive this to be helpful. We additionally discuss the barriers perceived by the providers and offer suggestions and future research to realize the potential benefits to using these data in clinical encounters.

(JMIR Mhealth Uhealth 2019;7(1):e9929) doi:10.2196/mhealth.9929

KEYWORDS

physicians; primary care; APRN; nurse practitioners; technology

Introduction

Background

Almost 50% of US adults report having one or more chronic diseases [1]. Chronic disease care accounted for 86% of health care spending in the United States in 2014 [2]. Health risk behaviors, including physical inactivity and poor nutrition, are cited as highly attributable causes for the illness and early death associated with chronic diseases [3]. With less than 50% of US adults meeting physical activity guidelines, and less than 25% of US adults meeting nutrition guidelines [3], behavioral aspects of health are considered a key component to target to reduce chronic disease risk and prevalence. Tracking health indicators is one behavior that is promoted to help address these healthy behavior choices.

Self-care through tracking health indicators has been shown to be successful in the management of many chronic diseases. Patient self-tracking, or recording of health indicators at home, has been used in a variety of situations, including the prediction of events such as migraines, weight control, physical activity patterns, and self-management of blood pressure and blood glucose [4-7]. The popularity of self-tracking has also grown, with almost 70% of US adults tracking a health indicator such as weight, diet, exercise, or health symptoms for themselves or for another individual [8], with similar rates in other countries [9]. In the past, much of this self-tracking has been done through the use of paper diaries [8]. Recently, the use of technology to more accurately track health indicators has increased; technologies such as consumer-grade sensor devices (eg, Fitbit) and Web- or mobile phone-based apps (eg, MyFitnessPal) that allow patients to track physical activity, diet, sleep, and a variety of other factors have proliferated [9]. In fact, the US Food and Drug Administration recently approved Apple Inc's smartwatch for monitoring the heart (through electrocardiography) and to detect atrial fibrillation [10]. Using these technologies has been associated with positive health outcomes across a wide range of conditions and behaviors, such as diet, physical activity, weight management, and mental health [11,12]. With this increased use of technologies for self-tracking of health behaviors, there are many implications for use of these technology-generated data in the clinical setting.

Much of the previous research has focused on consumer perspectives of these technologies [13]. Studies have examined the benefits to individuals who are healthy and simply want to track and improve their overall lifestyle [14-19]. As for individuals who have a chronic condition (eg, asthma, depression, diabetes), the evidence suggests that tracking may be beneficial [20-24]; however, the technologies' use in the clinical setting remains limited. Research on primary care providers' perceptions of information technologies that have been developed to be used as a diary for a specific illness is also limited, with only a handful of studies identified [25-29]. Additionally, 1 study was conducted on perceptions of data displays among both health care providers and laypersons [30]. Much of the research has hypothesized only about provider perspectives based on past experiences with other technologies [31,32]. Understanding primary care providers' perceptions may yield knowledge that can result in better design and use of these technologies for monitoring and managing patients' health, and improving their health outcomes.

Objectives

The objectives of this pilot study were to determine primary care physicians' and advanced practice registered nurses' (APRNs) use and perceptions of health tracking technologies. We included APRNs in this study because there is a well-documented shortage of primary care physicians, and APRNs often fulfill the role of a primary care health care provider for many patients [33]. We also sought providers' perceptions of the usefulness of these technologies on a variety of health issues. Additionally, we examined whether there were differences in perceptions by provider technology use status, as well as any differences between physicians and APRNs.

Methods

Study Design and Participants

In a cross-sectional study taking place from July to September 2016, using a random number table, we selected a random sample of 300 primary care physicians and 300 APRNs from office and hospital settings from the entire state of Michigan, USA, from a list purchased from DMD Marketing Corporation (an approved American Medical Association database licensee; Rosemont, IL, USA) and asked them to participate in a mail survey. We mailed providers a study packet that included a welcome letter, the survey, a self-addressed stamped return envelope, and a US $5 gift card to a national coffee chain. Follow-up postcards were sent to individuals who had not returned their original survey after 2 weeks. Past studies have demonstrated that mail surveys provide the best response rate of physicians [34-36]; however, on both the survey and the postcard follow-up, we provided a link to an online version of the survey (Qualtrics, Provo, UT, USA). The study was approved by the Michigan State University Institutional Review Board.

Data Collection

We developed the survey through a literature review of providers' perceptions regarding various technologies and patients bringing information to the visits [25,37]. From these studies, we developed items reflecting the themes from the results. We then conducted a small pilot study of the survey with 3 physicians and 2 nurse practitioners. The feedback they provided required us to shorten the survey and, if possible, to provide a larger incentive. The final survey included 25 questions in total pertaining to providers' use of these technologies (as defined above, consumer-grade sensor devices, such as Fitbit, and Web- or mobile phone-based apps, such as MyFitnessPal), their patients' use of these technologies, their perceptions of the usefulness of these technologies, and demographic questions. Figure 1 provides the definition of consumer-grade sensor devices and Web- or mobile phone-based apps that the providers were given. We also asked providers about their personal use of these technologies. These were answered on a 5-point Likert-type scale (ranging from "strongly agree" to "strongly disagree"). An additional 9 questions

pertained to the usefulness of technologies in specific contexts (eg, physical activity, diet, sleep, medication adherence, goal setting). Demographic and organizational characteristics were also included.

Statistical Analysis

We calculated descriptive statistics for all variables of interest. We conducted an exploratory factor analysis on the 19 perception statements of the data collected by technologies, barriers to use, and benefits of technologies. We identified 5 factors among 16 of the 19 statements; the remaining statements did not associate with a factor. Table 1 shows the factor analysis loadings. Reliability analysis revealed moderate to high reliability for all 5 factors (alpha≥.69).

We conducted Mann-Whitney *U* tests to determine whether the means of the 5 factors and individual perceptions questions were different for user versus nonuser and physician versus APRN comparisons. We used multinomial logistic regression using 95% confidence intervals to evaluate relationships for the likelihood of recommendation of technologies and perceptions of barriers and benefits. Independent variables for analyses included user status and job status. Dependent variables included likelihood of recommendation of technologies and perceptions of barriers and benefits. For the purpose of logistic regression, we condensed the perception scale categories to three—(1) agreement ("strongly agree" and "agree"), (2) neither agree nor disagree, and (3) disagreement ("strongly disagree" and "disagree")—and treated them as categorical data within the analysis. All analyses were conducted in IBM SPSS 24.0 statistical software (IBM Corporation).

Figure 1. The definitions of consumer-grade sensor devices and Web- or mobile phone-based apps that health care providers were given in the survey.

Table 1. Exploratory factor analysis loadings.

Perception statements	Factors				
	Data review	Provider trust of data	Patient's interest in technologies	Security and liability	Perceived patient barriers
There are worthwhile health benefits from reviewing patient-tracked data	.443[a]	.117	.564	.012	.027
Data from these apps and devices help me manage my patients' visits	.579[a]	.093	.511	.014	–.081
I want the data from these apps and devices to link with the electronic medical record system	.882[a]	.061	.020	–.131	–.052
I want the data from these apps and devices to link with the patient portal	.860[a]	.055	.015	–.111	–.023
I don't trust patient-reported data from these apps and devices[b]	.109	.874[a]	.124	–.136	–.076
I don't trust the data these apps and devices provide[b]	.107	.868[a]	.152	–.151	–.025
My patients are not familiar with tracking using these devices[b]	–.115	–.128	.541[a]	–.135	–.455
The technologies available to my patients are not useful[b]	.018	.475	.644[a]	–.205	.061
My patients are not interested in using technology to track behaviors or health[b]	–.083	.086	.732[a]	–.015	–.084
There is no value to me if my patients use these types of devices[b]	.268	.368	.646[a]	–.061	.140
I am concerned about security and privacy of the data collected from these devices and apps	–.077	–.058	–.111	.789[a]	.032
I am concerned about liability issues when it comes to recommending these devices and apps	–.095	–.107	–.084	.881[a]	.088
I am concerned about liability issues when viewing data as part of an electronic medical record system from these devices and apps	–.083	–.130	.480	.887[a]	.091
My older patients have a harder time with technology	–.058	.021	–.147	–.007	.811[a]
Not everyone has sufficient technological literacy to use these devices and apps	–.046	–.081	.012	.038	.800[a]
Not everyone has sufficient access to these devices and apps	–.105	–.032	.048	.184	.785[a]
Patients do not want to share their data with me because they don't want me to know the truth about their health[c]	.295	–.344	–.131	.340	.056
I don't get reimbursed for reviewing these data[c]	.086	–.323	.374	–.070	.245
Patients need to have enough detail to make reviewing tracked data worthwhile[c]	.238	–.120	.232	.001	.362

[a]Statement that loaded under the factor.

[b]Reverse coded.

[c]Removed from analysis because it did not fall within a factor.

Results

Sample Description

Respondents (189/562, 33.6% response rate; 38 packets were returned unopened) were primarily female (133/189, 70.4%), white (163/183, 89.1%), and between the ages of 35 and 54 years (96/188, 51.1%). Of the 189 completed surveys, 15 were completed using the Qualtrics link. The sample was divided between physicians (82/189, 43.4%) and APRNs (107/189, 56.6%), and 49.2% of respondents (93/189) reported using these types of technologies (as defined above) themselves. When asked about the typical insurance coverage their patients had, the sample reported that 41.9% of their patients had private insurance, 31.7% had Medicare, 27.5% had Medicaid, and 8.5% were uninsured (mean responses). Table 2 provides more detailed demographic information.

Provider Perceptions

The 5 factors from the factor analysis of perceptions questions were (1) data review (alpha=.78), defined as providers perceiving these data to be useful in patient encounters and wanting the data to be available through the electronic medical record system (EMR); (2) provider trust of the data (alpha=.79), including questions on the trustworthiness of the data from these technologies; (3) patients' interest in the technologies (alpha=.69), defined as the providers' perceptions of how interested in these technologies they believed their patients to be; (4) security and liability (alpha=.9), defined as the providers' perceptions of the data from these technologies; and (5) perceived patient barriers (alpha=.85), including the providers' perceptions of their patients' age, technology literacy, and access to these technologies. In the analysis of factors overall and individual questions pertaining to factors, notable findings included significant differences (*P* value range from <.001 to .02) between users and nonusers for data review, provider trust of data, and patients' interest in technologies. Additionally, we found only one difference by job status: APRNs were more interested in the linking of patient data to patient portals than physicians were (*P*=.02). Table 3 provides more detailed results of overall perceptions.

Table 2. Respondents' demographic characteristics.

Characteristics	Overall (n=189)	Physicians (n=82)	Advanced practice registered nurses (n=107)
Sex, n (%)			
Male	56 (29.6)	52 (63)	4 (3.7)
Female	133 (70.4)	30 (36)	103 (96.3)
Age (years), n (%)			
25-34	18 (9.6)	6 (7)	12 (11.2)
35-54	96 (51.1)	48 (59)	48 (44.9)
55-64	69 (36.7)	27 (33)	42 (39.3)
≥65	5 (2.7)	0 (0)	5 (4.7)
Race/ethnicity, n (%)			
White	163 (89.1)	66 (85)	97 (92.4)
African American	3 (1.6)	1 (1)	2 (1.9)
Asian	11 (6.0)	9 (12)	2 (1.9)
Native Hawaiian or Pacific Islander	1 (0.5)	1 (1)	0 (0)
American Indian or Alaskan Native	2 (1.1)	1 (1)	1 (1.0)
Other	3 (1.6)	0 (0)	3 (2.9)
User status, n (%)			
Users	93 (49.2)	32 (39)	61 (57.0)
Nonusers	96 (50.8)	50 (61)	46 (43.0)
Insurance, mean (SD)			
Private	41.9 (22.8)	45.7 (20.3)	38.2 (24.6)
Medicare	31.7 (18.5)	32.7 (13.3)	30.5 (22.8)
Medicaid	27.5 (22.0)	20.0 (17.9)	34.9 (28.8)
Uninsured	8.5 (12.3)	5.6 (5.5)	11.2 (15.9)

Table 3. Provider recommendations of tracking technologies and perceptions of technologies by job title and user status[a].

Independent variable	Overall (n=189)	User (n=93)	Nonuser (n=96)	Physician (n=82)	Advanced practice registered nurses (n=107)
Recommendation for..., n (%)					
Physical activity	124 (65.6)	79 (845)	45 (47)	52 (63)	72 (67.3)
Diet	118 (62.4)	77 (83)	41 (43)	49 (60)	69 (64.5)
Sleep	37 (19.6)	26 (28)	11 (12)	14 (17)	23 (21.5)
Perceptions of..., mean (SD)					
Data review	3.3 (0.7)	3.5 (0.7)[b] *P*<.001	3.2 (0.7)	3.3 (0.8)	3.4 (0.7)
There are worthwhile health benefits from reviewing patient-tracked data	3.9 (0.8)	4.0 (0.8)[b] *P*=.009	3.8 (0.8)	3.8 (0.9)	4.0 (0.7)
Data from these apps and devices help me manage my patients' visits	3.3 (1.0)	3.5 (0.9)[b] *P*=.002	3.0 (1.0)	3.3 (1.0)	3.2 (1.0)
I want the data from these apps and devices to link with the electronic medical record system	3.0 (1.0)	3.3 (1.0)[b] *P*=.005	2.8 (1.0)	2.9 (1.1)	3.1 (1.0)
I want the data from these apps and devices to link with the patient portal	3.2 (1.0)	3.4 (0.9)[b] *P*=.003	3.0 (1.0)	3.0 (1.1)[c] *P*=.02	3.4 (0.9)
Provider trust of data	3.5 (0.8)	3.6 (0.7)[b] *P*=.007	3.4 (0.8)	3.4 (0.8)	3.6 (0.7)
I don't trust patient-reported data from these apps and devices[d]	3.4 (0.8)	3.5 (0.8)	3.4 (0.8)	3.3 (0.8)	3.5 (0.8)
I don't trust the data these apps and devices provide[d]	3.6 (0.8)	3.7 (0.8)[b] *P*=.001	3.4 (0.8)	3.4 (0.8)	3.7 (0.8)
Patient's interest in technologies	3.6 (0.6)	3.7 (0.5)[b] *P*<.001	3.4 (0.6)	3.6 (0.6)	3.6 (0.6)
My patients are not familiar with tracking using these devices[d]	3.1 (0.9)	3.3 (1.0)	3.0 (0.9)	3.3 (0.9)	3.0 (0.9)
The technologies available to my patients are not useful[d]	3.8 (0.7)	4.0 (0.7)[b] *P*=.001	3.7 (0.7)	3.8 (0.8)	3.8 (0.6)
My patients are not interested in using technology to track behaviors or health[d]	3.5 (0.9)	3.7 (0.8)[b] *P*=.02	3.4 (0.9)	3.5 (0.9)	3.5 (0.9)
There is no value to me if my patients use these types of devices[d]	3.8 (0.8)	4.0 (0.7)[b] *P*=.001	3.6 (0.9)	3.8 (0.9)	3.9 (0.7)
Security and liability	2.8 (0.9)	2.7 (0.9)	2.9 (0.9)	2.7 (1.0)	2.9 (0.8)
I am concerned about security and privacy of the data collected from these devices and apps	2.8 (1.0)	2.7 (1.0)	2.9 (1.0)	2.7 (1.0)	2.9 (1.0)
I am concerned about liability issues when it comes to recommending these devices and apps	2.7 (1.0)	2.6 (1.0)	2.8 (1.0)	2.6 (1.1)	2.8 (0.9)
I am concerned about liability issues when viewing data as part of an electronic medical record system from these devices and apps	3.0 (1.1)	2.9 (1.1)	3.1 (1.1)	2.9 (1.2)	3.1 (1.0)
Perceived patient barriers	3.9 (0.7)	3.9 (0.7)	3.9 (0.7)	3.9 (0.7)	3.9 (0.7)
My older patients have a harder time with technology	3.9 (0.9)	3.9 (0.9)	3.9 (0.8)	4.0 (0.8)	3.8 (0.9)
Not everyone has sufficient technological literacy to use these devices and apps	3.9 (0.8)	4.0 (0.8)	3.9 (0.8)	3.9 (0.9)	3.9 (0.7)

Independent variable	Overall (n=189)	User (n=93)	Nonuser (n=96)	Physician (n=82)	Advanced practice registered nurses (n=107)
Not everyone has sufficient access to these devices and apps	4.0 (0.8)	4.0 (0.8)	4.0 (0.8)	3.9 (0.8)	4.1 (0.7)

[a]Perceptions scale: 1 = strongly disagree, 2 = disagree, 3 = neither agree nor disagree, 4 = agree, 5 = strongly agree.

[b]Significantly different from nonusers; *P*<.05.

[c]Significantly different from advanced practice registered nurses; *P*<.05.

[d]Question was reverse coded.

Table 4. Odds ratios (ORs) and 95% CIs for comparisons of user status and job title (N=189).

Independent variable	User vs nonuser[a], OR (95% CI)	Physician vs advanced practice registered nurses[b], OR (95% CI)
Recommendation for...		
Physical activity	6.4[c] (3.2-12.8)	0.8 (0.5-1.5)
Diet	6.5[c] (3.3-12.7)	0.8 (0.5-1.5)
Sleep	3.0[c] (1.4-6.5)	0.8 (0.4-1.6)
Perceptions of...		
Data review[d]	5.6[c] (1.7-19.1)	0.5 (0.2-1.4)
Provider trust of data[d]	1.8 (0.5-6.2)	0.5 (0.1-1.8)
Patient's interest in technologies[e]	7.6 (0.9-67.9)	1.8 (0.3-10.5)
Security and liability[e]	1.6 (0.8-3.3)	1.3 (0.6-2.6)
Perceived patient barriers[e]	2.5 (0.5-13.3)	1.0 (0.2-4.4)

[a]Reference group is nonusers.

[b]Reference group is advanced practice registered nurses.

[c]Statistically significant by 95% CI.

[d]Reference group is agreement.

[e]Reference group is disagreement.

Overall, providers found review of data to be useful and data to be trustworthy. Providers also perceived that their patients were interested in using these technologies. Security and liability issues with use of data and technologies were not perceived to be barriers to use, but there were concerns about barriers to patient use.

Provider Recommendations

The majority of the providers had recommended these technologies to their patients for physical activity (124/189, 65.6%) and diet (118/189, 62.4%). Additionally, providers who were technology users themselves were 6.4 times more likely to recommend devices and apps to their patients for physical activity tracking, 6.5 times more likely for diet tracking, and 3.0 times more likely for sleep tracking than nonusers. Users were also 5.6 times more likely than nonusers to perceive these technologies as useful in data review for their patients. We found no significant differences between physicians and APRNs (Tables 3 and 4).

Perceptions of Usefulness

The providers perceived varying levels of usefulness of these technologies for specific issues, rated on 5-point scales from "not at all useful" to "extremely useful." Many of the providers thought that these technologies were very useful or extremely useful for tracking physical activity (103/185, 55.6%), tracking diet (91/185, 49.2%), tracking vital signs (80/183, 43.7%), and goal setting (79/183, 43.2%; Table 5). However, the providers perceived that these technologies were not at all or slightly useful for sleep (92/181, 50.8%), smoking (94/178, 52.8%), mental states (108/182, 59.3%), and alcohol or drug use (110/176, 62.5%).

Table 5. Perceived usefulness of technologies (N=189).

	Not at all or slightly useful, n (%)	Neutral, n (%)	Very or extremely useful, n (%)
Physical activity (n=185)	22 (12.4)	60 (32.5)	103 (55.1)
Diet (n=185)	27 (16.8)	64 (35.2)	91 (48.0)
Vital signs (n=183)	49 (26.3)	54 (29.8)	80 (43.9)
Goal setting (n=183)	31 (18.0)	73 (39.3)	79 (42.7)
Medication adherence (n=182)	67 (34.5)	56 (31.5)	59 (34.0)
Smoking (n=178)	94 (52.0)	56 (30.0)	28 (18.0)
Sleep (n=181)	92 (50.2)	60 (32.8)	29 (16.9)
Alcohol or drug use (n=176)	110 (61.9)	43 (23.4)	23 (14.7)
Mental states (n=182)	108 (59.1)	55 (29.1)	19 (11.8)

Discussion

Principal Findings

This study examined health care providers' perceptions of consumer-grade, off-the-shelf behavioral tracking technologies. The responses of physicians and APRNs were almost the same. Overall, providers who personally owned or used these technologies were more likely to recommend them for tracking of physical activity, diet, and sleep. Additionally, providers who used these types of technologies were more likely to see the data as useful. Overwhelmingly, the providers who had personal experiences recommended these devices to their patients, indicating that once providers see and understand the data the technologies can provide, they can better counsel their patients in how these devices can help in lifestyle behavior change. We found only one difference between physicians and APRNs in their perceptions of the technologies, pertaining to connections to the patient portal.

Our study demonstrated that many providers strongly agreed or agreed that these technologies have benefits for their patients. Results revealed perceptions of worthwhile benefits to reviewing patient-tracked health data and that data could help in managing patient visits. Past research suggested that patient-generated health data may lead to better communication between the patient and provider, help set goals, and discover patients' habits and preferences [27]. Additionally, some studies showed that, when patients brought information to their visits, better health outcomes were achieved [25,38,39].

Overall, providers had positive perceptions of trusting the data that these devices provide. In contrast, previous research contended that physicians perceived the data from these devices to be unreliable [40]. However, Nundy et al found that, overall, these data are more trustworthy than self-report, with providers perceiving that some patients misrepresent their activity to please the providers [27]. In previous research regarding health technologies, Health Insurance Portability and Accountability Act of 1996 rules and regulations regarding data security were major concerns among providers [26,28,29]. However, we found that most providers were not concerned with these issues. Yet, if these data were synthesized and standardized for inclusion into the medical record, data security could become more of a concern [26]. Given, that the data do not currently become part of the medical record, there is very limited liability for the providers regarding these data. This also may be because providers are looking at the data as supplemental, as a way to help their patients increase healthy behaviors [41,42]. However, technology is going to continue be to marketed to consumers rather than to primary care providers. These entities—providers, insurers, and regulators—are going to have to consider the infrastructure of information, data infrastructure, and policies concerning data security and privacy as more and more of these technologies are introduced.

Prior work has demonstrated that providers perceive patients' lack of access to these technologies as a barrier to recommending them to patients [26-29], and our results support these findings. Our research demonstrated that provider-perceived barriers to recommending these technologies included older patients, technical literacy, and financial costs. However, the costs of these technologies are decreasing, ease of use is improving for certain populations, and some insurance companies may try to cover costs for these types of technologies [17].

We found no major differences in perceptions of these technologies between physicians and APRNs. While, to our knowledge, no previous studies have compared perceptions of these technologies between different types of health care providers, some studies examined perceptions of different health-related technologies (eg, telemedicine or EMRs). Physicians tended to be concerned with costs and perceived productivity [27,28,31], whereas nurses were concerned with how much effort went into learning the technologies and the support available for their integration and use in practice [31]. This supports APRNs wanting to link these data to the patient portal. Previous studies have found differences between physicians and APRNs [43]. Future research should examine whether these differences exist within organizations to explore whether differences emerge as these technologies become more commonplace.

Overall, the providers perceived these technologies to be the most useful for tracking physical activity, diet, vital signs, and goal setting. They viewed these technologies not to be effective for monitoring sleep, which is one of the benefits of using many of these consumer devices and apps. This could be because many devices have sleep tracking as a secondary function [44].

The providers also perceived technologies as not being useful for tracking mental states, alcohol or drug use, and smoking, which could be of concern, since some projects are seeking to use mobile devices to help with these areas [45-47]. We surveyed only primary care and family practice health care providers, which could be a potential reason for the perceived lack of usefulness for tracking mental states, alcohol or drug use, and smoking, as some of these issues could be referred to specialists. However, some patients may see only a primary care or family practice health care provider, and they consult with them for all health issues.

Limitations

This study had several limitations. First, as with any survey, recall bias may have been a factor. Second, this was a survey of physicians and APRNs in Michigan; this study should be replicated to include a US national sample of providers, including a broader range of providers. Third, the response rate was low (33.6%), but previous research has shown a downward trend of response rates when surveying health care providers [48]. Other studies have demonstrated that, in general, most survey respondents are more likely to participate if they are interested in the topic [49]. Fourth, our sample was majority white and women, who have been shown in past research to participate more in surveys [50-53]. Fifth, to ensure that the health care providers understood the class of technologies we were interested in, we asked them specifically about physical activity, diet, and sleep. This may have lowered the perceived usefulness of other ways in which these technologies can be used. However, the technologies used for the other issues we asked them about are in themselves not as popular or well known. As this sector of health care continues to grow, future research should examine these areas in more depth.

Conclusion

Our survey results have implications for providers, technology developers, patients, and insurers. Once providers have first-hand experience with technologies, they understand how to interpret the data better. Technology developers and manufacturers should continue to test the validity and reliability of their devices and apps to provide the credibility that providers expect. Our results also demonstrated that, if insurers would provide reimbursement for these types of technologies, the cost barrier could be reduced. On the other hand, payers also need to reimburse the providers for time to review the data. Additionally, being able to access these data through the EMR is perceived as an effective way to view the data. This does reinforce the finding that there must be a way to incorporate the data into routine medical encounters. For patients, this could be a way to communicate about their health care preferences and priorities to their provider. Makers of these devices may want to consider advertising directly to providers. Our findings suggested that those providers who already own a device or use a technology were more likely to recommend them to their patients and found positive outcomes through patient use.

This work demonstrated that primary care physicians and APRNs have an overall positive perception of consumer-grade off-the-shelf technologies that track individual health behaviors. Providers may serve as the gatekeepers for use of these technologies in improving health care, as their actual use could drive interest, acceptance, and possible better health outcomes. These providers, as trusted sources of health information, could be advocates for use of these behavioral health tracking technologies and help realize the public health benefits that could come from their wider adoption.

Conflicts of Interest

None declared.

References

1. Ward BW, Schiller JS, Goodman RA. Multiple chronic conditions among US adults: a 2012 update. Prev Chronic Dis 2014 Apr 17;11:E62 [FREE Full text] [doi: 10.5888/pcd11.130389] [Medline: 24742395]
2. Gerteis J, Izrael D, Deitz D, LeRoy L, Ricciardi R, Miller T, et al. Multiple chronic conditions chartbook: 2010 medical expenditure panel survey data. Rockville, MD: Agency for Healthcare Research and Quality; 2014. URL: https://www.ahrq.gov/sites/default/files/wysiwyg/professionals/prevention-chronic-care/decision/mcc/mccchartbook.pdf [accessed 2018-11-28] [WebCite Cache ID 744n9VwfU]
3. Centers for Disease Control and Prevention. Sortable risk factors and health indicators. 2016. URL: https://sortablestats.cdc.gov/ [accessed 2018-11-28] [WebCite Cache ID 74GiB07ik]
4. Burke LE, Warziski M, Starrett T, Choo J, Music E, Sereika S, et al. Self-monitoring dietary intake: current and future practices. J Ren Nutr 2005 Jul;15(3):281-290. [Medline: 16007557]
5. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. J Med Internet Res 2013;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
6. Giffin NJ, Ruggiero L, Lipton RB, Silberstein SD, Tvedskov JF, Olesen J, et al. Premonitory symptoms in migraine: an electronic diary study. Neurology 2003 Mar 25;60(6):935-940. [Medline: 12654956]
7. Spring B, Duncan JM, Janke EA, Kozak AT, McFadden HG, DeMott A, et al. Integrating technology into standard weight loss treatment: a randomized controlled trial. JAMA Intern Med 2013 Jan 28;173(2):105-111 [FREE Full text] [doi: 10.1001/jamainternmed.2013.1221] [Medline: 23229890]

8. Fox S, Duggan M. The diagnosis difference: a portrait of the 45% of US adults living with chronic health conditions. Washington, DC: Pew Research Center; 2013 Nov 26. URL: http://www.pewinternet.org/wp-content/uploads/sites/9/media/Files/Reports/2013/PewResearch_DiagnosisDifference.pdf [accessed 2018-11-21] [WebCite Cache ID 746I4VBaF]
9. 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]
10. Krueger AC. Re: DEN180044. Silver Spring, MD: Food & Drug Administration; 2018 Sep 11. URL: https://www.accessdata.fda.gov/cdrh_docs/pdf18/DEN180044.pdf [accessed 2018-11-28] [WebCite Cache ID 74H00HNi3]
11. Mohr DC, Zhang M, Schueller SM. Personal sensing: understanding mental health using ubiquitous sensors and machine learning. Annu Rev Clin Psychol 2017 May 08;13:23-47. [doi: 10.1146/annurev-clinpsy-032816-044949] [Medline: 28375728]
12. Swan M. The quantified self: fundamental disruption in big data science and biological discovery. Big Data 2013 Jun;1(2):85-99. [doi: 10.1089/big.2012.0002] [Medline: 27442063]
13. Loos JR, Davidson EJ. Wearable health monitors and physician-patient communication: the physician's perspective. New York, NY: IEEE; 2016 Presented at: 49th Hawaii International Conference on System Sciences (HICSS); Jan 5-8, 2016; Koloa, HI, USA p. 3389-3399. [doi: 10.1109/HICSS.2016.422]
14. Choe EK, Lee NB, Lee B, Pratt W, Kientz JA. Understanding quantified-selfers' practices in collecting and exploring personal data. New York, NY: ACM Press; 2014 Presented at: SIGCHI Conference on Human Factors in Computing Systems; Apr 26-May 1, 2014; Toronto, ON, Canada p. 1143-1152. [doi: 10.1145/2556288.2557372]
15. Mackert M, Mabry-Flynn A, Champlin S, Donovan EE, Pounders K. Health literacy and health information technology adoption: the potential for a new digital divide. J Med Internet Res 2016 Oct 04;18(10):e264 [FREE Full text] [doi: 10.2196/jmir.6349] [Medline: 27702738]
16. Mcfedries P. Tracking the quantified self [technically speaking]. IEEE Spectr 2013 Aug;50(8):24-24. [doi: 10.1109/mspec.2013.6565555]
17. Patel MS, Asch DA, Volpp KG. Wearable devices as facilitators, not drivers, of health behavior change. JAMA 2015 Feb 3;313(5):459-460. [doi: 10.1001/jama.2014.14781] [Medline: 25569175]
18. Swan M. Emerging patient-driven health care models: an examination of health social networks, consumer personalized medicine and quantified self-tracking. Int J Environ Res Public Health 2009 Feb;6(2):492-525 [FREE Full text] [doi: 10.3390/ijerph6020492] [Medline: 19440396]
19. Swan M. Sensor mania! The internet of things, wearable computing, objective metrics, and the quantified self 2.0. J Sens Actuator Netw 2012 Nov 08;1(3):217-253. [doi: 10.3390/jsan1030217]
20. Bartholomew LK, Gold RS, Parcel GS, Czyzewski DI, Sockrider MM, Fernandez M, et al. Watch, Discover, Think, and Act: evaluation of computer-assisted instruction to improve asthma self-management in inner-city children. Patient Educ Couns 2000 Feb;39(2-3):269-280. [Medline: 11040726]
21. Carlbring P, Ekselius L, Andersson G. Treatment of panic disorder via the internet: a randomized trial of CBT vs. applied relaxation. J Behav Ther Exp Psychiatry 2003 Jun;34(2):129-140. [Medline: 12899896]
22. Christensen H, Griffiths KM, Jorm AF. Delivering interventions for depression by using the internet: randomised controlled trial. BMJ 2004 Jan 31;328(7434):265 [FREE Full text] [doi: 10.1136/bmj.37945.566632.EE] [Medline: 14742346]
23. Lange A, van de Ven JP, Schrieken B. Interapy: treatment of post-traumatic stress via the internet. Cogn Behav Ther 2003;32(3):110-124. [doi: 10.1080/16506070302317] [Medline: 16291543]
24. Neville LM, O'Hara B, Milat A. Computer-tailored physical activity behavior change interventions targeting adults: a systematic review. Int J Behav Nutr Phys Act 2009;6:30 [FREE Full text] [doi: 10.1186/1479-5868-6-30] [Medline: 19490649]
25. Chung C, Cook J, Bales E, Zia J, Munson SA. More than telemonitoring: health provider use and nonuse of life-log data in irritable bowel syndrome and weight management. J Med Internet Res 2015;17(8):e203 [FREE Full text] [doi: 10.2196/jmir.4364] [Medline: 26297627]
26. Mancuso PJ, Thompson M, Tietze M, Kelk S, Roux G. Can patient use of daily activity monitors change nurse practitioner practice? J Nurse Pract 2014 Nov;10(10):787-793.e4. [doi: 10.1016/j.nurpra.2014.09.002]
27. Nundy S, Lu CE, Hogan P, Mishra A, Peek ME. Using patient-generated health data from mobile technologies for diabetes self-management support: provider perspectives from an academic medical center. J Diabetes Sci Technol 2014 Jan 1;8(1):74-82. [doi: 10.1177/1932296813511727] [Medline: 24876541]
28. Schneider T, Panzera AD, Martinasek M, McDermott R, Couluris M, Lindenberger J, et al. Physicians' perceptions of mobile technology for enhancing asthma care for youth. J Child Health Care 2014 Nov 26. [doi: 10.1177/1367493514556555] [Medline: 25427556]
29. Berkowitz CM, Zullig LL, Koontz BF, Smith SK. Prescribing an app? Oncology providers' views on mobile health apps for cancer care. JCO Clin Cancer Inform 2017 Nov(1):1-7. [doi: 10.1200/CCI.17.00107] [Medline: 29113731]
30. Beaudin JS, Intille SS, Morris ME. To track or not to track: user reactions to concepts in longitudinal health monitoring. J Med Internet Res 2006;8(4):e29 [FREE Full text] [Medline: 17236264]

31. McGinn CA, Grenier S, Duplantie J, Shaw N, Sicotte C, Mathieu L, et al. Comparison of user groups' perspectives of barriers and facilitators to implementing electronic health records: a systematic review. BMC Med 2011;9:46 [FREE Full text] [doi: 10.1186/1741-7015-9-46] [Medline: 21524315]
32. Whitten P, Holtz B. Provider utilization of telemedicine: the elephant in the room. Telemed J E Health 2008 Nov;14(9):995-997. [doi: 10.1089/tmj.2008.0126] [Medline: 19035815]
33. Petterson SM, Liaw WR, Phillips RL, Rabin DL, Meyers DS, Bazemore AW. Projecting US primary care physician workforce needs: 2010-2025. Ann Fam Med 2012;10(6):503-509 [FREE Full text] [doi: 10.1370/afm.1431] [Medline: 23149526]
34. James KM, Ziegenfuss JY, Tilburt JC, Harris AM, Beebe TJ. Getting physicians to respond: the impact of incentive type and timing on physician survey response rates. Health Serv Res 2011 Feb;46(1 Pt 1):232-242 [FREE Full text] [doi: 10.1111/j.1475-6773.2010.01181.x] [Medline: 20880042]
35. Cunningham CT, Quan H, Hemmelgarn B, Noseworthy T, Beck CA, Dixon E, et al. Exploring physician specialist response rates to web-based surveys. BMC Med Res Methodol 2015;15:32 [FREE Full text] [doi: 10.1186/s12874-015-0016-z] [Medline: 25888346]
36. Cook DA, Wittich CM, Daniels WL, West CP, Harris AM, Beebe TJ. Incentive and reminder strategies to improve response rate for internet-based physician surveys: a randomized experiment. J Med Internet Res 2016 Sep 16;18(9):e244 [FREE Full text] [doi: 10.2196/jmir.6318] [Medline: 27637296]
37. Murray E, Lo B, Pollack L, Donelan K, Catania J, White M, et al. The impact of health information on the internet on the physician-patient relationship: patient perceptions. Arch Intern Med 2003 Jul 28;163(14):1727-1734. [doi: 10.1001/archinte.163.14.1727] [Medline: 12885689]
38. Menachemi N, Powers TL, Brooks RG. The role of information technology usage in physician practice satisfaction. Health Care Manage Rev 2009;34(4):364-371. [doi: 10.1097/HMR.0b013e3181a90d53] [Medline: 19858921]
39. Weiner M, Biondich P. The influence of information technology on patient-physician relationships. J Gen Intern Med 2006 Jan;21 Suppl 1:S35-S39 [FREE Full text] [doi: 10.1111/j.1525-1497.2006.00307.x] [Medline: 16405708]
40. West P, Giordano R, Van Kleek M, Shadbolt N. The quantified patient in the doctor's office: challenges & opportunities. New York, NY: ACM; 2016 Presented at: CHI Conference on Human Factors in Computing Systems; May 7-12, 2016; San Jose, CA, USA p. 3066-3078. [doi: 10.1145/2858036.2858445]
41. De Angelis G, Davies B, King J, McEwan J, Cavallo S, Loew L, et al. Information and communication technologies for the dissemination of clinical practice guidelines to health professionals: a systematic review. JMIR Med Educ 2016 Nov 30;2(2):e16 [FREE Full text] [doi: 10.2196/mededu.6288] [Medline: 27903488]
42. Sawesi S, Rashrash M, Phalakornkule K, Carpenter JS, Jones JF. The impact of information technology on patient engagement and health behavior change: a systematic review of the literature. JMIR Med Inform 2016;4(1):e1 [FREE Full text] [doi: 10.2196/medinform.4514] [Medline: 26795082]
43. Holtz B, Krein S. Understanding nurse perceptions of a newly implemented electronic medical record system. J Technol Hum Serv 2011 Oct;29(4):247-262. [doi: 10.1080/15228835.2011.639931]
44. Holtz B, Krein SL, Bentley DR, Hughes ME, Giardino ND, Richardson CR. Comparison of Veteran experiences of low-cost, home-based diet and exercise interventions. J Rehabil Res Dev 2014;51(1):149-160 [FREE Full text] [doi: 10.1682/JRRD.2013.04.0088] [Medline: 24805901]
45. Cohn AM, Hunter-Reel D, Hagman BT, Mitchell J. Promoting behavior change from alcohol use through mobile technology: the future of ecological momentary assessment. Alcohol Clin Exp Res 2011 Dec;35(12):2209-2215 [FREE Full text] [doi: 10.1111/j.1530-0277.2011.01571.x] [Medline: 21689119]
46. Das S, Tonelli M, Ziedonis D. Update on smoking cessation: e-cigarettes, emerging tobacco products trends, and new technology-based interventions. Curr Psychiatry Rep 2016 May;18(5):51. [doi: 10.1007/s11920-016-0681-6] [Medline: 27040275]
47. 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;42(6):505-512. [doi: 10.1037/a0024485]
48. McLeod CC, Klabunde CN, Willis GB, Stark D. Health care provider surveys in the United States, 2000-2010: a review. Eval Health Prof 2013 Mar;36(1):106-126. [doi: 10.1177/0163278712474001] [Medline: 23378504]
49. Groves RM. Nonresponse rates and nonresponse bias in household surveys. Public Opin Q 2006;70(5):646-675. [doi: 10.1093/poq/nfl033]
50. Moore DL, Tarnai J. Evaluating nonresponse error in mail surveys. In: Groves RM, Dillman DA, Eltinge JL, Little RJA, editors. Survey Nonresponse. New York, NY: Wiley & Sons; 2002:197-211.
51. Groves RM, Singer E, Corning A. Leverage-saliency theory of survey participation: description and an illustration. Public Opin Q 2000;64(3):299-308. [Medline: 11114270]
52. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007 Sep;17(9):643-653. [doi: 10.1016/j.annepidem.2007.03.013] [Medline: 17553702]
53. Partin MR, Malone M, Winnett M, Slater J, Bar-Cohen A, Caplan L. The impact of survey nonresponse bias on conclusions drawn from a mammography intervention trial. J Clin Epidemiol 2003 Sep;56(9):867-873. [Medline: 14505772]

Abbreviations

APRN: advanced practice registered nurse
EMR: electronic medical record system

Edited by G Eysenbach; submitted 24.01.18; peer-reviewed by M Faber, A Zeleke, B Dooley; comments to author 17.08.18; revised version received 19.10.18; accepted 20.10.18; published 22.01.19.

Please cite as:
Holtz B, Vasold K, Cotten S, Mackert M, Zhang M
Health Care Provider Perceptions of Consumer-Grade Devices and Apps for Tracking Health: A Pilot Study
JMIR Mhealth Uhealth 2019;7(1):e9929
URL: https://mhealth.jmir.org/2018/1/e9929/
doi:10.2196/mhealth.9929
PMID:30668515

Original Paper

Assessing the Quality of Mobile Phone Apps for Weight Management: User-Centered Study With Employees From a Lebanese University

Marco Bardus[1], MA, PhD; Ahmed Ali[1], MD, MPH; Farah Demachkieh[1,2], BSN, MPH; Ghassan Hamadeh[3], MD

[1]Department of Health Promotion and Community Health, Faculty of Health Sciences, American University of Beirut, Beirut, Lebanon

[2]SANAD - The Home Hospice Organization of Lebanon, Beirut, Lebanon

[3]Department of Family Medicine, Faculty of Medicine, American University of Beirut, Beirut, Lebanon

Corresponding Author:
Marco Bardus, MA, PhD
Department of Health Promotion and Community Health
Faculty of Health Sciences
American University of Beirut
Van Dyck - Room 302
PO Box 11-0236 Riad El-Solh
Beirut, 1107 2020
Lebanon
Phone: 961 1 350 000 ext 4694
Fax: 961 1 744 470
Email: marco.bardus@gmail.com

Abstract

Background: Evaluating the quality of mobile health apps for weight loss and weight management is important to understand whether these can be used for obesity prevention and treatment. Recent reviews call for more research on multidimensional aspects of app quality, especially involving end users, as there are already many expert reviews on this domain. However, no quantitative study has investigated how laypersons see popular apps for weight management and perceive different dimensions of app quality.

Objective: This study aimed to explore how laypersons evaluate the quality of 6 free weight management apps (*My Diet Coach*, *SparkPeople*, *Lark*, *MyFitnessPal*, *MyPlate*, and *My Diet Diary*), which achieved the highest quality ratings in a related and recent expert review.

Methods: A user-centered study was conducted with 36 employees of a Lebanese university. Participants enrolled in the study on a rolling basis between October 2016 and March 2017. Participants were randomly assigned an app to use for 2 weeks. App quality was evaluated at the end of the trial period using the Mobile App Rating Scale user version (uMARS). uMARS assesses the dimensions of *engagement*, *functionality*, *aesthetics*, *information*, and *subjective quality* on 5-point scales. Internal consistency and interrater agreement were examined. The associations between uMARS scores and users' demographic characteristics were also explored using nonparametric tests. Analyses were completed in November 2017.

Results: Overall, the 6 apps were of moderately good quality (median uMARS score 3.6, interquartile range [IQR] 0.3). The highest total uMARS scores were achieved by *Lark* (mean 4.0 [SD 0.5]) and *MyPlate* (mean 3.8 [SD 0.4]), which also achieved the highest subjective quality scores (*Lark*: mean 3.3 [SD 1.4]; *MyPlate*: mean 3.3 [SD 0.8]). *Functionality* was the domain with the highest rating (median 3.9, IQR 0.3), followed by *aesthetics* (median 3.7, IQR 0.5), *information* (median 3.7, IQR 0.1), and *engagement* (median 3.3, IQR 0.2). *Subjective quality* was judged low (median 2.5, IQR 0.9). Overall, *subjective quality* was strongly and positively related ($P<.001$) with total uMARS score (ρ=.75), *engagement* (ρ=.68), *information*, and *aesthetics* (ρ=.60) but not *functionality* (ρ=.40; $P=.02$). Higher *engagement* scores were reported among healthy ($P=.003$) and obese individuals ($P=.03$), who also showed higher total uMARS ($P=.04$) and *subjective quality* ($P=.05$) scores.

Conclusions: Although the apps were considered highly functional, they were relatively weak in engagement and subjective quality scores, indicating a low propensity of using the apps in the future. As engagement was the subdomain most strongly associated with subjective quality, app developers and researchers should focus on creating engaging apps, holding constant the functionality, aesthetics, and information quality. The tested apps (in particular *Lark* and *MyPlate*) were perceived as more

engaging and of higher quality among healthy, obese individuals, making them a promising mode of delivery for self-directed interventions promoting weight control among the sampled population or in similar and comparable settings.

(JMIR Mhealth Uhealth 2019;7(1):e9836) doi:10.2196/mhealth.9836

KEYWORDS

mobile apps; weight loss; physical activity; healthy diet; workplace; mHealth

Introduction

Background

Mobile health (mHealth) apps offer cost-efficient and effective strategies to prevent noncommunicable diseases such as obesity or diabetes [1], as these technologies can reach millions of users. According to the 2017 mHealth App Economics report, there are more than 350,000 health apps available in online stores [2], a market worth US $25 billion in 2017 [3] and estimated to reach US $31 billion by 2020 [4]. mHealth apps are generally designed for chronically ill people (56%), fitness enthusiasts (33%), and physicians (32%) [4], with users downloading them with the aim to monitor their fitness and track foods as well as to manage chronic conditions [5]. A recent study specifically evaluating the market of weight management apps in 10 different countries [6] identified 28,905 unique apps that focus on physical activity (34%); diet (31%); and on tracking exercise, calorie intake, and body weight (23%) [6].

Although the mHealth app market is expected to expand in the next 3 years [7], recent market research reports show a decline in app usage [4]. Some qualitative studies show that users stop using apps because of hidden costs, increased data entry burden [8], and low perceived engagement [9]. Engagement with an app is generally associated with sustained app usage [1], but it has also been associated with positive changes in physical activity [10,11] and diet [12], fundamental behaviors to obtain an optimal weight management. Understanding which apps are perceived engaging and of good quality is important to develop effective public health strategies addressing these problems [3]. The more people use the apps they like, the more likely people will perform the desired behaviors.

Are mHealth apps effective? Several recent systematic reviews suggest that mobile phone apps are effective in promoting dietary self-regulation [13] and weight management [14-20]. Despite lacking evidence-based content [6,21], health apps can be used as stand-alone delivery modes in *self-directed* weight loss interventions [22,23] or as supplemental components of complex interventions. Some studies employing researcher-developed apps [24] or popular calorie counting apps (eg, *MyFitnessPal* [25,26]) in combination with face-to-face delivery modes showed generally larger effects compared with interventions using the apps as standalone [27-29].

How do these apps work? According to several app audits or reviews, mobile phone apps include features that can trigger cognitive processes underpinning effective behavior change strategies or techniques [30-35], combining principles derived from self-determination theory [22,23] and persuasive technology [36,37]. For example, apps may include messages or notifications that remind users about their weight goals and provide positive feedback or reinforcements for achieving those goals. In a recent review of 23 popular weight management apps [30], researchers found that most apps included several change techniques that are commonly employed in effective behavior change interventions. The most frequently identified change techniques were self-monitoring of behavior (20/23, 87%), self-monitoring and goal setting of outcomes (both 19/23, 83%), feedback on outcomes (17/23, 74%), feedback on behavior (16/23, 70%), and goal setting of behavior (13/23, 57%) [30]. Although research demonstrated the efficacy of these techniques in influencing behavior, available evaluations of app quality cannot demonstrate app efficacy. Assessing app quality has become an important stream of research, with several authors arguing for the need to improve the quality evaluation and the need to use standardized tools and systematic approaches [38]. However, expert app evaluations or reviews do not take into account the point of view of end users. Little is known about how end users perceive the apps and in what terms they judge their quality.

In a recent review on app quality assessment methods [39], the authors emphasized the need to use multidimensional tools to comprehensively determine the quality of mobile phone apps, which should also include end users' viewpoints. This is because the views of researchers and end users tend to diverge. On one side, researchers focus on aspects related to theoretical and evidence-based content [38,39]. For example, in the aforementioned expert app review [30], the authors judged the 23 apps as highly functional but poor in information quality, lamenting the absence of references to evidence-based content. At the same time, their quality ratings were not significantly associated with the 5-star ratings derived from Google Play and iTunes stores, suggesting a potential gap between the wisdom of the crowds and of the experts [30]. App store ratings cannot be entirely trusted as these ratings can be piloted through reviews and ratings provided by humans or bots paid by the same developer companies [40]. On the other side, developers tend to focus on usability and aesthetic aspects, such as design, ease of use, and customizability, as some qualitative studies demonstrate that these aspects are particularly appreciated by end users [8,9,41].

One of the most comprehensive and multidimensional tools to evaluate app quality is the Mobile App Rating Scale (MARS). Developed by Stoyanov et al for expert reviews [42], the MARS has also been developed and validated for end users [43]. The MARS and the user version of the Mobile App Rating Scale, uMARS are multidimensional as they encompass the domains of *engagement*, *functionality*, *aesthetics*, and *information*, which are used to estimate an *objective* app quality dimension (calculated as an average score of the aforementioned domains), based on objective features and characteristics of an app. Each

domain consists of a set of items, assessed on 5-point scales. The engagement domain includes 5 items: *entertainment*, *interest*, *customization*, *interactivity*, and *target group*. Functionality includes 4 items: *performance*, *ease of use*, *navigation*, and *gestural design*. Aesthetics includes 3 items: *layout*,*graphics*, and *visual appeal*. Information includes 4 core items: *quality*, *quantity*, *visual* information, as well as *credibility* of the source of information. The MARS scale includes 2 additional items: *accuracy of app description* and *goals* (ie, Does app have specific, measurable, and achievable goals specified in app store description or within the app itself?). The latter items, in fact, require additional information that a lay user might not easily find while using the app. Finally, both scales have also a *subjective quality* domain, which includes 4 items: Would you recommend this app to people who might benefit from it?; How many times do you think you would use this app in the next 12 months, if it was relevant to you?; Would you pay for this app?; and What is your overall star rating of the app? Due to the third item, it can be assumed that the higher the subjective quality score, the more likely the users would use the app in the future; however, the instrument does not include a measure of actual behavior (eg, "How many times have you used this app in the past day or week"). The MARS and uMARS tools are available from the respective MARS [42] and uMARS [43] development studies.

The MARS tool, generalized to primary prevention apps [44], has been used in several expert reviews of apps for a variety of behaviors such as drink driving [45], sustainable food consumption [46], medication adherence [47], mental health and mindfulness [48], quality of life [49], rheumatoid arthritis [50], weight loss related to smoking cessation [51], and weight management [30]. The user version, originally tested on 2 harm minimization and affect management apps [43], assessed the apps according to the same domains. The only differences between the 2 tools are wording of the questions and the number of items assessing the information domain. The uMARS use has been documented in research protocols of trials addressing type 2 diabetes [52], health-related quality of life [53], pneumococcal disease [54], and breastfeeding [55]. However, to the best of our knowledge, the uMARS tool has not been used to quantitatively evaluate commercially available weight management apps. In addition, little is known about what users believe are important app characteristics, that is, app quality dimensions and how these dimensions relate to the overall app quality. Furthermore, according to the leading author of the scale (Stoyanov S, personal communication, November 2017), the items belonging to each domain were logically grouped, but no MARS or uMARS studies to date appear to have evaluated the relationships among different app quality dimensions.

Aims of the Study

In response to the call for more research on app quality evaluations from end users [39], the overarching goal of this study was to explore how laypersons evaluate the quality of a set of weight management apps, which experts considered of high quality in a recent review [30]. Specifically, this study aimed to (1) test the uMARS within a set of weight management apps; (2) understand which dimensions of app quality contribute the most to the overall app quality and how functionality, aesthetics, engagement, and information dimensions are related to subjective quality (as proxy of future app use); and (3) explore the associations between uMARS scales and users' characteristics.

Methods

App Selection

A user experience study was used to examine the perceived quality and usability of selected apps and identify which apps achieve the best quality scores, which could be used in further studies with the same target population (employees of an academic institution). The units of analysis of this study were derived from a recent review of mobile phone apps for weight management [30]. In the cited review, only 6 out of the 23 apps reviewed (Table 1) scored above the median point of the MARS scale (3 out of 5), which is the median value of a 5-point scale. This value has been considered the minimum threshold of acceptability in the study by Mani et al [56].

Participants and Procedures

Following recommendations from user experience and usability testing literature [57,58], we aimed to recruit 5 to 6 evaluators per app (30-36 participants). Participants were employees (faculty and staff) of the American University of Beirut, who were recruited through social media postings and email invitations (the research team obtained a list of randomly selected email addresses).

Table 1. List of apps used in the study, sorted by total Mobile App Rating Scale score, with app store information.

App name	Total MARS[a] score[b]	Google Play rating[c] (n)	iTunes rating[c] (n)
My Diet Coach	4.6	4.6 (20,115)	4.6 (6040)
SparkPeople	4.4	4.4 (30,453)	4.6 (3677)
Lark	4.1	4.1 (2940)	4.1 (4294)
MyFitnessPal	3.9	4.6 (1,701,093)	4.7 (621,127)
MyPlate	3.5	4.6 (18,085)	4.6 (18,688)
My Diet Diary	3.4	4.1 (18,415)	4.2 (1280)

[a]MARS: Mobile App Rating Scale.

[b]Derived from the expert review by Bardus et al [30].

[c]Average 5-star rating and total number of ratings based on all versions of the app, as of November 15, 2017.

Interested employees submitted an informed consent and completed a Web-based eligibility survey. Inclusion criteria were participants aged 18 to 65 years, employees of the university, and owning either Android or iPhone devices. After enrollment and after signing an informed consent, which included all study schedules and requirements, participants completed a Web-based sociodemographic and behavioral baseline survey. Then, they were randomly assigned to use 1 of the apps for 2 weeks. A member of the research team helped each participant install the assigned app and verified that it was correctly installed and functioning. The same member of the research team encouraged participants to use the app at least daily for the duration of 2 weeks. At the end of this study period, they were invited to complete a final Web-based app evaluation survey. They received US $10 to complete each survey. The study was approved by the local institutional review board (reference number FHS.MB.01) and was conducted between October 2016 and March 2017; analyses were completed in November 2017.

Measures

Background Characteristics

Background characteristics of the users included sociodemographic (age, gender, marital status, education, income, and number of working hours), health-related, and behavioral factors (perceived health status, height and weight, and physical activity assessed through the International Physical Activity Questionnaire-short form) [59]. App usage characteristics included operative system (Android or iOS) and previous experience with mHealth apps (for physical activity, diet, or weight tracking).

Quantitative Outcomes

App quality was evaluated employing the uMARS tool [43], which includes 20 items, as described in the introduction. The items are grouped into 4 *objective* subdomains: *engagement* (5 items), *functionality* (4), *aesthetics* (3), *information* (4), and 1 additional domain of *subjective quality* (4). *Subjective quality* scale includes 4 items that assess the intention to use the app in the future (ie, "Would you recommend this app to people who might benefit from it?" and "How many times do you think you would use this app in the next 12 months if it was relevant to you?"), propensity to pay for it ("Would you pay for this app?"), and an overall 5-star rating ("What is your overall star rating of the app?"), which reflects the way app stores rate the apps. All uMARS items are assessed through 5-point scales. Subscales are computed by averaging the respective domain items. A total uMARS score is calculated by averaging all subdomains, whereas *subjective quality* is calculated by averaging its related subitems. In the source study, the uMARS tool showed good internal consistency (Cronbach alpha=.90) and good test-retest reliability [43].

Data Analyses

Survey data were summarized using descriptive statistics. Background characteristics were kept continuous (age), dichotomous (gender), or categorical (height and weight were used to compute body mass index, BMI). Following the International Physical Activity Questionnaire scoring protocol, physical activity was categorized as high, moderate, or low [59]. For uMARS items, answers categorized by users as "don't know/not applicable" were coded as missing. Missing value analysis was performed to estimate the frequency and level of missingness and determine the best strategy to address the issue (eg, multiple imputation [MI] and listwise deletion). Internal consistency (Cronbach alpha) was interpreted as excellent (≥.90), good (.80-.89), acceptable (.70-.79), questionable (.60-.69), poor (.50-.59), and unacceptable (<.50) [44].

As each app was evaluated by different groups of users, traditional interrater reliability (IRR) indices (ie, intraclass correlation coefficients, ICCs), reported in MARS and uMARS development studies, were not applicable [60,61]. To ensure that ratings could be aggregated, we evaluated interrater agreement (IRA) following literature recommendations [62,63], using 3 families of indices: James et al's $r_{WG(J)}$ [64,65] (based on multiple null distributions) [66], Brown et al's $a_{WG(J)}$ [67], and the adjusted average deviation index $A_{DMJ(adj)}$ [68]. IRA was established with pragmatic and theoretical cut-off points such as for the $r_{WG(J)}$: no agreement (<.29), weak (.30-.49), moderate (.50-.69), strong (.70-.89), and very strong (>.90) [64,65]; $a_{WG(J)}$: not acceptable (<.59), weak (.60-.69), moderate (.70-.79), and strong agreement (>.80) [67]; and $A_{DMJ(adj)}$: agreement above .80 [68]. Strong agreement was considered when all indices were consistently indicating an acceptable level of agreement.

In addition to the arithmetic mean of each uMARS score, we calculated a response data–based weighted mean (WDMEAN) [69]. The WDMEAN allows to incorporate individual raters' disagreements as it is calculated as the sum of each individual score multiplied by its weight, which is a function of the distance of the individual response from the unweighted group mean. This aggregation approach has been employed in organizational and management literature to summarize opinions from key informants who may not share the same knowledge about the object of study [70,71] and have some expected disagreement [69,70,72]. Unweighted and weighted mean scores (range: 1-5) were expressed as percent scores. The scale midpoint (3, converted in percent, assuming that 1=0%, 5=100%, and 3=50%) was considered the minimum level of acceptability, as reported in the study by Mani et al [56]. The WDMEAN, in presence of full agreement, would correspond to the arithmetic mean.

Considering the small sample size and the nature of the scores (which might be prone to non-normal distribution), associations among and with uMARS domain scores were examined by inspecting Spearman rho (ρ) coefficients. Total and uMARS subdomains were associated with subjective quality, as the associations among uMARS subdomains are not considered meaningful [43] or interpretable (Stoyanov S, personal communication, November 2017). Given the multiple tests, *P* values were corrected for type 1 error [73]. Mann-Whitney and Kruskal-Wallis (K-W) tests examined differences in continuous variables. Due to the exploratory nature of the study, no inferential statistics were attempted. All analyses were performed with IBM SPSS Statistics v.24 for Mac.

Results

Participant Recruitment

Invitations were sent to 600 randomly selected email addresses, and additional 145 employees were recruited through social media postings. Out of 745 potentially interested employees, 44 provided informed consent and 5 were ineligible. The remaining 39 employees successfully enrolled in the study. Moreover, 36 of them completed the app evaluations and were included in the analyses. Their characteristics are reported in Table 2. Employees were on average 36 years old (SD 10.8), mostly female (24/36, 67%), married (19/36, 53%), with a graduate-level university degree (16/36, 44%), earned less than US $2000 per month (17/36, 47%), and worked on average 48 hours per week (SD 11.9). The majority reported being in very good or excellent health status (16/36, 44%), normal weight (17/36, 47%), or overweight (16/36, 44%), and moderately active (28/36, 78%), spending on average 6.6 hours per day (SD 2.4) sitting. Most users owned an iOS device (21/36, 58%), and some had previously used apps for tracking physical activity (22/36, 60%), diet (8/36, 23%), or weight (4/36, 11%). A total of 6 participants had previously used 1 of the reviewed apps (*MyFitnessPal*). Group allocation was not associated with any background characteristic.

App Quality Evaluation

Of the 36 users, 14 (39%) provided complete data covering 91% of values across the 20 uMARS items. The highest proportion of missingness was in the 3 *information* items (*credibility of source*: 39%; *visual information*: 25%; and *quantity of information*: 22%) and in 1 *engagement* item (*customization*: 19%). As missing was completely at random (Little's missing completely at random test: χ^2_{264}=251.8; *P*=.69), MI was employed. We generated 10 complete datasets [74,75] and ran the analyses with both incomplete and complete datasets to ensure comparability of results. For clarity and accuracy, all uMARS scores presented here are based on pooled means and variance estimates obtained from the MI datasets.

Internal consistency and IRA estimates are reported in Multimedia Appendix 1. Overall, Cronbach alpha values varied across the uMARS subdomains, being acceptable for *engagement* (alpha=.75) and *aesthetics* (alpha=.71), questionable for *functionality* (alpha=.61), poor for *information* (alpha=.51), and good for *subjective quality* (alpha=.88). Within each app, alphas were good for *subjective quality* (median .82, range .74 [*My Diet Diary*] to .93 [*Lark*]), acceptable for *engagement* (median .71, range .46 [*My Diet Coach*] to .93 [*MyFitnessPal*]) and *aesthetics* (median .70, range .42 [*Lark*] to .86 [*SparkPeople*]), and unacceptable for *information* (median .23, range .15 [*SparkPeople*] to .46 [*MyPlate*]). Negative alpha values were found among *engagement* and *information* items (*SparkPeople* and *MyFitnessPal* groups, respectively), indicating negative correlations among those items. IRA indices suggested overall agreement among users in most subdomains and for most apps. Moderate to strong agreement was found in *functionality* and *aesthetics* (all apps), whereas low agreement was found in *engagement* (*MyFitnessPal* and *My Diet Diary*), *information* (*My Diet Diary*, *MyPlate*, and *SparkPeople*), and *subjective quality* (*Lark*, *MyFitnessPal*, and *My Diet Diary*).

Table 2. Characteristics of study participants according to app group and total sample (n=36).

Participants' characteristics	*Lark* (n=7)	*MyFitnessPal* (n=6)	*My Diet Coach* (n=6)	*My Diet Diary* (n=5)	*MyPlate* (n=6)	*SparkPeople* (n=6)	Total sample (n=36)	*P* value
Sociodemographics								
Age (years), mean (SE)	39.7 (5.3)	41.5 (3.8)	29.8 (2.6)	31.2 (4.5)	38.7 (4.4)	31.5 (4.1)	35.6 (1.8)	.14
Gender (female), n (%)	2 (29)	3 (50)	5 (83)	4 (80)	5 (83)	5 (83)	24 (67)	.19
Marital status, n (%)								.95
Single	1 (14)	2 (33)	2 (33)	1 (20)	2 (33)	3 (50)	11 (31)	
Engaged or in a relationship	1 (14)	1 (17)	1 (17)	2 (40)	1 (17)	0 (0)	6 (17)	
Married	5 (71)	3 (50)	3 (50)	2 (40)	3 (50)	3 (50)	19 (53)	
Education, n (%)							.06	
High school (secondary)	0 (0)	1 (17)	0 (0)	0 (0)	0 (0)	0 (0)	1 (3)	
Bachelor	3 (43)	0 (0)	0 (0)	4 (81)	1 (17)	2 (33)	10 (28)	
Master	3 (43)	2 (40)	3 (60)	0 (0)	5 (83)	3 (50)	16 (44)	
PhD	1 (14)	3 (60)	2 (40)	1 (20)	0 (0)	1 (17)	8 (22)	
Income (n=33), n (%)								.57
<$US 2000	3 (43)	2 (33)	3 (50)	3 (60)	3 (50)	3 (50)	17 (47)	
$US 2001 to $US 4000	0 (0)	2 (33)	1 (17)	1 (20)	2 (33)	3 (50)	9 (25)	
>US $4000	2 (29)	2 (33)	2 (33)	1 (20)	0 (0)	0 (0)	7 (19)	
Working hours per week (n=35), mean (SE)	46.7 (3.1)	43.3 (2.1)	45.8 (2.4)	38.0 (9.6)	45.0 (4.1)	35.0 (7.7)	42.8 (2.0)	.65
Health and behavioral characteristics								
Health status, n (%)								.49
Poor or fair	3 (43)	2 (33)	3 (50)	0 (0)	0 (0)	2 (33)	10 (28)	
Good	1 (14)	2 (33)	0 (0)	2 (40)	4 (67)	1 (17)	10 (28)	
Very good or excellent	3 (43)	2 (33)	3 (50)	3 (60)	2 (33)	3 (50)	16 (44)	
BMI[a] category, n (%)								.32
Normal weight	1 (14)	3 (50)	3 (50)	3 (60)	2 (33)	5 (83)	17 (47)	
Overweight	4 (57)	2 (33)	3 (50)	2 (40)	4 (67)	1 (17)	16 (44)	
Obese and morbidly obese	2 (29)	1 (17)	0 (0)	0 (0)	0 (0)	0 (0)	3 (8)	
Activity level, n (%)[b]								.32
High	3 (43)	0 (0)	2 (33)	1 (20)	2 (33)	0 (0)	8 (22)	
Moderate	4 (57)	6 (100)	4 (67)	4 (80)	4 (67)	6 (100)	28 (78)	
Sitting time (hours per day; n=35), mean (SE)	7.1 (0.8)	6.7 (0.3)	8.4 (1.4)	6.8 (0.1)	4.4 (1.2)	6.2 (0.6)	6.6 (0.4)	.14
Mobile phone use and mobile health (mHealth) app use								
Operative system (iOS), n (%)	3 (43)	5 (83)	5 (83)	2 (40)	4 (67)	2 (33)	21 (58)	.29
Past experience with mHealth apps (n=35)[c], n (%)								
Used apps to track physical activity	3 (43)	6 (100)	4 (67)	2 (40)	4 (67)	2 (33)	21 (60)	.18
Used apps to track diet	3 (43)	0 (0)	3 (50)	0 (0)	2 (33)	0 (0)	8 (23)	.09
Used apps to monitor weight	0 (0)	1 (17)	2 (33)	0 (0)	1 (17)	0 (0)	4 (11)	.34
Never used mHealth apps	1 (14)	0 (0)	2 (33)	3 (60)	2 (33)	4 (67)	12 (34)	.12
Use of listed apps in the past 6 months (n=6)[c], n (%)								
MyFitnessPal	2 (29)	1 (17)	1 (17)	0 (0)	2 (33)	0 (0)	6 (17)	.64

[a]BMI: body mass index.

[b]Categorization based on the International Physical Activity Questionnaire scoring protocol [59].

[c]Multiple choice questions. *P* values represent the significance level of chi-square test (categorical variable) or Kruskal-Wallis test (continuous variables).

The unweighted, WDMEANs, and percent scores are presented in Table 3. Unweighted and WDMEANs were practically the same, with the former being generally lower than the latter. *Information* was the domain with the largest difference between unweighted and WDMEAN (1.3%), followed by *engagement* and *functionality* (both 1%), *aesthetics* and total uMARS score (0.1%). *Subjective quality* scores were also similar, with the highest difference in *Lark* (−2.4%).

Overall, all apps scored above the minimum threshold for acceptability (50%) in the total uMARS score and its main 4 subdomains. *Functionality* was the highest rated domain (median 3.9, interquartile range [IQR] 0.3), followed by *aesthetics* (median 3.7, IQR 0.5), *information* (median 3.7, IQR 0.1), and *engagement* (median 3.3, IQR 0.2). The *subjective quality* score was low (median 2.5, IQR 0.9). The scores are presented in the boxplot below (Figure 1). Only 2 apps (*MyPlate* and *Lark*) scored above the median thresholds in both uMARS and subjective quality scores.

After applying the Bonferroni correction for *P* values (*P*=.01), *subjective quality* was strongly and positively related (*P*<.001) with total uMARS score (ρ=.75), *engagement* (ρ=.68), *information*, and *aesthetics* (ρ=.60) and not significantly related with *functionality* (ρ=.40; *P*=.02).

Associations With Users' Characteristics

Correlations with users' background characteristics are reported in Table 4. After applying the appropriate *P* value corrections for multiple correlation tests [73], good health status was associated with *engagement*, total uMARS, and *subjective quality*; being obese with total uMARS score; and use of *Lark* with *functionality* and *information*. Very good or excellent health status was negatively related to *engagement*; use of *SparkPeople* was negatively related to *information*. K-W tests revealed significant differences across health status groups in *engagement* (χ^2_2=11.9; *P*=.003), total uMARS (χ^2_2=9.4; *P*=.009), and *subjective quality* (χ^2_2=8.1; *P*=.02). Participants in good health status had higher median scores than those of the other 2 groups. Similarly, the 3 BMI categories (normal, overweight, and obese) scored significantly different in *engagement* (χ^2_2=6.8; *P*=.03), *functionality* (χ^2_2=6.1; *P*=.05), total uMARS score (χ^2_2=6.6; *P*=.04), and *subjective quality* (χ^2_2=6.11; *P*=.05). Obese individuals had higher median scores than those of the other 2 groups. Finally, K-W tests showed significasent differences among app groups in *information* (χ^2_5=14.4, *P*=.01) and total uMARS score (χ^2_5=12.4; *P*=.03). Users of *Lark* reported larger median *information* and total uMARS scores than the other apps. In *Lark*, *subjective quality* was positively associated with *engagement* (ρ=.87; *P*=.007) and total app quality (ρ=.90; *P*=.006). In *SparkPeople*, *subjective quality* was positively related to *information* (ρ=.97; *P*<.001).

Table 3. Comparison of user-based unweighted and weighted user version of the Mobile App Rating Scale scores.

App quality domains	Mean (SD)	Percent mean score	WDMEAN[a]	Percent WDMEAN score
Engagement				
Lark	3.41 (0.84)	68.2	3.37	67.4
MyFitnessPal	3.34 (1.22)	66.8	3.40	68.0
My Diet Coach	3.32 (0.58)	66.4	3.44	68.8
My Diet Diary	2.83 (0.83)	56.6	2.86	57.2
MyPlate	3.36 (0.53)	67.2	3.39	67.8
SparkPeople	3.05 (0.39)	61.0	3.17	63.4
Functionality				
Lark	4.32 (0.53)	86.4	4.39	87.8
MyFitnessPal	3.94 (0.53)	78.8	4.00	80.0
My Diet Coach	3.82 (0.51)	76.4	3.80	76.0
My Diet Diary	3.64 (0.61)	72.8	3.63	72.6
MyPlate	4.04 (0.49)	80.8	4.17	83.4
SparkPeople	3.45 (0.54)	69.0	3.49	69.8
Aesthetics				
Lark	3.98 (0.74)	79.6	3.99	79.8
MyFitnessPal	3.61 (0.49)	72.2	3.64	72.8
My Diet Coach	3.72 (0.65)	74.4	3.85	77.0
My Diet Diary	3.40 (0.55)	68.0	3.35	67.0
MyPlate	4.00 (0.42)	80.0	4.00	80.0
SparkPeople	3.17 (0.81)	62.0	3.10	62.0
Information				
Lark	4.24 (0.60)	84.8	4.31	86.2
MyFitnessPal	3.70 (0.73)	74.0	3.79	75.8
My Diet Coach	3.56 (0.64)	71.2	3.57	71.4
My Diet Diary	3.61 (0.53)	72.2	3.60	72.0
MyPlate	3.70 (0.79)	74.0	3.76	75.2
SparkPeople	3.03 (0.87)	60.6	3.10	62.0
Total score				
Lark	3.98 (0.50)	79.2	3.96	79.2
MyFitnessPal	3.65 (0.55)	74.0	3.70	74.0
My Diet Coach	3.60 (0.43)	71.6	3.58	71.6
My Diet Diary	3.37 (0.38)	65.8	3.29	65.8
MyPlate	3.78 (0.40)	76.2	3.81	76.2
SparkPeople	3.17 (0.45)	64.2	3.21	64.2
Subjective quality				
Lark	3.25 (1.40)	65.0	3.37	67.4
MyFitnessPal	2.70 (1.04)	54.0	2.73	54.6
My Diet Coach	2.20 (0.76)	44.0	2.20	44.0
My Diet Diary	2.25 (0.66)	45.0	2.24	44.8
MyPlate	3.30 (0.84)	66.0	3.27	65.4
SparkPeople	2.08 (0.68)	41.6	2.08	41.6

[a]WDMEAN: response data–based weighted mean [69].

Figure 1. Boxplots of user version of the Mobile App Rating Scale subdomains and subjective quality with scatterplot representing each app.

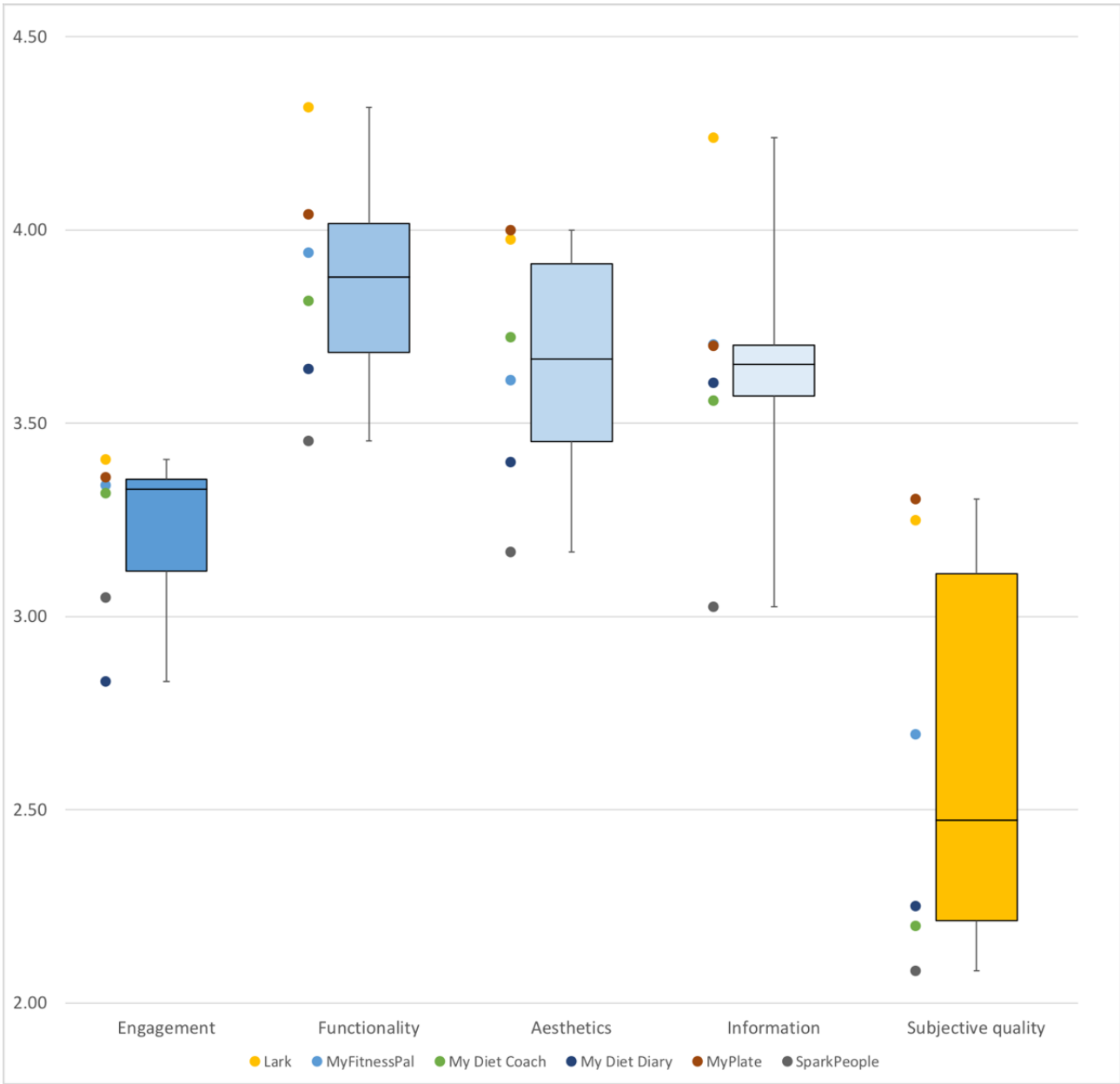


Table 4. Correlations between user version of the Mobile App Rating Scales and users' background characteristics.

Participants' characteristics	User version of the Mobile App Rating Scales					
	Engagement	Functionality	Aesthetics	Information	Total score	Subjective quality
Sociodemographics						
Age (years)	−0.11	0.34[a]	−0.03	0.17	0.09	0.29
Gender: female	−0.14	−0.05	−0.17	−0.18	−0.15	−0.16
Marital status						
Single	0.24	−0.19	0.06	−0.16	0.04	−0.06
Engaged	−0.05	−0.04	−0.16	0.17	0.02	−0.04
Married	−0.18	0.21	0.06	0.02	−0.05	0.09
Education						
High school	0.28	0.10	0.10	0.08	0.19	0.12
Bachelor	−0.11	−0.01	−0.06	0.03	−0.01	−0.06
Master	−0.05	−0.16	0.06	−0.09	−0.09	−0.19
PhD	0.06	0.16	−0.06	0.04	0.04	0.24
Income						
<US $2000	0.14	0.09	−0.03	0.06	0.13	−0.11
<US $3000	0.00	−0.23	−0.03	−0.07	−0.07	0.06
<US $4000	−0.30	−0.12	−0.08	−0.06	−0.18	−0.20
>US $4000	−0.02	0.19	−0.05	0.18	0.05	0.31
Working hours per week	−0.10	0.17	0.15	0.12	0.10	0.08
Health and behavioral characteristics						
Health status						
Poor or fair	0.01	0.02	−0.03	−0.07	−0.05	−0.22
Good	0.54[b]	0.23	0.36[a]	0.34[a]	0.50[b]	0.48[b]
Very good or excellent	−0.49[b]	−0.23	−0.30	−0.25	−0.40[a]	−0.24
Body mass index						
Normal weight	0.11	−0.29	−0.11	−0.28	−0.17	0.01
Overweight	−0.32	0.08	−0.11	0.13	−0.07	−0.23
Obese	0.38[a]	0.37[a]	0.38[a]	0.27	0.43[b]	0.40[a]
Activity level: high	0.17	−0.02	0.07	0.11	0.09	−0.05
Sitting time (hours per day)	0.12	−0.11	−0.06	−0.06	−0.03	0.04
Mobile phone use and mobile health (mHealth) app use						
Mobile operative system: iOS	0.17	−0.01	0.25	0.18	0.18	0.22
Past experience with mHealth apps						
Used apps to track physical activity	0.08	−0.13	−0.06	0.16	0.01	0.08
Used apps to track diet	−0.16	0.26	0.05	0.22	0.10	−0.13
Used apps to monitor weight	0.13	0.32	0.05	0.20	0.20	0.01
Never used mHealth apps	−0.03	−0.07	−0.03	−0.29	−0.14	−0.05
App used in the study						
Used *Lark*	0.08	0.43[b]	0.35[a]	0.47[b]	0.42[a]	0.24
Used *MyFitnessPal*	0.08	0.05	−0.09	0.07	0.04	0.05
Used *My Diet Coach*	0.06	−0.07	0.03	−0.16	−0.09	−0.21

Participants' characteristics	User version of the Mobile App Rating Scales					
	Engagement	Functionality	Aesthetics	Information	Total score	Subjective quality
Used *My Diet Diary*	0.17	0.02	−0.12	−0.03	0.04	−0.01
Used *MyPlate*	0.10	0.11	0.24	0.19	0.22	0.30
Used *SparkPeople*	−0.12	−0.33[a]	−0.30	−0.47[b]	−0.37[a]	−0.24

[a]*P*<.05.

[b]*P*<.001. With Bonferroni correction, the significance value becomes *P*<.0003.

Discussion

Principal Findings

This is the first study that explored how laypersons evaluated the quality of free and popular mobile phone apps for weight management, using the uMARS tool [43]. The tool showed acceptable internal consistency levels in most subdomains, except for *information* (alpha=.51). Heterogeneity in alpha values was found within each app group. In 2 cases (*SparkPeople* for *engagement* and *MyFitnessPal* for *information*), alphas assumed negative values, which indicate small, negative correlations among the items in those subscales and lack of consistency. The internal consistencies we found are below those reported in the uMARS source study [43] and below the levels commonly recommended by the literature, suggesting large measurement errors [76]. Low alphas might be because of the number of items and sample size [77]. In addition, users might have had different interpretations of the items as some IRA indices pointed to low or no agreement within *engagement* items (*MyFitnessPal* and *My Diet Diary*), *information* (*My Diet Diary*, *MyPlate*, and *SparkPeople*), and *subjective quality* (*Lark*, *MyFitnessPal*, and *My Diet Diary*). Although IRA does not imply reliability [62,63], low agreement suggests a large degree of subjectivity in evaluating the apps, which can be expected, as the users are supposed to be free to have their own opinions about the apps, based on their own characteristics and needs.

Furthermore, large item nonresponse rates were registered in the *information* domain (22%-39%). Some users might have misunderstood these items or might not have known how to answer, thus leaving them blank. The missing information might explain the poor consistency and low agreement estimates in this specific domain. Unfortunately, the uMARS source study does not provide solutions in case of poor internal consistency or low agreement [43], and other studies employing uMARS did not report such issues [54,55]. To account for these limitations, we calculated the WDMEAN [69], an approach that allowed to retain all items. Eventually, the unweighted and weighted means were very similar, suggesting that applying the uMARS scoring protocol can still yield robust results. Nevertheless, the uMARS tool should be generalized to weight management apps, with larger user populations. We also recommend exploring users' perceptions about the items including qualitative methodologies such as the *think aloud method* [78].

In this study, we employed the WDMEAN approach to estimate the responses from our key informants who were asked to apply the uMARS tool without previous training. To the best of our knowledge, no uMARS and MARS studies have used this approach, employing users who have undergone some level of training. This is the first study that utilizes the tool for users. By employing the WDMEAN method, it is possible to estimate app quality while accounting for the respondents' potential disagreements, hence providing a *truer* average score, which accounts for the response of each individual [69]. On the contrary, the arithmetic mean can be influenced by extreme values (either very low or very high scores), and at the same time, it might reduce the intrinsic variability among raters' ratings. The WDMEAN approach can be applied to many other studies, with small samples, in which researchers are interested in estimating scores while accounting for the agreement or disagreement among raters.

The second objective was to understand which app quality dimensions (ie, engagement, functionality, aesthetics, and information) contributed the most to the overall app quality score. All apps scored high in *functionality*, followed by *aesthetics*, *information*, and *engagement*. This is consistent with some qualitative research suggesting that users appreciate functional and aesthetic characteristics [8,9,41]. This is also consistent with the findings reported in the expert review, upon which this study is based, as the apps were deemed highly functional and with limited information quality [30]. However, *engagement*, *aesthetics*, and *information* appeared to be strongly related with *subjective quality*, which includes questions that indicate the propensity of using the apps in the future ("Would you recommend...," "Would you pay," "How many times would you use it...?," and "What is the overall star rating?"). This might indicate that users might not engage with these apps regardless of their good functional features. This is consistent with findings from qualitative studies, which show that users might stop using an app not because of technical features but rather because of low engagement or hidden costs [8,9]. Another important consideration was that in our study, *subjective quality* was only weakly correlated with *functionality* (ρ=.40; *P*=.02). Conversely, *engagement* had the strongest correlation with *subjective quality* (ρ=.68; *P*<.001). This might indicate that app engagement can play an important role in achieving sustained app usage [10,12]; however, future studies should be conducted to establish whether a causal link between engagement and future app use exists.

The third objective was to explore the associations between uMARS scales and users' characteristics. In this sample, we found that obese users and those in good health status provided higher app quality ratings in *engagement*, total uMARS, and *subjective quality*. In other words, healthy, obese individuals

perceived these apps particularly engaging and of high quality. As engagement is related to app usage [1], these individuals might be more likely to use the apps in the future. These findings are particularly suggestive, as these popular weight management apps (in particular, *Lark*, *MyFitnessPal*, and *MyPlate*) may be used in interventions addressing obesity prevention (healthy volunteers) and treatment (obese) [9]. Future research could test whether these apps, which had demonstrated having high behavior change potential [30], can effectively influence behavior and promote weight loss among overweight or obese individuals. This study informed the development of a self-directed weight control intervention, which targets the same population (clinical trial registry: NCT03321331).

Limitations

The results of this study need to be interpreted bearing in mind its limitations. A major limitation is the design (noncrossover). For feasibility reasons (budget and time constraints), we could not ask all users to evaluate each app, hence allowing us to calculate IRR using ICC indices. To overcome this limitation, we employed methodological solutions that have never been employed in similar studies (ie, IRA estimates [60,62] and WDMEAN [69]). These solutions allowed us to ensure the robustness of the responses obtained from the employees recruited in this study. This solution is pragmatic and allows to be applied in real-life scenarios, whereby research study participants might not be willing or able to dedicate more time to the study. Moreover, users evaluated the free version of the apps used for 2 weeks. Ratings might have differed if they had used the *pro* versions with additional functionalities. App evaluation might also be influenced by actual app use and by the amount of time spent on each app. As the authors of the expert review noted [30], some apps prompt different feedback and unlock features only after repeated use. We instructed participants to use the apps at least daily for 2 weeks, but we did not assess actual app use. Another limitation is the sampling of this study as we had access to a convenience sample of employees from an academic institution in Lebanon, who voluntarily agreed to participate. Although we found correlations with health status and BMI categories, this study might not be generalizable to the entire population and to other cultural contexts and settings, as we recruited mostly female, educated, and healthy individuals. The small sample size is also another limitation; however, the size was based on pragmatic considerations and aligned with recommendations from the heuristic evaluation literature [57,58]. Larger samples should investigate whether these findings hold truth in different segments of the population. It will be practical to focus studies on specific segments of the population to increase the accuracy of the findings. Nevertheless, we believe the results are generalizable to similar academic institutions in Lebanon or in the Middle East region or who have similar employee populations, although the tested apps are available internationally. Another limitation is the use of self-reported data and self-administered Web-based surveys that are prone to missing data. We used Web-based tools because we wanted to avoid interviewer bias and we did not want to interfere with the users' evaluations of the apps. We wanted the users to test the apps *in the wild* for 2 weeks, without specialized training, which is usually a prerequisite of expert reviews. We could have used interviewers to reduce data entry mistakes or inconsistencies, but we opted for self-administered Web-based forms to avoid interviewer bias. A related limitation is the presence of large amounts of missing data in some of the subdomains of the uMARS scale (eg *information* domain), which forced us to apply caution when interpreting the results. Although we employed modern techniques to deal with missing data, we cannot make strong assumptions on the reasons for the missing responses backed on data, as the instrument (Web-based survey) did not capture comments related to the uMARS scale. We recommend that future studies investigate how users respond to the survey and how they apply the answers. We have already suggested that qualitative techniques such as the *think aloud method* [78] could be applied to understand the thought processes that people use when answering questionnaires. These techniques would allow to identify potential pitfalls in the scale, hence improving its validity across cultures and sample populations.

Conclusions

Across the 6 popular and free weight management apps analyzed in this study, *functionality* is the quality dimension that laypersons valued the most. However, *engagement* was strongly associated with *subjective quality*, a dimension that includes future app use. The higher the subjective quality and engagement, the more likely users might use the app. App developers and public health professionals should ensure that an app is both functional and engaging so that users will be more likely to use it. Future longitudinal studies are needed to ascertain this connection.

The tested apps (in particular *Lark* and *MyPlate*) were perceived as more engaging and of higher quality among healthy, obese individuals, making them promising modes of delivery for obesity prevention and treatment interventions.

From a methodological standpoint, the uMARS tool is a practical and feasible tool that can be used to assess app quality by laypersons without specialized training. However, further research is needed to establish its validity in the domain of weight management.

Acknowledgments

The authors would like to thank all employees who participated in the study; Ms Fida Abou Hassan and Ms Hiba Al-Shami, who, respectively, provided help in the planning of the study and implementation; and Ms Laya Samaha who provided feedback on this paper.

This study was funded by the University Research Board of the American University of Beirut (grant number 103182). The publication of this manuscript was partially supported by the Faculty of Health Sciences (MB's seed fund) and by individual contributions (MB and GH).

Authors' Contributions

MB conceptually designed and supervised the study implementation, performed the analyses, drafted the manuscript, and incorporated all feedback from the coauthors. AA provided intellectual input to the study design, assisted in the implementation of the study, performed preliminary analyses, edited and provided feedback on the different versions of the manuscript. FD assisted in the conduct of the study, provided intellectual input to the study and manuscript. GH provided intellectual input to the design and execution of the study and edited and provided feedback on the different versions of the manuscript. All authors read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Internal consistency and interrater agreement (IRA) indices for the user version of the Mobile App Rating Scale scores.

[PDF File (Adobe PDF File), 24KB - mhealth_v7i1e9836_app1.pdf]

References

1. 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]
2. Research 2 Guidance. Mhealth app developer economics 2017 Internet. Berlin, Germany: research2guidance; 2017. URL: https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/ [accessed 2018-11-29] [WebCite Cache ID 6wOhZ1qY6]
3. The Lancet. Does mobile health matter? Lancet 2017 Nov 18;390(10109):2216. [doi: 10.1016/S0140-6736(17)32899-4] [Medline: 29165258]
4. Research 2 Guidance. Mhealth app developer economics 2016 Internet. Berlin, Germany: research2guidance; 2016. URL: https://research2guidance.com/product/mhealth-app-developer-economics-2016/ [accessed 2018-01-12] [WebCite Cache ID 6wOhjVDEV]
5. 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]
6. Nikolaou CK, Lean ME. Mobile applications for obesity and weight management: current market characteristics. Int J Obes (Lond) 2017 Dec;41(1):200-202. [doi: 10.1038/ijo.2016.186] [Medline: 27780974]
7. Research 2 Guidance. mHealth App Market Sizing 2015-2020 Internet. Berlin, Germany: research2guidance; 2015. URL: http://research2guidance.com/product/mhealth-app-market-sizing-2015-2020/ [accessed 2018-11-29] [WebCite Cache ID 74HxJhKRi]
8. Alnasser AA, Alkhalifa AS, Sathiaseelan A, Marais D. What overweight women want from a weight loss app: a qualitative study on arabic women. JMIR Mhealth Uhealth 2015 May 20;3(2):e41 [FREE Full text] [doi: 10.2196/mhealth.4409] [Medline: 25993907]
9. Tang J, Abraham C, Stamp E, Greaves C. How can weight-loss app designers' best engage and support users? A qualitative investigation. Br J Health Psychol 2015 Feb;20(1):151-171. [doi: 10.1111/bjhp.12114] [Medline: 25130682]
10. Hoj TH, Covey EL, Jones AC, Haines AC, Hall PC, Crookston BT, et al. How do apps work? An analysis of physical activity app users' perceptions of behavior change mechanisms. JMIR Mhealth Uhealth 2017 Aug 03;5(8):e114 [FREE Full text] [doi: 10.2196/mhealth.7206] [Medline: 28778846]
11. Turner-McGrievy GM, Beets MW, Moore JB, Kaczynski AT, Barr-Anderson DJ, Tate DF. Comparison of traditional versus mobile app self-monitoring of physical activity and dietary intake among overweight adults participating in an mHealth weight loss program. J Am Med Inform Assoc 2013 May 01;20(3):513-518 [FREE Full text] [doi: 10.1136/amiajnl-2012-001510] [Medline: 23429637]
12. West J, Belvedere L, Andreasen R, Frandsen C, Hall P, Crookston B. Controlling your "app"etite: how diet and nutrition-related mobile apps lead to behavior change. JMIR Mhealth Uhealth 2017 Jul 10;5(7):e95 [FREE Full text] [doi: 10.2196/mhealth.7410] [Medline: 28694241]
13. Semper HM, Povey R, Clark-Carter D. A systematic review of the effectiveness of smartphone applications that encourage dietary self-regulatory strategies for weight loss in overweight and obese adults. Obes Rev 2016 Dec;17(9):895-906. [doi: 10.1111/obr.12428] [Medline: 27192162]

14. Aguilar-Martínez A, Solé-Sedeño JM, Mancebo-Moreno G, Medina FX, Carreras-Collado R, Saigí-Rubió F. Use of mobile phones as a tool for weight loss: a systematic review. J Telemed Telecare 2014 Sep;20(6):339-349. [doi: 10.1177/1357633X14537777] [Medline: 24875928]
15. Bardus M, Smith JR, Samaha L, Abraham C. Mobile phone and web 2.0 technologies for weight management: a systematic scoping review. J Med Internet Res 2015 Nov 16;17(11):e259 [FREE Full text] [doi: 10.2196/jmir.5129] [Medline: 26573984]
16. Quelly SB, Norris AE, DiPietro JL. Impact of mobile apps to combat obesity in children and adolescents: a systematic literature review. J Spec Pediatr Nurs 2016 Jan;21(1):5-17. [doi: 10.1111/jspn.12134] [Medline: 26494019]
17. Riaz S, Sykes C. Are smartphone health applications effective in modifying obesity and smoking behaviours? A systematic review. Health Technol 2015 Mar 25;5(2):73-81. [doi: 10.1007/s12553-015-0104-4]
18. Flores Mateo G, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015 Nov 10;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
19. Bardus M, Smith JR, Samaha L, Abraham C. Mobile and Web 2.0 interventions for weight management: an overview of review evidence and its methodological quality. Eur J Public Health 2016 Aug;26(4):602-610. [doi: 10.1093/eurpub/ckw090] [Medline: 27335330]
20. Schippers M, Adam PC, Smolenski DJ, Wong HT, de Wit JB. A meta-analysis of overall effects of weight loss interventions delivered via mobile phones and effect size differences according to delivery mode, personal contact, and intervention intensity and duration. Obes Rev 2017 Dec;18(4):450-459. [doi: 10.1111/obr.12492] [Medline: 28187246]
21. Pagoto S, Schneider K, Jojic M, DeBiasse M, Mann D. Evidence-based strategies in weight-loss mobile apps. Am J Prev Med 2013 Nov;45(5):576-582. [doi: 10.1016/j.amepre.2013.04.025] [Medline: 24139770]
22. Tang JC, Abraham C, Greaves CJ, Nikolaou V. Self-directed interventions to promote weight loss: a systematic review and meta-analysis. Health Psychol Rev 2016 Sep;10(3):358-372. [doi: 10.1080/17437199.2016.1172979] [Medline: 27091296]
23. Tang J, Abraham C, Greaves C, Yates T. Self-directed interventions to promote weight loss: a systematic review of reviews. J Med Internet Res 2014 Feb 19;16(2):e58 [FREE Full text] [doi: 10.2196/jmir.2857] [Medline: 24554464]
24. Fukuoka Y, Gay CL, Joiner KL, Vittinghoff E. A novel diabetes prevention intervention using a mobile app: a randomized controlled trial with overweight adults at risk. Am J Prev Med 2015 Aug;49(2):223-237 [FREE Full text] [doi: 10.1016/j.amepre.2015.01.003] [Medline: 26033349]
25. Hartman SJ, Nelson SH, Cadmus-Bertram LA, Patterson RE, Parker BA, Pierce JP. Technology- and phone-based weight loss intervention: pilot RCT in women at elevated breast cancer risk. Am J Prev Med 2016 Dec;51(5):714-721 [FREE Full text] [doi: 10.1016/j.amepre.2016.06.024] [Medline: 27593420]
26. Ipjian ML, Johnston CS. Smartphone technology facilitates dietary change in healthy adults. Nutrition 2017 Jan;33:343-347. [doi: 10.1016/j.nut.2016.08.003] [Medline: 27742102]
27. Laing BY, Mangione CM, Tseng C, Leng M, Vaisberg E, Mahida M, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. Ann Intern Med 2014 Nov 18;161(10 Suppl):S5-12 [FREE Full text] [doi: 10.7326/M13-3005] [Medline: 25402403]
28. Jospe MR, Roy M, Brown RC, Williams SM, Osborne HR, Meredith-Jones KA, et al. The effect of different types of monitoring strategies on weight loss: a randomized controlled trial. Obesity (Silver Spring) 2017 Dec;25(9):1490-1498 [FREE Full text] [doi: 10.1002/oby.21898] [Medline: 28703448]
29. Nollen NL, Mayo MS, Carlson SE, Rapoff MA, Goggin KJ, Ellerbeck EF. Mobile technology for obesity prevention: a randomized pilot study in racial- and ethnic-minority girls. Am J Prev Med 2014 Apr;46(4):404-408 [FREE Full text] [doi: 10.1016/j.amepre.2013.12.011] [Medline: 24650843]
30. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int J Behav Nutr Phys Act 2016 Mar 10;13:35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
31. Direito A, Dale LP, Shields E, Dobson R, Whittaker R, Maddison R. Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? BMC Public Health 2014 Jun 25;14:646 [FREE Full text] [doi: 10.1186/1471-2458-14-646] [Medline: 24965805]
32. Conroy DE, Yang C, Maher JP. Behavior change techniques in top-ranked mobile apps for physical activity. Am J Prev Med 2014 Jun;46(6):649-652. [doi: 10.1016/j.amepre.2014.01.010] [Medline: 24842742]
33. Cowan LT, Van Wagenen SA, Brown BA, Hedin RJ, Seino-Stephan Y, Hall PC, et al. Apps of steel: are exercise apps providing consumers with realistic expectations?: a content analysis of exercise apps for presence of behavior change theory. Health Educ Behav 2013 Apr;40(2):133-139. [doi: 10.1177/1090198112452126] [Medline: 22991048]
34. Lyons EJ, Lewis ZH, Mayrsohn BG, Rowland JL. Behavior change techniques implemented in electronic lifestyle activity monitors: a systematic content analysis. J Med Internet Res 2014 Aug 15;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]
35. Azar KM, Lesser LI, Laing BY, Stephens J, Aurora MS, Burke LE, et al. Mobile applications for weight management: theory-based content analysis. Am J Prev Med 2013 Nov;45(5):583-589. [doi: 10.1016/j.amepre.2013.07.005] [Medline: 24139771]

36. Fogg B. The Future of Persuasion is Mobile. In: Fogg B, Eckles D, editors. Mobile persuasion: 20 perspectives on the future of behavior change. Stanford, California: Stanford Captology Media; 2007:5-15.
37. Matthews J, Win KT, Oinas-Kukkonen H, Freeman M. Persuasive technology in mobile applications promoting physical activity: a systematic review. J Med Syst 2016 Mar;40(3):72. [doi: 10.1007/s10916-015-0425-x] [Medline: 26748792]
38. BinDhim NF, Hawkey A, Trevena L. A systematic review of quality assessment methods for smartphone health apps. Telemed J E Health 2015 Feb;21(2):97-104. [doi: 10.1089/tmj.2014.0088] [Medline: 25469795]
39. 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]
40. Zhu H, Xiong H, Ge Y, Chen E. Discovery of ranking fraud for mobile apps. IEEE Trans Knowl Data Eng 2015 Jan;27(1):74-87. [doi: 10.1109/TKDE.2014.2320733]
41. Solbrig L, Jones R, Kavanagh D, May J, Parkin T, Andrade J. People trying to lose weight dislike calorie counting apps and want motivational support to help them achieve their goals. Internet Interv 2017 Mar;7:23-31 [FREE Full text] [doi: 10.1016/j.invent.2016.12.003] [Medline: 28286739]
42. 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]
43. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the mobile application rating scale (uMARS). JMIR Mhealth Uhealth 2016 Jun 10;4(2):e72 [FREE Full text] [doi: 10.2196/mhealth.5849] [Medline: 27287964]
44. Domnich A, Arata L, Amicizia D, Signori A, Patrick B, Stoyanov S, et al. Development and validation of the Italian version of the Mobile Application Rating Scale and its generalisability to apps targeting primary prevention. BMC Med Inform Decis Mak 2016 Dec 07;16:83 [FREE Full text] [doi: 10.1186/s12911-016-0323-2] [Medline: 27387434]
45. Wilson H, Stoyanov SR, Gandabhai S, Baldwin A. The quality and accuracy of mobile apps to prevent driving after drinking alcohol. JMIR Mhealth Uhealth 2016 Aug 08;4(3):e98 [FREE Full text] [doi: 10.2196/mhealth.5961] [Medline: 27502956]
46. Sullivan RK, Marsh S, Halvarsson J, Holdsworth M, Waterlander W, Poelman MP, et al. Smartphone apps for measuring human health and climate Change co-benefits: a comparison and quality rating of available apps. JMIR Mhealth Uhealth 2016 Dec 19;4(4):e135 [FREE Full text] [doi: 10.2196/mhealth.5931] [Medline: 27993762]
47. 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]
48. Rickard N, Arjmand H, Bakker D, Seabrook E. Development of a mobile phone app to support self-monitoring of emotional well-being: a mental health digital innovation. JMIR Ment Health 2016 Nov 23;3(4):e49 [FREE Full text] [doi: 10.2196/mental.6202] [Medline: 27881358]
49. Zini F, Reinstadler M, Ricci F. Increasing Quality of Life Awareness with Life-Logging. In: Giokas K, Bokor L, Hopfgartner F, editors. eHealth 360°. Cham: Springer; 2017:282-293.
50. Grainger R, Townsley H, White B, Langlotz T, Taylor WJ. Apps for people with Rheumatoid arthritis to monitor their disease activity: a review of apps for best practice and quality. JMIR Mhealth Uhealth 2017 Feb 21;5(2):e7 [FREE Full text] [doi: 10.2196/mhealth.6956] [Medline: 28223263]
51. Patel R, Sulzberger L, Li G, Mair J, Morley H, Shing MN, et al. Smartphone apps for weight loss and smoking cessation: quality ranking of 120 apps. N Z Med J 2015 Sep 04;128(1421):73-76. [Medline: 26370762]
52. Goyal S, Lewis G, Yu C, Rotondi M, Seto E, Cafazzo JA. Evaluation of a behavioral mobile phone app intervention for the self-management of type 2 diabetes: randomized controlled trial protocol. JMIR Res Protoc 2016 Aug 19;5(3):e174 [FREE Full text] [doi: 10.2196/resprot.5959] [Medline: 27542325]
53. Mareva S, Thomson D, Marenco P, Estal Muñoz V, Ott C, Schmidt B, et al. Study protocol on ecological momentary assessment of health-related quality of life using a smartphone application. Front Psychol 2016;7:1086 [FREE Full text] [doi: 10.3389/fpsyg.2016.01086] [Medline: 27486425]
54. Panatto D, Domnich A, Gasparini R, Bonanni P, Icardi G, Amicizia D, et al. An eHealth project on invasive Pneumococcal disease: comprehensive evaluation of a promotional campaign. J Med Internet Res 2016 Dec 02;18(12):e316 [FREE Full text] [doi: 10.2196/jmir.6205] [Medline: 27913372]
55. White BK, Martin A, White JA, Burns SK, Maycock BR, Giglia RC, et al. Theory-based design and development of a socially connected, gamified mobile app for men about breastfeeding (milk man). JMIR Mhealth Uhealth 2016 Jun 27;4(2):e81 [FREE Full text] [doi: 10.2196/mhealth.5652] [Medline: 27349756]
56. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. JMIR Mhealth Uhealth 2015 Aug 19;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]
57. Nielsen J. Heuristic evaluation. In: Usability Inspection Methods. New York, NY, USA: John Wiley & Sons, Inc; 1994:25-62.
58. Roher C. Nielsen Norman Group. 2014 Oct 12. When to Use Which User-Experience Research Methods Internet URL: https://www.nngroup.com/articles/which-ux-research-methods/ [accessed 2018-01-12] [WebCite Cache ID 6wOh4sv4x]
59. IPAQ Research Committee. 2011. Background - International Physical Activity Questionnaire Internet URL: https://sites.google.com/site/theipaq/scoring-protocol [accessed 2018-01-12] [WebCite Cache ID 6wOi78iYU]

60. Hallgren KA. Computing inter-rater reliability for observational data: an overview and tutorial. Tutor Quant Methods Psychol 2012;8(1):23-34 [FREE Full text] [Medline: 22833776]
61. Gwet K. Handbook of Inter-Rater Reliability: The Definitive Guide to Measuring the Extent of Agreement Among Multiple Raters, 3rd Edition. 3rd edition. Gaithersburg, MD: Advanced Analytics, LLC; 2012.
62. O'Neill TA. An overview of interrater agreement on likert scales for researchers and practitioners. Front Psychol 2017;8:777 [FREE Full text] [doi: 10.3389/fpsyg.2017.00777] [Medline: 28553257]
63. LeBreton JM, Senter JL. Answers to 20 questions about interrater reliability and interrater agreement. Organ Res Methods 2007 Nov 28;11(4):815-852. [doi: 10.1177/1094428106296642]
64. James LR, Demaree RG, Wolf G. Estimating within-group interrater reliability with and without response bias. J Appl Psychol 1984;69(1):85-98. [doi: 10.1037/0021-9010.69.1.85]
65. James L, Demaree R, Wolf G. rwg: an assessment of within-group interrater agreement. J Appl Psychol 1993;78(2):306-309. [doi: 10.1037/0021-9010.78.2.306]
66. Biemann T, Cole MS, Voelpel S. Within-group agreement: on the use (and misuse) of rWG and rWG(J) in leadership research and some best practice guidelines. Leadersh Q 2012 Feb;23(1):66-80. [doi: 10.1016/j.leaqua.2011.11.006]
67. Brown RD, Hauenstein NM. Interrater agreement reconsidered: an Alternative to the rwg indices. Organ Res Methods 2016 Jun 29;8(2):165-184. [doi: 10.1177/1094428105275376]
68. Lohse-Bossenz H, Kunina-Habenicht O, Kunter M. Estimating within-group agreement in small groups: a proposed adjustment for the average deviation index. Eur J Work Organ Psychol 2013 Jan 25;23(3):456-468. [doi: 10.1080/1359432X.2012.748189]
69. Van Bruggen GH, Lilien G, Kacker M. Informants in organizational marketing research: why use multiple informants and how to aggregate responses. J Mark Res 2002 Nov;39(4):469-478. [doi: 10.1509/jmkr.39.4.469.19117]
70. Wagner SM, Rau C, Lindemann E. Multiple informant methodology: a critical review and recommendations. Sociol Methods Res 2010 Jun 16;38(4):582-618. [doi: 10.1177/0049124110366231]
71. Leuffen D, Shikano S, Walter S. Measurement and data aggregation in small-n social scientific research. Eur Polit Sci 2012 Mar 30;12(1):40-51. [doi: 10.1057/eps.2012.8]
72. Woehr DJ, Loignon AC, Schmidt PB, Loughry ML, Ohland MW. Justifying aggregation with consensus-based constructs. Organ Res Methods 2015 May 05;18(4):704-737. [doi: 10.1177/1094428115582090]
73. Curtin F, Schulz P. Multiple correlations and Bonferroni's correction. Biol Psychiatry 1998 Oct 15;44(8):775-777. [Medline: 9798082]
74. Barnes SA, Lindborg SR, Seaman JW. Multiple imputation techniques in small sample clinical trials. Stat Med 2006 Jan 30;25(2):233-245. [doi: 10.1002/sim.2231] [Medline: 16220515]
75. Graham JW, Olchowski AE, Gilreath TD. How many imputations are really needed? Some practical clarifications of multiple imputation theory. Prev Sci 2007 Sep;8(3):206-213. [doi: 10.1007/s11121-007-0070-9] [Medline: 17549635]
76. Tavakol M, Dennick R. Making sense of Cronbach's alpha. Int J Med Educ 2011 Jun 27;2:53-55 [FREE Full text] [doi: 10.5116/ijme.4dfb.8dfd] [Medline: 28029643]
77. Sijtsma K. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. Psychometrika 2009 Mar;74(1):107-120 [FREE Full text] [doi: 10.1007/s11336-008-9101-0] [Medline: 20037639]
78. Perski O, Blandford A, Ubhi H, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. BMC Med Inform Decis Mak 2017 Dec 28;17(1):25 [FREE Full text] [doi: 10.1186/s12911-017-0422-8] [Medline: 28241759]

Abbreviations

BMI: body mass index
ICC: intraclass correlation coefficient
IQR: interquartile range
IRA: interrater agreement
IRR: interrater reliability
K-W: Kruskal-Wallis
MARS: Mobile App Rating Scale
mHealth: mobile health
MI: multiple imputation
uMARS: user version of the Mobile App Rating Scale
WDMEAN: response data–based weighted mean

Edited by G Eysenbach; submitted 14.01.18; peer-reviewed by E Lyons, S Stoyanov; comments to author 11.07.18; revised version received 15.09.18; accepted 24.09.18; published 23.01.19.

Please cite as:
Bardus M, Ali A, Demachkieh F, Hamadeh G
Assessing the Quality of Mobile Phone Apps for Weight Management: User-Centered Study With Employees From a Lebanese University
JMIR Mhealth Uhealth 2019;7(1):e9836
URL: https://mhealth.jmir.org/2019/1/e9836/
doi:10.2196/mhealth.9836
PMID:30672742

Original Paper

Effectiveness of a Multimodal Digital Psychotherapy Platform for Adult Depression: A Naturalistic Feasibility Study

Enitan T Marcelle[1], BA; Laura Nolting[2], BS; Stephen P Hinshaw[1,3], PhD; Adrian Aguilera[4,5], PhD

[1]Department of Psychology, University of California, Berkeley, Berkeley, CA, United States

[2]Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, United States

[3]Department of Psychiatry, University of California, San Francisco, San Francsco, CA, United States

[4]School of Social Welfare, University of California, Berkeley, Berkeley, CA, United States

[5]Department of Psychiatry, San Francisco General Hospital, University of California, San Francisco, San Francisco, CA, United States

Corresponding Author:
Enitan T Marcelle, BA
Department of Psychology
University of California, Berkeley
Tolman Hall
Berkeley, CA,
United States
Phone: 1 2164968894
Email: emarcelle@berkeley.edu

Abstract

Background: Although psychotherapy is one of the most efficacious and effective treatments for depression, limited accessibility to trained providers markedly limits access to care. In an attempt to overcome this obstacle, several platforms seeking to provide these services using digital modalities (eg, video, text, and chat) have been developed. However, the use of these modalities individually poses barriers to intervention access and acceptability. Multimodal platforms, comprising those that allow users to select from a number of available modalities, may be able to provide a solution to these concerns.

Objective: We aimed to investigate the preliminary effectiveness of providing psychotherapy through a multimodal digital psychotherapy platform. In addition, we aimed to examine differential responses to intervention by gender, self-reported physical health status, and self-reported financial status, as well as how prior exposure to traditional face-to-face psychotherapy affected the effectiveness of a multimodal digital psychotherapy intervention. Finally, we aimed to examine the dose-response effect.

Methods: Data were collected from a total of 318 active users of BetterHelp, a multimodal digital psychotherapy platform. Data on physical health status, financial status, and prior exposure to psychotherapy were obtained using self-report measures. Effectiveness was determined by the extent of symptom severity change, which was measured using the Patient Health Questionnaire at Time 1 (time of enrollment) and Time 2 (3 months after enrollment). Intervention dosage was measured as the sum of individual therapist-user interactions across modalities.

Results: Depression symptom severity was significantly reduced after the use of the multimodal digital psychotherapy intervention (P<.001). Individuals without prior traditional psychotherapy experience revealed increased improvement after intervention (P=.006). We found no significant dose-response effect of therapy, nor significant differences in outcomes across gender, self-reported financial status, and self-reported physical health status.

Conclusions: Users of BetterHelp experienced significantly reduced depression symptom severity after engaging with the platform. Study findings suggest that this intervention is equally effective across gender, self-reported financial status, and self-reported physical health status and particularly effective for individuals without a history of psychotherapy. Overall, study results suggest that multimodal digital psychotherapy is a potentially effective treatment for adult depression; nevertheless, experimental trials are needed. We discuss directions for future research.

(JMIR Mhealth Uhealth 2019;7(1):e10948) doi:10.2196/10948

KEYWORDS

cognitive therapy; depression; digital health; live chat; mHealth; mental health; text messaging; video; mobile phone

Introduction

Major depressive disorder is a commonly occurring condition [1,2] associated with a multitude of adverse health outcomes [3,4] and is projected to be the second leading cause of disability worldwide by the year 2020 [5,6]. Development and dissemination of efficacious and accessible treatment for the disorder are an area of increasing importance. Psychotherapy has been shown to be one of the most efficacious psychosocial treatments for depression [7,8] and works by teaching patients cognitive strategies that enable them to target and manage undesirable thoughts, habits, and emotions underlying presenting depressive symptoms [9]. Although the literature has demonstrated the need for psychotherapy in the successful treatment of depression [10-13], limited geographical access to trained professionals remains one of the most significant barriers to traditional face-to-face care [14]. Telemental health, or the use of digital technology to provide long-distance clinical mental health care, is a field rapidly growing in response to issues of access [15-18].

Multimodal digital therapy platforms, that is, platforms that offer multiple modes of digital communication, hold promise in overcoming persisting barriers to intervention access and acceptability. In the United States, 1 in every 4 adults is in need of counseling, although only 13.4% of adults report receiving services [19]. This treatment gap is seen around the world [20] and is driven by a combination of stigma surrounding mental health support-seeking behavior [21] and limited geographical access to providers [14]. Digital psychotherapeutic interventions have been shown to increase intervention reach by providing efficacious alternatives to traditional therapy. These digital interventions increase both ease of access and anonymity of service seeking. To date, digital therapy has largely centered around the internet- (eg, videoconference or chat) and mobile (eg, short message service [SMS] text messaging or phone)-based modalities.

Extant literature demonstrates marked demographic differences in modality access and acceptance. Internet-based interventions have been shown to be efficacious [22], with a 2009 meta-analysis of 12 studies finding a mean effect size of d= 0.41 for internet-based psychological treatments for adult depression [23]. This effect size was increased (d=0.61) after excluding standalone interventions and considering only those that included support or guidance from a therapist. While evidence is mixed [24], research suggests that increased guidance improves the efficacy of internet-based psychotherapy [25]. Although such interventions have been found to be efficacious, barriers to accessible, reliable, and consistent internet connection keep these interventions beyond the reach of traditionally underserved populations [18]. In response to this persisting digital divide, researchers have investigated internet-free, mobile phone- and SMS text message-based therapies as a way of continuing to increase access to care [26]. Mobile phone access is considerably more ubiquitous than internet access [15,27], and the development of mobile platforms has enabled mental health professionals to provide care to an even greater percentage of individuals in need. In addition, mobile phone-based psychotherapy for adult depression has been demonstrated to be efficacious [28], yet this workaround does not come without its own set of limitations. Specifically, younger adults feel more comfortable communicating and building relationships remotely [29], although older adults prefer communication modalities that can mirror more traditional, in-person interactions, such as video calls. A multitude of studies have examined and demonstrated the efficacy and effectiveness of single- or even dual-modality digital psychotherapy platforms [22,30-34], but we know of no existing work investigating the effectiveness of a multimodal platform. We propose a multimodal psychotherapy platform that allows users to choose from internet-based video or live chat and internet-free, mobile SMS text message or phone therapy interchangeably as a potential solution to this demographic variability in intervention usability and effectiveness.

In addition to the modality, other important predictors of digital psychotherapy outcomes include demographic characteristics as well as level of engagement. The existing literature in face-to-face psychotherapy has focused on differences in outcomes between men and women, as well as on differences across ages and socioeconomic status. Despite growing product innovation, the characteristics and demographics of populations for whom digital psychotherapies do and do not work remains unknown. Although evidence of gender effects on psychotherapy outcomes is mixed, research suggests that women report an increased symptom reduction after traditional psychotherapy [35-39]. Yet research on Web-based psychotherapy suggests that men report an increased symptom reduction compared with women. Given higher rates of the stigma surrounding therapy-seeking behaviors in men [40], it has been suggested that the anonymity afforded by Web-based platforms may be an explanation for this latter finding. Research examining the effect of age on psychotherapy outcomes also remains largely inconclusive. Some evidence suggests that younger clients may respond more quickly and report greater posttreatment improvement in psychotherapy, although older adults are more likely to adhere to treatment [7,41,42]. The literature focusing on broader socioeconomic predictors of traditional psychotherapy acceptance and response has focused on economic status, physical health status, and experience with therapy as key predictive variables, finding positive correlations between outcomes and these factors. We speculate that prior counseling experience and higher financial or health status will predict improved digital psychotherapy outcomes as well. Although a number of researchers have examined the effect of engagement, or dose-response effect of psychotherapy, in traditional settings [43], considerably fewer studies have examined this relationship in the context of digital psychotherapy [44-46]. We additionally aimed to examine the dose-response effect of our digital psychotherapy platform.

The aim of this feasibility study was to investigate the initial effectiveness of delivering psychotherapy via BetterHelp, a multimodal internet- and mobile-based psychotherapy service provider. Bowen et al [47] defined a feasibility study as any study aiming to “determine whether an intervention is appropriate for further testing.” In line with this definition, this naturalistic and quasi-experimental investigation aimed to examine individuals’ responses to BetterHelp to generate useful

data to guide and justify future randomized controlled trials. Effectiveness studies are defined as those that investigate the extent to which an intervention does more good than harm when administered in a "real-word" setting, as opposed to in ideal and highly controlled conditions [48,49]. In line with our aim of investigating the effectiveness of a multimodal method of delivering psychotherapy for adult patients with depression, this work defined the intervention effectiveness as a significant reduction in the depressive symptom severity among individuals using BetterHelp. We hypothesized that engagement with BetterHelp will significantly reduce depression symptom severity. In addition, in this exploratory study, we aimed to investigate the ways in which multimodal digital psychotherapy outcomes vary by subpopulation (gender, age, financial and physical status, and prior therapy experience) as well as by the level of engagement.

Methods

Participants

In total, 318 BetterHelp clients (of whom 254, 79.9%, were females), recruited from a larger pool of active BetterHelp users, participated in this study. BetterHelp users are individuals aged ≥18 years seeking to improve their quality of life. Users aged <18 years or under the care of a legal guardian were excluded from BetterHelp participation. Furthermore, individuals with thoughts of hurting themselves or others, those in urgent crisis or emergency situations, those diagnosed with a severe mental illness or advised to be in psychological supervision or psychiatric care, and those required to undergo therapy or counseling either by a court order or by any other authority were also excluded. BetterHelp users who met the eligibility criteria to participate in this study were invited to participate. Inclusion criteria for this study included a self-endorsed primary concern of feelings of overwhelming sadness, grief, or depression. Study participants were excluded from participation if preintervention levels of depression fell below mild clinical significance, that is, a score of <5 on the Patient Health Questionnaire (PHQ-9) [50], or if they had not engaged with BetterHelp for a minimum of 90 days. Users with preintervention depression in the mild to severe range, (ie, PHQ-9 scores ≥5 and ≤27) were included (Textbox 1). A total of 1148 BetterHelp users met all the eligibility criteria and were invited to participate. Our observed response rate of 27.70% (318/1148) is comparable to response rates seen in other Web-based survey studies [51]. The racial and ethnic makeup of this sample, as well as comorbid mental or physical health concerns, is unfortunately unknown as the BetterHelp platform does not currently collect this information. Ages of participants meeting the eligibility criteria ranged from 19 to 72 (mean 33.27 [SD 11.29]) years. At baseline, of 318, 119 (37.4%) participants met the criteria for mild depression, 91 (28.6%) for moderate depression, 75 (23.6%) for moderately severe depression, and 33 (10.4%) for severe depression [50,52,53] (Table 1).

Textbox 1. Inclusion and exclusion criteria for study participation.

Inclusion criteria

- Age ≥18 years
- BetterHelp user for at least 90 days
- Baseline Patient Health Questionnaire score ≥5 and ≤27
- Self-endorsed primary concerns of feelings of overwhelming sadness, grief, or depression

Exclusion criteria

- Age <18 years
- Under the care of a legal guardian
- Thoughts of hurting self or others
- Severe mental illness
- Advised to be in psychological supervision or psychiatric care, or required to undergo therapy or counseling either by a court order or by any other authority
- BetterHelp user for <90 days
- Baseline Patient Health Questionnaire score <5

Table 1. Preintervention and postintervention Patient Health Questionnaire (PHQ-9) scores by the diagnostic category (N=318).

PHQ-9 diagnostic category	PHQ-9 score	Preintervention, n (%)	Postintervention, n (%)
Minimal depression	0-4	0 (0)	63 (19.8)
Mild depression	5-9	119 (37.4)	141 (41.2)
Moderate depression	10-14	91 (28.6)	66 (20.8)
Moderately severe depression	15-19	75 (23.6)	40 (12.6)
Severe depression	20-27	33 (10.4)	18 (5.7)

Measures

Patient Health Questionnaire

The PHQ-9 [50,52,53] is a 10-item, self-report measure inquiring about the presence of depressive symptoms in the previous 2 weeks. It is used in clinical practice to monitor depression symptoms and severity [54]. The measure probes how often a respondent has been bothered by specific problems, takes only a few minutes to complete, and is scored on a 4-point Likert scale ranging from 0 (Not At All) to 3 (Nearly Every Day). The scale has demonstrated high internal consistency (Cronbach alpha=.86 to .89) as well as excellent test-retest reliability (r=0.84) [52]. Participants completed this measure at baseline and follow-up.

Working Alliance Inventory-Short Revised

The Working Alliance Inventory-Short Revised (WAI-SR) [55] is a 12-item measure assessing the quality of the therapeutic relationship. This measure of therapeutic alliance assesses the following domains: (1) agreement on the tasks of therapy; (2) agreement on the goals of therapy; and (3) development of an affective bond. The measure has consistently demonstrated good reliability (alpha>.80) as well as good convergent validity (r>0.64); it was administered at follow-up to assess user rapport with therapist.

Prior Exposure to Therapy

Before beginning therapy, a binary measure was administered to assess prior exposure to psychotherapy. BetterHelp users were asked the question "Have you ever been in counseling or therapy before?" and probed to reply with either "Yes" or "No."

Self-Reported Physical Health Status and Financial Status

At baseline, BetterHelp users were asked to rate their current physical health and current financial status. Responses were scored on a 3-point Likert scale ranging from Good to Poor.

Procedure

Intervention Description

The BetterHelp psychotherapy platform is currently the largest multimodal digital psychotherapy platform available worldwide [56]. BetterHelp utilizes a preference-based approach in which users can use any and all combinations of text, video, chat, or phone communication over the course of psychotherapy, as they choose. BetterHelp procedures are as follows: before beginning therapy, clients are asked to complete questionnaires probing symptom levels, personal history, and motivation for seeking therapy. Although BetterHelp counselors vary in approach (ie, cognitive behavioral therapy, acceptance and commitment therapy, etc), each BetterHelp counselor is a required to have attained a PhD, PsyD, Marriage Family Therapist, Licensed Clinical Social Worker, Licensed Professional Counselor, or Licensed Master Social Worker-level license to practice. BetterHelp's algorithm then matches clients with an available BetterHelp counselor who best fits their objectives, counselor preferences, and needs. Preferred modality of communication is not taken into account in BetterHelp's algorithm, as BetterHelp users can utilize any form of communication at any time, and all BetterHelp counselors are required to make themselves available to provide therapy through the client's chosen modality. After the match is made, BetterHelp provides client and counselor with a dedicated "room" in which all communication takes place. Video, live chat, and phone sessions require advanced scheduling, while SMS text message exchanges do not. Multimedia Appendices 1-4 display BetterHelp platform as well as its video and live chat invitation and interface.

Data Collection

After engaging with the BetterHelp platform for 3 months, users received a notification inviting them to participate in an ongoing research study. This 3-month time window was selected to mirror the existing dose-response research in psychotherapy, suggesting that >50% of patients are able to respond after 12.7 sessions of weekly psychotherapy [57]. BetterHelp users were allowed 2 weeks to respond to prompts inviting them to participate in research. Users who did not respond within the allotted 2 weeks were excluded from study participation. Respondents were asked to repeat the PHQ-9 as well as complete the WAI-SR. All respondents to the research invitation completed all questionnaires. BetterHelp data analysts sent relevant data to authors in a deidentified data file.

Data Analysis

All analyses were conducted in SPSS Version 25 (IBM Corp Released 2017. IBM SPSS Statistics for Macintosh, Version 25.0) and R version 2.13.1 (R Development Core Team (2011). To examine the effectiveness of delivering psychotherapy through a multimodal digital platform, we examined the percentage of users exhibiting established markers of clinical improvement, as determined by change in PHQ-9 scores [53]. We first examined the percentage of users demonstrating clinically significant improvement. We additionally examined the percentage of users demonstrating a partial response to intervention and the percentage of users qualifying as being in remission after intervention use. We then examined changes in the symptom severity among BetterHelp users from pre- to posttreatment using paired samples t test. Because of our study design (ie, comparing mean PHQ-9 scores of BetterHelp users across 2 time-points), we determined paired samples t test to be the most appropriate method of analysis [58]. Our study design and the nature of our dependent and independent variables led us to determine a one-way analysis of covariance (ANCOVA), covarying baseline symptom severity, as the most appropriate means testing the differential effects of age and gender on therapy outcomes. To examine the effect of financial status, health status, and prior exposure to therapy on treatment outcomes, we conducted a second one-way ANCOVA, again adjusting for baseline symptom severity. The effect of dosage on treatment outcomes was examined using probit dose-response regression, a method used across the extant dose-response literature to predict the amount or dose of treatment needed to achieve a desired response or effect [43].

Ethics Statement

This study was conducted in accordance with the Institutional Review Board and Office for the Protection of Human Subjects

at the University of California, Berkeley. All subjects were informed of the risks and benefits of participating in the study and gave electronic informed consent to participate via the BetterHelp platform. Participants were provided with access to the BetterHelp platform and were not additionally compensated or incentivized to participate in any way. The research study was approved by the Institutional Review Board at the University of California, Berkeley.

Results

Descriptive Statistics

Table 2 describes the sociodemographic characteristics (age, gender, health or financial status, and prior exposure to therapy) of the study sample. Overall, of 318 participants, 72 (22.6%) rated their financial status as poor, 167 (52.5%) as fair, and 49 (15.4%) as good. In addition, 24 (7.5%) participants rated their health status as poor, 165 (51.9%) as fair, and 118 (37.1%) as good; furthermore, 91 (28.6%) participants did not have prior exposure to counseling or therapy. Table 3 displays the modality usage in the BetterHelp user sample.

Table 2. Sociodemographic characteristics of BetterHelp users.

Characteristic	BetterHelp users, n (%)
Gender	
Female	254 (79.9)
Male	64 (20.1)
Age (years)	
18-34	200 (62.6)
35-49	89 (27.7)
>50	29 (9.1)
Physical health status	
Poor	24 (7.5)
Fair	165 (51.9)
Good	118 (37.1)
Financial status	
Poor	72 (22.6)
Fair	167 (52.5)
Good	49 (15.4)
Prior counseling experience	
No	91 (28.6)
Yes	216 (67.9)

Table 3. Modality usage of BetterHelp users.

Modality usage	BetterHelp user, n (%)
Short message service text message user	318 (100)
Live chat user	53 (16.7)
Phone user	37 (11.6)
Video user	1 (0.3)
Single modality user	236 (74.2)
Dual modality user	74 (23.3)
Tri modality user	7 (2.2)
All modality user	1 (0.3)

Overall Effectiveness of BetterHelp

In this study, of 318 participants, 120 (37.8%) demonstrated a clinically significant improvement and 194 (62%) demonstrated a partial response (as defined by at least a 5-point score reduction on the PHQ-9 and a postintervention score ≤9, respectively [53]) after engaging with BetterHelp for 3 months; in addition, 63 (19.8%) participants qualified as being in remission (as defined by a postintervention score <5) by Time 2. Paired samples *t* test results revealed a significant decrease in symptom severity posttreatment with effect size in the medium range (pretreatment: mean 12.57 [SD 5.35]; posttreatment: mean 9.36 [SD 5.51]; t_{317}=10.80; *P*<.001, one-tailed; Cohen *d*=0.61; Figure 1). Mean PHQ-9 pretreatment scores reflected moderate levels of current depression, whereas mean posttreatment scores reflected mild levels of current depression, as determined by PHQ-9 clinical cutoffs [52]. As shown in Table 1, after using BetterHelp for 3 months, of 318 participants, 63 (19.8%) met the criteria for minimal depression, 141 (41.2%) for mild depression, 66 (20.8%) for moderate depression, 40 (12.6%) for moderately severe depression, and 18 (5.7%) for severe depression [52,53].

Demographic Influence on Outcomes

A one-way ANCOVA examining the differential effects of age and gender on therapy outcomes, covarying baseline symptom severity, revealed no significant differences. Outcomes did not significantly differ across age ($F_{45,242}$=0.98; *P*=.51) and gender ($F_{1,242}$=.092; *P*=.76).

Socioeconomic and Environmental Influence on Outcomes

A one-way ANCOVA examining effects of financial status, health status, and prior exposure to therapy on treatment outcomes, covarying baseline symptom severity, revealed a significant effect of prior exposure to therapy on treatment outcome ($F_{1,260}$=7.531; *P*=.006). Individuals with prior therapy exposure experienced significantly fewer gains after treatment compared with individuals without prior exposure (Figure 2). Treatment outcomes did not significantly differ across participant financial status ($F_{2,260}$=1.563; *P*=.21) or health status ($F_{2,260}$=1.575; *P*=.21).

Dose-Response Effect

Treatment dosage was measured as the sum of individual therapist-user interactions across modalities (text message, phone, video, and live chat). Response was measured as a binary variable indicating the presence or absence of clinical improvement, as defined by a PHQ-9 score change of ≥5 points. Mean number of interactions in this study was 125.3 (SD 392.5). Following the methodology laid out by Howard et al in the probit model [59], the dose corresponded to the log of the number of total interactions, to reduce skew. Because participants who did not interact with BetterHelp were not included in the sample, taking the log of 0 was not a possibility. Results revealed no significant dose-response effect (*P*=.14), which was maintained after adding baseline severity into the model.

Post-Hoc Analyses

To adjust PHQ-9 scores for regression to the mean, an ANCOVA in which the difference between baseline and posttreatment scores was regressed onto the mean-centered baseline score was used [58,60]. These additional analyses also revealed a significant mean difference (*P*<.001).

To assess the hypothesis that differences in treatment gains between individuals with and without prior therapy experience may be driven by differences in therapeutic alliance associated with prior therapy exposure, an independent samples *t* test was performed. Cronbach alpha for the 12 WAI-SR items was high (alpha=.946). The results revealed no significant effect of prior therapy experience on WAI-SR total scores (*P*=.55).

Figure 1. Overall Patient Health Questionnaire (PHQ-9) pre-post change. Error bars represent SEs.

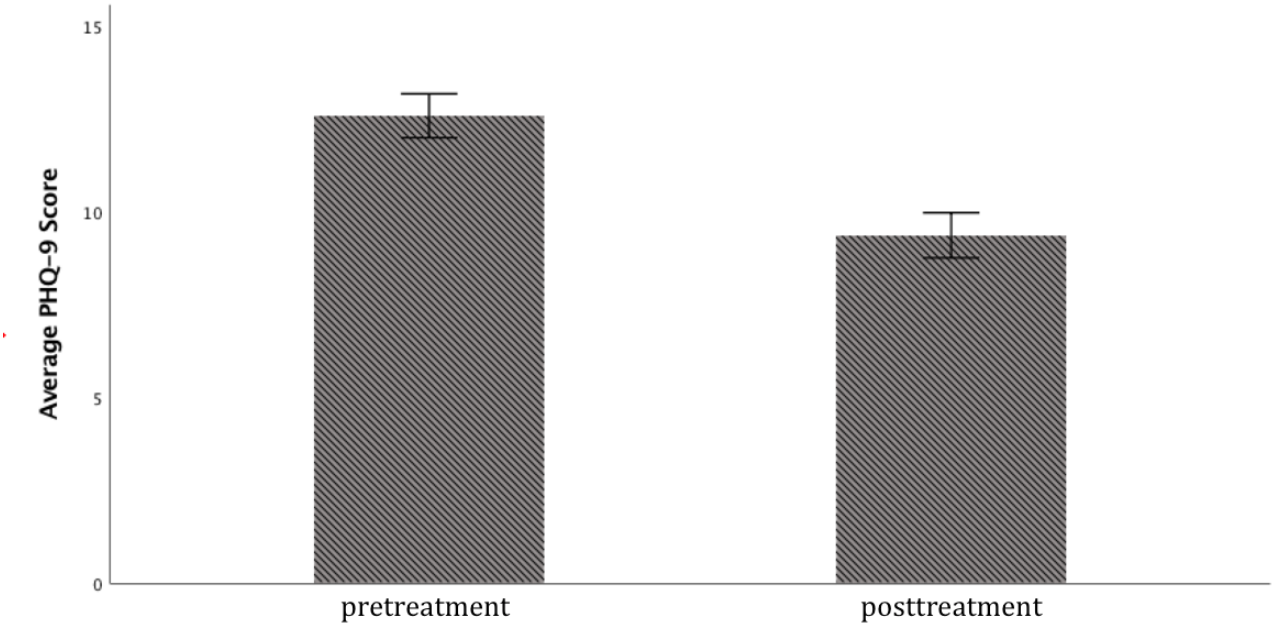
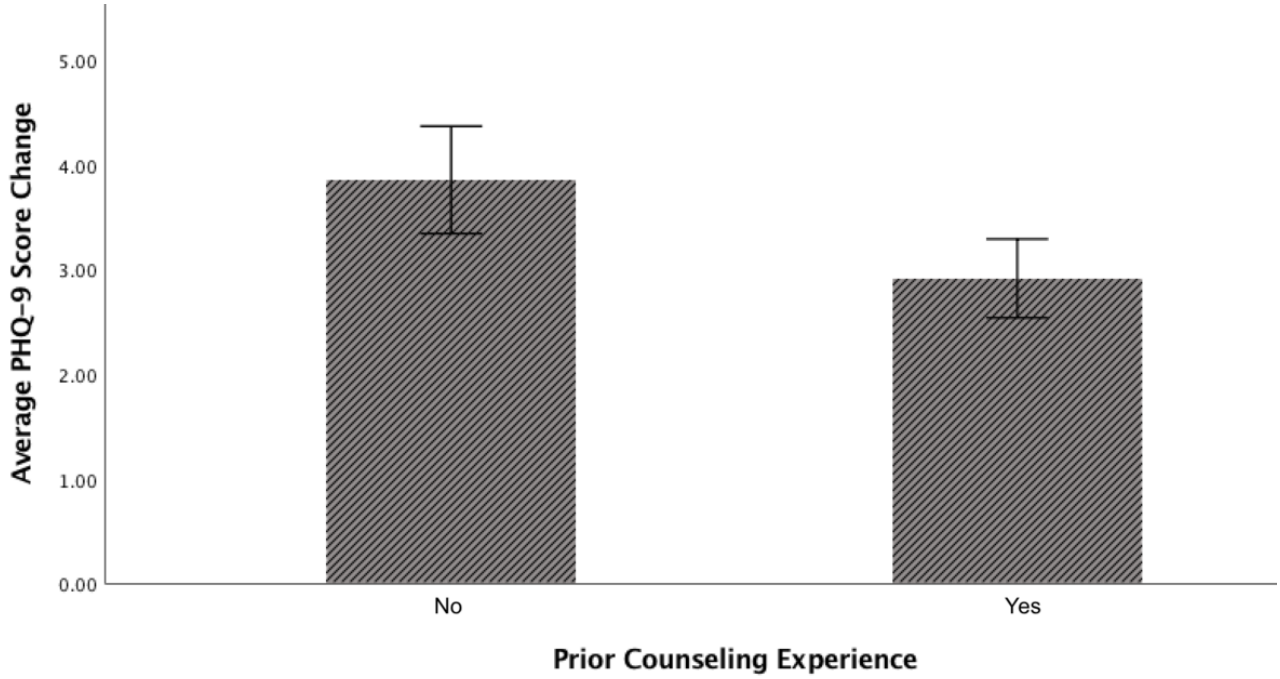


Figure 2. Difference in Patient Health Questionnaire (PHQ-9) score change by prior counseling experience status. Error bars represent SEs.



Discussion

Principal Findings

Our results indicate that multimodal digital psychotherapy may be an effective treatment for adult depression. Given our preliminary results demonstrating the effectiveness of this method of psychotherapy dissemination, we believe that multimodal digital psychotherapy may hold promise in overcoming some existing issues of psychotherapy access. Hence, continued research is needed. Users of a multimodal digital psychotherapy platform experienced significantly improved self-reported symptoms after engaging with BetterHelp, with 37.8% (120/318) users experiencing clinically significant improvement in depressive symptoms within 3 months. Moreover, no significant associations were found between changes in depressive symptoms and sociodemographic variables, including age, gender, and self-reported financial or physical health status. We suspect that increased accessibility and flexibility provided by a preference-based multimodal platform may be driving these latter findings [61,62]. Although this study cannot establish causality, our findings suggest that multimodal digital psychotherapy may be an effective solution to reduce existing barriers to accessible and preference-based digital psychotherapy.

Participants who had previously engaged in traditional face-to-face psychotherapy showed significantly less symptomatic improvement compared with those who had not. Among several potential explanations, we postulate that individuals with prior therapy experience may present with more chronic, complex, comorbid, or treatment-resistant forms of depression requiring a higher level of care [63-65]. Post-hoc analyses did not reveal any significant association between therapeutic relationship quality (as measured by the WAI) and prior therapy engagement, suggesting that this finding is not merely related to differences in therapist rapport and alliance linked to prior treatment exposure. Future research should seek to investigate differences in outcomes across diagnostic categories, taking into account commonly comorbid conditions known to influence response to treatment such as generalized anxiety disorder, attention deficit/hyperactivity disorder, and panic disorder. Baseline evaluation of the presence and severity of additional diagnoses may ensure that individuals with comorbid conditions are matched with an appropriately specialized therapist.

No general dose-response effect was detected in this study. Although this is not the first study to find nonsignificant dose-response effects of psychotherapy, we speculate that our existing measure of dosage may disguise existing dose-response effects, given that the quantity and significance of content shared in a single message may vary greatly by individual. Future research may seek to examine effects of word count as opposed to the message count, though due to user privacy concerns, researchers were not able to access these data at this time. Furthermore, the number of days spent interacting with the BetterHelp platform may prove to be a more valid measure of engagement than total interactions.

All study participants in this preference-based study utilized the SMS text message modality. A majority of participants utilized only one therapy modality over the course of the intervention, with only a single participant making use of all 4 modalities. We hypothesize that this pattern of use is driven by the increased convenience and flexibility provided by SMS text message-based therapy (ie, a previously scheduled appointment is not needed to utilize the SMS text message-based modality, whereas an appointment is needed to utilize live chat-, phone-, or video-based therapy). Future work will seek to test this hypothesis by obtaining qualitative data from participants regarding motivations behind modality choices, as well as to elucidate the added benefit of a multimodality psychotherapy

platform (ie, using a randomized "text-only" vs "multimodal" design).

Limitations

This exploratory investigation contains several limitations of note. First and foremost, we lacked a randomly assigned control group. Future investigations with appropriate controls, as well as a more comprehensive demographic and diagnostic screening, will enable a more valid approach. In addition, like many survey-based studies, the results of this work may be influenced by sample bias. It is possible that individuals who had notably positive or negative experiences with BetterHelp are those who chose to respond to our prompt to participate in research. We are further unable to investigate the effect of the type of psychotherapy provided (ie, cognitive behavioral therapy, acceptance and commitment therapy, etc) on outcomes. However, it is worth noting that the existing literature examining differential effects of different types of psychotherapies for depression does not suggest that psychotherapy type has a significant effect on outcomes in traditional psychotherapeutic settings [8]. In this study, financial and physical heath status were measured by self-report; such subjective indicators may not be fully valid. Future work may seek to utilize additional objective and sensitive measures of financial and physical health status to provide further insight. Finally, in this study, we lacked data on the potentially crucial moderator variables of race, ethnicity, and gender nonconformity.

Conclusions

Major depressive disorder is a pervasive and debilitating condition from which many individuals are unable to recover due to lack of accessible and appropriate treatment. The existing literature has demonstrated the use of digital technology as a feasible solution to this growing and widespread dilemma. This study examined the effect of a multimodal digital psychotherapy platform for the treatment of depression in adults. We proposed that a multimodal platform through which users can dynamically select from multiple modes of digital communication throughout therapy may be an effective method of delivering psychotherapy to adults with depression. Our study results demonstrate the initial effectiveness of such a model, with users experiencing significant symptom reduction after the intervention. This feasibility study's preliminary demonstration of such a platform's effectiveness provides an important first step in understanding the potential benefits of such a model of mental health care delivery. Given these results, subsequent research will aim to investigate the unique benefits of a preference-based multimodal platform compared with in-person, single- or bimodal psychotherapy dissemination.

Acknowledgments

ETM is supported by the National Science Foundation Graduate Research Fellowship under grant #DGE 1752814. The publication is made possible, in part, by support from the Berkeley Research Impact Initiative sponsored by the University of California, Berkeley, Library.

Conflicts of Interest

ETM is a former consultant for BetterHelp. LN is a community and support manager for BetterHelp.

Multimedia Appendix 1

BetterHelp video session invitation.

[PNG File, 157KB - mhealth_v7i1e10948_app1.png]

Multimedia Appendix 2

BetterHelp live chat invitation.

[PNG File, 53KB - mhealth_v7i1e10948_app2.png]

Multimedia Appendix 3

BetterHelp live chat interface (user perspective).

[PNG File, 84KB - mhealth_v7i1e10948_app3.png]

Multimedia Appendix 4

BetterHelp messaging platform.

[PNG File, 156KB - mhealth_v7i1e10948_app4.png]

References

1. Weissman M, Bland RC, Canino GJ, Faravelli C, Greenwald S, Hwu HG, et al. Cross-national epidemiology of major depression and bipolar disorder. JAMA 1996;276(4):293-299. [Medline: 8656541]
2. Kessler R, Bromet EJ. The epidemiology of depression across cultures. Annu Rev Public Health 2013;34:119-138 [FREE Full text] [doi: 10.1146/annurev-publhealth-031912-114409] [Medline: 23514317]
3. Chaudhury P, Deka K, Chetia D. Disability associated with mental disorders. Indian J Psychiatry 2006 Apr;48(2):95-101 [FREE Full text] [doi: 10.4103/0019-5545.31597] [Medline: 20703393]
4. Druss BG, Hwang I, Petukhova M, Sampson NA, Wang PS, Kessler RC. Impairment in role functioning in mental and chronic medical disorders in the United States: results from the National Comorbidity Survey Replication. Mol Psychiatry 2009 Jul;14(7):728-737 [FREE Full text] [doi: 10.1038/mp.2008.13] [Medline: 18283278]
5. Sayers J. The world health report 2001 - Mental health: new understanding, new hope. Bulletin of the World Health Organisation 2001;79(11) [FREE Full text] [doi: 10.1590/S0042-96862001001100014]
6. Ustün TB, Ayuso-Mateos JL, Chatterji S, Mathers C, Murray CJL. Global burden of depressive disorders in the year 2000. Br J Psychiatry 2004 May;184:386-392. [Medline: 15123501]
7. Dobson KS. A meta-analysis of the efficacy of cognitive therapy for depression. J Consult Clin Psychol 1989 Jun;57(3):414-419. [Medline: 2738214]
8. Cuijpers P, van Straten A, Andersson G, van Oppen P. Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. J Consult Clin Psychol 2008 Dec;76(6):909-922. [doi: 10.1037/a0013075] [Medline: 19045960]
9. Rupke SJ, Blecke D, Renfrow M. Cognitive therapy for depression. Am Fam Physician 2006 Jan 01;73(1):83-86 [FREE Full text] [Medline: 16417069]
10. Al-Harbi KS. Treatment-resistant depression: therapeutic trends, challenges, and future directions. Patient Prefer Adherence 2012;6:369-388 [FREE Full text] [doi: 10.2147/PPA.S29716] [Medline: 22654508]
11. Rush AJ. STAR*D: what have we learned? Am J Psychiatry 2007 Feb;164(2):201-204. [doi: 10.1176/ajp.2007.164.2.201] [Medline: 17267779]
12. Blackburn IM, Bishop S, Glen AI, Whalley LJ, Christie JE. The efficacy of cognitive therapy in depression: a treatment trial using cognitive therapy and pharmacotherapy, each alone and in combination. Br J Psychiatry 1981 Sep;139:181-189. [Medline: 7317698]
13. Rush A, Beck AT, Kovacs M, Hollon S. Comparative efficacy of cognitive therapy and pharmacotherapy in the treatment of depressed outpatients. Cogn Ther Res 1977 Mar;1(1):17-37. [doi: 10.1007/BF01173502]
14. Thomas K, Ellis AR, Konrad TR, Holzer CE, Morrissey JP. County-level estimates of mental health professional shortage in the United States. Psychiatr Serv 2009 Oct;60(10):1323-1328. [doi: 10.1176/ps.2009.60.10.1323] [Medline: 19797371]
15. Aguilera A. Digital Technology and Mental Health Interventions: Opportunities and Challenges. Arbor 2015 May 12;191(771):a210. [doi: 10.3989/arbor.2015.771n1012]
16. Yuen E, Goetter EM, Herbert JD, Forman EM. Challenges and opportunities in internet-mediated telemental health. Professional Psychology: Research and Practice 2012;43(1):1-8. [doi: 10.1037/a0025524]
17. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med 2007 Mar;37(3):319-328. [doi: 10.1017/S0033291706008944] [Medline: 17112400]
18. Barak A, Grohol JM. Current and Future Trends in Internet-Supported Mental Health Interventions. Journal of Technology in Human Services 2011 Jul;29(3):155-196. [doi: 10.1080/15228835.2011.616939]
19. Backhaus A, Agha Z, Maglione ML, Repp A, Ross B, Zuest DD, et al. Videoconferencing psychotherapy: a systematic review. Psychol Serv 2012 May;9(2):111-131. [doi: 10.1037/a0027924] [Medline: 22662727]
20. Kohn R, Saxena S, Levav I, Saraceno B. The treatment gap in mental health care. Bull World Health Organ 2004 Nov;82(11):858-866 [FREE Full text] [Medline: 15640922]
21. Mohr DC, Ho J, Duffecy J, Baron KG, Lehman KA, Jin L, et al. Perceived barriers to psychological treatments and their relationship to depression. J Clin Psychol 2010 Apr;66(4):394-409 [FREE Full text] [doi: 10.1002/jclp.20659] [Medline: 20127795]
22. Berrouiguet S, Baca-García E, Brandt S, Walter M, Courtet P. Fundamentals for Future Mobile-Health (mHealth): A Systematic Review of Mobile Phone and Web-Based Text Messaging in Mental Health. J Med Internet Res 2016 Dec 10;18(6):e135 [FREE Full text] [doi: 10.2196/jmir.5066] [Medline: 27287668]
23. Andersson G, Cuijpers P. Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 2009;38(4):196-205. [doi: 10.1080/16506070903318960] [Medline: 20183695]
24. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn Behav Ther 2018 Jan;47(1):1-18. [doi: 10.1080/16506073.2017.1401115] [Medline: 29215315]
25. Baumeister H, Reichler L, Munzinger M, Lin J. The impact of guidance on Internet-based mental health interventions — A systematic review. Internet Interventions 2014 Oct;1(4):205-215. [doi: 10.1016/j.invent.2014.08.003]
26. Lindhiem O, Bennett CB, Rosen D, Silk J. Mobile technology boosts the effectiveness of psychotherapy and behavioral interventions: a meta-analysis. Behav Modif 2015 Nov;39(6):785-804 [FREE Full text] [doi: 10.1177/0145445515595198] [Medline: 26187164]

27. Zickhur K, Smith A. Pew Research Center: Internet & Technology. 2013 Aug 26. Home Broadband 2013 URL: http://www.pewinternet.org/2013/08/26/home-broadband-2013/ [accessed 2018-11-23] [WebCite Cache ID 748fTFjIV]
28. Watts S, Mackenzie A, Thomas C, Griskaitis A, Mewton L, Williams A, et al. CBT for depression: a pilot RCT comparing mobile phone vs. computer. BMC Psychiatry 2013 Feb 07;13:49 [FREE Full text] [doi: 10.1186/1471-244X-13-49] [Medline: 23391304]
29. Thayer SE, Ray S. Online communication preferences across age, gender, and duration of Internet use. Cyberpsychol Behav 2006 Aug;9(4):432-440. [doi: 10.1089/cpb.2006.9.432] [Medline: 16901247]
30. Mohr D, Tomasino KN, Lattie EG, Palac HL, Kwasny MJ, Weingardt K, et al. IntelliCare: An Eclectic, Skills-Based App Suite for the Treatment of Depression and Anxiety. J Med Internet Res 2017 Dec 05;19(1):e10 [FREE Full text] [doi: 10.2196/jmir.6645] [Medline: 28057609]
31. Aguilera A, Bruehlman-Senecal E, Demasi O, Avila P. Automated Text Messaging as an Adjunct to Cognitive Behavioral Therapy for Depression: A Clinical Trial. J Med Internet Res 2017 Dec 08;19(5):e148 [FREE Full text] [doi: 10.2196/jmir.6914] [Medline: 28483742]
32. Furber G, Jones GM, Healey D, Bidargaddi N. A comparison between phone-based psychotherapy with and without text messaging support in between sessions for crisis patients. J Med Internet Res 2014 Oct 08;16(10):e219 [FREE Full text] [doi: 10.2196/jmir.3096] [Medline: 25295667]
33. Wright J, Wright AS, Albano AM, Basco MR, Goldsmith LJ, Raffield T, et al. Computer-assisted cognitive therapy for depression: maintaining efficacy while reducing therapist time. Am J Psychiatry 2005 Jun;162(6):1158-1164. [doi: 10.1176/appi.ajp.162.6.1158] [Medline: 15930065]
34. Hull TD, Mahan K. A Study of Asynchronous Mobile-Enabled SMS Text Psychotherapy. Telemed J E Health 2017 Dec;23(3):240-247. [doi: 10.1089/tmj.2016.0114] [Medline: 27797646]
35. Ogrodniczuk JS, Piper WE, Joyce AS, McCallum M. Effect of patient gender on outcome in two forms of short-term individual psychotherapy. J Psychother Pract Res 2001;10(2):69-78 [FREE Full text] [Medline: 11264331]
36. Ogrodniczuk J, Staats H. [Psychotherapy and gender: do men and women require different treatments?]. Z Psychosom Med Psychother 2002;48(3):270-285. [Medline: 12136448]
37. Ogrodniczuk J. Differences in Men'S and Women'S Responses to Short-Term Group Psychotherapy. Psychotherapy Research 2004 Jun 01;14(2):231-243. [doi: 10.1093/ptr/kph019]
38. Kirshner L, Genack A, Hauser ST. Effects of gender on short-term psychotherapy. Psychotherapy: Theory, Research & Practice 1978;15(2):158-167. [doi: 10.1037/h0085856]
39. Frank E, Carpenter LL, Kupfer DJ. Sex differences in recurrent depression: are there any that are significant? Am J Psychiatry 1988 Jan;145(1):41-45. [doi: 10.1176/ajp.145.1.41] [Medline: 3337291]
40. Addis ME, Mahalik JR. Men, masculinity, and the contexts of help seeking. Am Psychol 2003 Jan;58(1):5-14. [Medline: 12674814]
41. Wetherell J, Petkus AJ, Thorp SR, Stein NB, Chavira SA, Campbell-Sills L, et al. Age differences in treatment response to a collaborative care intervention for anxiety disorders. Br J Psychiatry 2013 Jul;203(1):65-72 [FREE Full text] [doi: 10.1192/bjp.bp.112.118547] [Medline: 23580378]
42. Hobbs MJ, Mahoney AEJ, Andrews G. Integrating iCBT for generalized anxiety disorder into routine clinical care: Treatment effects across the adult lifespan. J Anxiety Disord 2017 Oct;51:47-54. [doi: 10.1016/j.janxdis.2017.09.003] [Medline: 28926805]
43. Hansen NB, Lambert MJ, Forman EM. The psychotherapy dose-response effect and its implications for treatment delivery services. Clinical Psychology: Science and Practice 2002 Sep;9(3):343. [doi: 10.1093/clipsy.9.3.329]
44. van Drongelen A, Boot CR, Hlobil H, Twisk JW, Smid T, van der Beek AJ. Evaluation of an mHealth intervention aiming to improve health-related behavior and sleep and reduce fatigue among airline pilots. Scand J Work Environ Health 2014 Nov;40(6):557-568 [FREE Full text] [doi: 10.5271/sjweh.3447] [Medline: 25121620]
45. Mattila E, Lappalainen R, Välkkynen P, Sairanen E, Lappalainen P, Karhunen L, et al. Usage and Dose Response of a Mobile Acceptance and Commitment Therapy App: Secondary Analysis of the Intervention Arm of a Randomized Controlled Trial. JMIR Mhealth Uhealth 2016 Jul 28;4(3):e90 [FREE Full text] [doi: 10.2196/mhealth.5241] [Medline: 27468653]
46. Evans W, Nielsen PE, Szekely DR, Bihm JW, Murray EA, Snider J, et al. Dose-response effects of the text4baby mobile health program: randomized controlled trial. JMIR Mhealth Uhealth 2015 Jan 28;3(1):e12 [FREE Full text] [doi: 10.2196/mhealth.3909] [Medline: 25630361]
47. Bowen D, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
48. Singal AG, Higgins PDR, Waljee AK. A primer on effectiveness and efficacy trials. Clin Transl Gastroenterol 2014 Jan 02;5:e45 [FREE Full text] [doi: 10.1038/ctg.2013.13] [Medline: 24384867]
49. Kim SY. Efficacy versus Effectiveness. Korean J Fam Med 2013 Jul;34(4):227 [FREE Full text] [doi: 10.4082/kjfm.2013.34.4.227] [Medline: 23904951]
50. Depression Guideline Team. Michigan Medicine: University of Michigan. 2011 Aug. UMHS Depression Guideline Update URL: http://www.med.umich.edu/1info/FHP/practiceguides/depress/depress.pdf [accessed 2018-11-23] [WebCite Cache ID 748jjuvZE]

51. Nulty D. The adequacy of response rates to online and paper surveys: what can be done? Assessment & Evaluation in Higher Education 2008 Jun;33(3):301-314. [doi: 10.1080/02602930701293231]
52. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]
53. Kroenke K, Spitzer R. The PHQ-9: A New Depression Diagnostic and Severity Measure. Psychiatric Annals 2002 Sep 01;32(9):509-515. [doi: 10.3928/0048-5713-20020901-06]
54. Löwe B, Unützer J, Callahan CM, Perkins AJ, Kroenke K. Monitoring depression treatment outcomes with the patient health questionnaire-9. Med Care 2004 Dec;42(12):1194-1201. [Medline: 15550799]
55. Munder T, Wilmers F, Leonhart R, Linster HW, Barth J. Working Alliance Inventory-Short Revised (WAI-SR): psychometric properties in outpatients and inpatients. Clin Psychol Psychother 2010;17(3):231-239 [FREE Full text] [doi: 10.1002/cpp.658] [Medline: 20013760]
56. BetterHelp. Convenient, affordable, private online counseling, anytime, anywhere URL: https://www.betterhelp.com/ [accessed 2018-05-01] [WebCite Cache ID 6z6Er5RHX]
57. Ross A, Willson V. Paired Samples T-Test. In Basic and Advanced Statistical Tests. Rotterdam: SensePublishers; 2017:17-19.
58. Hedberg EC, Ayers S. The power of a paired t-test with a covariate. Soc Sci Res 2015 Mar;50:277-291 [FREE Full text] [doi: 10.1016/j.ssresearch.2014.12.004] [Medline: 25592936]
59. Howard KI, Kopta SM, Krause MS, Orlinsky DE. The dose-effect relationship in psychotherapy. Am Psychol 1986 Feb;41(2):159-164. [Medline: 3516036]
60. Barnett A, van der Pols JC, Dobson AJ. Regression to the mean: what it is and how to deal with it. Int J Epidemiol 2005 Feb;34(1):215-220. [doi: 10.1093/ije/dyh299] [Medline: 15333621]
61. Reid DJ, Reid FJM. Text or talk? Social anxiety, loneliness, and divergent preferences for cell phone use. Cyberpsychol Behav 2007 Jun;10(3):424-435. [doi: 10.1089/cpb.2006.9936] [Medline: 17594267]
62. Forgays D, Hyman I, Schreiber J. Texting everywhere for everything: Gender and age differences in cell phone etiquette and use. Computers in Human Behavior 2014 Feb;31:314-321. [doi: 10.1016/j.chb.2013.10.053]
63. Hamilton KE, Dobson KS. Cognitive therapy of depression: pretreatment patient predictors of outcome. Clin Psychol Rev 2002 Jul;22(6):875-893. [Medline: 12214329]
64. Schlagert HS, Hiller W. The predictive value of early response in patients with depressive disorders. Psychother Res 2017 Dec;27(4):488-500. [doi: 10.1080/10503307.2015.1119329] [Medline: 26776721]
65. Kessing LV, Hansen MG, Andersen PK, Angst J. The predictive effect of episodes on the risk of recurrence in depressive and bipolar disorders - a life-long perspective. Acta Psychiatr Scand 2004 May;109(5):339-344. [doi: 10.1046/j.1600-0447.2003.00266.x] [Medline: 15049770]

Abbreviations

ANCOVA: analysis of covariance
PHQ-9: Patient Health Questionnaire
SMS: short message service
WAI-SR: Working Alliance Inventory-Short Revised

Edited by G Eysenbach; submitted 02.05.18; peer-reviewed by J Lundgren, W Tumuhimbise; comments to author 07.06.18; revised version received 08.10.18; accepted 26.10.18; published 23.01.19.

Please cite as:
Marcelle ET, Nolting L, Hinshaw SP, Aguilera A
Effectiveness of a Multimodal Digital Psychotherapy Platform for Adult Depression: A Naturalistic Feasibility Study
JMIR Mhealth Uhealth 2019;7(1):e10948
URL: http://mhealth.jmir.org/2019/1/e10948/
doi:10.2196/10948
PMID:30674448

Original Paper

The Implementation of an Innovative Hydration Monitoring App in Care Home Settings: A Qualitative Study

Alison Steven[1], RN, BSc (Hons), PGCRM, MSc, PhD; Gemma Wilson[1], BSc, MSc, PhD; Lesley Young-Murphy[2], RGN, HV, BSc (Hons) Midwifery, BSc (Hons) Community Health, PhD

[1]Department of Nursing, Midwifery and Health, Faculty of Health and Life Sciences, Northumbria University Newcastle, Newcastle upon Tyne, United Kingdom

[2]NHS North Tyneside Clinical Commissioning Group, North Shields, United Kingdom

Corresponding Author:
Alison Steven, RN, BSc (Hons), PGCRM, MSc, PhD
Department of Nursing, Midwifery and Health
Faculty of Health and Life Sciences
Northumbria University Newcastle
Coach Lane Campus, Room B121
Benton
Newcastle upon Tyne, NE77XA
United Kingdom
Phone: 44 191215 ext 6483
Email: alison.steven@northumbria.ac.uk

Abstract

Background: In response to marked concern regarding inadequate fluid intake recording in care homes, an innovative mobile hydration app was collaboratively developed. “Hydr8” aimed to facilitate accurate recording and communication of residents’ fluid intake and ultimately increase care quality and patient safety.

Objective: The aim of this study was to examine the implementation of Hydr8 in a sample of care homes in one area in England.

Methods: The principles of Realist Evaluation and Action research were drawn upon throughout the study. Overall, 5 care homes participated in this study, 3 interview-only sites and 2 case-study sites, where interviews and observations were conducted at 3 time-points. Furthermore, 28 staff members participated, including care staff, management, a registered nurse, and administrative staff.

Results: Findings suggest that Hydr8 benefits practice, enhancing the understanding of hydration and person-centered care and improving staff communication. However, technical glitches hindered the seamless embedding of Hydr8 into everyday practice, and enthusiasm for long-term use was dependent on the resolution of issues. In addition, Hydr8 heightened perceptions of personal accountability, and while managers viewed this as positive, some staff members were apprehensive. However, individuals were enthusiastic about the long-term use and potential of Hydr8.

Conclusions: Utilizing the findings of this study to further develop and adapt Hydr8 indicates the long-term use of Hydr8 as promising. Although perceptions of Hydr8 were primarily positive, setbacks in its implementation and use created difficulties in normalizing the solution into everyday practice. This study highlights the need for education related to hydration practice and a change of infrastructure in care home settings to implement technical solutions and changes to care.

(JMIR Mhealth Uhealth 2019;7(1):e9892) doi:10.2196/mhealth.9892

KEYWORDS

education; frail elderly; mobile apps; patient safety; residential facilities; water-electrolyte balance

Introduction

Background

Hydration management is recognized as essential to older adults’ care, with age-related variations increasing vulnerability to dehydration risk [1]. Dehydration in older adults is a patient safety concern and clinically associated with stroke, diabetes, influenza, constipation, respiratory infection, gastroenteritis, urinary tract infection, delirium, seizure, risk of falling, and

mortality [1-3]. A UK-based study found over a third of older adults dehydrated on hospital admission [4].

Despite dehydration being largely preventable, care homes reportedly fail to consistently provide adequate fluids to residents [5,6]. From data obtained under freedom of information laws, it was found that 1158 care home residents in the United Kingdom suffered dehydration-related deaths between 2003 and 2012 [7]. In an analysis of death certificates, it was reported that dehydration was either the leading cause of death or a contributory factor [7].

Fundamental issues affecting hydration management are the recording of information and encouraging fluid intake. Charts for recording hydration elements such as fluid intake, fluid output, or fluid balance are frequently used with the aim of capturing fluid status and assisting deficit identification. The accurate recording of fluid-balance information is fundamental to safe care [8]; however, while monitoring fluid balance may be viewed as a simple task, completion of records is notoriously inadequate or inaccurate [9,10]. Research investigating the completion of fluid-balance charts in hospital wards found none were completed appropriately [11]. Staff shortages, lack of training, and lack of time were cited as reasons for incomplete and inaccurate charts [11]. In addition, further research highlighted problems with fluid-balance records due to a lack of communication between a hospital ward health care team and a lack of awareness and education of the importance of fluid status, especially among staff members most often completing records [8].

While hospitals have similar basic features across the globe, a care home in the United Kingdom is a residential setting in which older adults typically live in single rooms with on-site care services [12]. Care can either be paid for personally or by either the National Health Service or the local government. Care home staff requirements are regulated as part of the Health and Social Care Act 2008 and comprise largely of “care” assistants, professionally qualified nurses, and management staff.

The Hydration Solution App

An innovative mobile hydration app “Hydr8” was developed in response to concerns regarding older adults’ hydration management and poor completion of records. Issues with the numeracy skills of some care home staff members, together with nonstandardization of recording the cup or vessel size were also considered during the coproduction of the app. While many hydration apps exist in the general market, these tend to target individuals inputting their own hydration levels. Hydr8 is an app specifically developed to be used in a health care setting by health professionals inputting hydration data for residents in a care home. In addition, unlike many existing apps, Hydr8 enables personalization to individual needs (eg, safety requirements such as thickened fluids) and preferences (eg, residents’ likes and dislikes). Furthermore, a clinical commissioning group (CCG), software development company, and care home managers worked collaboratively to coproduce Hydr8. During the development phase, a focus group approach was used to involve patients, relatives, and care home staff. This approach enabled discussion and contributions to be made regarding the appearance of the app after which a hard copy of the initial design was taken back to care homes for further comments by other staff members, patients, and relatives.

As discussed, inaccurate recording of hydration information has various implications on patients’ safety. Hydr8 aims to (1) facilitate accurate recording and communication of residents’ fluid intake; (2) automate fluid recordings and maximize the use of accessible technology; (3) enable care home staff to see cumulative totals for each resident’s intake; (4) be time-efficient, thus, releasing staff to engage in more care and leadership; (5) enable individualized care; and (6) improve awareness of the importance of hydration. Personalization to individual needs (eg, safety requirements such as thickened fluids) and preferences (eg, residents’ likes and dislikes) also increases the likelihood of maintaining hydration.

To ensure appropriate individual targets, volumes were calculated per the existing CCG and care home policy. This involved a base calculation of 30 mL fluid per kg body weight, with the addendum of 1500 mL per day as a minimum for older people [13]. This base calculation was then tailored to individually assessed needs through discussions with medical staff (ie, general practitioner or medical consultant) and other clinical staff (eg, registered nurses, dietitians, or allied health professionals involved in the individuals care). These discussions took into consideration individual health conditions, comorbidities, and treatment regimes.

Hydr8 comprises two core parts: the back-system accessed through a Web browser and a tablet-based app. Both components are accessed through username and password. The back-system permits users to add or remove residents from the app and allows them to view data across various time periods. In addition, the back-system provides opportunities for health professionals, including doctors (eg, general practitioners) or registered nurses, to access this information in real time while off-site. The app displays personalized breakdowns of fluid intake including the current daily level, last time fluids were given, and an overview of fluid intake covering the previous 7 days. These factors are visually illustrated through a body outline that fills with water as recorded fluid intake increases, with adequacy levels indicated in red, amber, and green (Figure 1). These colored levels (daily and 7-day levels; Figures 1 and 2) act as a visual signal and warning to staff.

Hydr8 enables further personalization by allowing the input of residents’ photograph, their likes and dislikes, and information on choking hazards, which is displayed using a pop-up notification. Hydr8 sends an alert when residents fall below optimum levels of fluid intake.

We aimed to explore and evaluate the pilot implementation of Hydr8 in care homes, with a particular focus on the operationalization of the system, impact on care provision, and the development needs of staff.

Figure 1. A screenshot from Hydr8 showing fluid intake information. Source: "Hydr8" Brochure produced by Elaros, North Tyneside Clinical Commissioning Group and the Academic Health Science Network North East and North Cumbria.

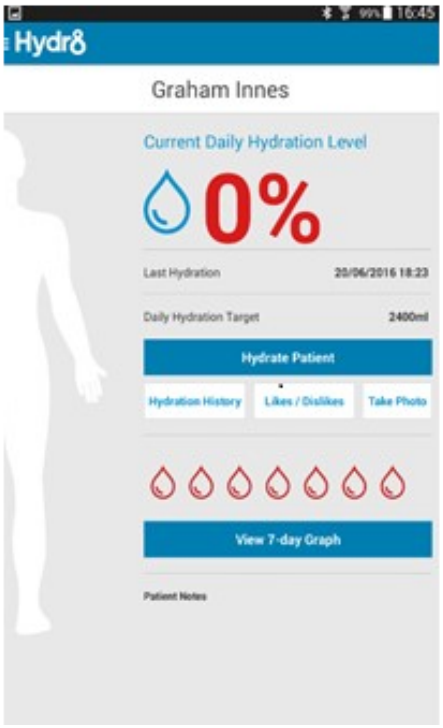


Figure 2. A screenshot from Hydr8 showing input screens. Source: "Hydr8" Brochure produced by Elaros, North Tyneside Clinical Commissioning Group and the Academic Health Science Network North East and North Cumbria.

Methods

Study Design

Researching the implementation and impact of new practices or interventions in health care is problematic given the complex, context-bound nature of everyday care [14-16]. Due to the multitude of things that can influence variations in practice, traditional quantitative methods are not adequate to discern and understand the impact of complex interventions and new initiatives, be they educational or technological [16]. This difficulty is viewed as analogous to that encountered in the evaluation of complex interventions [17-19]. This inquiry, therefore, drew on the principles of Realistic Evaluation [16], which emphasizes the role of context, taking into account, for example, differing organizational settings, workforce, teams, and sociopolitical issues [20]. Akin to action research [21], interim findings were feedback to the CCG and app developers on a frequent basis to continuously develop and improve Hydr8. In addition, the Normalization Process Theory was used as a lens through which to explore the embedding and "normalizing" of Hydr8 into everyday practice. A qualitative design was utilized, encompassing observations and interviews.

Ethical Considerations

This project was approved by the Faculty of Health and Life Sciences Research Ethics Committee at Northumbria University.

Study Population

Data were gathered from care home sites within one CCG locality in the North of England. In this study, 5 care homes participated: 3 interview-only sites and 2 case-study sites. A sixth site declined to participate. Data collection at interview-only sites consisted of semistructured interviews at one time-point, whereas at case-study sites, it consisted of observations and semistructured interviews at 3 separate time-points.

In the study locality, the care staff age profile ranged from 19 to 60 years. Currently, for employment, UK care home staff are required to have a minimum of "National Care Certificate" qualification [22] or are obliged to work toward this within the first 6 months of employment. Depending on their role and length of time in employment, existing staff may hold levels 2-4 of the previously used "National Vocational Qualifications," or more recent Qualifications and Credit Framework, level 2 Diploma in Care. From January 1, 2018, these previous qualifications were replaced by the "Regulated Qualifications Framework" [23].

Data Collection

Once care home area management had given written consent for each home to be approached, GW and AS met with local management at each site to provide a study overview, discuss the study process, and disseminate information to staff. Participant information sheets and reply slips were left in a communal area of each home. If staff members were happy to participate, they were asked to leave a reply slip containing their details in a sealed box provided. This ensured anonymity of responses. Returned reply slips were collected after 7 days, and a time was arranged to return and collect data.

Semistructured interviews were conducted with staff in a quiet location in the care home. Before interviews began, participants were encouraged to ask questions about the study and then sign a consent form. Participants were advised they could withdraw from the study at any point. Interviews explored the use of the system in everyday practice, its ease and relevance, perceptions of purpose, worth, value and impact, and perceptions of development needs (Textbox 1).

In case-study sites, observations were also conducted at 3 time-points: around 1 month, 5 months, and 8 months after using the app. This allowed a continued examination of its use and changes in use over the study period. GW and AS observed and took notes quietly in the corner of a room, watching Hydr8 being used. Observations lasted up to an hour and focused on the use and usability of the system; the normalization of the system as part of everyday practice; visible impacts on care provision and outcomes; and potential education, development, and training needs (Textbox 2).

Some staff also made spontaneous comments that were recorded. Only the staff that had provided informed, written consent were observed. Semistructured interviews were undertaken at 3 time-points following the approach used at interview-only sites but with the addition of questions regarding the observations.

Data Analysis

All interviews were recorded digitally and transcribed verbatim. Observation field notes and interview transcripts were analyzed first by each member of the research team using thematic analysis [24] and facilitated by NVIVO 10 software (QSR International Pty Ltd). The thematic analysis aims to extract themes and subthemes from interview data highlighting patterns within the dataset [24]. Specifically, the analysis followed the 6 steps of conducting the thematic analysis: familiarizing yourself with the data, generating initial codes, searching for themes, reviewing themes, defining and naming themes, and producing the report [24]. Data and initial coding were compared and discussed with the wider research team to challenge, refine, and confirm emerging findings and ensure they were rooted in the original data. In line with action research principles [21], interim findings were intermittently fed back to the development team. Realist evaluation [16] and the normalization process theory [14] were drawn upon throughout.

Textbox 1. Interview schedule.

Introduction

- Seek verbal consent, answer any questions, and explain recording device

Opening: Prompts provided and examples sought throughout

- Have you worked here long?
- What's your job role?
- Do you have a lot of input with the residents?
- With what they eat or drink?
- Are you involved in recording what they drink or eat?

The Hydr8 system: Prompts provided and examples sought throughout

- Are you aware of the new Hydr8 system?
- Do you use it? or Does everyone know about it? or Who uses it?
- Can you tell me a bit about your experiences of using it?
- What was it like to use the first time? or Did it take time to get used to?
- Were you shown how to use it? or How was it to learn to use?
- What happened if you got stuck?
- What is it like to carry about?

Perceived impact on care provision and outcomes: Prompts provided and examples sought throughout

- Do you use the app instead of other monitoring tools, or as a duplicate?
- How does it fit in with other tasks or practice?
- Does it make a difference? Has it changed anything?
- To your work
- To the residents
- To other staff
- Has anything changed since you started using it?

Embedding: Prompts provided and examples sought throughout

- What do you think about it?
- Is it relevant to your job?
- Do you see a point to it?
- What did everyone think about it?

Textbox 2. Observation sheet.

The use and usability of the system
How long does it take to fill in?
Do they fill it in when they give drink or when drink is finished?
Do they fill it in easily?
Any technical difficulties presented with device or app?
Any usability issues with device or app?
Any verbal or visible frustration with the device or app?
Do different members of staff use the app differently?
Are they filled in enough or correctly?
Any obvious facilitators or barriers?
Normalization of the system as part of everyday practice
Who uses the app?
Complete every time? More than one entry per time?
Do they use in addition to other balance charts?
Do staff automatically record information on app or is it a second thought?
Any obvious facilitators or barriers?
Does the completion of the app seem to work well with other tasks or does it get in the way?
Impacts on care provision and outcomes
Discuss app or hydration with resident when completing?
Does it seem to affect the amount of fluid given to residents?
Do members of staff ask resident questions about adequate or inadequate hydration levels over the week?
Used differently with any residents?
Are drinks given appropriately?
Potential education, development, and training needs of staff
Any discussions about the app between staff or staff and residents or residents?
Help give by staff to other staff completing this?
Do staff or residents seem to understand its importance?

Results

Data Collection Statistics

In total, 10 interviews (3 at site 1, 2 at site 2, and 5 at site 3) were conducted at interview-only sites. Table 1 shows data collection for the case-study sites (n=2). Observations provided contextual understanding that helped situate and make sense of interview findings (Table 1).

Overall, 28 participants took part in interviews over 5 sites. Of these, 21 were care staff members, 5 from management, 1 administrative assistant, and 1 registered nurse. Care staff members were most frequently interviewed, as they were the staff predominantly tasked with monitoring hydration and, thus, mainly used Hydr8.

Findings

Four interrelated themes emerged: knowledge of hydration, fitting into established systems of care, surveillance, and future gazing.

Knowledge of Hydration

A positive outcome of Hydr8 was the impact on care home staffs' knowledge and understanding of hydration. Evidently, visual illustrations displaying fluid intake were more meaningful than paper-based charts.

> *It means less on paper* [P006/care assistant]
>
> *If you're looking on the app you can think "oh-well actually…he didn't drink that one, he could do with a bit more." So you are pushing fluids with that particular person […] you wouldn't if it was on paper because you wouldn't realise, but now that it is visual, giving you the push* [P025/care assistant]

Table 1. Number of observations and interviews conducted over the 3 data collection points for the case-study sites.

Case-study sites	Visits (n)			
	Visit 1	Visit 2	Visit 3	Total
Site 1				
Observations conducted	1	0	0	1
Interviews conducted	4	2	3	9
Site 2				
Observations conducted	1	0	0	1
Interviews conducted	4	3	2	9

Through using the app, staff also gained a heightened awareness of individual preferences and individual differences in fluid intake.

> *You can put in their likes and dislikes if, like, they would prefer a drink. So there's like some ladies that just like a cup of tea or milk. There's others who, like, quite prefer a colder drink. So, it knows.* [P001/care assistant]

> *It calculates, everyone's difference. Like, weight, size, and how much they sort of, should, need.* [P001/care assistant]

In addition, there were recognized changes in practice due to heighted awareness of the importance of contextual factors and individual differences, increasing person-centered care.

> *If it is warm, obviously, the staff are aware and I've heard them say, "it's warm today, we'll get some extra drinks out." Or juice as opposed to a cup of tea* [P003/management]

> *The carers are a bit more involved…it's down to height and weight, medical history […] It's quite interesting for the carers to see that certain residents need more fluids, and other residents need less* [P003/management]

However, Hydr8 also had unintended consequences with some staff "frightened" (P002/care assistant) of overhydrating residents. At the time of the study, Hydr8 did not record fluid output or compute fluid balance; therefore, percentage data could show residents >100% recommended intake and, notwithstanding clarifications, this caused some anxiety. Feedback to the commissioners and developers regarding this issue prompted consideration of the future development of the app to include fluid output.

Fitting Into Established Systems of Care

The normalization of new technology into everyday practice is an important consideration in implementation. A number of technical issues, glitches, and knock-on effects emerged and impacted the embedding of Hydr8 into routine care. In the short-term, during the course of the study, Wi-Fi connectively was often poor, which was time-consuming for users and often resulted in the delayed record completion.

> *There is a lot of loading that you don't have with paperwork […] It's just that you can't wait around for ages every time you want to record something* [P004/care assistant]

> *It freezes, it skips, it jumps, it doesn't load. The Wi-Fi connection keeps coming off and doesn't connect back up to the Wi-Fi* [P011/care assistant]

> *We use it upstairs, but the problem is through the [Wi-Fi] signal, we cannot get a signal upstairs* [P020/care assistant]

These faults led to time taken away from other duties and fueled staff frustration. In addition, Hydr8 repeatedly "froze," resulting in care homes not being able to use the app for long periods of time.

> *There's been a few times where it just crashes and it has been saying, unfortunately Hydr8 has stopped* [P004/management]

Feedback of these issues to the commissioners resulted in Wi-Fi boosters being provided to care homes with limited connectivity. While some system errors created problems, connectivity issues related to poor Wi-Fi prevailed and undoubtedly impacted the normalization of Hydr8 into daily practice. Other factors affecting normalization related to the implementation being integral to the ongoing development process. For example, during the study, most care homes (n=4) continued to complete paper-based charts to ensure no data could be lost. This duplication of information necessitated additional staff time, and many care staff members were unaware that the duplication was a short-term measure; therefore, Hydr8 was often viewed as an additional task in an already demanding workload.

> *I think they'll love the app once the paperwork goes* [P023/management]

Furthermore, given this was a developmental phase a limited number of tablets were supplied to each care home (n=2), which was perceived as insufficient.

> *They haven't necessarily been able to record at that moment in time, because somebody else has been using* [P003/management]

Daily routines were affected by time spent searching for devices, and participants were not always able to input data when they needed to. To manage the technical glitches, duplication, and lack of tablets, participants developed "workarounds." Workarounds included carrying information on paper for uploading later, thus, enabling continued recording of data

despite the issues experienced. These extra activities also impacted the embedding of Hydr8 into everyday working.

Surveillance

An interesting and unanticipated finding was participants' perceptions that Hydr8 may function as a method of surveillance for management and external agencies. There were also apprehensions that external agencies may not fully understand data produced. Hydr8 potentially heightened accountability, and there was anxiety regarding the way in which the data were presented, how these may appear to others, and potential for increased individual accountability.

> *When [the external agency] come in, they do go through paperwork and bits and I don't know how they are going to react with having to go through this* [P014/care assistant]
>
> *It looks like we flooded them* [P011/care assistant]
>
> *[The manager] can keep an eye on it as well. So, if, like, someone has missed a drink or something…he can come up straight away and say, "look, why hasn't this one had a drink for 3, 4 hours?"* [P026/care assistant]

However, management staff viewed the accountability potential and the possibility to remotely access records as beneficial.

> *Having that accountability is important* [P003/management]
>
> *I've been over the moon with being able to observe from the office. The board of directors have actually been sitting in Harrogate observing* [P016/management]
>
> *I look at it from just after lunch every day. And I sit and go through it. And as soon as I see the deficit, a concern or a problem, I'm out and I want to know why* [P023/management]

Hydr8 was advantageous due to the possibility to view data from a 7-day period, and to do this remotely, thus, increasing the potential for communication between stakeholders.

Future Gazing

Respondents often talked about "technology" as a concept, and individuals often discussed the inevitability of technology becoming an integral part of their future roles.

> *It's the next, sort of, generation* [P001/care assistant]
>
> *It's definitely the way forward* [P003/management]

However, design changes such as the ability to edit inputted data and increased flexibility were repeatedly raised by participants and felt to be imperative to ensure long-term use.

> *They're not editable either. I know they are on the back end, but it means that the carer makes a typing error – there's nothing they can do* [P004/management]
>
> *[It needs to be] as flexible as a piece of paper* [P014/care assistant]

Although conditional on the elimination of technical setbacks and connectivity issues, participants were enthusiastic about the future use of Hydr8

> *I just think it is a brilliant idea if it all runs smoothly and works* [P016/management]
>
> *If it was working properly and it wasn't getting stuck, it would be brilliant. So much easier* [P026/care assistant]

Furthermore, enhancements and additional functions were deemed necessary for the long-term use of Hydr8. These included "output" (P011/care assistant) and "food charts" (P023/management). Individuals also suggested the inclusion of a "24-hour personal care record" (P027/management) and additions to render Hydr8 suitable for residents with "dementia" (P023/management) or those at the "end of life stage" (P027/management). Participants felt such improvements would improve person-centered care and were enthusiastic about using Hydr8 in the future. All of these issues were feedback to the development team in a timely manner

Discussion

Principal Findings

Specific benefits of the Hydr8 app and solution include heightened staff understanding of hydration, increased person-centered care, and enhanced communication. However, participants also proposed additions and enhancements that would further improve Hydr8.

Hydr8 increased staff awareness and understanding of individual and contextual factors in hydration management. The importance of staff education to avoid dehydration has been highlighted in the literature [8,25-27], and systems such as Hydr8 could offer additional opportunities for work-based education relevant to the client or patient group being cared for. Information recorded using Hydr8 reflected the importance of changes in culture regarding nutrition and hydration practice and a need for a person-centered approach in recording fluid preferences and individual needs [1,27]. Understanding individual differences is an essential part of hydration management when encouraging older adults to drink more [12,27], and it was apparent that Hydr8 data were more meaningful and individual compared with traditional paper records. Indeed, the visual "KANBAN" (which means signboard or billboard in Japanese) [28] type signal given by the body shape and red, amber, or green (RAG) app display appeared to heighten staff awareness. The additional "backroom" facility, allowing managers to see the RAG rating at a glance, provides further overarching assurance, and the use of these levels of visual alert together with staff and manager monitoring may offer a certain level of "mistake proofing." However, one unintended consequence, linked to the recording of fluid intake only, was increased anxiety felt by some staff regarding overhydration. This indicates the need for future developments of Hydr8 to include output and fluid-balance calculation and further preparation and education for staff.

Hydr8 enabled fluid intake information to be communicated more effectively given multiple individuals (with permission) could view data charts covering a 7-day period and could do so

remotely. Management valued this function, as it improved their longitudinal awareness of fluid intake. Hydration could be charted over days, allowing greater sensitivity to gradual dehydration, thereby adding a further quality and safety check into the care system. Indeed, in a recent literature review, Oates and Price [27] concluded that hydration should be a collective responsibility and management also noted the increased staff accountability Hydr8 offered. Hydr8 aimed to be efficient and release staff to engage in more care and leadership activities; however, technical and implementation difficulties increased time spent recording fluid intake. One disadvantage of paper-based fluid-balance charts is that input can be time-consuming [8]; therefore, it was imperative that Hydr8 be time-efficient to make it a more "attractive" option and engender "buy-in."

In this study, the Hydr8 system did not appear to become completely routinized or "normalized" into daily practice [26]. There was some coherence in the understanding of the goals and aims of Hydr8 and some "buy-in" by staff (illustrated by the future gazing and knowledge enhancement). However, some participants were apprehensive, perceiving Hydr8 as a potential staff surveillance and monitoring tool; this unease was heightened by the technical difficulties that resulted in recording inaccuracies. These apprehensions and staff not being fully aware of the iterative, developmental nature of the "pilot" implementation project may have limited the buy-in (or complete cognitive participation) by staff [26,29]. Furthermore, the "fit" of the Hydr8 system into existing skill sets and working practices (collective action) was hampered by the technical difficulties experienced, which disrupted the use of the app [26,29]. Despite the introduction of Wi-Fi boosters into some care homes, technical difficulties persisted because of poor Wi-Fi connectivity. With further development, these issues can be resolved, and the use of Hydr8 may result in time savings and staff being freed up for other duties. Furthermore, from this study, the importance of collaborating with software developers and companies who have an insight into, and understanding of, the complexities of the health and social care sector has emerged. This, however, remains a hypothesis and the implementation of new working practices does not always follow a preconceived logic [15]; therefore, further research is necessary to ascertain the consequences, intended or unintended, of the use of a refined Hydr8 system.

The findings illustrate the importance of technology being embedded in practice routines and culture. The implementation of technology is not simply about the device itself but the many connected sociomaterial "things" being introduced into existing social practices [30]. Introducing a new practice that is not sufficiently refined or tested may result in participants disengaging or expressing unfavorable opinions, as in this study. However, it could be argued that new practices (systems or technology) cannot be comprehensively developed before some level of implementation takes place, be it through small-scale implementation and/or piloting. Indeed, it is this period of testing and trying out that allows unforeseen issues and consequences to emerge and be resolved. Thus, the issue here was not the "piloting" in and of itself, but the need, perhaps, for much greater engagement of the care staff in the cocreation of Hydr8. Greater collaborative engagement of this section of stakeholders may have resulted in them being much more alert to emerging issues and may have raised their tolerance and allowed them to develop more complete cognitive participation, to see beyond the short-term disadvantages, specifically the duplication of information, limited tablets, and Wi-Fi connectivity issues.

This study mirrored aspects of action research [21] by investigating the implementation while also feeding back into the developmental process. As this inquiry was undertaken and resulting from the feedback of findings, commissioners and developers are working on developing this app by adding further elements, such as fluid output and nutrition, all into one app. Although there are only a small number of residents in which urine output is accurately measured in nursing homes (and during the study this was recorded for residents using traditional charts), clearly this is important in other settings and as part of the future development, adoption, and spread of the Hydr8 system.

Limitations

One limitation of this study is that the sample was small and restricted to a specific geographical area; therefore, it is not representative of the wider population, and care must be taken when extrapolating the findings. Although not generalizable, these findings have some transferability [31]. Technical issues reduced the use of Hydr8 during this pilot study, limiting its use and preventing observations as part of data collection; however, this in itself was an important aspect of the developmental enquiry. The technical issues negatively impacted the use and effectiveness of the Hydr8 app, and technical functionality is necessary before further implementing the app in care homes. This study has highlighted the importance of monitoring ongoing technical issues during wider implementation and, as a result, the clinical commissioning group has engaged a third party that identified and rectified technical coding issues, which were at the heart of some of the problems encountered. Rectifying these issues will enable more seamless use of the app and transfer of data in the future.

Staff interviewed were those who volunteered to participate on the day, and this was the deciding factor in the numbers involved. In addition, some sites (n=2) only implemented the use of the Hydr8 app in specific parts of the care home, and the researchers were not aware of it until data collection took place. Thus, demographic details were not collected from the individual staff, and this is acknowledged as a limitation for inclusion in any future research.

Future Considerations

Based on the findings of this pilot evaluation, Hydr8 will be further developed and evaluated. The focus of further study needs to encompass multiple aspects of use, including normalization into a daily routine, technical issues experienced, information needed on implementation, residents' perceptions, and participants' content and design suggestions.

A future longitudinal study is planned and will incorporate additional collection and analysis of long-term quality and safety outcomes. "Backroom" quantitative data regarding the aspects of app usage and individual resident recordings are constantly

being collected by the system, and this has been ongoing since the initial implementation and piloting. While these data are visible to care homes and CCG, these are yet to be analyzed, and these would form part of planned longitudinal research. This will not only allow further assessment of its use but also include economic evaluation and residents' perceptions of hydration management before and during the use of Hydr8. The development of plans and materials for staff preparation, education, and training for further roll out of the system is ongoing by the CCG. Such plans include investigation of peer-to-peer education, use of Hydr8 champions, and both Web-based and traditional, paper-based materials.

In addition, the study reported in this paper highlights the need for ongoing research into the human factors involved in the implementation and normalization of this system, including staff education regarding hydration and information technology literacy and individual perceptions and behaviors of residents and those of relatives and visitors. While this study took place in the United Kingdom, these issues regarding health and social care economies and delivery of best care to aging populations are of global concern.

Conclusions

This developmental inquiry highlights the potential benefits of utilizing this electronic hydration monitoring solution in the care home setting. Specifically, the use of Hydr8 increased understanding of hydration practice and improved communication of fluid intake data; furthermore, individuals were enthusiastic about its future use in the care home settings. The developmental process led to issues being highlighted and changes being implemented during the process. However, further considerations need to be taken into account for future implementation, namely, design and technical difficulties and staff education in the care home setting. Hydr8, with the necessary amendments highlighted in this study, has the potential to effectively improve the quality and safety of care.

Acknowledgments

We acknowledge the funding provided by the UK Academic Health Science Network (AHSN) for the North East and North Cumbria. We express our thanks to the participants who devoted their time to take part in this project; we hope we have represented their views with fairness. In addition, we acknowledge the important contribution of our late, dear colleague Dr Deborah Goodhall, who was key in undertaking the literature review that was the foundation for this project.

Conflicts of Interest

LYM was the lead applicant on the Academic Health Sciences Network bid, which led to the development of the Hydr8 app and piloting it in nursing homes—nursing homes in the locality agreed to be cited as part of that bid. The initial company that developed the app was chosen following a mini tender exercise. The app is now (at the time of submission for publishing) being commercialized by Elaros, a partner organization that North Tyneside CCG is now part of which is in line with the AHSN aspirations.

References

1. Mentes J. Oral hydration in older adults: greater awareness is needed in preventing, recognizing, and treating dehydration. Am J Nurs 2006 Jun;106(6):40-9; quiz 50. [Medline: 16728843]
2. Sheehy CM, Perry PA, Cromwell SL. Dehydration: biological considerations, age-related changes, and risk factors in older adults. Biol Res Nurs 1999 Jul;1(1):30-37. [doi: 10.1177/109980049900100105] [Medline: 11225294]
3. Lecko C. Improving hydration: an issue of safety. Nursing and Residential Care 2008 Mar;10(3):149-150. [doi: 10.12968/nrec.2008.10.3.28593]
4. El-Sharkawy AM, Watson P, Neal KR, Ljungqvist O, Maughan RJ, Sahota O, et al. Hydration and outcome in older patients admitted to hospital (The HOOP prospective cohort study). Age Ageing 2015 Nov;44(6):943-947 [FREE Full text] [doi: 10.1093/ageing/afv119] [Medline: 26316508]
5. Begum M, Johnson CS. A review of the literature on dehydration in the institutionalized elderly. e-SPEN, the European e-Journal of Clinical Nutrition and Metabolism 2010 Feb;5(1):e47-e53. [doi: 10.1016/j.eclnm.2009.10.007]
6. Campbell N. Dehydration: why is it still a problem? Nurs Times 2011;107(22):12-15. [Medline: 21761784]
7. Riley-Smith B. The Telegraph. 2013 Dec 01. More than a thousand care home residents die thirsty URL: https://www.telegraph.co.uk/news/health/news/10487305/More-than-a-thousand-care-home-residents-die-thirsty.html [accessed 2018-11-22] [WebCite Cache ID 7475gAw3R]
8. Vincent M, Mahendiran T. Improvement of fluid balance monitoring through education and rationalisation. BMJ Qual Improv Rep 2015;4(1) [FREE Full text] [doi: 10.1136/bmjquality.u209885.w4087] [Medline: 26893885]
9. Boylan A, Brown P. Student observations. Fluid balance charts. Nurs Times 1985;81(20):35-38. [Medline: 3846969]
10. Bennett C. "At A Glance" Fluid Balance Bar Chart. London: NHS Institute for Innovation and Improvement; 2010.
11. Reid J, Robb E, Stone D, Bowen P, Baker R, Irving S, et al. Improving the monitoring and assessment of fluid balance. Nurs Times 2004 May 18;100(20):36-39. [Medline: 15176276]
12. UK Legislation. 2000. Care Standards Act 2000 URL: https://www.legislation.gov.uk/ukpga/2000/14 [WebCite Cache ID 6wL7yzMiS]

13. Chernoff R. Meeting the nutritional needs of the elderly in the institutional setting. Nutr Rev 1994 Apr;52(4):132-136. [Medline: 8028818]
14. May CR, Mair F, Finch T, MacFarlane A, Dowrick C, Treweek S, et al. Development of a theory of implementation and integration: Normalization Process Theory. Implement Sci 2009 May 21;4:29 [FREE Full text] [doi: 10.1186/1748-5908-4-29] [Medline: 19460163]
15. Pawson R. The science of evaluation: a realist manifesto. London: SAGE; 2013.
16. Pawson R, Tiley N. Realistic evaluation. London: SAGE; 1997.
17. Anderson R. New MRC guidance on evaluating complex interventions. BMJ 2008 Oct 22;337:a1937. [doi: 10.1136/bmj.a1937] [Medline: 18945728]
18. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M, Medical Research Council Guidance. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ 2008 Sep 29;337:a1655 [FREE Full text] [doi: 10.1136/bmj.a1655] [Medline: 18824488]
19. A Framework for development and evaluation of RCTs for complex interventions to improve health. London: Medical Research Council; 2000.
20. Partlett M, Hamilton D. In: Tawney D, editor. Curriculum Evaluation Today: Trends and Implications: A Second Collection of Papers from Members of the Schools Council Project Evaluators Group on Aspects of Their Work. London: Macmillan; 1976.
21. Reason P, Bradbury H. Handbook of action research: Participative inquiry and practice. London: SAGE; 2001.
22. Skills for Health. The Care Certificate URL: http://www.skillsforhealth.org.uk/standards/item/216-the-care-certificate [accessed 2018-11-22] [WebCite Cache ID 7477XTTTF]
23. UK Government. 2015 Sep 17. Regulated Qualifications Framework: a postcard URL: https://www.gov.uk/government/publications/regulated-qualifications-framework-a-postcard [accessed 2018-11-22] [WebCite Cache ID 74784Hg1r]
24. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
25. Abdallah L, Remington R, Houde S, Zhan L, Melillo KD. Dehydration reduction in community-dwelling older adults: perspectives of community health care providers. Res Gerontol Nurs 2009 Jan;2(1):49-57. [doi: 10.3928/19404921-20090101-01] [Medline: 20077993]
26. Ashurst A. Hydration, hydration, hydration. Nursing and Residential Care 2011 Jul;13(7):311-311. [doi: 10.12968/nrec.2011.13.7.311]
27. Oates L, Price CI. Clinical assessments and care interventions to promote oral hydration amongst older patients: a narrative systematic review. BMC Nurs 2017;16(1):4 [FREE Full text] [doi: 10.1186/s12912-016-0195-x] [Medline: 28104998]
28. Lage Junior M, Godinho Filho M. Variations of the kanban system: Literature review and classification. International Journal of Production Economics 2010 May;125(1):13-21. [doi: 10.1016/j.ijpe.2010.01.009]
29. Care Quality Commission. 2011. Dignity and nutrition inspection programme: National overview URL: https://www.cqc.org.uk/sites/default/files/documents/20111007_dignity_and_nutrition_inspection_report_final_update.pdf [accessed 2018-11-22] [WebCite Cache ID 7479UI0Xp]
30. Binder T, De Michelis G, Ehn P, Jacucci G, Linde P, Wagner I. Design Things. Cambridge: MIT Press; 2011.
31. Seale C. Quality in Qualitative Research. Qualitative Inquiry 2016 Jun 29;5(4):465-478. [doi: 10.1177/107780049900500402]

Abbreviations

CCG: clinical commissioning group
RAG: red, amber, or green

Edited by G Eysenbach; submitted 14.03.18; peer-reviewed by S Bevc, A Wong, M Danilovich; comments to author 11.07.18; revised version received 26.09.18; accepted 07.10.18; published 29.01.19.

Please cite as:
Steven A, Wilson G, Young-Murphy L
The Implementation of an Innovative Hydration Monitoring App in Care Home Settings: A Qualitative Study
JMIR Mhealth Uhealth 2019;7(1):e9892
URL: http://mhealth.jmir.org/2019/1/e9892/
doi:10.2196/mhealth.9892
PMID:30694207

Original Paper

Exploring Community Smokers' Perspectives for Developing a Chat-Based Smoking Cessation Intervention Delivered Through Mobile Instant Messaging: Qualitative Study

Tzu Tsun Luk[1], BNurs; Sze Wing Wong[1], MCouns; Jung Jae Lee[1], PhD; Sophia Siu-Chee Chan[1], PhD; Tai Hing Lam[2], MD; Man Ping Wang[1], PhD

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

[2]School of Public Health, The University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Man Ping Wang, PhD
School of Nursing, The University of Hong Kong
21 Sassoon Road, Pokfulam
Hong Kong,
China (Hong Kong)
Phone: 852 39176636
Email: mpwang@hku.hk

Abstract

Background: Advances in mobile communication technologies provide a promising avenue for the delivery of tobacco dependence treatment. Although mobile instant messaging (IM) apps (eg, WhatsApp, Facebook messenger, and WeChat) are an inexpensive and widely used communication tool, evidence on its use for promoting health behavior, including smoking cessation, is scarce.

Objective: This study aims to explore the perception of using mobile IM as a modality to deliver a proposed chat intervention for smoking cessation in community smokers in Hong Kong, where the proportion of smartphone use is among the highest in the world.

Methods: We conducted 5 focus group, semistructured qualitative interviews on a purposive sample of 15 male and 6 female current cigarette smokers (age 23-68 years) recruited from the community in Hong Kong. All interviews were audiotaped and transcribed. Two investigators independently analyzed the transcripts using thematic analyses.

Results: Participants considered mobile IM as a feasible and acceptable platform for the delivery of a supportive smoking cessation intervention. The ability to provide more personalized and adaptive behavioral support was regarded as the most valued utility of the IM–based intervention. Other perceived utilities included improved perceived psychosocial support and identification of motivator to quit. In addition, participants provided suggestions on the content and design of the intervention, which may improve the acceptability and usability of the IM–based intervention. These include avoiding health warning information, positive messaging, using former smokers as counselors, and adjusting the language style (spoken vs written) according to the recipients' preference.

Conclusions: This qualitative study provides the first evidence that mobile IM may be an alternative mobile health platform for the delivery of a smoking cessation intervention. Furthermore, the findings inform the development of a chat-based, IM smoking cessation program being evaluated in a community trial.

(JMIR Mhealth Uhealth 2019;7(1):e11954) doi:10.2196/11954

KEYWORDS

chat intervention; instant messaging; mHealth; mobile phone; social media; smoking cessation; tobacco dependence; WhatsApp

Introduction

Burgeoning mobile communication technologies have provided a promising means to deliver mobile health (mHealth) intervention, including behavioral change treatment [1,2]. The widespread use of mobile devices has rendered mobile smoking cessation program a scalable measure to combat tobacco use, the leading modifiable cause of morbidity and mortality

worldwide [3]. Commonly studied mHealth interventions for treating tobacco use include short message service (SMS) text messaging support [4], social media (eg, Facebook and Twitter) [5], and smartphone apps [6]. Substantial research has found that SMS text messaging-based interventions are effective in promoting quitting, though the evidence base for other mHealth cessation modalities is still developing [4,7]. Some SMS text messaging programs personalized the intervention contents according to the smokers' characteristics and allowed 2-way communication wherein a recipient can text keywords, such as "crave," to receive on-demand support [8,9]. However, it remains unknown if chat messaging interventions, which feature more personalized and 2-way communication between smokers and treatment providers, may further improve smoking cessation outcomes.

Mobile instant messaging (IM) apps (eg, WhatsApp and WeChat), which allow the exchange of text, emojis, voice messages, and multimedia files freely through the internet, have rapidly replaced SMS as the most widely used mobile communication tool [10]. We found that health information exposure from IM was associated with healthier behaviors, including more frequent physical activity and less smoking, suggesting IM may be an alternative and potentially effective mHealth tool for delivering behavioral interventions [11]. Despite a growing interest in integrating IM into health care, emerging research has largely focused on using IM to promote clinical patient management and interprofessional communication [12-14]. The effectiveness of using IM for promoting smoking cessation and other health behaviors remains largely untested. Recent reviews identified only 2 related studies in the literature [12,14]. The first study is a randomized controlled trial we piloted, which found that an intervention using WhatsApp social group was effective in preventing smoking relapse in recent quitters [15]. Another study evaluated a WhatsApp-based physical activity program, which showed small beneficial effects on physical fitness and cardiovascular risk [16].

Hong Kong is the most developed and westernized city of China with stringent tobacco control measures. The daily cigarette smoking rate of 10.0% in 2017 is among the lowest in the world [17]. Novel interventions that motivate and assist the remaining smokers to quit can be one of the solutions to reduce the smoking rate to 5% for implementing tobacco endgame policies [18]. Leveraging the extensive smartphone penetration in the local population (88.6% in 2017) [17], we proposed to develop a chat-based smoking cessation intervention delivered through mobile IM to be evaluated in a community-based trial. Given the lack of similar prior research, we conducted a formative qualitative study to inform the content and design of the intervention in our target population. This study aims to explore the perception of the proposed mobile IM intervention for smoking cessation in community Chinese smokers.

Methods

The Proposed Intervention

The proposed mobile IM intervention for smoking cessation consists of 2 major components. The first component is a chat-based intervention wherein a smoker can chat individually (one-to-one) with a trained counselor through IM, who will provide real-time behavioral support to help the smoker achieve abstinence. The second component consists of messages on information about smoking cessation regularly sent by the counselor through IM. This study determined the contents of both intervention components.

Study Design and Participants

We conducted this qualitative study in a purposive sample of Chinese community smokers in Hong Kong using semistructured, focus group interviews. The eligibility criteria include the following: (1) being a Hong Kong resident aged ≥18 years; (2) ability to communicate in Cantonese; (3) currently smoking cigarettes (at least weekly); and (4) having used IM apps installed on a smartphone. Subjects who were physically or mentally unable to communicate were excluded. To increase the heterogeneity of the sample, we purposefully selected subjects of different sex, age group, and smoking pattern (daily/nondaily). The Public Opinion Programme (POP) of the University of Hong Kong, one of the leading local survey agencies, was commissioned to recruit subjects by telephone contact to potentially eligible households. Persons who answered the phone were asked if there was a residing household member that met the purposive sampling criteria. If so, the member was then invited to participate in the focus group interview. Subject recruitment ceased after data saturation.

This study was approved by the Institutional Review Board of the University of Hong Kong/Hospital Authority Hong Kong West Cluster (UW 17-206). We obtained informed written consent from all participants before the interview began.

Study Procedures

All interviews were conducted in the evening in a quiet meeting room arranged by the POP at the University of Hong Kong. Before the interviews began, participants completed a one-page questionnaire on their sociodemographic characteristics and pattern of cigarette smoking.

A smoking cessation research nurse and an experienced counselor conducted the focus group interviews. No other personnel apart from participants and researchers were present throughout the interviews. The focus group discussion began with the interviewer asking participants to describe their smoking patterns and previous quit attempt as opening questions. Then, following an interview guide, the interviewer introduced the 2 components of the proposed mobile IM intervention for smoking cessation and asked open-ended interview questions (eg, what is your view on using mobile instant messaging for smoking cessation? What is your suggestion on the intervention content to strengthen your motivation to quit?). Probes (eg, "tell me more…") were used where appropriate to illicit more in-depth responses. All interviews were conducted in Cantonese and audiorecorded with participants' consent. Each participant was reimbursed HK $150 (HK $7.8 ≈ US $1) in cash for their time and travel expenses.

Data Analyses

The audio records were transcribed verbatim. Interim analyses were performed following each focus group to refine the interview questions. Two investigators independently analyzed the transcripts following the principles of thematic analyses, as described by Braun and Clarke [19]. Researchers first familiarized themselves with the data by reading all the transcripts line-by-line to generate initial thoughts on the data. Next, all passages pertaining to the research question were coded. Codes that shared similar meaning were then clustered into subcategories. Broader themes were then searched, reviewed, and defined by collating conceptually similar subcategories. Discrepancies in the coding decision were resolved by reanalysis of the transcript and discussion with an additional investigator, who verified the consistency and coherence in thematic codes [20]. All analyses were performed in the original Cantonese. Selected interview excerpts were translated into English for reporting and back-translated for checking by the researchers who were bilingual (Cantonese and English).

Results

Sample Characteristics

We interviewed 21 smokers in 5 focus groups (about 4 participants each) from May to June 2017. Each interview lasted about 1 hour. The mean (SD, range) age of participants was 48.2 (17.1, 23-68) years and 71.4% (15/21) were male (Table 1). More than half (14/21, 67%) were daily smokers.

Thematic analyses identified 2 major themes from the data. These included (1) the perceived utility of IM for smoking cessation and (2) recommendation about the content and delivery of the proposed intervention.

Perceived Utility of Instant Messaging for Smoking Cessation

Participants found using mobile IM for smoking cessation novel and considered IM as a feasible and acceptable platform for the delivery of cessation advice and support to smokers. Some participants preferred receiving counseling support from IM to the phone as IM is neither restricted by time nor smokers' availability.

> *WhatsApp is better than telephone, you can choose when to reply and not being disturbed by telephone calls* [#15, male, 38 years]

Under the theme of perceived utility, we identified 4 subthemes on strategies or techniques for promoting quitting using the IM intervention as discussed by participants: (1) providing personalized behavioral support; (2) providing psychosocial support; (3) identifying motivators to quit; and (4) serving as an information center.

Providing Personalized Behavioral Support

Participants felt that IM could facilitate 2-way communication by which counselors could learn about the characteristics of smokers and provide more personally meaningful feedback to support their quitting. The personalized intervention could be based on age, sex, and other characteristics of smokers such as values and emotional needs.

> *It is very important to discuss with the smokers in a personalized context, like according to the smokers' family background and working environment…and provide tailored response.* [#7, male, 40s]

> *You can provide more specific feedback based on their characteristic like sex. For example, females may care more about their appearance…and you can tell how smoking would affect their skin or teeth… I would be attracted to such information.* [#19, female, 29 years]

> *All of us have different emotional needs at different age, which may be related to study, work or family…if the messages can target and soothe my distress, then I may be able to skip that cigarette.* [#12, female, 58 years]

The importance of providing personally relevant information, which smokers could relate to, was further highlighted by participants.

> *Some people do not have a child…sending messages related to family and that sort of thing to them would be meaningless.* [#7, male, 40s]

> *I am already in my sixties…I wouldn't care much about my look…so information like how smoking affects the skin would not work on me…* [#20, male, 66 years]

Many participants reported habitual smoking in their daily living such as taking a cigarette while commuting to work or immediately after a meal. They conceived that a counselor can appreciate their daily smoking routine and then provide more timely IM support to help them break their smoking habit.

> *The counselor can understand your pattern of smoking and then schedule the message accordingly.* [#14, female, 53 years]

> *You can text me at 1:30 pm…right after my lunch…ask about my progress and encourage me not to smoke that cigarette…* [#11, female, 66]

For those who preferred to quit progressively, participants considered IM useful for a counselor to guide their quitting through an iterative process of goal setting and follow-up. The counselor can help smokers monitor their cigarette consumption and keep track of their target.

> *You can set a target, like first cutting down my daily intake from 1 pack to 5 cigarettes, and then check my progress in WhatsApp. Then, you can set another target…* [#19, female, 29 years]

Table 1. Sample characteristics (N=21).

Characteristics	Participants, n (%)
Mean age (SD), years	48.2 (17.1)
Male	15 (71)
Marital status	
Unmarried	8 (38)
Married	10 (48)
Others	3 (14)
Education level	
Lower secondary (US grade 7-9)	1 (5)
Upper secondary (US grade 10-12)	9 (43)
Tertiary	11 (52)
Employment status	
Employed	11 (52)
Unemployed	2 (10)
Retired or students or housekeepers	8 (38)
Monthly personal income (in HK $; HK $7.8 ≈ US $1)	
≤$9999	6 (30)
$10,000-$19,999	8 (40)
≥$20,000	6 (30)
Having a child	10 (48)
Daily smoking (vs nondaily smoking)	14 (67)
Number of cigarettes per day	
≤10	14 (67)
11-20	6 (29)
>21	1 (5)
Previous quit attempts	
0	8 (40)
1	5 (25)
>2	7 (35)
Intention to quit	
Planning to quit within 1 month	0 (0)
Planning to quit between 1 and 6 months	5 (24)
Not planning to quit in 6 months	16 (76)
Number of instant messages sent and received per day	
1-20	7 (35)
21-50	6 (30)
51-100	4 (20)
>100	3 (15)

Providing Psychosocial Support

Regular receipt of messages from a counselor via IM was considered by participants as a continuing source of psychosocial support for their quit attempts. They indicated that simple messages, like greeting or reminders that they were quitting, are helpful in creating a sense of being backed by someone who cares about their quitting and encourage them to continue through the quitting process.

> *I am the kind of person who needs support from others to quit smoking…I relapsed in my last quit attempt*

> *because no one acknowledged my efforts…if somebody can keep reminding and motivating me to quit through those messages, I think it would work on me.* [#9, male, 42 years]
>
> *I want to quit but have found it very difficult as all of my colleagues smokes…having someone encouraging and following up on my progress can push me harder to stop smoking* [#19, female, 29 years]

In addition, participants suggested that the intervention can empower their significant others, like their family members, through IM to provide additional, potentially stronger psychosocial support for their quitting.

> *You might as well approach my son through WhatsApp and ask him to relay the messages to me…it would be warmer to hear face-to-face from my son saying "Daddy, will you please smoke less?"…* [#15, male, 38 years]

Identifying Motivators to Quit

Participants emphasized the importance of having a reason to quit or reduce their cigarette consumption based on their personal values, which may be related to their health, image (as smoking is severely denormalized in Hong Kong), and family. They remarked a counselor can help them identify their values (often described as a "weakness point") through IM, which can be exploited to strengthen their commitment to become smoke-free.

> *I think personalized messaging is paramount to finding a reason that can prompt the smoker to stop smoking.* [#4, male, 29 years]
>
> *I believe everyone has a trigger point for quitting, and through chatting in WhatsApp you can find the trigger point and motivate the smokers to quit.* [#19, female, 29 years]
>
> *Through conversation you will be able to identify and "prick" the smoker's weakness point to stop smoking…we all have one.* [#7, male, 40s]

Serving as an Information Center

Some participants noted that IM could act as an accessible platform where they can obtain knowledge about methods of quitting (eg, how to use smoking cessation medications) and inquire about other information related to tobacco use.

> *WhatsApp can provide a channel for smokers to ask about methods to quit…like how to quit progressively, or how to use the nicotine gum and patch.* [#20, male, 66 years]
>
> *There are tobacco products other than cigarettes like electronic cigarettes or IQOS…WhatsApp allows the smokers to obtain more information about them…* [#17, male, 23 years]

Furthermore, they thought they could refer back to the previous conversation with the counselor for information about quitting in times of need such as during episodes of craving.

Design Consideration of the Intervention

Participants provided suggestions on the content of the intervention and methods of delivery to optimize the proposed IM smoking cessation support program. These suggestions fell into 4 broad categories as follows: (1) avoid health warnings; (2) positive messaging; (3) former smokers as counselors; and (4) text language.

Avoid Health Warnings

When asked about health information to be included in the proposed intervention, participants overwhelmingly disapproved messages related to the negative consequence of smoking. They referred to such information as unnecessary, unhelpful, and aversive.

> *I am aware of the risk of smoking posed, like those labelled on the cigarette packages, and it doesn't work on me…I would feel annoyed to be reminded of these warnings [3, male, 25].*
>
> *I believe 9.9 out of 10 smokers know the harms of smoking; all of us can list them…You can provide such information only when we ask for it.* [#8, female, 25]

In addition, participants suggested health warning information may be counterproductive and may trigger their desire to smoke as a gesture of defiance.

> *I would feel disgusted if someone keep telling me about the adverse effect of smoking…and I might be prompted to smoke…like saying "it's none of your business."* [#9, male, 42]

Positive Messaging

Following their criticism on health warning information, participants suggested a counselor should motivate smokers using more positive, nonjudgmental messages, like reframing the harm of smoking as a benefit of quitting.

> *Warning is an outdated way to motivate smokers to quit…those who dread the risk of smoking would have mostly quitted already. I think a gentler approach like reward and encouragement would be better.* [#15, male, 38]
>
> *I think information like how smoking would disrupt your family well-being or lead to early death is rather negative…you can present them in an opposite manner, like telling how quitting could promote happiness or family harmony…instilling them hope.* [#21, male, 53]

In addition, participants welcomed messages about the immediate, positive impact of quitting, such as recovery of health or amount of money saved from not buying cigarettes. They believe these messages could validate and reinforce their commitment to quit.

> *Messages about what would happen to my body, like how my lung function would improve minutes, hours and days after I stop smoking cigarette, would encourage me to keep it up…as I wouldn't want to ruin my progress…* [#4, male, 25]

Former Smokers as Counselors

To some participants, having a former smoker as a counselor to coach their quitting is considered helpful in strengthening the intervention. They suggested the quit advice coming from a person who successfully quit is more convincing. In addition, they believed that former smokers could better empathize with the smokers' history of smoking and the challenge of making a quit attempt.

> *It is like we are in the same boat…The counsellor can relate to my feeling better.* [#7, male, 40s]
>
> *Smokers and nonsmokers are different…it would be easier to communicate with a counsellor who smoked before, as he or she can understand my struggles and thoughts…* [#12, female, 58]

Text Language

Responses varied among participants about the style of text language (spoken language vs written language), which appeared to be governed by their personal preference on the closeness of their relationships with a counselor. Some participants recommended messages in spoken language, describing them as more casual and amiable for a closer relationship with the counselor.

> *To me is spoken language…you can tell from the tone the emotion and attitude of the person you are talking with. Written language feels less lively.* [#7, male, 40s]
>
> *Spoken language makes you feel like chatting with a friend.* [#20, male, 66]

Other participants suggested that the use of written language to be more appropriate for maintaining a more formal relationship with the counselor and considered health information in written language more authentic.

> *written language is better for a more distanced relation to the counselor, which seems more professional…* [#8, female, 25]
>
> *I prefer written language…health advice in formal tone sounds more credible.* [#4, male, 25]

Discussion

Principal Findings

This qualitative study explored the perception of using mobile IM for promoting smoking cessation in community cigarette smokers in Hong Kong. Participants considered IM as a feasible and acceptable modality of delivering smoking cessation support to smokers. They brainstormed how IM could facilitate their quitting, including individual behavioral change techniques (BCTs), which promotes their motivation to quit (eg, identify reasons for wanting to stop smoking), self-control ability (goal setting), use of adjuvant cessation aids (advise on stop-smoking medication), and other supportive BCTs that features interaction (elicit and answer questions about smoking cessation) [21]. Furthermore, they provided suggestions on the content and design of the intervention, which may improve its acceptability and usability.

The ability to provide more personalized behavioral support based on the individual characteristics of smokers was considered by participants as the primary merit of the proposed chat intervention for smoking cessation using IM. A formative study for developing mHealth programs for smoking cessation in the United States also reported similar findings [22]. Crucially, personalization requires personal information, which is often difficult to gather during brief clinical interaction at the baseline owing to time constraint. Through IM, a counselor can continuously learn about the characteristic of smokers, monitor their progress, and tailor the type and intensity of behavioral change intervention that targets different aspects of the quitting process. In addition, IM may facilitate behavioral phenotyping such that the counselor can determine how smokers respond to failure (eg, slip or relapse during quitting) and then apply appropriate behavioral techniques to address their reactions [23].

Practical cessation support apart, the participants valued receiving external emotional support or a feeling of "being cared for" during their quit attempt through IM. Moreover, mechanistic evaluations of effective automated SMS text messaging programs for smoking cessation found smokers' perceived psychosocial support as a key mediator for achieving abstinence, even when they acknowledged that messages were coming from a computer rather than a real person [24,25]. Through true person-to-person interaction and delivery of more personalized and tailored cessation support using IM, our proposed chat intervention may potentially improve their perception of psychosocial support and, thus, cessation outcomes relative to other existing SMS text messaging-based interventions, which deliver more static computer-generated responses to end users [26].

Participants suggested that the IM intervention could help clarify their personal value or "weakness point," which was regarded as crucial motivators to quit smoking; this provides a case for adopting value-based counseling models such as motivational interviewing (MI) and acceptance and commitment therapy (ACT) [27,28]. In the context of smoking cessation, MI aims to help smokers clarify and deal with their indecision to making a quit attempt, whereas ACT aims to maximize their psychological flexibility against the negative experiences associated with smoking cessation. Both MI and ACT attempt to strengthen the smokers' commitment to quit based on their values. Preliminary results from emerging research have suggested that ACT delivered through a smartphone platform is feasible and promising to promote quitting [29,30]. Taken together, the ACT may be a useful behavioral change model that can be incorporated into and enhance the IM intervention.

This formative study conducted in our target end users has informed the content and design of our proposed intervention. The chat intervention now adopts the counseling model of the ACT and includes all BCTs suggested by participants, which will be personalized and tailored to the need indicated by smokers. As participants showed personal preference on the text language style, the counselor would also adjust accordingly to facilitate a therapeutic relation with smokers [31]. Based on the participants' comments on health warning information and positive messages, we also modified regular messages that were

used in our previous smoking cessation trials [32,33]. Specifically, we removed loss-framed messages (eg, harm of smoking) and added gain-framed messages (eg, the short-term benefit of quitting on health) and messages that encourage smokers to quit/maintain abstinence for their values. Currently, the intervention is being evaluated in an ongoing randomized controlled trial for Chinese community smokers in Hong Kong (NCT03182790).

Compared with smartphone apps for smoking cessation, IM may be more scalable and market-ready, as many mobile IM apps are freely available and widely used. However, our proposed IM program, which requires a counselor to deliver the chat messaging intervention to recipients, may incur greater operation cost than other SMS text messaging interventions, which can be delivered through an automatic computer response system. Nevertheless, as artificial intelligence and related techniques, such as natural language processing, continue to mature, chatbots can be developed to simulate the conversation made by a human, effectively reducing the manpower and cost of chat-based intervention for smoking cessation [34].

Although this study focused on using IM for smoking cessation, it may be relevant to similar intervention for treating other health risk behaviors, such as alcohol use disorder and physical inactivity, and other interactive digital health intervention, including Chatbot. Intervention studies that implement our proposed chat-based IM intervention can generate real-world text data (dialogues between smokers and counselors), which are necessary for training the chatbots to recognize the need of smokers.

Limitations

This study has several limitations. First, as it focused on the use of mobile IM for smoking cessation, nonsmartphone owners were excluded. The views expressed during the interview were specific to current cigarette smokers who owned a smartphone, who are typically younger, and have a higher socioeconomic position than the general population [35]. Second, as participants were recruited from the community in Hong Kong, our findings may not be applicable to clinical and non-Chinese populations. Third, the qualitative study was relatively small (N=21), although the study endpoint was determined by data saturation, which became apparent during the fourth focus group and confirmed in the fifth focus group. Finally, owing to the lack of funding, member checking (actively involving participants in verifying the results) was not conducted. Nevertheless, 2 investigators independently coded and analyzed the transcripts to improve the trustworthiness of the findings.

Conclusions

This formative qualitative study in Chinese community smokers provides a case for developing and evaluating smoking cessation interventions delivered through IM, which may be an alternative mHealth modality for treating tobacco dependence. The findings inform the development of a chat-based smoking cessation intervention delivered through mobile IM, which is being evaluated in an ongoing randomized controlled trial for Chinese community smokers in Hong Kong.

Acknowledgments

The authors thank the study participants for their participation in this study and the POP of the University of Hong Kong for arranging the interviews. This study is not supported by any grant or funding.

Authors' Contributions

TTL, SSCC, THL, and MPW conceived the study. JJL, a qualitative researcher, provided guidance on the conduct of the study. TTL and SWW conducted the interviews. TTL, SWW, and MPW analyzed the data. TTL wrote the first draft of the manuscript.

Conflicts of Interest

None declared.

References

1. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
2. Roess A. The Promise, Growth, and Reality of Mobile Health - Another Data-free Zone. N Engl J Med 2017 Nov 23;377(21):2010-2011. [doi: 10.1056/NEJMp1713180] [Medline: 29116869]
3. World Health Organization. Mobile health (mHealth) for tobacco control. URL: http://www.who.int/tobacco/mhealth/en/ [accessed 2018-08-16] [WebCite Cache ID 71i1A1swW]
4. Scott-Sheldon LAJ, Lantini R, Jennings EG, Thind H, Rosen RK, Salmoirago-Blotcher E, et al. Text Messaging-Based Interventions for Smoking Cessation: A Systematic Review and Meta-Analysis. JMIR Mhealth Uhealth 2016;4(2):e49 [FREE Full text] [doi: 10.2196/mhealth.5436] [Medline: 27207211]
5. Naslund JA, Kim SJ, Aschbrenner KA, McCulloch LJ, Brunette MF, Dallery J, et al. Systematic review of social media interventions for smoking cessation. Addict Behav 2017 Oct;73:81-93. [doi: 10.1016/j.addbeh.2017.05.002] [Medline: 28499259]

6. Haskins BL, Lesperance D, Gibbons P, Boudreaux ED. A systematic review of smartphone applications for smoking cessation. Transl Behav Med 2017 Jun;7(2):292-299 [FREE Full text] [doi: 10.1007/s13142-017-0492-2] [Medline: 28527027]
7. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016;4:CD006611. [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]
8. Free C, Knight R, Robertson S, Whittaker R, Edwards P, Zhou W, et al. Smoking cessation support delivered via mobile phone text messaging (txt2stop): a single-blind, randomised trial. Lancet 2011 Jul 2;378(9785):49-55 [FREE Full text] [doi: 10.1016/S0140-6736(11)60701-0] [Medline: 21722952]
9. Abroms LC, Boal AL, Simmens SJ, Mendel JA, Windsor RA. A randomized trial of Text2Quit: a text messaging program for smoking cessation. Am J Prev Med 2014 Sep;47(3):242-250. [doi: 10.1016/j.amepre.2014.04.010] [Medline: 24913220]
10. Church K, de Oliveira R. What's up with whatsapp?: comparing mobile instant messaging behaviors with traditional SMS. New York: ACM Digital Library; 2013 Presented at: 15th International Conference on Human-computer Interaction with Mobile Devices and Services; August 27-30, 2013; Munich p. 352-361 URL: http://doi.org/10.1145/2493190.2493225 [doi: 10.1145/2493190.2493225]
11. Shen C, Wang MP, Wan A, Viswanath K, Chan SSC, Lam TH. Health information exposure from information and communication technologies and its associations with health behaviors: Population-based survey. Prev Med 2018 Aug;113:140-146. [doi: 10.1016/j.ypmed.2018.05.018] [Medline: 29787775]
12. Kamel Boulos M, Giustini D, Wheeler S. Instagram and WhatsApp in Health and Healthcare: An Overview. Future Internet 2016 Jul 26;8(3):37 [FREE Full text] [doi: 10.3390/fi8030037]
13. Mars M, Scott RE. WhatsApp in Clinical Practice: A Literature Review. Stud Health Technol Inform 2016;231:82-90. [Medline: 27782019]
14. Giordano V, Koch H, Godoy-Santos A, Dias BW, Esteves SPR, Labronici P. WhatsApp Messenger as an Adjunctive Tool for Telemedicine: An Overview. Interact J Med Res 2017 Jul 21;6(2):e11 [FREE Full text] [doi: 10.2196/ijmr.6214] [Medline: 28733273]
15. Cheung YTD, Chan CHH, Lai CJ, Chan WFV, Wang MP, Li HCW, et al. Using WhatsApp and Facebook Online Social Groups for Smoking Relapse Prevention for Recent Quitters: A Pilot Pragmatic Cluster Randomized Controlled Trial. J Med Internet Res 2015;17(10):e238 [FREE Full text] [doi: 10.2196/jmir.4829] [Medline: 26494159]
16. Muntaner-Mas A, Vidal-Conti J, Borràs PA, Ortega FB, Palou P. Effects of a Whatsapp-delivered physical activity intervention to enhance health-related physical fitness components and cardiovascular disease risk factors in older adults. J Sports Med Phys Fitness 2017;57(1-2):90-102. [doi: 10.23736/S0022-4707.16.05918-1] [Medline: 26364690]
17. Census and Statistics Department. Thematic household survey report no. 64: Pattern of smoking, and personal computer and internet penetration. Hong Kong Special Administrative Region: Census and Statistics Department; 2018. URL: https://www.statistics.gov.hk/pub/B11302642018XXXXB0100.pdf [accessed 2018-08-16] [WebCite Cache ID 71i1HL3UP]
18. Wang MP, Wang X, Lam TH, Viswanath K, Chan SS. The tobacco endgame in Hong Kong: public support for a total ban on tobacco sales. Tob Control 2015 Mar;24(2):162-167. [doi: 10.1136/tobaccocontrol-2013-051092] [Medline: 24046209]
19. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006 Jan;3(2):77-101 [FREE Full text] [doi: 10.1191/1478088706qp063oa]
20. Yardley L. Demonstrating validity in qualitative psychology. In: Smith JA, editor. Qualitative Psychology: A Practical Guide to Research Methods. London: SAGE Publications; 2008:235-251.
21. Michie S, Hyder N, Walia A, West R. Development of a taxonomy of behaviour change techniques used in individual behavioural support for smoking cessation. Addict Behav 2011 Apr;36(4):315-319. [doi: 10.1016/j.addbeh.2010.11.016] [Medline: 21215528]
22. Hartzler AL, BlueSpruce J, Catz SL, McClure JB. Prioritizing the mHealth Design Space: A Mixed-Methods Analysis of Smokers' Perspectives. JMIR Mhealth Uhealth 2016 Aug 05;4(3):e95 [FREE Full text] [doi: 10.2196/mhealth.5742] [Medline: 27496593]
23. Kangovi S, Asch DA. Behavioral Phenotyping in Health Promotion: Embracing or Avoiding Failure. JAMA 2018 Dec 22;319(20):2075-2076. [doi: 10.1001/jama.2018.2921] [Medline: 29710244]
24. Douglas N, Free C. 'Someone batting in my corner': experiences of smoking-cessation support via text message. Br J Gen Pract 2013 Nov;63(616):e768-e776 [FREE Full text] [doi: 10.3399/bjgp13X674459] [Medline: 24267860]
25. Hoeppner BB, Hoeppner SS, Abroms LC. How do text-messaging smoking cessation interventions confer benefit? A multiple mediation analysis of Text2Quit. Addiction 2017 Apr;112(4):673-682. [doi: 10.1111/add.13685] [Medline: 27943511]
26. Naughton F. Commentary on Hoeppner et al. (2017): Have we overlooked the importance of feeling cared for in digital smoking cessation interventions? Addiction 2017 Apr;112(4):683-684. [doi: 10.1111/add.13744] [Medline: 28261986]
27. Rollnick S, Miller W, Butler C. Motivational Interviewing in Health Care: Helping Patients Change Behavior. New York: Guilford Publications, Inc; 2007.
28. Hayes SC, Wilson KG, Strosahl KD. Acceptance and Commitment Therapy: An Experiential Approach to Behavior Change. New York: Guilford Press; 2004.

29. Bricker JB, Mull KE, Kientz JA, Vilardaga R, Mercer LD, Akioka KJ, et al. Randomized, controlled pilot trial of a smartphone app for smoking cessation using acceptance and commitment therapy. Drug Alcohol Depend 2014 Oct 1;143:87-94. [doi: 10.1016/j.drugalcdep.2014.07.006] [Medline: 25085225]
30. Bricker JB, Copeland W, Mull KE, Zeng EY, Watson NL, Akioka KJ, et al. Single-arm trial of the second version of an acceptance & commitment therapy smartphone application for smoking cessation. Drug Alcohol Depend 2016 Nov 04;170:37-42. [doi: 10.1016/j.drugalcdep.2016.10.029] [Medline: 27870987]
31. Ireland ME, Slatcher RB, Eastwick PW, Scissors LE, Finkel EJ, Pennebaker JW. Language style matching predicts relationship initiation and stability. Psychol Sci 2011 Jan;22(1):39-44. [doi: 10.1177/0956797610392928] [Medline: 21149854]
32. Chan SSC, Wong DCN, Cheung YTD, Leung DYP, Lau L, Lai V, et al. A block randomized controlled trial of a brief smoking cessation counselling and advice through short message service on participants who joined the Quit to Win Contest in Hong Kong. Health Educ Res 2015 Aug;30(4):609-621 [FREE Full text] [doi: 10.1093/her/cyv023] [Medline: 26116584]
33. Weng X, Wang MP, Suen YN, Li WHC, Wu Y, Cheung DYT, et al. Comparing different intensities of active referral to smoking cessation services in promoting smoking cessation among community smokers: a study protocol of a cluster randomized controlled trial. BMC Public Health 2018 Jul 04;18(1):830 [FREE Full text] [doi: 10.1186/s12889-018-5782-1] [Medline: 29973191]
34. Dreyer L, Torey V. Are Chat Bots the Next mHealth Frontier?. 2016 Oct 5. URL: https://smokefree.gov/mhealthmonitor/are-chat-bots-next-mhealth-frontier [accessed 2018-08-16] [WebCite Cache ID 71i3oNF0L]
35. Luk TT, Wang MP, Shen C, Wan A, Chau PH, Oliffe J, et al. Short version of the Smartphone Addiction Scale in Chinese adults: Psychometric properties, sociodemographic, and health behavioral correlates. J Behav Addict 2018 Dec 01;7(4):1157-1165. [doi: 10.1556/2006.7.2018.105] [Medline: 30418073]

Abbreviations

ACT: acceptance and commitment therapy
BCT: behavioral change techniques
IM: instant messaging
mHealth: mobile health
MI: motivational interviewing
POP: Public Opinion Programme
SMS: short message service
US: upper secondary

Edited by G Eysenbach; submitted 16.08.18; peer-reviewed by K Mull, M Marcolino, M Mars; comments to author 27.09.18; revised version received 24.10.18; accepted 04.11.18; published 31.01.19.

Please cite as:
Luk TT, Wong SW, Lee JJ, Chan SSC, Lam TH, Wang MP
Exploring Community Smokers' Perspectives for Developing a Chat-Based Smoking Cessation Intervention Delivered Through Mobile Instant Messaging: Qualitative Study
JMIR Mhealth Uhealth 2019;7(1):e11954
URL: https://mhealth.jmir.org/2019/1/e11954/
doi:10.2196/11954
PMID:30702431

Original Paper

Text Messaging to Enhance Mindfulness-Based Smoking Cessation Treatment: Program Development Through Qualitative Research

Claire A Spears[1], PhD; Sharrill A Bell[1], MPH; Charlayne A Scarlett[1], BS; Natalie K Anderson[2], MA; Cherell Cottrell-Daniels[1], MPH; Sadaf Lotfalian[2], MA; Maitreyi Bandlamudi[1], BS; Amanda Grant[1], BS; Anna Sigurdardottir[1], BS; Brittani P Carter[1], MPH; Lorien C Abroms[3], ScD; David W Wetter[4], PhD

[1]Division of Health Promotion and Behavior, Georgia State University School of Public Health, Atlanta, GA, United States

[2]Department of Psychology, Catholic University of America, Washington, DC, United States

[3]Prevention and Community Health, Milken Institute School of Public Health, George Washington University, Washington, DC, United States

[4]Center for Health Outcomes and Population Equity, University of Utah and Huntsman Cancer Institute, Salt Lake City, UT, United States

Corresponding Author:
Claire A Spears, PhD
Division of Health Promotion and Behavior
Georgia State University School of Public Health
140 Decatur Street SE
Suite 612
Atlanta, GA
United States
Phone: 1 404 413 9335
Email: cspears@gsu.edu

Abstract

Background: Mindfulness-based programs show promise for promoting smoking cessation in diverse populations. Mobile health strategies could increase treatment engagement and in-the-moment support, thus enhancing the effects of mindfulness-based smoking cessation interventions. However, most mobile health programs have been developed without sufficient input from the target population.

Objective: By eliciting input from the target population, predominantly low socioeconomic status (SES) African American adult smokers, throughout the development of an SMS (short message service) text messaging program that teaches mindfulness for smoking cessation, we aimed for the resulting program to be optimally effective and consistent with participants' needs and preferences.

Methods: Two qualitative studies (N=25) were conducted with predominantly low SES, African American adult smokers. In Study 1 (initial qualitative input; n=15), participants engaged in focus groups to provide suggestions for program development. In Study 2 (abbreviated trial; n=10), participants received a 1-week version of the SMS text messaging program and provided feedback through in-depth interviews.

Results: In Study 1, participants suggested that the SMS text messaging program should be personalized and interactive (ie, involve two-way messaging); provide strategies for coping with cravings and recovering from smoking lapses; involve relatively short, to-the-point messages; and include pictures. In Study 2, participants were highly engaged with the texts, indicated that the program was useful, and provided additional suggestions for improvement.

Conclusions: Eliciting feedback from the target population throughout the intervention development process allowed for iterative revisions to increase feasibility, acceptability, and effectiveness. Overall, SMS text messaging appears to be a feasible, appealing way to provide in-the-moment personalized support and encourage mindfulness among low-income African American smokers.

(JMIR Mhealth Uhealth 2019;7(1):e11246) doi:10.2196/11246

KEYWORDS

mobile phone; low socioeconomic status; qualitative; short message service text messaging; smoking cessation

Introduction

Tobacco use is the leading cause of preventable morbidity and mortality in the United States [1]. Although most smokers indicate interest in quitting, the rates of successful smoking cessation are quite low (eg, based on 2015 data, only 7.4% of adult smokers had quit in the past year [2]). Adults with low socioeconomic status (SES) exhibit disproportionately high rates of smoking, often have greater difficulty quitting, and consequently experience profound tobacco-related health disparities [2-6]. Compared to other racial or ethnic groups, African Americans tend to have greater difficulty quitting smoking and higher incidence and mortality rates for diseases associated with smoking [2-4,7,8]. There is an urgent need to develop accessible, evidence-based smoking cessation interventions targeting disparity populations, including low SES and African American smokers.

Mindfulness is defined as purposeful, present-focused attention with an attitude of acceptance and nonjudgment [9-11]. Mindfulness-based interventions show promise for improving smoking cessation outcomes in diverse populations [12-17]. Moreover, mindfulness may be particularly beneficial in low SES and racially or ethnically diverse populations. Practicing a nonjudgmental, compassionate attitude toward oneself could be especially useful for people from marginalized backgrounds [18], and mindfulness practice (eg, mindfully focusing on one's breath) does not require a high level of education or resources. We found that low SES, predominantly African American adults perceived the ability to practice mindfulness on their own (regardless of availability of external resources) as empowering and beneficial to both mental and physical health [19]. Regardless of the target population, between-session mindfulness practice (eg, sitting meditation, gentle yoga, mindful awareness of breath) is thought to be integral in producing benefits [10]. However, participants do not always practice mindfulness in their daily lives [20].

Mobile health (mHealth) technology presents unique opportunities for mindfulness research and intervention, and cell phone use has become ubiquitous in the United States [21]. Between-session short message service (SMS) text messaging might be an effective means to encourage participants to use mindfulness techniques in moments when they need them most, thus enhancing treatment effectiveness. Furthermore, data suggest that adults from racial or ethnic minority backgrounds and those with lower education tend to text more frequently than Caucasians and those with higher education [22]. Thus, SMS text messaging might be an appropriate way to target smoking cessation in low SES and African American smokers.

There is substantial evidence supporting the efficacy of SMS text message-based smoking cessation interventions (for reviews, see [23,24,25]); however, these interventions have not focused on mindfulness. Thus, the current qualitative studies were designed to inform the development of SMS text messages to be used in a mindfulness-based smoking cessation intervention. Participants from low SES backgrounds were targeted with the goal of creating a SMS text messaging program that is engaging, acceptable, and optimally effective in this population. Researchers have highlighted the critical need to elicit feedback from the target population at the outset and throughout the process of developing mHealth interventions [26,27]. Unfortunately, mHealth programs are often developed with insufficient input from the target population [28] or lack of attention to evidence-based strategies [29].

Abroms et al [26] recommended steps for developing and pretesting text messaging-based health promotion programs were followed. First, we defined the population of interest as low SES adult smokers (with a large proportion of African Americans represented given the racial distribution of the cities where the intervention was developed) and conducted focus groups (Study 1) to inform the initial development of the intervention. Study 1 assessed participants' preferences regarding timing and content of SMS text messages. Next, we designed the initial SMS text message library, with a focus on evidence-based mindfulness and cognitive-behavioral strategies for quitting smoking. SMS text messages were intended to remind participants of specific strategies in times of need, as well as to encourage daily mindfulness practice (which could increase facility of mindfulness techniques in high-risk and other situations). We developed a brief, 1-week version of the program to be pretested with a separate group of participants (Study 2) and then used participant feedback to further improve the program. A longer version of the program will then be evaluated in a randomized controlled trial, with continued revisions as needed. Through gathering in-depth qualitative input from the target population throughout intervention development, our overarching aim is for the resulting program to be optimally effective and consistent with participants' needs and preferences.

Methods

Study 1: Initial Qualitative Input

Participants

Participants were recruited from the Washington, DC metropolitan area through flyers at local community health centers and community centers and shelters and through Craigslist. The inclusion criteria were as follows: age 18-65 years (individuals aged >65 years were excluded given lower use of texting in this population [22]); current smoker with history of ≥5 cigarettes per day for the past year (verified in-person with expired carbon monoxide, CO, ≥6ppm); motivated to quit within the next 30 days; valid home address in the greater DC area; functioning telephone number; can speak, read, and write in English; and marginal or adequate health literacy (determined using the Rapid Estimate of Adult Literacy in Medicine [30]). The exclusion criteria were as follows: contraindication for nicotine patch, active substance abuse or dependence, regular use of tobacco products other than cigarettes, current use of tobacco cessation medications, pregnancy or lactation, or a household member enrolled in the study. These inclusion and exclusion criteria were chosen in order to maintain consistency with eligibility criteria for the later randomized controlled trial of the SMS text messaging intervention to be developed.

Procedures

After completing eligibility screening and providing informed consent, participants completed a background questionnaire to indicate demographic information and mobile phone usage. Focus groups were approximately 90 minutes each, and the facilitator asked participants about their level of interest in SMS text messaging to help them quit smoking; suggestions for the content of messages (including reactions to sample SMS text messages); preferences for structure and timing of messages; and other ideas for making the program more helpful and user-friendly.

Group sessions were audio-recorded with a digital voice recorder and transcribed verbatim. Transcripts were managed and coded using NVivo 10 software (QSR International). The data coding and analysis followed both inductive and deductive approaches [31,32]. The first author (who facilitated the focus groups) and 2 separate coders (trained research team members) each reviewed the transcripts to develop an initial set of themes from the interview topics and the conceptual framework, with additional codes identified from concepts that emerged from the data. After developing the initial coding scheme, the 2 coders each independently coded one focus group and met as a group with the first author to review inconsistencies and further refine the coding scheme. The remaining transcripts were independently coded, with regular meetings to discuss the coding process and discrepancies, maintain consistency over time, and refine the coding scheme as needed. Discrepancies were resolved through group discussion, with final decisions made by the first author.

Study 2: Qualitative Feedback After Text Messaging Trial

Participants

Recruitment and inclusion and exclusion criteria were identical to those in Study 1, with 2 exceptions. Study 2 was conducted in the Atlanta, Georgia metropolitan area, and eligible participants needed to have a cell phone with SMS text messaging capability and an unlimited SMS text message plan.

Procedures

After completing informed consent, participants completed a brief questionnaire inquiring about demographic information and mobile phone usage. Because participants were not necessarily expected to have experience with mindfulness, research staff provided a definition of mindfulness, discussed ways to practice mindfulness in daily life, and informed participants that the SMS text messages would involve mindfulness. Participants provided the researchers with their mobile phone numbers in order to receive SMS text messages over the following week. As participants in the later intervention study would receive SMS text messages for 7 weeks, each of the 7 days in this pilot study was designed to mimic each week of the planned intervention study (ie, on day 3 of this pilot, participants received the same number and types of messages that they would receive during week 3 of the intervention; this allowed us to collect information on various types of messages that would be sent at different stages of the quitting process). The research team explained this to participants and asked them to imagine that Day 5 was their quit day (corresponding to week 5 of the intervention). Participants received 2 messages on Day 1, 3 on Day 2, 4 on Day 3, 5 on Day 4, 6 on Day 5, 4 on Day 6, and 3 on Day 7. In addition, participants could text CRAVE, STRESS, or SLIP at any point during the week to receive additional SMS text message support for coping with cravings, stress, or smoking lapses, respectively.

Mindfulness techniques incorporated into the 1-week period were mindful breathing, mindful eating, mindful stretching, and mindful awareness of body sensations (including experiences of craving). These are core practices of Mindfulness-Based Addiction Treatment for smoking cessation [17] and were adapted as brief, in-the-moment practices (eg, “Stand up and take a deep breath. Move your arms slowly in circles and then stretch your whole body from side to side, noticing how you feel”). Participants were asked to keep a log of their responses to and suggestions for improving the SMS text messages and to bring this log to their next appointment to facilitate the recall of their experiences during the interview.

After the 1-week period, participants returned for an in-person individual interview, which was audio-recorded and transcribed verbatim. The interview guide inquired about participants’ experiences receiving the SMS text messages, aspects that were helpful and not helpful, suggestions for improving the SMS text messages, suggestions regarding the number and timing of messages, experiences using (or not using) the keywords, and any concerns about receiving and sending SMS text messages for smoking cessation (including any privacy concerns). The general analytic approach was identical to that in Study 1.

Results

Study 1

Participant Characteristics

A total of 15 adult smokers participated in the focus groups (3 groups of 4 participants each and 1 group of 3 participants). The majority of participants were women (12/15, 80%), African American (13/15, 87%), and reported a total annual income of <$18,000 (9/15, 60%). The mean age was 44.3 (SD 11.6) years. Participants smoked 14.3 cigarettes per day on average (SD 9.8) and most (10/15, 67%) reported smoking within 30 minutes of waking. All participants indicated that they currently owned a mobile phone with an unlimited SMS text message plan. On average, participants indicated that they receive 14.9 (SD 8.7) texts per day and send 13.7 (SD 9.1) texts per day.

Qualitative Results

Major themes were the perceived acceptability or helpfulness of SMS text messages for smoking cessation; concerns about text messaging; suggestions about the content of messages; and suggestions about the format of messages.

Perceived Acceptability or Helpfulness of Text Messages for Smoking Cessation

Overall, participants were open to the idea of SMS text messaging for smoking cessation, and most were enthusiastic. Some participants were skeptical that texts would help them

quit in the absence of other support but believed that SMS text messages in addition to in-person treatment and nicotine patches would be helpful. Participants thought texts would be helpful because "it gives you something else to do instead of smoking a cigarette" and "you could be ready to light a cigarette, and see the text message…and put [the cigarette] down." They noted that texts could help remind them of their goal to quit smoking; for example: "I think it's a positive thing to get text messages so my awareness stay up. A lot of times I say I'll quit smoking, and I forget that I said that." Participants also liked the perceived privacy of text messaging, discerning this form of communication to be discrete; for example: "I was thinking the whole texting is good too, 'cause it's private….It's between you and what's on your phone. No one has to hear your response."

Concerns About Text Messaging for Smoking Cessation

Concerns About a One-Size-Fits-All Approach

The most common concern was that SMS text messaging might work for some people and not others ("different strokes for different folks"), mostly based on personal SMS text messaging habits. One person said, "I don't even read my texts half the time." Another suggested that an SMS text messaging program "is for someone who really uses their text messages."

Concerns About Too Many Texts

Participants noted that too many texts and requests to answer questions via text could be "annoying," arrive at inconvenient times, or be ignored by users (eg, "It's going to start bothering people;" "If I had to do a lot of texting and a lot of questions I'd have to answer, I'd be a little frustrated"). One person noted, "I wouldn't answer a lot of [texts]. I would probably block it out…if I'm at a party and I'm drinking."

Suggestions About the Content of Text Messages

Regarding the content of SMS text messages, common themes included suggestions for personalization; strategies for managing cravings; encouragement for coping with lapses; and visual messages.

Personalization

A common theme was that the program be highly personalized. For example, one participant indicated, "It has to be personalized to their life, day by day." Participants suggested personalization in terms of the content of text messages (eg, reminders for personal reasons for quitting: "Maybe in the beginning, you list why you want to do the program in the first place and it will periodically text you…'Remember, you wrote this down'"). They also suggested personalization of the text schedule (eg, "individually set your hours").

Specific Strategies for Managing Cravings

Participants were interested in receiving texts with specific strategies to cope with cravings (eg, "I want to hear a way to cope with [cravings] and something else that I could do instead of smoking a cigarette to make the craving go away"). Another participant noted, "For me, if I'm craving…I have a really hard time thinking. So, [the text could give] a very solid list of things I can do to make the craving go away, and I don't have to think of it myself." Another said, "I like the suggestions. You know, try to do something, play a game, take your mind off it, call somebody, I like that."

Encouragement for Coping with Lapses

Participants noted that they would appreciate encouragement to forgive themselves and get back on track after smoking lapses. One participant suggested messages like "Nobody is perfect. It's hard to get off cigarettes….Don't beat yourself up! You made a mistake but you've got more lives in this game!" Another indicated that it would be helpful to hear "Brush yourself off, get back on the bike, keep going, try again. Don't beat yourself up. Alright, you messed up, but try it again. See what happens next time."

Visual Text Messages

In addition to traditional SMS text messages, participants suggested including images. One participant noted, "I like the text pictures…that would just stick in my mind. But words, I have to read it for a while. [Pictures] that will just pop out at you would be nice. And it would do more good than reading a message." Another said, "If I'm stressed, I don't know what would calm me down. Something visual though, something I could look at. Something positive…something relaxing." While some suggested positive images (eg, "flowers," "something funny"), others suggested graphic images of health effects of smoking (eg, "hole in someone's throat from smoking").

Suggestions About Format of Messages

Regarding SMS text message format, participants suggested that the messages be as short as possible. They also suggested interactive two-way messaging and frequent messaging.

Interactive Text Messages

Participants indicated that in addition to receiving texts, they wanted to be able to initiate and respond to texts. Example quotes included "If you get a text message, then you have to somehow respond to it, or log something so you feel like you are invested;" and "I think if every time we want to smoke a cigarette, we should grab the phone and text you." Another participant said, "If you're going to try to help people through text messages, let them hit you when they need you. You have to be available when I need you…Like right now I'm thinking about smoking a cigarette…Let me pick up this phone, and see what [the text] could say to me. I'm texting you."

Number and Timing of Text Messages

Participants overwhelmingly suggested that SMS text messages be sent frequently (eg, "every hour or maybe every half hour;" or "at least 4 to 5 times a day"). Participants indicated, "It has to be frequent and it has to be repetitious…if you're serious about quitting," and "If you get enough of them throughout the day, it'll make you think twice before you light up another cigarette." Participants preferred a graded (vs fixed) message schedule, in which they would receive more frequent SMS text messages on and around their quit date.

Initial Design of Text Message Library

The SMS text messaging intervention was designed to encourage mindfulness-based and cognitive-behavioral strategies for smoking cessation. Based on suggestions from focus groups,

SMS text messages were designed to be personalized (eg, participants' first names were used, and they were reminded of their personal reasons for quitting and the amount of money they would save after various time periods of quitting) and interactive. Participants were asked questions (eg, "There are people, places, and things that make you want to smoke. What are your top 3 triggers to smoke?" "Good morning, David! Would you like to try a mindfulness exercise?," and "Have you smoked today?"), with automated responses based on participants' replies. Participants could also initiate texts by sending keywords CRAVE, STRESS, or SLIP for support. Texts after participants' smoking lapses promoted self-compassion and encouragement for getting back on track. Picture messages were incorporated (eg, nature images with captions like "Breathe" and "Today is a new day"). A graded message schedule was used, such that the frequency of texts increased around the quit date and then gradually reduced after the quit date.

Study 2

Participant Characteristics

A total of 10 adult smokers (5 women and 5 men; separate from those in Study 1) participated in Study 2. The majority (8/10, 80%) of smokers were African American individuals; 1 was a Caucasian, 1 reported ≥1 race, and 1 was a Hispanic person. Most (7/10, 70%) participants reported a total annual income of <$18,000. The mean age was 44.9 (SD 9.7) years. Participants smoked 21.3 (SD 6.9) cigarettes per day on average and most (8/10, 80%) reported smoking within 30 minutes of waking. Excluding 1 participant who indicated sending and receiving 800 texts per day, participants indicated that they receive 29.1 (SD 19.7) and send 18.2 (SD 10.2) texts per day on average.

Engagement With Text Messages

All participants answered at least some of the questions that were asked of them via text. All participants texted the system back (at least once) when they were not specifically prompted to answer a question (eg, by responding "thank you" after messages, although they were not specifically asked to do so). The number of these texts ranged from 1 to 24 (mean 11.6, SD 8.5). Of the 10 participants, 7 texted a keyword (CRAVE, STRESS, or SLIP) at least once. Among the participants who used keywords, the number of keyword texts ranged from 1 to 11 (mean 3.6, SD 3.4). Accordingly, 6 participants texted CRAVE, 2 texted STRESS, and 4 texted SLIP.

Overall, participants' SMS text message responses during the 1-week period indicated that the texts were perceived as helpful. For example, when a text suggested taking a walk to relax, 1 man texted, "I wanted a cigarette this morning so instead of smoking I walked to the store it was a good walk it made me forget about the cigarette." When encouraged to seek support from friends and family, 1 man responded, "I have talked about it with my sister she told me don't be hard on myself just brush myself off and start over again so she is being very supportive." Several participants responded to text questions to indicate their smoking triggers (eg, "stress," "people," "being angry," "after I eat"), strategies that might work for them to cope with cravings (eg, "toothpicks," "brushing teeth," "gum"), and ways that they will reward themselves when they quit (eg, "a slice of New York cheesecake," "buy a new pair of winter boots," "go to the movies"). Participants often texted positive responses including "Thank you for that inspiring message" and "That is a real motivational thing to say to myself." Participants also indicated positive experiences with the mindfulness exercises (eg, by indicating that they were "peaceful" and "relaxing;" saying "I really needed that, thanks").

Qualitative Results

Major themes were the perceived acceptability or helpfulness of SMS text messages for smoking cessation, most helpful aspects of texts, and suggestions for improvement.

Perceived Acceptability or Helpfulness of Text Messages for Smoking Cessation

Overall, participants noted positive experiences receiving the SMS text messages, and several indicated that the texts helped them to quit or reduce their smoking over the week-long period (eg, "I like them because they were straight to the point. They helped me quit smoking;" "I cut back a whole lot"). Examples of participants using the texts to avoid or reduce their smoking are "I just grabbed my phone instead of a cigarette" and "It does help when the text message comes, you know? Couple times I got a text and dumped the whole cigarette out the window."

Regarding the most helpful aspects of the SMS text messaging program, participants commonly noted the positive, encouraging tone; mindfulness; social support; strategies for avoiding or coping with triggers; picture messages; the opportunity to text keywords; and personalization. Participants dismissed any concerns about privacy (eg, other people seeing their texts or giving their phone number to the research team).

Most Helpful Aspects of Text Message Program

Positive Tone

Participants appreciated the positive, encouraging tone. One participant indicated, "When you're on these streets, there's so much negative and then when you hear something positive through a phone, it just makes you wanna do good." Another said, "Basically [the text] was like, 'Mistakes happen, you know? But keep pushing and try to get back on track.' That was helpful."

Mindfulness

Participants found value in messages encouraging mindfulness. One participant said, "The mindful text message was pretty helpful. Take a minute and just be in that minute right there. I did that when I was driving one day and just took a deep breath, not smoking a cigarette. It makes you feel a little good about yourself right quick." Another said, "I loved it. I'd only read about mindfulness as far as mindful eating, but I never really considered applying it to my entire life, and it's really simple. You just stop take a deep breath, stretch and it really does help. Something so easy can be so helpful." Another said that when stressed, the most helpful message would be "that mindful one. Take a moment, take a breath… stretch." Overall, participants found it relatively easy to engage in these types of mindfulness practices in the moment (mindful breathing was most often mentioned in this context). Several appreciated learning to focus

on the present (eg, "don't worry about the past or the future, just be in this moment") and described incorporating regular mindfulness practice into their daily lives (eg, "Every day. I do about 3-4 minutes").

Social Support

Although participants knew that the program was automated, they often described a sense of social support from the SMS text messages. Example quotes were "It's like having a friend who texts you when you are feeling stressed or having a feeling like you want to smoke" and "It's like having a little quitting coach that encouraged you along and helps you out when you need help."

Strategies for Avoiding or Coping with Triggers

Participants noted that texts encouraging them to remove triggers ("helped me get rid of my ashtray, cigarette lighters, and all of that") or use specific strategies to cope with cravings were useful. One participant said, "[The text messages] were very helpful… the substitute, you know, gum or candy, or meditating like I was saying. Finding someone to call if I'm craving."

Picture Text Messages

In general, people liked the picture messages (though some participants suggested including graphic images that were not included in the program; see "Tough Love and Graphic Warnings" below). One person said, "[The pictures] motivated me and gave me a reason to stop smoking." Picture messages were described as "peaceful" and "inspirational." One participant noted that he was charged extra by his wireless provider for downloading pictures, but he decided to download them anyway because they were helpful.

Keywords

Participants appreciated the opportunity to initiate interaction by texting CRAVE, STRESS, or SLIP. One participant said, "It was really helpful, and being able to say, 'I want to smoke, somebody help me!'" Another said, "It's very helpful to have somebody you can rely on. You know, if you're craving, you put 'crave,' it's gonna send a message or picture or exercise, or something to do… tell you to do stuff except smoking." Participants noted the positive tone of the responses to keywords: "When I did text SLIP or CRAVE it didn't make me feel bad. It wasn't condemning. It just gave me more encouragement."

Personalization

Participants appreciated the texts reminding them of how much money they would save when they quit smoking (based on their individual smoking habits and price per pack) and personal reasons for quitting. One participant noted, "It gave me a reason to want to stop smoking… about my family, I want to see my grandkids grow up, I want to go on jogs, I don't want cancer."

Suggestions for Improvement

Some participants indicated that there were not enough SMS text messages, and some said that they came at inconvenient times. Suggestions for improvement included sending more frequent SMS text messages; more personalization; incorporating religion or spirituality; reminding participants of the CRAVE, STRESS, and SLIP keywords; and incorporating "tough love."

More Frequent Text Messages

The most common suggestion was to include more messages, especially around the quit day. Example quotes were "I would have liked to receive more messages, because it just made me feel better about not smoking" and "I would love to have more of them, because I still have them and I'm going to go over them." Participants suggested sending texts every 15, 30, or 60 minutes, highlighting the need for constant reminders throughout the day (eg, "it will keep reminding me… because when you smoke, you don't just smoke one now, and one every two to three hours, some people smoke back to back to back.").

More Personalization

Participants often indicated that tailored timing of messages to their personal schedule would be useful (eg, timing around work schedules and high-risk times of day). They also suggested sending personally motivating pictures (eg, "Personal pictures. That is gonna really make it hit home" and "I would have liked to see pictures of my kids.").

Incorporating Religion or Spirituality

Several participants mentioned religion or spirituality and suggested including messages like "Pray," "Let go and let God," and "Have faith in God" or incorporating "scriptures from the Bible."

Reminders of Keywords

Some participants indicated forgetting about the keywords, what the keywords were (CRAVE, STRESS, SLIP) or what they meant. For example, 1 participant indicated that he would be more likely to use them if the program would "repeat [the keywords] over and over again, and explain to me what they mean."

"Tough Love" and "Scare Tactics"

Some participants indicated that the messages should be more direct (eg, "Just don't smoke"), "tougher," or include "tough love, not soft love." One person suggested, "Be hard on me." Some participants suggested including graphic images to elicit fear. One participant explained, "Pictures of people who have holes in their necks… those pictures are scary, but they have to make the public aware…. Scare tactics are a good thing." Another suggested sending a "picture of a burned-up lung."

Discussion

Principal Findings

Two qualitative studies with predominantly low-income African American adult smokers provided in-depth information about how and why SMS text messaging might be useful for smoking cessation, with specific suggestions for improving interventions for this population. Based on suggestions from Study 1 (initial qualitative input), the SMS text messaging program was designed to be personalized and interactive; provide mindfulness-based and cognitive-behavioral strategies for coping with cravings and recovering from lapses; involve relatively short, to-the-point messages; and include pictures. In Study 2

(abbreviated trial), participants who received a 1-week version of the program were highly engaged with the texts, indicated that the program was useful, and provided additional suggestions for improvement. Eliciting feedback from the target population throughout the intervention development process has allowed us to iteratively revise the program to increase feasibility, acceptability, and effectiveness. Overall, SMS text messaging appears to be a feasible, appealing way to provide in-the-moment personalized support and encourage mindfulness among low-income African American smokers.

Comparison With Prior Work

Several themes fit well with findings from other qualitative studies of SMS text messaging for smoking cessation. For example, the personalization of SMS text messages was deemed to be important among young adult smokers [33], pregnant smokers [34], and a clinical sample of adult smokers recruited through an emergency department [35]. Quotations from the current sample echo those from smokers in other studies suggesting the importance of tailoring messages to individual schedules, preferences, and motivations [33,35]. Giving participants more choice in the content and timing of messages may enhance their sense of control [33] and reduce the likelihood that people become frustrated or inconvenienced by the texts. A related issue is the potential for SMS text messages to inadvertently stimulate craving for cigarettes [34,36,37]. Although this did not come up in this study, allowing participants to control the number and timing of SMS text messages could help prevent such unintended effects. Messages could also be tailored to personal preferences regarding message tone (eg, positive encouragement vs "tough love"). As in this study, some but not all participants in Bock et al's [33] study advocated for the use of "scare tactics."

Notably, although participants were aware that the SMS text messaging system was automated, many noted the social support and accountability that the texts offered, much like a "friend" or "coach." This is a common theme across SMS text messaging programs, which have been described as "somebody holding your hand" [37] or as though "they were really on my side" [35]. Participants often attribute human characteristics to digital programs, noting a sense of accountability and feeling "shame," "guilt," or "a need to obtain 'forgiveness'" after lapsing [36]. Greater personalization of texts may foster the sense of a "relationship," and research is needed to determine optimal ways to promote adaptive "relationships" with mHealth programs.

Although there are a number of commonalties between this study and extant research on SMS text messaging for smoking cessation, several unique aspects are noted. First, our program places central focus on encouraging mindfulness during the process of quitting smoking. Overall, the concept of mindfulness and related SMS text messages were well received in this sample of low-income, predominantly African American adults. Given that the majority of research on mindfulness has been conducted with higher income, primarily Caucasian samples, this study adds to the growing literature on perceived acceptability and utility of mindfulness interventions in lower income, African American samples [19,38,39]. Participants in mindfulness treatment studies often do not practice mindfulness on their own as much as instructed [20]. For example, rates of between-session mindfulness practice were relatively low in a recent trial of Mindfulness-Based Addiction Treatment for adult smokers [17]. SMS text messaging could be a novel way to promote daily mindfulness practice during the process of quitting, which might enhance smoking cessation outcomes.

Second, participants noted that religious and spiritual factors motivate them to quit smoking, and we added SMS text messages to address participants' suggestion to include spiritualty. Incorporating spirituality and religious coping, as well as integrating with faith-based organizations, have been identified as culturally important factors for tailoring health promotion interventions in African American communities [40,41]. In fact, religious coping and religious support may help to buffer the effects of stress on tobacco use among African Americans [42]. However, appeals to organized religion may be off-putting for people without religious affiliation, and a more general focus on spirituality in the context of health promotion may be more broadly accepted by participants [43]. Thus, we created messages incorporating spirituality without reference to any specific religion (eg, "Look to your spiritual beliefs for comfort and strength. You might pray or read/listen to something that is inspirational to you"). Such messages might also fit well in the context of mindfulness treatment. For example, in a qualitative study of mindfulness among low-income, primarily African American adults, participants noted practicing mindfulness through religious or spiritual experiences (eg, while praying, listening to religious music, or reading sacred texts [19]). Although mindfulness-based smoking cessation treatment is typically taught in a secular context, it may be worthwhile to encourage individuals to incorporate spirituality in any way that is helpful.

Third, the inclusion of pictures is a unique feature that was added explicitly based on participants' suggestions. Texts containing pictures may be more vivid and memorable than those containing only words. Moreover, pictures that do not require extensive reading may be especially appealing and effective for adults with lower levels of education and reading ability. More research is needed to continue to develop and evaluate pictures that are especially motivating for smoking cessation. As participants in this study suggested, SMS text messaging programs might even collect personally motivating pictures from participants and send them back at high-risk times.

Limitations

The current studies are limited by small samples of predominantly low-income African American smokers living in urban areas, and larger-scale trials are needed to evaluate the efficacy of SMS text messaging to enhance mindfulness-based smoking cessation treatment. Small sample sizes are typical of qualitative research, and the current research focuses on an underserved population at risk for tobacco-related health disparities. However, given our focus on this specific population and additional exclusion criteria (eg, excluding those with active co-occurring substance abuse or dependence), results may not generalize to other populations. Furthermore, larger studies that experimentally test the efficacy of various approaches (eg,

whether greater personalization leads to better engagement and quit rates) are needed to evaluate whether certain treatment components work better than others.

Conclusions

Based on iterative user feedback, SMS text messaging appears to be an acceptable intervention for smoking cessation among predominantly low-income African American adults. Suggestions for SMS text messaging smoking cessation programs in this population include sending texts that are short, to the point, and interactive (with reminders of available keywords); providing in-the-moment strategies for coping with cravings; including visual messages; incorporating spirituality; and using personalization (eg, reminders of personal motivation for quitting and tailoring timing and frequency of messages). Although some of these strategies have been discussed in extant studies (eg, importance of personalization), others (particularly the use of visual messages and spirituality) will uniquely inform our future work on SMS text messaging with low-income African American smokers. Participants disagreed on whether the program should include "scare tactics," and based on mixed evidence and potential negative consequences such as increased anxiety leading to stronger urges to smoke [44], as well as lower compatibility with mindfulness-based approaches, we have not included scare tactics in our current program. However, many of our participants' suggestions could be relatively easily integrated into existing SMS text messaging programs and inform SMS text messaging as an adjunct to evidence-based in-person services (eg, cognitive behavioral treatment [45]).

SMS text messaging may have great potential for encouraging mindfulness and other strategies for smoking cessation, in addition to fostering a sense of social support. This mode of intervention, which is relatively low cost and available to participants 24/7, could be particularly useful for underserved populations with lower access to evidence-based smoking cessation services. Continued research is needed to further improve, evaluate, and (if efficacious) more widely disseminate mindfulness-based SMS text messaging interventions for smoking cessation.

Acknowledgments

This work was supported by the National Center for Complementary and Integrative Health (K23AT008442). DWW is supported by the Huntsman Cancer Foundation, the National Cancer Institute (P30CA042014), and the National Center for Advancing Translational Sciences (UL1TR001067). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Conflicts of Interest

LCA has stock in Welltok, Inc, and receives royalties from the licensing of Text2Quit to Welltok, Inc. The other authors have no conflicts of interest to declare.

References

1. U.S. Department of Health and Human Services. The Health Consequences of Smoking-50 Years of Progress: A Report of the Surgeon General. Atlanta (GA): U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014.
2. Babb S, Malarcher A, Schauer G, Asman K, Jamal A. Quitting Smoking Among Adults - United States, 2000-2015. MMWR Morb Mortal Wkly Rep 2017 Jan 06;65(52):1457-1464 [FREE Full text] [doi: 10.15585/mmwr.mm6552a1] [Medline: 28056007]
3. Centers for Disease Control and Prevention. Quitting smoking among adults--United States, 2001-2010. MMWR Morb Mortal Wkly Rep 2011 Nov 11;60(44):1513-1519 [FREE Full text] [Medline: 22071589]
4. U.S. National Cancer Institute. A sociological approach to addressing tobacco-related health disparities. National Cancer Institute Tobacco Control Monograph 22. NIH Publication No. 17-CA-8035A. U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute 2017.
5. Drope J, Liber AC, Cahn Z, Stoklosa M, Kennedy R, Douglas CE, et al. Who's still smoking? Disparities in adult cigarette smoking prevalence in the United States. CA Cancer J Clin 2018 Mar;68(2):106-115 [FREE Full text] [doi: 10.3322/caac.21444] [Medline: 29384589]
6. Jamal A, Phillips E, Gentzke AS, Homa DM, Babb SD, King BA, et al. Current Cigarette Smoking Among Adults - United States, 2016. MMWR Morb Mortal Wkly Rep 2018;67(2):53-59 [FREE Full text] [doi: 10.15585/mmwr.mm6702a1] [Medline: 29346338]
7. Fagan P, Moolchan ET, Lawrence D, Fernander A, Ponder PK. Identifying health disparities across the tobacco continuum. Addiction 2007 Oct;102 Suppl 2:5-29. [doi: 10.1111/j.1360-0443.2007.01952.x] [Medline: 17850611]
8. Irvin Vidrine J, Reitzel LR, Wetter DW. The role of tobacco in cancer health disparities. Curr Oncol Rep 2009 Nov;11(6):475-481 [FREE Full text] [Medline: 19840525]
9. Baer RA, Smith GT, Hopkins J, Krietemeyer J, Toney L. Using self-report assessment methods to explore facets of mindfulness. Assessment 2006 Mar;13(1):27-45. [doi: 10.1177/1073191105283504] [Medline: 16443717]
10. Kabat-Zinn J. Full catastrophe living: Using the wisdom of your body and mind to face stress, pain, and illness. New York: Delacourt; 1990.

11. Kabat-Zinn J. Wherever you go, there you are: Mindfulness and meditation in everyday life. New York: Hyperion; 1994.
12. Brewer JA, Mallik A, Babuscio TA, Nich C, Johnson HE, Deleone CM, et al. Mindfulness training for smoking cessation: results from a randomized controlled trial. Drug Alcohol Depend 2011;119(1-2):72-80 [FREE Full text] [doi: 10.1016/j.drugalcdep.2011.05.027] [Medline: 21723049]
13. Davis J, Fleming M, Bonus K, Baker T. A pilot study on mindfulness based stress reduction for smokers. BMC Complement Altern Med 2007 Jan 25;7:2 [FREE Full text] [doi: 10.1186/1472-6882-7-2] [Medline: 17254362]
14. Davis J, Goldberg S, Anderson M, Manley A, Smith S, Baker T. Randomized trial on mindfulness training for smokers targeted to a disadvantaged population. Subst Use Misuse 2014 Apr;49(5):571-585 [FREE Full text] [doi: 10.3109/10826084.2013.770025] [Medline: 24611852]
15. Davis J, Manley A, Goldberg S, Smith S, Jorenby D. Randomized trial comparing mindfulness training for smokers to a matched control. J Subst Abuse Treat 2014 Sep;47(3):213-221 [FREE Full text] [doi: 10.1016/j.jsat.2014.04.005] [Medline: 24957302]
16. Davis J, Mills D, Stankevitz K, Manley A, Majeskie M, Smith S. Pilot randomized trial on mindfulness training for smokers in young adult binge drinkers. BMC Complement Altern Med 2013 Sep 03;13:215 [FREE Full text] [doi: 10.1186/1472-6882-13-215] [Medline: 24006963]
17. Vidrine JI, Spears CA, Heppner WL, Reitzel LR, Marcus MT, Cinciripini PM, et al. Efficacy of mindfulness-based addiction treatment (MBAT) for smoking cessation and lapse recovery: A randomized clinical trial. J Consult Clin Psychol 2016;84(9):824-838 [FREE Full text] [doi: 10.1037/ccp0000117] [Medline: 27213492]
18. Witkiewitz K, Greenfield B, Bowen S. Mindfulness-based relapse prevention with racial and ethnic minority women. Addict Behav 2013 Dec;38(12):2821-2824 [FREE Full text] [doi: 10.1016/j.addbeh.2013.08.018] [Medline: 24018224]
19. Spears CA, Houchins S, Bamatter W, Barrueco S, Hoover DS, Perskaudas R. Perceptions of Mindfulness in a Low-income, Primarily African American Treatment-Seeking Sample. Mindfulness (N Y) 2017 Dec;8(6):1532-1543. [Medline: 29333200]
20. Vettese L, Toneatto T, Stea J, Nguyen L, Wang J. Do mindfulness meditation participants do their homework? And does it make a difference? A review of the empirical evidence. Journal of Cognitive Psychotherapy 2009;23(3):198-225.
21. Pew Research Center. 2018. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [WebCite Cache ID 6zycrGIuw]
22. Smith A. Pew Research Center. 2011. Americans and text messaging URL: http://www.pewinternet.org/2011/09/19/americans-and-text-messaging/ [accessed 2018-06-07] [WebCite Cache ID 6zycTUE04]
23. Whittaker R, McRobbie H, Bullen C, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2016 Apr 10;4:CD006611. [doi: 10.1002/14651858.CD006611.pub4] [Medline: 27060875]
24. Whittaker R, McRobbie H, Bullen C, Borland R, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2012;11:CD006611. [doi: 10.1002/14651858.CD006611.pub3] [Medline: 23152238]
25. Armanasco AA, Miller YD, Fjeldsoe BS, Marshall AL. Preventive Health Behavior Change Text Message Interventions: A Meta-analysis. Am J Prev Med 2017 Mar;52(3):391-402. [doi: 10.1016/j.amepre.2016.10.042] [Medline: 28073656]
26. Abroms LC, Whittaker R, Free C, Mendel Van Alstyne J, Schindler-Ruwisch JM. Developing and Pretesting a Text Messaging Program for Health Behavior Change: Recommended Steps. JMIR Mhealth Uhealth 2015 Dec 21;3(4):e107 [FREE Full text] [doi: 10.2196/mhealth.4917] [Medline: 26690917]
27. Bock BC, Rosen RK, Barnett NP, Thind H, Walaska K, Foster R, et al. Translating Behavioral Interventions Onto mHealth Platforms: Developing Text Message Interventions for Smoking and Alcohol. JMIR Mhealth Uhealth 2015 Feb 24;3(1):e22 [FREE Full text] [doi: 10.2196/mhealth.3779] [Medline: 25714907]
28. McCurdie T, Taneva S, Casselman M, Yeung M, McDaniel C, Ho W, et al. mHealth consumer apps: the case for user-centered design. Biomed Instrum Technol 2012;Suppl:49-56. [doi: 10.2345/0899-8205-46.s2.49] [Medline: 23039777]
29. Abroms LC, Lee Westmaas J, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. Am J Prev Med 2013 Dec;45(6):732-736 [FREE Full text] [doi: 10.1016/j.amepre.2013.07.008] [Medline: 24237915]
30. Davis TC, Crouch MA, Long SW, Jackson RH, Bates P, George RB, et al. Rapid assessment of literacy levels of adult primary care patients. Fam Med 1991 Aug;23(6):433-435. [Medline: 1936717]
31. Ryan GW, Bernard HR. Techniques to identify themes. Field Methods 2003;15:85-109.
32. Ayres L, Kavanaugh K, Knafl KA. Within-case and across-case approaches to qualitative data analysis. Qual Health Res 2003 Jul;13(6):871-883. [doi: 10.1177/1049732303013006008] [Medline: 12891720]
33. Bock BC, Heron KE, Jennings EG, Magee JC, Morrow KM. User preferences for a text message-based smoking cessation intervention. Health Educ Behav 2013 Apr;40(2):152-159. [doi: 10.1177/1090198112463020] [Medline: 23086555]
34. Naughton F, Jamison J, Sutton S. Attitudes towards SMS text message smoking cessation support: a qualitative study of pregnant smokers. Health Educ Res 2013 Oct;28(5):911-922 [FREE Full text] [doi: 10.1093/her/cyt057] [Medline: 23640985]
35. Grau LE, Pham T, O'Leary T, Weiss J, Toll B, Bernstein SL. Smokers' Perspectives on Texting for Tobacco Dependence Treatment: A Qualitative Analysis. Nicotine Tob Res 2017 Mar 01;19(3):307-313. [doi: 10.1093/ntr/ntw184] [Medline: 27613935]

36. Brandt CL, Dalum P, Thomsen TT. “I miss the care even though I know it's just a machine”: an explorative study of the relationship between an Internet-based smoking cessation intervention and its participants. Health Informatics J 2013 Sep;19(3):233-243. [doi: 10.1177/1460458212470572] [Medline: 23981397]
37. Douglas N, Free C. 'Someone batting in my corner': experiences of smoking-cessation support via text message. Br J Gen Pract 2013 Nov;63(616):e768-e776 [FREE Full text] [doi: 10.3399/bjgp13X674459] [Medline: 24267860]
38. Burnett-Zeigler IE, Satyshur MD, Hong S, Yang A, T Moskowitz J, Wisner KL. Mindfulness based stress reduction adapted for depressed disadvantaged women in an urban Federally Qualified Health Center. Complement Ther Clin Pract 2016 Nov;25:59-67. [doi: 10.1016/j.ctcp.2016.08.007] [Medline: 27863611]
39. Dutton MA, Bermudez D, Matas A, Majid H, Myers NL. Mindfulness-Based Stress Reduction for Low-Income, Predominantly African American Women With PTSD and a History of Intimate Partner Violence. Cogn Behav Pract 2013 Feb 01;20(1):23-32 [FREE Full text] [doi: 10.1016/j.cbpra.2011.08.003] [Medline: 24043922]
40. Kreuter MW, Lukwago SN, Bucholtz RDDC, Clark EM, Sanders-Thompson V. Achieving cultural appropriateness in health promotion programs: targeted and tailored approaches. Health Educ Behav 2003 Apr;30(2):133-146. [doi: 10.1177/1090198102251021] [Medline: 12693519]
41. Campbell MK, Hudson MA, Resnicow K, Blakeney N, Paxton A, Baskin M. Church-based health promotion interventions: evidence and lessons learned. Annu Rev Public Health 2007;28:213-234. [doi: 10.1146/annurev.publhealth.28.021406.144016] [Medline: 17155879]
42. Belgrave FZ, Johnson J, Nguyen A, Hood K, Tademy R, Clark T, et al. Stress and tobacco use among African-American adolescents: the buffering effect of cultural factors. J Drug Educ 2010;40(2):173-188. [doi: 10.2190/DE.40.2.e] [Medline: 21133330]
43. Joseph RP, Ainsworth BE, Mathis L, Hooker SP, Keller C. Incorporating religion and spirituality into the design of community-based physical activity programs for African American women: a qualitative inquiry. BMC Res Notes 2017 Oct 23;10(1):506 [FREE Full text] [doi: 10.1186/s13104-017-2830-3] [Medline: 29058603]
44. Hastings G, Stead M, Webb J. Fear appeals in social marketingtrategic and ethical reasons for concern. Psychol Market 2004;21(11):961-986.
45. Fiore MC, Jaen CR, Baker TB, Bailey WC, Benowitz NJ, Curry SJ. Treating Tobacco Use and Dependence: Update. Rockville, MD: U.S. Department of Health and Human Services (USDHHS), Public Health Service (PHS); 2008.

Abbreviations

mHealth: mobile health
SES: socioeconomic status
SMS: short message service

Edited by G Eysenbach; submitted 06.06.18; peer-reviewed by B Bock, J Magee; comments to author 02.08.18; revised version received 24.09.18; accepted 10.10.18; published 07.01.19.

Please cite as:
Spears CA, Bell SA, Scarlett CA, Anderson NK, Cottrell-Daniels C, Lotfalian S, Bandlamudi M, Grant A, Sigurdardottir A, Carter BP, Abroms LC, Wetter DW
Text Messaging to Enhance Mindfulness-Based Smoking Cessation Treatment: Program Development Through Qualitative Research
JMIR Mhealth Uhealth 2019;7(1):e11246
URL: https://mhealth.jmir.org/2019/1/e11246/
doi:10.2196/11246
PMID:30617043

Original Paper

Predicting Attrition in a Text-Based Nutrition Education Program: Survival Analysis of Text2BHealthy

Stephanie K Grutzmacher[1], MS, PhD; Ashley L Munger[2], MS, PhD; Katherine E Speirs[3], MS, PhD; Yassaman Vafai[4], MPH, PhD; Evan Hilberg[1], MS, MPH; Erin Braunscheidel Duru[5], MHS, RD, LDN; Laryessa Worthington[5], MS, RD, LDN; Lisa Lachenmayr[5], MS

[1]School of Biological and Population Health Sciences, College of Public Health and Human Sciences, Oregon State University, Corvallis, OR, United States

[2]Department of Child and Family Studies, California State University, Los Angeles, Los Angeles, CA, United States

[3]Department of Family Studies and Human Development, Norton School of Family and Consumer Sciences, University of Arizona, Tucson, AZ, United States

[4]Department of Family Science, School of Public Health, University of Maryland, College Park, MD, United States

[5]Maryland Food Supplement Nutrition Education, University of Maryland, College Park, Columbia, MD, United States

Corresponding Author:
Stephanie K Grutzmacher, MS, PhD
School of Biological and Population Health Sciences
College of Public Health and Human Sciences
Oregon State University
118 Milam Hall
Corvallis, OR, 97331
United States
Phone: 1 3019080779
Email: stephanie.grutzmacher@oregonstate.edu

Abstract

Background: Text-based programs have been shown to effectively address a wide variety of health issues. Although little research examines short message service (SMS) text messaging program characteristics that predict participant retention and attrition, features of SMS text message programs, such as program duration and intensity, message content, and the participants' context, may have an impact. The impact of stop messages—messages with instructions for how to drop out of an SMS text message program—may be particularly important to investigate.

Objective: The aim of this study was to describe attrition from Text2BHealthy, a text-based nutrition and physical activity promotion program for parents of low-income elementary school children, and to determine the impact of message content and number of stop messages received on attrition.

Methods: Using data from 972 parents enrolled in Text2BHealthy, we created Kaplan-Meier curves to estimate differences in program duration for different SMS text message types, including nutrition, physical activity, stop, and other messages. Covariates, including rurality and number of stop messages received, were included.

Results: Retention rates by school ranged from 74% (60/81) to 95.0% (132/139), with an average retention rate of 85.7% (833/972) across all schools. Program duration ranged from 7 to 282 days, with a median program duration of 233 days and an average program duration of 211.7 days. Among those who dropped out, program duration ranged from 7 to 247 days, with a median program duration of 102.5 days. Receiving a stop message increased the probability of attrition compared with receiving messages about nutrition, physical activity, or other topics (hazard ratio=51.5, 95% CI 32.46-81.7; $P<.001$). Furthermore, each additional stop message received increased the probability of attrition (hazard ratio=10.36, 95% CI 6.14-17.46; $P<.001$). The degree of rurality also had a significant effect on the probability of attrition, with metropolitan county participants more likely to drop out of the program than rural county participants. The interaction between SMS text message type and total number of stop messages received had a significant effect on attrition, with the effect of the number of stop messages received dependent on the SMS text message type.

Conclusions: This study demonstrates the potential of SMS text message programs to retain participants over time. Furthermore, this study suggests that the probability of attrition increases substantially when participants receive messages with instructions

for dropping out of the program. Program planners should carefully consider the impact of stop messages and other program content and characteristics on program retention. Additional research is needed to identify participant, programmatic, and contextual predictors of program duration and to explicate the relationship between program duration and program efficacy.

(JMIR Mhealth Uhealth 2019;7(1):e9967) doi:10.2196/mhealth.9967

KEYWORDS

text messaging; retention; diet, food, and nutrition; food assistance; parents; survival analysis

Introduction

Background

Using text messages, also known as short message service (SMS), to deliver or supplement health interventions has increased in popularity in recent years. SMS text messages are an appealing mode of program delivery largely because they can be easily and inexpensively used to reach a broad audience [1]. In the United States, approximately 95% of adults own a cell phone [2] and 81% of US adults with cell phones use SMS text messages [3]. SMS text messages may also be a useful tool for accessing underserved populations and addressing health disparities. Low-income and minority populations use cell phones at rates that equal or exceed those of their higher-income and white counterparts [2].

Text-based programs have been successfully used to address a wide variety of health issues such as weight loss [4], smoking cessation [5], diabetes management [6], and sexual health [7]. Exposure to health-related SMS text messages has also been effective in promoting participants' adherence to a program and maintaining healthy behavior changes [8]. A meta-analysis investigating the efficacy of SMS text message–based health promotion interventions found that the overall weighted mean effect size on health outcomes among the 19 randomized controlled trials included in analysis was d=0.329 (95% CI 0.274-0.385; P<.001), indicating a small effect across studies [9].

Although many studies have shown that retention is generally high in SMS text message programs [8,10-12], retention rates also vary quite widely. The aforementioned meta-analysis found that mean retention at follow-up among the included studies was 86% [9]. In another meta-analysis, which investigated the efficacy of weight management programs that incorporated SMS text messages, the retention rate at the postintervention stage among the 14 included studies ranged from 46% to 96% [13].

Understanding when and why participants tend to drop out of SMS text message programs is imperative for effective program planning. There is some evidence that when participants drop out, they do so early in the program; Coa and Patrick found that among those who dropped out of a diet and physical activity SMS text message program, 65% (54/83; 28% of all users) did so within the first 2 weeks [14]. As SMS text message programs vary widely in content, frequency, relevance, tone, theoretical underpinnings, and other characteristics, we do not know whether timing of dropout is related to these characteristics or whether participants are simply tired of SMS text message programming over time. Knowing typical timing of dropout can help researchers prioritize content delivery and develop evaluation plans that maximize time with the majority of participants. Though there are several factors that could potentially impact attrition, such as message content and frequency, these factors remain unexamined, making it difficult to provide guidance concerning program development and features that hinder retention.

Objectives

The purpose of this study was to describe attrition from Text2BHealthy, an SMS text message-based healthy eating and physical activity promotion program. Due to the lack of existing research about attrition from SMS text message programs, certain features of Text2BHealthy thought to be most likely to influence attrition were selected for examination. For example, in the program, several types of message content were delivered, including messages about nutrition, physical activity, and *stop messages*. Stop messages provide participants with instructions for withdrawing from the program. Due to a concern of the program funding agency about participants without SMS text messaging in their data plans incurring costs to receive program messages, stop messages were a required feature of the program. It is unknown whether and how much such messages increase attrition. In addition, the Text2BHealthy program was tailored to particular elementary schools located throughout the state of Maryland. These schools varied widely in terms of the degree of rurality of the area. Participants in rural areas might be less likely to drop out of a text-based health promotion program as rural areas tend to have fewer health services and programs available [15].

In this study, we sought to discern (1) how long participants remained in the program before dropping out; (2) whether particular types of messages, particularly stop messages, increased the likelihood that participants will drop out of the program; (3) whether the number of stop messages received increased the likelihood that participants will drop out of the program; (4) whether school rurality was associated with attrition; and (5) whether the impact of the number of stop messages received differs by message type.

Methods

Text2BHealthy Program

Text2BHealthy is a Maryland Supplemental Nutrition Assistance Program-Education (SNAP-Ed) nutrition and physical activity promotion program delivered by SMS text messages to parents of elementary school children [16]. Parents received message content tailored to their children's school and local community. Program participants received 2 to 3 messages per week. Messages provided information and actionable nudges about

nutrition and physical activity as well as a variety of other related content. Although Text2BHealthy did not typically solicit responses to most messages, participants were occasionally asked to respond to simple evaluation questions or set goals via text response. Participants were able to remove themselves from the program at any time by texting *stop* to the program short code or responding to any message with a message that included *stop*. As required by the program funding agency, SMS text messages informing participants about how to leave the program were sent approximately every 6 weeks.

Participants and Recruitment

Participants included in this study were 972 parents of children attending a selection of low-income elementary schools with youth SNAP-Ed programs in Maryland and participating in the Text2BHealthy program. Parents were recruited through school events and program promotional items sent home at the beginning of the school year. They were enrolled by either providing their phone number to program staff or enrolling themselves through a keyword texted to a short code. Data for this study came from the September 2012 to June 2013 program year, which includes enrolled parents from 10 elementary schools in 5 Maryland counties and Baltimore City. Parents could enroll in and drop out of Text2BHealthy at any point during the program year.

Variables and Measures

The outcome of interest for this study was attrition from the Text2BHealthy program (ie, survival time in days). Attrition data were recorded by the Web platform that was used to send messages to participants. When participants sent an SMS text message to the program phone number indicating that they wanted to be removed from the program, the Web platform would automatically and instantly remove them from the list of participants to receive SMS text messages and record the date and time that they left the program. Program enrollment date and dropout date were used to calculate program duration (ie, survival time) in days.

Survival analyses were conducted to examine differences in participant attrition from the program. The primary predictor for survival time, in days, was message type. Message type was created by coding the last SMS text message received (either the last message before dropout or the last message sent during the program year) into 5 categories: nutrition; physical activity; stop messages describing how to drop out of the program; stop messages combined with content about nutrition or physical activity; and a variety of other content that included evaluation questions, goal-setting, and healthy event notifications and other messages not explicitly addressing nutrition and physical activity actions (see Table 1).

Several covariates were included in the models, including rurality and the number of stop messages received during the program year. Rurality was determined for each school using the United States Department of Agriculture's 2013 Rural-Urban Continuum Codes (RUCC; the year the data were collected) for the county where the school was located. These codes range from 1 to 9, with 1 to 3 indicating metropolitan areas and 4 to 9 indicating nonmetropolitan areas [17].

Statistical Analysis

Frequencies were used to determine retention rates. Mean and median program durations were calculated to determine how long participants remained in the program before dropping out. Kaplan-Meier curves were created to estimate differences in participant attrition for different message types. Additional models that included total number of stop messages received, RUCC, and interactions between message type and total number of stop messages received were run. Using the survival package, Cox proportional hazards models were fit to estimate hazard ratios. A graphical inspection of the residuals was done to test the proportional hazards assumption. Pairwise comparisons were conducted to determine the effect of the number of stop messages by message type. All analyses were run in R Statistical Computing Package (v 3.2.1, R Foundation for Statistical Computing).

Table 1. Message types, frequency, and examples of the last message received.

Message content type	Messages sent, n	Message example
Nutrition	587	Some students made mango salsa in nutrition class last week. All children have a copy of the recipe in their backpack today. Give it a try!
Physical activity	170	It's going to be almost 50 degrees this afternoon! Enjoy some time outside with the kids after school. Take a walk with your family or play a game of catch!
Stop	55	Text2BHealthy is a free program from the University of Maryland. If this program does not fit into your text plan, or you no longer want messages reply STOP.
Nutrition with stop[a]	14	Broccoli is in season and $1.79 per pound at [local grocery store]. Steam or eat raw with low fat dip. Kids love broccoli! Msg & Data Rates May Apply. Reply STOP to quit.
Physical activity with stop[a]	16	It's December & it's warm outside! Children love the extra time to play outside before dinner. Msg & Data Rates May Apply. To quit receiving messages, reply STOP.
Other[b]	130	Take the family downtown this Saturday & Sunday for the Book Festival. Read, dance & hear from chefs & food experts.

[a]Nutrition with stop and physical activity with stop were combined because of the small number of messages in each category.

[b]Other message types represent a broad variety of content that could not be classified into 1 of the 5 predominant message content types. Other messages include evaluation questions and survey reminders, goal setting, and general community health event notifications.

Results

Descriptive Statistics

There were 972 participants during the 2012 to 2013 program year. Among the 10 participating schools, the average free and reduced meals (FARM) rate, indicating the percentage of students in the school receiving free or reduced price meals, was 80.9%. According to the 2013 RUCC [17], 7 out of 10 (70%) schools were located in metro counties, with populations of 1 million or more, whereas 1 school was located in a metro county with a population of fewer than 250,000, and 2 schools were located in a nonmetro county, with urban populations between 2500 and 19,999. Retention rates by school ranged between 74% (60/81) and 95.0% (132/139), with an average retention rate of 85.7% (833/972) across all schools. Program duration ranged from 7 to 282 days, with a median program duration of 233 days and average program duration of 211.7 days (see Table 2). In total, 14.3% (139/972) of participants dropped out of the program. Among those who dropped out, program duration ranged from 7 to 247 days, with a median program duration of 102.5 days.

Survival Analysis

Kaplan-Meier survival curves indicate differences in program attrition by SMS text message type (see Figure 1). A log rank test was used to examine differences in attrition among SMS text message types (χ^2_3=916.6; P<.001). Among those who did drop out of the program, participants were more likely to drop out after having received a stop message, followed by a stop message paired with nutrition or physical activity content and then by physical activity content alone. Participants were least likely to drop out after receiving a nutrition message. On the basis of these findings, Cox proportional hazards models were fit to estimate the effect of several covariates on program attrition.

Model 1 examines the predictive effect of SMS text message type on program attrition (see Table 3). As seen in Figure 1, stop messages are associated with a high probability of program attrition (hazard ratio=51.5, 95% CI 32.46-81.7; P<.001). The addition of covariates to the model attenuates this relationship; however, the magnitude of the effect of stop messages remains substantial in comparison with nutrition messages. In model 2, which examines SMS text message type and the total number of stop messages received, the hazard ratio decreases substantially to 10.36 (95% CI 6.14-17.46; P<.001). In addition, model 2 shows that the more stop messages a participant receives, the greater the probability of attrition. In model 3, rurality of the program site impacts attrition, as the addition of the RUCC increases the hazard ratio to 11.60 (95% CI 6.78-19.84; P<.001).

Table 2. School characteristics.

School (n=participants)	FARM[a] rate (%)	RUCC[b]	Retention rate (%)	Duration (days), mean (SD)	Duration (days), median
1 (participants, n=65)	97	1	86	202.31 (71.6)	234
2 (n=46)	63	6	85	223.93 (52.2)	241
3S[c] (n=99)	77	1	83	198.42 (64.2)	219
3E[d] (n=55)	77	1	80	198.91 (69.1)	219
4 (n=119)	91.2	1	84.9	215.45 (47.7)	233
5 (n=139)	77.7	3	86.3	213.12 (60.2)	243
6 (n=95)	99	1	97	229.38 (34.5)	240
7 (n=68)	75	1	85	183.25 (71.6)	226
8 (n=139)	92.4	1	95.0	230.14 (33.3)	240
9 (n=66)	73	1	76	195.03 (75.5)	221
10 (n=81)	63	6	74	213.89 (61.4)	227
Total (N=972)	80.9	—[e]	85.7	211.66 (58.8)	233

[a]FARM: free and reduced meals. Students are eligible for free school meals if household annual income falls below 130% of the federal poverty guidelines. Students are eligible for reduced price meals if household annual income falls between 130% and 185% of the federal poverty guidelines. FARM rates are school-level data that represent the entire population of each school, not the Text2BHealthy participant sample.

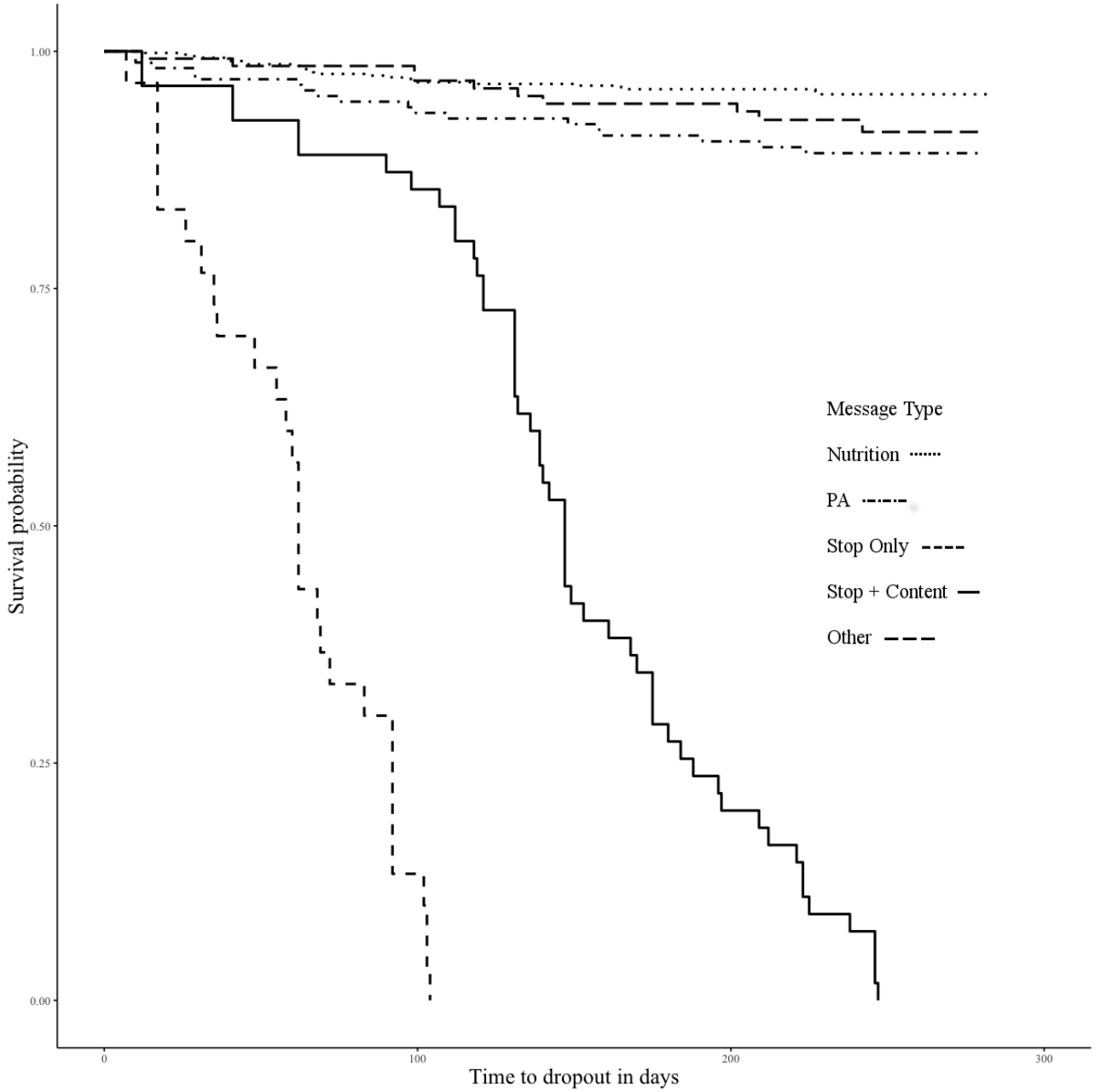
[b]RUCC: Rural-Urban Continuum Code. RUCC 1: fringe counties of metro areas of 1 million population or more (metro county); RUCC 3: counties in metro areas of fewer than 250,000 population (metro county); RUCC 6: urban population of 2500 to 19,999, adjacent to a metro area (nonmetro county).

[c]3S: Spanish language messages sent to participants at school 3. These 2 groups were kept separate because 3 participants from school 3 elected to receive both English and Spanish messages. These 2 groups were kept separate because 3 participants from school 3 elected to receive both English and Spanish messages.

[d]3E: English language messages sent to participants at school 3. These 2 groups were kept separate because 3 participants from school 3 elected to receive both English and Spanish messages.

[e]Not applicable.

Figure 1. Kaplan-Meier survival curves by message type.



Living in a metro county with populations under 250,000 (RUCC=3) is associated with higher probability of dropping out of the program (hazard ratio=4.27, CI 2.33-7.83; *P*<.001), whereas living in a nonmetro county with a population between 2500 and 19,999 people (RUCC=6) is associated with lower, but statistically insignificant probability of dropping out of the program (hazard ratio=0.63, CI 0.39-1.02; *P*>.05), respectively, compared with living in a metro county with populations greater than 1 million (RUCC=1).

In model 4, the final model in this analysis, interaction terms between the total number of stop messages and SMS text message types are included. The effect of the interaction term was explored through pairwise comparisons by message type (see Pairwise Comparisons by Message Type section). The addition of all the covariates and the interaction terms to the model leads to a 99% decrease in the effect of stop message only from model 1, and the association is no longer statistically significant (*P*>.05). The interaction between stop messages and the total number of stop messages received, however, has the largest effect on likelihood of dropping out of the program (hazard ratio=3.31, 95% CI 2.35-4.64; *P*<.001), suggesting a possible moderating effect of the total number of stop messages on the association between receiving stop messages and attrition rate in the program.

Table 3. Predictors of attrition by message type.

Variable[a]	Model 1[b] (95% CI)	*P* value	Model 2[c] (95% CI)	*P* value	Model 3[d] (95% CI)	*P* value	Model 4[e] (95% CI)	*P* value
Physical activity message	2.50 (1.36-4.58)	.003	0.76 (0.39-1.47)	.41	0.88 (0.45-1.74)	.71	0.17 (0.06-0.49)	.001
Stop message	51.5 (32.46-81.7)	<.001	10.36 (6.14-17.46)	<.001	11.60 (6.78-19.84)	<.001	0.43 (0.16-1.16)	.10
Other message	1.82 (0.87-3.80)	.11	1.00 (0.46-2.22)	.98	0.45 (0.19-1.08)	.07	0.21 (0.04-1.08)	.06
Total stop messages received	—[f]	—	0.04 (0.02-0.05)	<.001	0.03 (0.02-0.04)	<.001	0.02 (0.01-0.03)	<.001
County RUCC[g] of 3	—	—	—	—	4.27 (2.33-7.83)	<.001	3.05 (1.65-5.64)	<.001
County RUCC of 6	—	—	—	—	0.63 (0.39-1.02)	.06	0.45 (0.27-0.76)	.003
Physical activity[h] × total stop	—	—	—	—	—	—	1.76 (1.21-2.57)	.003
Stop message[h] × total stop	—	—	—	—	—	—	3.31 (2.35-4.64)	<.001
Other message[h] × total stop	—	—	—	—	—	—	1.43 (0.82-2.50)	.21

[a]Hazard ratios for cox proportional hazard models.

[b]Unadjusted model.

[c]Model 1+Total number of stop messages.

[d]Model 2+County RUCC.

[e]Model 3+Interaction term between message type and total number of stop messages.

[f]Not applicable.

[g]RUCC: Rural-Urban Continuum Code.

[h]Interaction term.

Table 4. Pairwise comparisons by message type for 2 levels of total stop messages received.

Pairwise comparison	3 stop messages, hazard ratios (95% CI)	6 stop messages, hazard ratios (95% CI)
Other versus nutrition	0.62 (0.44-0.89)	1.84 (0.89-2.55)
Physical activity versus nutrition	0.91 (0.73-1.23)	5.1 (1.68-8.64)
Physical activity versus other	1.47 (1.09-1.88)	2.74 (1.79-3.14)
Stop versus nutrition	15.44 (12.92-18.31)	557.6 (506.33-604.51)
Stop versus physical activity	16.89 (9.61-22.04)	110.1 (76.33-132.63)
Stop versus other	24.84 (12.3-37.5)	303.89 (230.45-374.69)

Pairwise Comparisons by Message Type

The effect of the number of stop messages differs by SMS text message type (see Table 4). To further explore these relationships, 2 different values representing the 10th and 90th percentile for total number of stop messages received were chosen. These values were used to estimate the hazard ratios for different pairwise comparisons of message types. The effect of the number of stop messages received had differing effects for different pairwise comparisons. Most notably, the impact of the number of stop messages was particularly pronounced when comparing the probability of dropout after receiving a stop message with the probability of dropping out after receiving another type of message. For example, the stop message and nutrition message comparison shows that receiving a stop message had a 15 times greater probability of resulting in a dropout than receiving a nutrition message with 3 stop messages received, but once 6 stop messages had been received, receiving a stop message had a 557 times greater probability of resulting in a dropout than a nutrition message. A similar pattern was observed when comparing stop and physical activity messages as well as stop and other types of messages.

Discussion

Principal Findings and Comparison With Previous Work

The objective of this study was to examine the effect of SMS text message type and number of stop messages received on attrition in Text2BHealthy, a text-based health promotion program. We found that overall attrition differed by message

type; in particular, sending a stop message substantially increased the risk of participants dropping out of the program compared with other types of messages, including nutrition, physical activity, and others. Although providing information about how to withdraw from a program might be necessary, the way in which this information is provided has important implications for program attrition.

Most participants did not drop out of Text2BHealthy; 85.7% (833/972) were retained through the end of the school year, indicating possible exposure to key messages over a lengthy period. This is consistent with other literature, including 2 meta-analyses that found a mean retention rate of 86% [9] and rates ranging from 46% to 96% [13] across a variety of text-based programs. However, depending on the participant burden in specific program designs, high retention might reflect that passively remaining in a text-based program is easier than taking action to drop out. Future research should move beyond looking at retention to also explore the extent of participants' active engagement (eg, opening, reading, and acting on SMS text messages received). In addition, program characteristics such as planned program duration, participant burden, frequency of messages, message timing, and difficulty of changing the targeted health behaviors may influence retention. Studies isolating these factors for comparison and meta-analyses that examine program characteristics are needed.

We also found that the median program duration for participants who dropped out of the program was high (102.5 days), meaning that many participants were exposed to a substantial amount of program content despite ultimately unenrolling from the program. This observation is inconsistent with research conducted by Coa and Patrick [14], showing that attrition tends to occur within the first 2 weeks of a program. Although it is unclear why results from these studies differ, program planners should assess the likelihood of attrition at various points in their own programs, take into account when program fatigue is likely to occur when determining program length, and consider overenrolling participants to limit the impact of attrition on program evaluation. SMS text message program formats are both effective in the short term [18,19] and beneficial in extending contact with participants beyond an initial intervention period [20], but more research is needed to examine the conditions under which participants tolerate longer program durations.

In examining the context of program participants, we found that parents living in rural counties were less likely to drop out than parents living in more metropolitan counties. This finding may be explained by the relative scarcity of health services and programs in rural areas. Participants in rural areas may be less likely to drop out of a text-based health promotion program, as rural areas tend to have fewer health services and programs available [15]. It is likely that many other contextual factors may impact attrition, but more research is needed to identify such factors and mitigate their unique impact in SMS text message programs. If limited program availability in rural areas is both a motivation to use SMS text message programs with isolated populations and an explanation of high retention, related characteristics of limited access to resources, health disparities, and isolation such as socioeconomic status, race or ethnicity, immigration status, basic literacy, and exposure to other programs should be examined in future research.

We found that the effect of the number of stop messages received during the program period differs by SMS text message type. In particular, receiving 6 stop messages results in greater probabilities of attrition for all SMS text message types than receiving 3 stop messages, suggesting a possible dose-response relationship between number of stop messages received and the likelihood of dropping out of the program. In addition, the interaction between stop messages and the total number of stop messages received yields the largest effect on attrition compared with the interaction between other message types and the number of stop messages received. Although we were unable to find any other research examining the effect of stop message receipt on attrition, these findings might echo dose-response observations in another study showing that number of messages received was associated with positive behavior changes related to weight management [20]. These findings are also consistent with previous research emphasizing the importance of message characteristics, including content and the number of SMS text messages received, in achieving high retention rates in SMS text message–based health promotion programs [21-23].

Program planners needing to send instructions on how to drop out of a program should consider limiting the number of times this information is provided, as it is possible that a greater frequency of such information has negative implications for program retention. In addition, a better understanding of the ways in which participants' characteristics impact probability of dropping out after receiving a stop message could also improve retention. In particular, certain participants might be more responsive to stop messages, such as those who join many text-based programs and those with limited facility with SMS text messaging who might drop out accidentally or misunderstand the intention of the stop message. Furthermore, future research may identify the best ways of informing participants about how to remove themselves from the program and illuminate which groups of participants may be expected to already know how to remove themselves from any SMS text message program.

Limitations

This study has a number of important limitations. First, we were unable to control for individual demographic characteristics that might explain differences in attrition. Second, although the data analysis accounts for different frequencies with which particular messages were sent, the relatively small number of messages combining nutrition content with a stop message and physical activity with a stop message signifies a small number of possible dropout events to observe, resulting in limited or inadequate statistical power. We, therefore, combined these 2 SMS text message types with stand-alone stop messages into 1 group for analysis, which may obscure differences that might have been detected with more observations. It is also possible that certain messages within a message category impacted attrition differentially, but because of somewhat varied content across schools, we were unable to examine attrition probabilities for each unique message. Third, although we are able to link attrition events to the most recently sent message, this does not

necessarily indicate that a participant chose to drop out of the program because of the content of this particular message. In addition, program data indicate only that messages were sent to a functioning cellular number, not whether participants read the messages they receive or whether the messages were impactful. Therefore, participant retention is not itself an indication of either participant engagement or program efficacy.

Conclusions

This study of attrition in the Text2BHealthy program demonstrates the potential of SMS text message programs to retain participants over a long program duration. In examining the patterns of attrition, we have provided evidence that the probability of attrition increases when participants receive SMS text messages with instructions about withdrawing from the program. Program planners should carefully consider how and how often to provide such information to minimize its effect on retention and determine other possible message content and characteristics that may undermine retention. Despite substantial progress in understanding best practices in SMS text message program design and implementation, more research is needed to determine participant, programmatic, and contextual predictors of program duration and attrition to mitigate their impact in SMS text message programs. Furthermore, the relationship between program duration and attrition and targeted behavioral outcomes also necessitates examination.

Acknowledgments

The authors wish to acknowledge the contributions of Kaitlyn Moberly, Lauren Messina, Lindsey Zemeir, and Nicole Finkbeiner in evaluating the program; Amy Bortnick and Sally Kamen in implementing the program; and Lisa Gonzales, Quinney Harris, Tammy Humberson, Jane Kostenko, Mary Rooks, Kristin Spiker, Haylee Staruk, Karen Turner, Joi Vogin, Lauren Williams, Michelle Wilson, Amy Wolpert, and Virgilia Zabala in supporting recruitment and evaluation activities and facilitating relationships with participating schools. Funding for the Supplemental Nutrition Assistance Program-Education program is provided by United States Department of Agriculture in cooperation with the Maryland Department of Human Services and the University of Maryland.

Authors' Contributions

SKG, ALM, KES, EBD, LW, and LL contributed to the formative evaluation, design, development, and implementation of the Text2BHealthy program, and SKG, ALM, KES, and LL contributed to the design and implementation of the broader Text2BHealthy program evaluation. SKG, KES, EBD, LW, and ALM recruited and enrolled participants in the program and evaluation. LW managed program recruitment activities, wrote and sent Text2BHealthy messages, managed the Web platform database, and facilitated the acquisition of data. SKG, EBD, and LL facilitated the acquisition of funding for the program and contributed to project administration, coordination, and supervision. SKG served as the principal investigator for Text2BHealthy evaluation research, and LL is the Director and principal investigator for the Maryland Supplemental Nutrition Assistance Program-Education program. SKG, ALM, and KES contributed to the conception and design of this study. SKG and YV extracted, cleaned, and coded data, whereas EH conducted data analyses. SKG directed data preparation and analysis, and SKG and ALM cowrote initial drafts of the paper. KES and YV contributed to multiple drafts, provided substantial edits and revisions, and organized references. SKG managed the review, revision, and copyediting processes. All authors reviewed and approved the final paper.

Conflicts of Interest

None declared.

References

1. Fjeldsoe BS, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-message service. Am J Prev Med 2009 Feb;36(2):165-173. [doi: 10.1016/j.amepre.2008.09.040] [Medline: 19135907]
2. Pew Research Center. Pew Research Center. 2017. Mobile Fact Sheet URL: http://www.pewinternet.org/fact-sheet/mobile/ [WebCite Cache ID 72NlmWuQK]
3. Pew Research Center. Cell phone activities in 2013 URL: http://www.pewinternet.org/2013/09/19/cell-phone-activities-2013/ [accessed 2018-10-10] [WebCite Cache ID 734XuHsJo]
4. Siopis G, Chey T, Allman-Farinelli M. A systematic review and meta-analysis of interventions for weight management using text messaging. J Hum Nutr Diet 2015 Feb;28(Suppl 2):1-15. [doi: 10.1111/jhn.12207] [Medline: 24480032]
5. Scott-Sheldon LA, Lantini R, Jennings EG, Thind H, Rosen RK, Salmoirago-Blotcher E, et al. Text messaging-based interventions for smoking cessation: a systematic review and meta-analysis. JMIR Mhealth Uhealth 2016;4(2):e49 [FREE Full text] [doi: 10.2196/mhealth.5436] [Medline: 27207211]
6. Nelson LA, Mayberry LS, Wallston K, Kripalani S, Bergner EM, Osborn CY. Development and Usability of REACH: A tailored theory-based text messaging intervention for disadvantaged adults with type 2 diabetes. JMIR Hum Factors 2016 Sep 08;3(2):e23 [FREE Full text] [doi: 10.2196/humanfactors.6029] [Medline: 27609738]
7. L'Engle KL, Mangone ER, Parcesepe AM, Agarwal S, Ippoliti NB. Mobile phone interventions for adolescent sexual and reproductive health: a systematic review. Pediatrics 2016 Sep;138(3). [doi: 10.1542/peds.2016-0884] [Medline: 27553221]

8. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010 Mar;32:56-69 [FREE Full text] [doi: 10.1093/epirev/mxq004] [Medline: 20354039]
9. 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]
10. Speirs KE, Grutzmacher SK, Munger AL, Messina LA. Recruitment and retention in an SMS-based health education program: lessons learned from Text2BHealthy. Health Informatics J 2016 Sep;22(3):651-658. [doi: 10.1177/1460458215577995] [Medline: 25916832]
11. 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]
12. Gazmararian JA, Elon L, Yang B, Graham M, Parker R. Text4baby program: an opportunity to reach underserved pregnant and postpartum women? Matern Child Health J 2014 Jan;18(1):223-232. [doi: 10.1007/s10995-013-1258-1] [Medline: 23494485]
13. Siopis G, Chey T, Allman-Farinelli M. A systematic review and meta-analysis of interventions for weight management using text messaging. J Hum Nutr Diet 2015 Feb;28(Suppl 2):1-15. [doi: 10.1111/jhn.12207] [Medline: 24480032]
14. Coa K, Patrick H. Baseline motivation type as a predictor of dropout in a healthy eating text messaging program. JMIR Mhealth Uhealth 2016 Sep 29;4(3):e114 [FREE Full text] [doi: 10.2196/mhealth.5992] [Medline: 27688034]
15. Douthit N, Kiv S, Dwolatzky T, Biswas S. Exposing some important barriers to health care access in the rural USA. Public Health 2015 Jun;129(6):611-620. [doi: 10.1016/j.puhe.2015.04.001] [Medline: 26025176]
16. Grutzmacher SK, Braunscheidel Duru E, Speirs KE, Worthington L, Munger AL, Lachenmayr LA. Using text messages to engage low-income parents in school-based nutrition education. J Hunger Environ Nutr 2017 Oct 09;13(3):1-5. [doi: 10.1080/19320248.2017.1364196]
17. United States Department of Agriculture. Economic Research Service. 2016. Rural-Urban Continuum Codes 2016 URL: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/ [WebCite Cache ID 72NmdQWKt]
18. Bacigalupo R, Cudd P, Littlewood C, Bissell P, Hawley MS, Buckley Woods H. Interventions employing mobile technology for overweight and obesity: an early systematic review of randomized controlled trials. Obes Rev 2012 Nov 20;14(4):279-291. [doi: 10.1111/obr.12006]
19. Shaw R, Bosworth H. Short message service (SMS) text messaging as an intervention medium for weight loss: a literature review. Health Informatics J 2012 Dec;18(4):235-250 [FREE Full text] [doi: 10.1177/1460458212442422] [Medline: 23257055]
20. Spark LC, Fjeldsoe BS, Eakin EG, Reeves MM. Efficacy of a text message-delivered extended contact intervention on maintenance of weight loss, physical activity, and dietary behavior change. JMIR Mhealth Uhealth 2015 Sep 15;3(3):e88 [FREE Full text] [doi: 10.2196/mhealth.4114] [Medline: 26373696]
21. Abroms LC, Whittaker R, Free C, Mendel VA, Schindler-Ruwisch JM. Developing and pretesting a text messaging program for health behavior change: recommended steps. JMIR Mhealth Uhealth 2015 Dec;3(4):e107 [FREE Full text] [doi: 10.2196/mhealth.4917] [Medline: 26690917]
22. Hingle M, Nichter M, Medeiros M, Grace S. Texting for health: the use of participatory methods to develop healthy lifestyle messages for teens. J Nutr Educ Behav 2013 Jan;45(1):12-19. [doi: 10.1016/j.jneb.2012.05.001] [Medline: 23103255]
23. Woolford SJ, Barr KL, Derry HA, Jepson CM, Clark SJ, Strecher VJ, et al. OMG do not say LOL: obese adolescents' perspectives on the content of text messages to enhance weight loss efforts. Obesity (Silver Spring) 2011 Dec;19(12):2382-2387 [FREE Full text] [doi: 10.1038/oby.2011.266] [Medline: 21869762]

Abbreviations

FARM: free and reduced meals
RUCC: Rural-Urban Continuum Codes
SMS: short message service
SNAP-Ed: Supplemental Nutrition Assistance Program-Education

Edited by G Eysenbach; submitted 28.05.18; peer-reviewed by B Bock, E Bergner, A Kotlo, J Miller; comments to author 17.08.18; revised version received 16.10.18; accepted 17.10.18; published 21.01.19.

Please cite as:
Grutzmacher SK, Munger AL, Speirs KE, Vafai Y, Hilberg E, Braunscheidel Duru E, Worthington L, Lachenmayr L
Predicting Attrition in a Text-Based Nutrition Education Program: Survival Analysis of Text2BHealthy
JMIR Mhealth Uhealth 2019;7(1):e9967
URL: https://mhealth.jmir.org/2019/1/e9967/
doi:10.2196/mhealth.9967
PMID:30664489

Original Paper

Better Ask Than Tell: Responses to mHealth Interrogative Reminders and Associations With Colorectal Cancer Screening Subsequent Uptake in a Prospective Cohort Intervention

Lea Hagoel[1,2*], PhD; Nili Stein[1*], MPH; Gad Rennert[1,2], MD, PhD; Efrat Neter[3*], PhD

[1]Department of Community Medicine and Epidemiology, Carmel Medical Center, Haifa, Israel

[2]Faculty of Medicine, Technion, Haifa, Israel

[3]Department of Behavioral Sciences, Ruppin Academic Center, Emeq Hefer, Israel

[*]these authors contributed equally

Corresponding Author:
Efrat Neter, PhD
Department of Behavioral Sciences
Ruppin Academic Center
Beit 3
Emeq Hefer, 4025000
Israel
Phone: 972 45 646 2678
Fax: 972 9 898 4034
Email: neter@ruppin.ac.il

Abstract

Background: Text message (short message service, SMS) interrogative reminders were adopted in population screening for the early detection of colorectal cancer (CRC).

Objective: This study aims to examine responses to text message (SMS) reminders and associate responses with senders' characteristics, message type (interrogative/declarative), and subsequent screening uptake.

Methods: We conducted a prospective cohort intervention. Text message (SMS) reminders to undergo CRC screening, randomized into interrogative and declarative phrasing, were sent to nonadherent 40,000 women and men (age 50-74 years) at CRC average risk. We analyzed recipient responses by message phrasing, recipient characteristics, and for content, the latter predicting subsequent CRC screening per program database.

Results: While interrogative text message (SMS) reminders elicited 7.67% (1475/19,227) responses, declarative ones elicited 0.76% (146/19,262) responses. Text message (SMS) responses were content analyzed and grouped into attitudes toward CRC screening (1237/1512, 81.8% positive) and intention to screen (1004/1512, 62.6%). Text message (SMS) respondents screened significantly more than nonrespondents after 6 months (415/1621, 25.6% vs 3322/36,868, 9.0%; $\chi 12=487.5$, $P<.001$); 1 year (340/1621, 21.0% vs 4711/36,868; $\chi 12=91.5$, $P<.001$); and 2 years (225/1621, 13.9% vs 3924/36,868; $\chi 12=16.9$, $P<.001$) following the reminders. In a multivariable logistic regression among text message (SMS) respondents, screening after 6 months was significantly predicted by older age, past sporadic screening, attitudes, and intentions.

Conclusions: Interrogative text message (SMS) reminders reached previously uninvolved sectors in the CRC target population—men, sporadic-screenees, and the "never-tested" before. This novel application resulted in a population-level, incrementally enhanced screening. Asking patients about their future health behavior may be relevant for enhancing other health behaviors in preventive medicine and clinical settings.

(JMIR Mhealth Uhealth 2019;7(1):e9351) doi:10.2196/mhealth.9351

KEYWORDS

adherence; colorectal cancer; cancer screening; health behaviors; interrogative reminders; short message service text messages

Introduction

The mortality rate from colorectal cancer (CRC) may be reduced following routine screening and early detection of the disease [1]. CRC screening rates, however, remain relatively low, and enhancement efforts result in a slow, cumulative, change [2]. Adherence to CRC screening is observed mainly among women and older adults; further interventions are also needed for men and younger individuals. Recommendations for innovative approaches to increase CRC screening rates advocate maintaining a "human connection" [3] with individuals in the target population. Reminders using an mobile health (mHealth) technique with attention to wording exemplify such an undertaking [4].

Short message service (SMS) text messaging emerged in 1992, and by 1995, it was a socially acceptable and widely used means of communication [3,5,6]. Since then, SMS text messaging graduated from a personal means of communication among friends and colleagues to a tool used by organizations to contact and inform target audiences [7,8]. SMS has been used (as pre- or postnotification reminders) in the health domain to improve response rates to mailed questionnaires [9], enhance appointment attendance [10,11], reduce posttreatment risk [12], adhere to medication [13], and promote self-management and risk reduction among patients with cardiovascular and coronary heart disease [14,15]. Some reviews indicated that SMS interventions are a robust means for effectively targeting health behavior changes; however, effects have been small to moderate [13,15].

A refined view of SMS is continuously evolving. Studies using the SMS method to remind individuals of recommended health behaviors often imply that this is a unidirectional communication channel. However, 2-way communication between a public agency and stakeholders has also been examined previously [16,17]; the authors inferred that SMS reminders enhanced dynamic feedback and change in health behaviors [17] and provided "information comparable to other modes" [18]. Moreover, the SMS use evoked a social context among recipients that was based on the rapport previously established between provider and health care target audiences; such a rapport is essential for long-term behavior changes [3] in health programs, including cancer early detection.

To date, while few studies have examined the effect of SMS text messages to promote participation in cancer screening [19], very few have focused on SMS text message wording for enhanced screening participation, which is important for reducing CRC-related mortality [2]. This study offers a novel, combined approach to enhance cancer screening through (1) minimal SMS text message reminders for routine CRC screening tests; (2) interrogative wording as reminders, translating a psychological technique to preventive medicine; (3) content analysis of addressees' responses as an interactive dimension; and (4) an objective outcome measure (test performance).

This study branched out of a 50,000 participant field experiment [4], which adapted the question-behavior-effect (QBE) [20] to the population level by using an mHealth tool. Reminders (to screen) were worded as questions or as statements, and either invoked a social comparison or not [21,22] and were sent as SMS text messages through mobile phones to a target audience of nonadherent individuals. *Asking* a question about a person's intention to carry out a health behavior (CRC screening, in this case) in an SMS reminder was found to be more effective than an SMS reminder *stating* CRC screening was advised. Multimedia Appendix 1 displays the original experimental conditions.

This work is an account of recipient *responses* to the SMS reminders in less adherent population sectors invited for CRC screening. The study posed the following research questions (RQs):

> RQ1: What characterized respondents to SMS reminders—demographic attributes, past screening participation, and the experimental condition?
>
> RQ2: What does the response content reveal about attitudes and intentions regarding CRC screening?
>
> RQ3: Are responses to reminders and their content associated with subsequent CRC screening participation?
>
> RQ4: Does the response to the SMS mediate between the experimental condition and CRC screening?

Methods

Participants

In 2013, 50,000 Israeli women and men were routinely invited by mail to screen under the National Israeli Colorectal Cancer Early Detection program [4]. Participants were randomly assigned to 5 equal groups. Individuals in 4 groups received one of 4 SMS versions, while the fifth (control) group received none. This analysis focused on 40,000 addressees in the experimental groups (Figure 1). The Internal Review Board approval number for this study is as follows: 021–26513, 5.5.13 [4].

Procedure

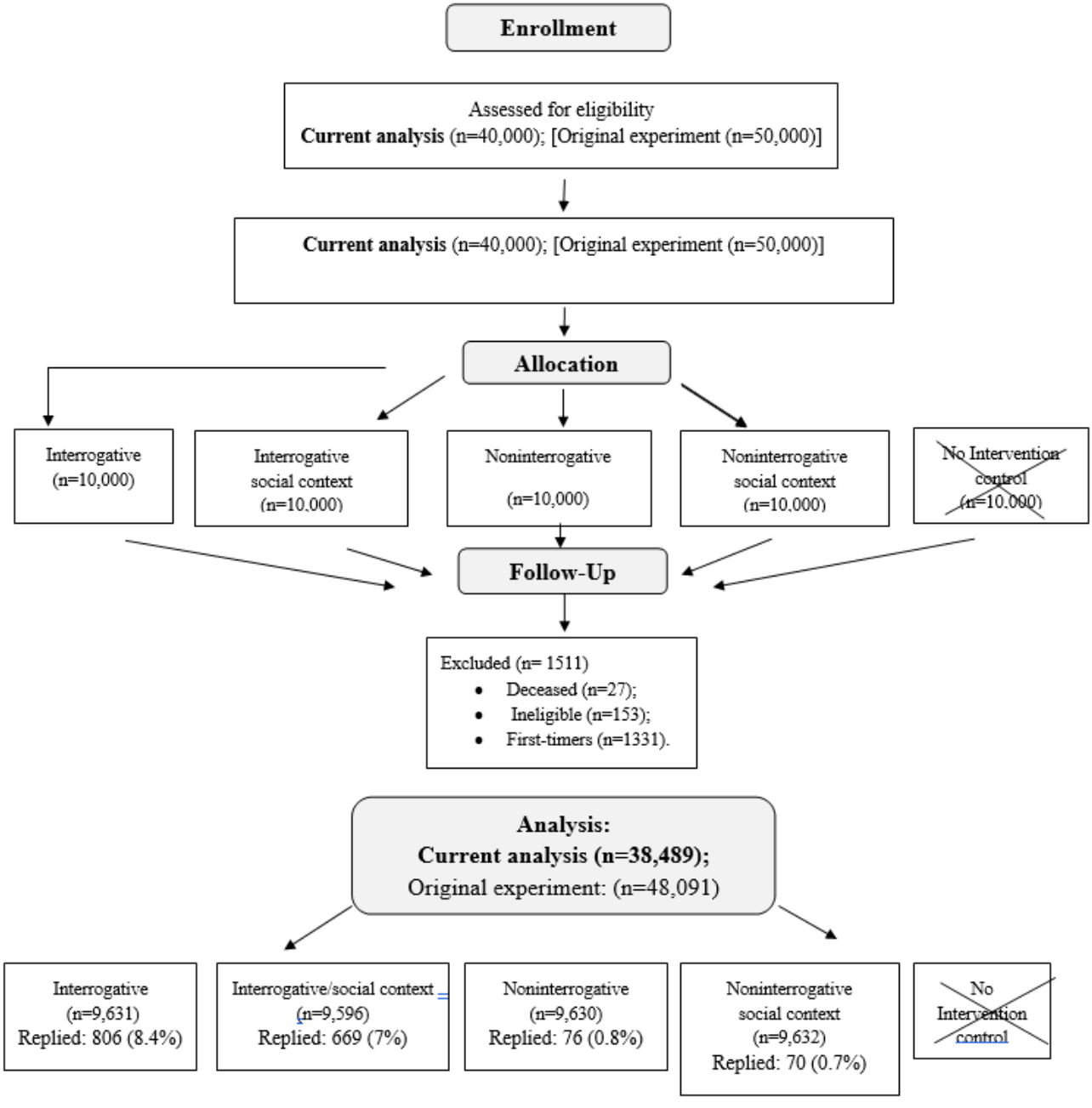
SMS text message reminders yielded responses that were analyzed as predictors of the subsequent fecal occult blood test (FOBT, recommended for individuals at average risk) performance. Demographic variables (age, gender, socioeconomic status, SES) and FOBT performance (past—2004-2012; subsequent—within 6 months, 1 year, and 2 years following the SMS text message) were retrieved from the program's computerized database.

Materials

Short Message Service Text Messages Wording

The brief SMS text messages (122-135 characters) varied in grammatical form (interrogative/noninterrogative): "...do you intend to mail-order an FOBT kit and be tested?" or "...it is important to mail-order a kit and be tested,"), and social comparison of performing FOBT ("as others your age do") [4]. Each version combined grammatical form with/without social comparison (Multimedia Appendix 1).

Figure 1. Flowchart of participants. The two boxes representing the control group are X-ed out, as the participants in the X-ed out boxes are not part of the current analyses.



Responses to the Short Message Service Text Messages

In this study, responses to the SMS were open SMS text messages.

Demographic Characteristics

We retrieved demographic characteristics of participants from the Health Maintenance Organization database and included age, gender, and SES. The SES was determined by the address of the neighborhood clinic insured members attend; members of this Health Maintenance Organization generally attend primary health clinics located in their residential neighborhood. The SES of the clinic's address was based on the classification by the Israeli Central Bureau of Statistics, which assigns an SES ranking to street addresses.

Design

This was a prospective cohort experiment. Initially, the experimental conditions (grammatical form and social comparison) and background variables were independent variables, while participants' responses constituted the dependent variable. In the content analysis stage, the responses, coded and grouped, served as the independent variable, while FOBT performance (past and subsequent) was the dependent variable.

Data Analysis

First, respondents were characterized by experimental condition (grammatical form and social comparison), demographic attributes, and FOBT past performance (sporadic/never). A statistical test of main effects and interaction between grammatical form and social comparison was conducted on responding.

Responses to the SMS were coded for content and length. Two researchers (EN and LH) looked for underlying concepts in an open coding followed by axial coding [23], and labeled categories. The identified categories were grouped into 2 new variables as follows: attitude toward CRC screening and the intention to perform FOBT.

Next, using χ^2 analyses, respondents were tested whether they held positive or negative attitudes toward CRC screening, and whether respondents' intentions toward FOBT performance differed by demographic attributes, past FOBT screening rates, and experimental conditions. Of note, SMS responses were excluded only if they were illegible or if they were returned by ineligible respondents. In addition, the prospective association of the valence of the attitudes and intentions to undergo FOBT at 3 endpoints (at 6 months, 1 year, and 2 years) was examined, as well as the prospective association of SMS response to the reminders (Yes or No) with the undergoing of FOBT at the 3 endpoints mentioned above; all used a chi-square analysis.

Then, a multivariable analysis predicting FOBT screening after 6 months was conducted; the predictors were demographic variables, past (sporadic/never) FOBT performance, experimental conditions (interrogative/declarative), and response valence.

Finally, to test the mediational effect of the experimental condition (*X*) on the FOBT performance (*Y*) by responding to the SMS (*M*), we computed the appropriate indirect effect. To account for the binary nature of *M* and *Y*, we specified these variables as categorical and estimated a model using probit link with ordinal mediator; we followed the example presented in Table 8.26 of Muthén et al [24]. Furthermore, background variables (age, sex, SES, and past FOBT behavior) were included in the model to account for their possible confounding with *M* or *Y*.

This path model was estimated using the Mplus software [25]. See Multimedia Appendix 2 for model specifications, Mplus syntax code, the conceptual and statistical model, and detailed results of the path analysis model. As a robustness check, we also ran the analyses using the percentile bootstrapping method to account for potential nonnormality of our estimates [26].

Results

Responding to Short Message Service Text Messages (RQ1)

An SMS text message response was returned by 4.21% (1621/38,489) of the participants out of SMS recipients. As shown (Figure 1), 7.67% (1475/19,227) responses followed the interrogative conditions, while 0.76% (146/19,262) followed the declarative ones. The grammatical form had a significant effect on response (odds ratio [OR] 11.481, 95% CI 9.059 to 14.551; *P*<.001], while social comparison and the interaction between grammatical form and social comparison did not (OR 0.920, 95% CI 0.664 to 1.275; *P*<.617 and OR 0.892, 95% CI 0.633 to 1.258; *P*<.512, respectively). Social comparison conditions were collapsed in subsequent analyses.

A comparison between respondents and nonrespondents indicated that among respondents, there were significantly more women, individuals of a higher SES, and past FOBT sporadic performers. The response rate did not differ by age (Table 1).

Content Analysis of Responses (RQ2)

The 1621 responses were read, repetitive themes were noted, and categories of responses were defined. Each response was coded accordingly. Researchers worked separately and mostly agreed; in a few cases with divergent judgments, a discussion led to an agreement. Table 2 presents categories and median length of the response field.

Table 1. A comparison between respondents and nonrespondents to the short message service text messages (N=38,489).

Characteristics	No response to SMS[a] (N=36,868), n (%)	SMS respondents (N=1621), n (%)	*P* value
Gender, women	18,776 (50.92)	866 (53.42)	.049
Age, >60 years	18,462 (50.08)	791 (48.79)	.31
Socioeconomic status			<.001
Low	9740 (26.42)	277 (17.09)	
Medium	16,849 (45.70)	695 (42.87)	
High	10,159 (27.56)	644 (39.73)	
Past Fecal Occult Blood Test testing			<.001
Sporadic	9862 (26.75)	608 (37.51)	
Never	27,006 (73.25)	1013 (62.49)	

[a]SMS: short message service.

Table 2. Short message service text message response categories (n=1621).

Content	Responses, n (%)	Code number	Median length of response field[a] (IQR[b])
Yes/OK	626 (38.6)	2	2 (2,6)
Please send me a kit	291 (18)	1	20 (14,31)
No/not interested	212 (13.1)	9	3 (2,7)
I underwent a colonoscopy	135 (8.3)	4	33 (26,43)
[illegible message]	88 (5.4)	7	10 (3,17)
I underwent the test under this program	85 (5.2)	3	25 (17,38)
I did not receive the invitation letter	63 (3.9)	10	18 (14,26)
I have a kit/will soon undergo the test	57 (3.5)	5	14 (9,24)
I have a question (regarding the test or CRC screening)	23 (1.4)	6	19 (11,43)
I was diagnosed with cancer (ie, ineligible for screening)	21 (1.3)	8	29 (17,44)
I underwent the test in a private clinic	13 (0.8)	12	39 (26,50)
Maybe (I'll undergo the test) OR I might undergo the test	7 (0.4)	11	11 (7,27)

[a]In characters, including spaces.

[b]Interquartile range.

The median length of the SMS response field was informative—short for simple messages (#2 and #9), longer, higher variability in elaborate responses (#3, #4, and #12), explaining why respondents did not perform FOBT at this particular time.

The content categories that were identified were then grouped into 2 new variables, relevant to the QBE framework (focusing on participants' intentions; see Multimedia Appendix 1).

The first grouped variable was "*Attitude* toward CRC screening." Responses that implied support for CRC screening were coded as positive; these responses included: (1) explanations why respondents did not perform the test following this intervention (eg, already had undergone the test within the program, #3; or in a private clinic; #12; or underwent a colonoscopy, #4); (2) procedural questions (eg, asking for information on how to obtain an FOBT kit; #6); or (3) clear expressions of positive attitudes such as "yes," "OK," "please send me the kit," "I will soon undergo the test" (such as in responses #1, #2, and #5), and leaning toward undergoing the test ("maybe"; #11). Thus, categories #1, #2, #3, #4, #5, #6, #11, and #12 were grouped as expressing a positive attitude. Categories #9 ("not interested": "no") and #10 ("did not receive an invitation") were grouped as expressing a negative attitude. Note that already having taken the test (#3), having taken the test in a private clinic (#12), or having done another test (#4) express a positive position toward early detection of CRC (not necessarily toward FOBT).

The second grouped variable was "*Intention* to perform FOBT." Categories #1, #2, #5, #6, and #11, with responses such as "please send me a kit" (#1), "yes," "will soon do it" (#5), "I have a question" (#6), and "I may take the test" (#11) were interpreted as conveying an intention to screen. Conversely, categories #3, #4, #9, #10, and #12 where participants reported that they had undertaken screening (either colonoscopy or FOBT; #3 and #4, and #12) were uninterested (#9) or did not receive the invitation, were coded as expressing a negative intention. Most respondents (1237/1512, 81.8% participants) expressed a positive attitude toward CRC screening, and 62.6% (1004/1512) expressed an intention to screen using the FOBT modality.

A bivariate analysis showed that both positive attitudes and intentions toward CRC screening were associated with age (younger), and with past FOBT sporadic uptake (see Tables 3 and 4): individuals aged 50-60 years expressed more positive attitudes toward CRC screening than individuals aged >60 years (χ_1^2=7.4, P=.006), and an intention to undergo FOBT more than others aged >60 years (χ_1^2=25.2, P<.001). Similarly, past sporadic performers expressed a more positive attitude than the never tested (χ_1^2=17.8, P<.001) and showed more intent to undergo FOBT (χ_1^2=14.983, P<.001); the majority of "never-tested" participants expressed positive attitudes (735/936, 78.5%) and intentions (587/936, 62.7%). Attitudes regarding CRC screening and intentions to undergo FOBT were similar and nonsignificant by gender and SES. Finally, receivers of interrogative SMS were not different from receivers of declarative SMS in their attitudes, yet they expressed more intentions to undergo FOBT (944/1407, 67.1% vs 60/105, 57.1%, respectively; χ_1^2=4.3, P=.037).

Table 3. Attitudes toward colorectal cancer screening by participants' background and past screening behavior (n=1512).

Characteristic	Positive (N=1237), n (%)	Negative (N=275), n (%)	*P* value
Gender			.58
Women	670 (81.3)	154 (18.7)	
Men	567 (82.4)	121 (17.6)	
Age (years)			.006
50-60	661 (84.4)	122 (15.6)	
Above 60	576 (79.0)	153 (21.0)	
Socioeconomic status[a]			.11
Low	196 (77.8)	56 (22.2)	
Medium	541 (83.7)	105 (16.3)	
High	497 (81.6)	112 (18.4)	
Past Fecal Occult Blood Test testing			<.001
Sporadic	502 (78.5)	74 (12.8)	
Never	735 (78.5)	201 (21.5)	
Experimental condition			.582
Declarative	88 (83.8)	17 (16.2)	
Interrogative	1149 (81.7)	258 (18.3)	

[a]N=1507, owing to missing data.

Table 4. Intentions to undergo Fecal Occult Blood Test by participants' background and past screening behavior (n=1512).

Characteristic	Yes (N=1004), n (%)	No (N=508), n (%)	*P* value
Gender			.67
Women	551 (66.9)	273 (33.1)	
Men	453 (65.8)	235 (34.2)	
Age			<.001
50-60	566 (72.3)	217 (27.7)	
Above 60	438 (60.1)	291 (39.9)	
Socioeconomic status[a]			.15
Low	168 (66.7)	84 (33.3)	
Medium	447 (69.2)	199 (30.8)	
High	386 (63.4)	223 (36.6)	
Past Fecal Occult Blood Test testing			<.001
Sporadic	417 (72.4)	159 (27.6)	
Never	587 (62.7)	49 (37.3)	
Experimental condition			.04
Declarative	60 (57.1)	45 (42.9)	
Interrogative	944 (67.1)	463 (32.9)	

[a]N=1507, owing to missing data.

Association Between Response Content and Undergoing Colorectal Cancer Screening (RQ3)

Valence in the 2 grouped variables significantly distinguished between SMS respondents, as it was associated with undergoing FOBT in the *6 months* following sending of the SMS text messages—30.5% (377/1237) participants expressing a positive attitude toward CRC screening tested within the next 6 months, compared with 7.3% (20/275) who expressed a negative attitude (χ_1^2=62.5, *P*<.001). Participants who expressed no intention to undergo FOBT underwent the test significantly less than those

who expressed an intention to test—11.4% (58/508), compared with 33.8% (339/1004), respectively (χ_1^2=87.0, *P*<.001). Similarly, 23.3% (288/1237) participants expressing a positive attitude toward CRC screening were tested *after 1 year*, compared with 12.4% (34/275) who expressed a negative attitude (χ_1^2=16.0, *P*<.001). Participants who expressed no intention to undergo FOBT underwent the test significantly less than those who expressed an intention to test—11.6% (59/508), compared with 26.2% (263/1004), respectively (χ_1^2=42.8, *P*<.001). *Two years* following the intervention, 15.7% (194/1237) participants who had expressed a positive attitude toward CRC screening were tested, compared with 5.5% (15/275) who had expressed a negative attitude (χ_1^2=19.8, *P*<.001). Participants who had expressed no intention to undergo FOBT underwent the test significantly less than those who expressed an intention to test—8.1% (41/508), compared with 16.7% (168/1004), respectively (χ_1^2=21.3, *P*<.001). Figures 2 and 3 display screening at 6, 12, and 24 months following reminders by attitude and intentions.

Though the interrogative conditions yielded 10 times more responses than the declarative conditions, participants who chose to respond, across experimental conditions, underwent FOBT more than nonrespondents after 6 months (415/1621, 25.60% vs 3322/36,868, 9.01%; χ_1^2=487.5, *P*<.001; Figure 4). The difference was significant after 1 year (340/1621, 20.97% vs 4711/36,868, 12.78%; χ_1^2=91.50, *P*<.001), and even after 2 years (225/1621, 13.88% vs 3924/36,868, 10.64%; χ_1^2=16.92, *P*<.001).

Then, a multivariable logistic regression on respondents (n=1507) was carried out, with FOBT performance after 6 months as the dependent variable. The predictors were demographic variables, past (sporadic/never) FOBT performance, the 2 grouped variables, *attitude* and *intention*, and the experimental condition. Age (older), past sporadic FOBT performance, attitude, and intention to test expressed in the SMS text message response had a significant effect (OR 1.421, 95% CI 1.097 to 1.840; *P*=.008; OR 3.271, 95% CI 2.540 to 4.213; *P*<.001; OR 2.166, 95% CI 1.204 to 3.894; *P*=.010; OR 2.817, 95% CI 1.909 to 4.156; *P*<.001, respectively).

Mediation Analysis: The Path Between the Experimental Condition (Short Message Service Text Message Type), Responding and Screening (RQ4)

The indirect effect (ie, the total natural indirect effect) of *X*, the experimental manipulation, on *Y* through *M*, was positive and significant (estimate=0.005, *P*<.000), while the pure natural direct effect was insignificant (estimate=–0.004, *P*=.083). The percentile bootstrapping method (with 1000 replicates) yielded similar results—the total natural indirect effect was 0.005 (95% CI 0.004 to 0.006), while the pure natural direct effect was –0.004 (95% CI –0.008 to 0.001). Figure 2 and Table 1 in Multimedia Appendix 2 present detailed results.

Figure 2. The Fecal Occult Blood Test uptake (at months) by attitude.

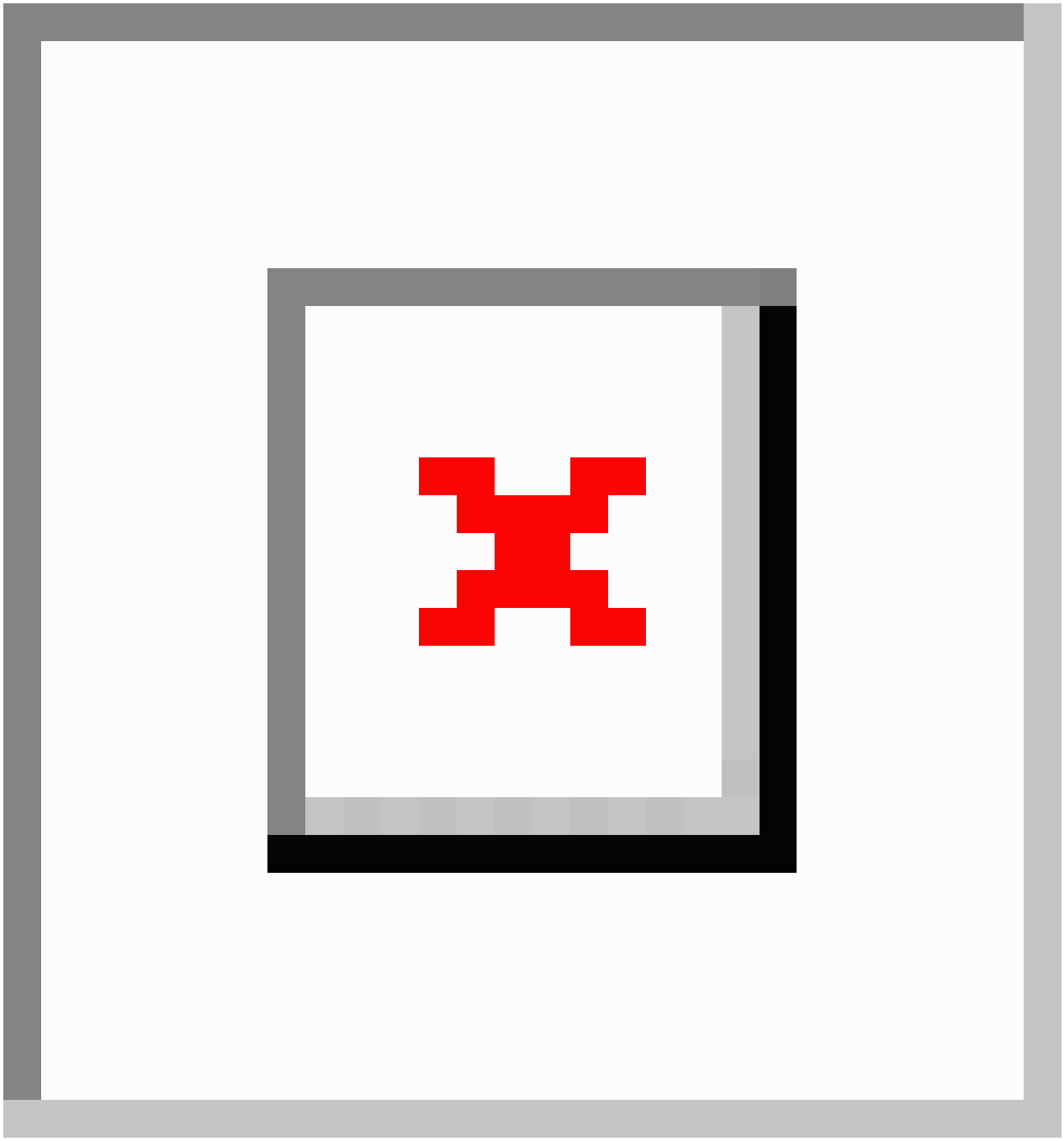


Figure 3. The Fecal Occult Blood Test uptake (at months) by intention.

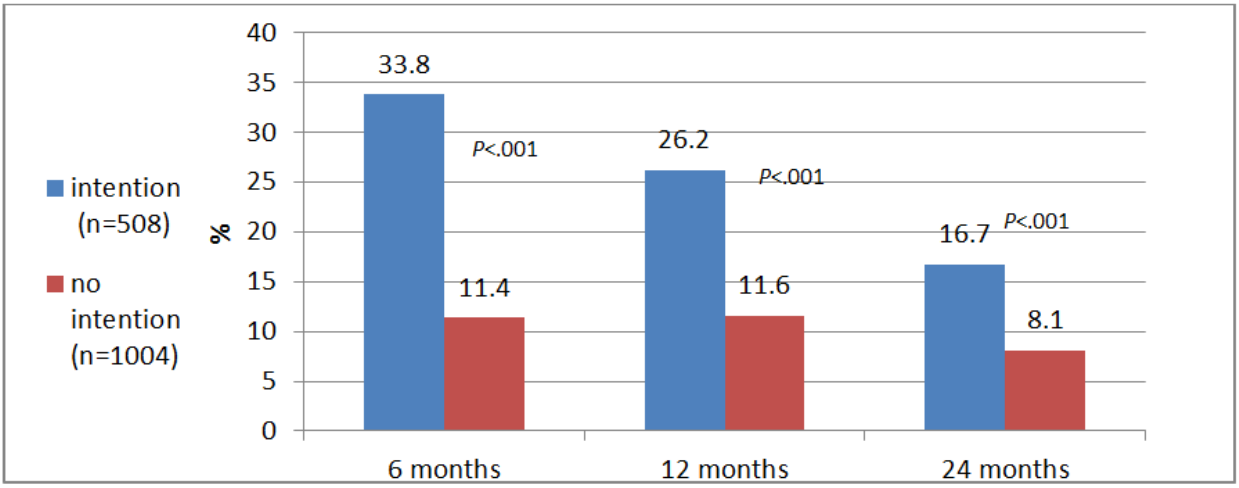
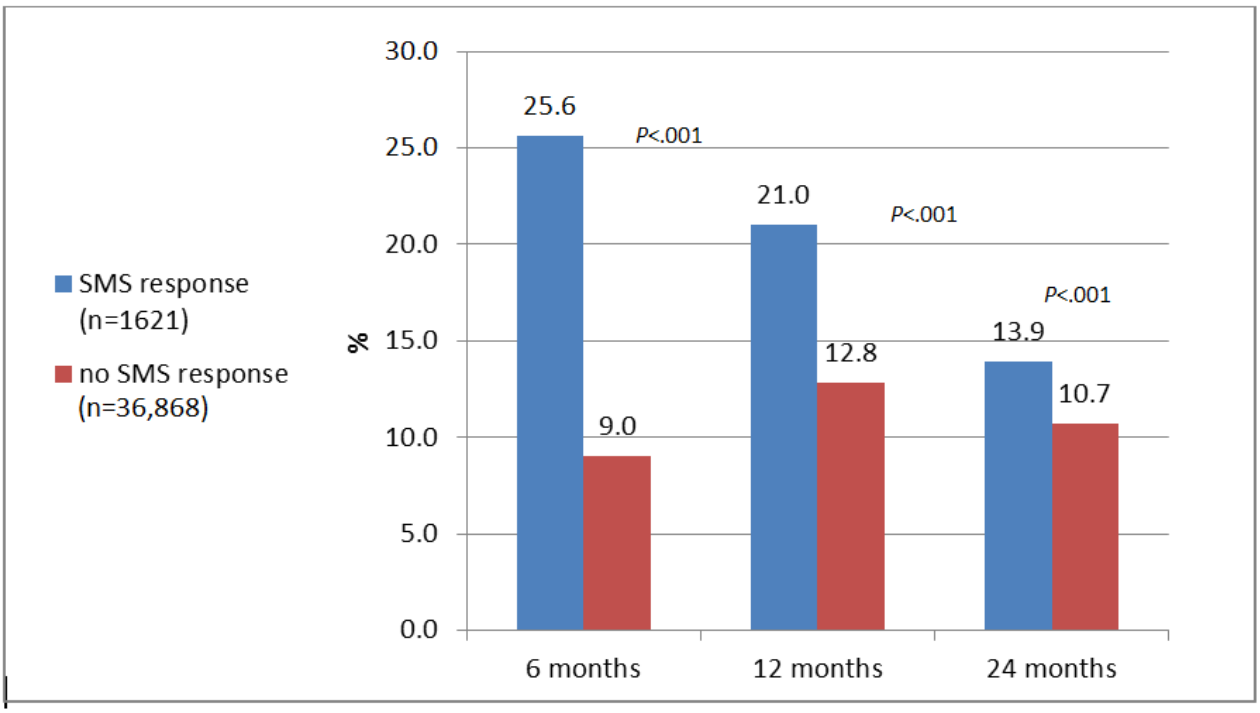


Figure 4. The Fecal Occult Blood Test uptake (at months) by response to short message service (SMS) text message.



Discussion

Principal Findings and Comparison With Prior Work

This analysis addressed responses to mobile phone SMS reminders to enhance the CRC screening participation among a nonadherent sector of the target population. Spontaneous, open SMS text message responses returned to the screening team uncovered another aspect of participants' characteristics, as related to their intention and their subsequent screening uptake.

SMS text message respondents in a nonadherent sector of the target population provided the following 3 indicators: the act of sending back an SMS text message, its content, and engaging in CRC screening. The main findings, discussed in this order, were as follows: (1) interrogative SMS text messages yielded more responses than did typical, declarative reminders (RQ1); (2) the act of responding was predictive of subsequent screening (after 6 months, 1 and 2 years; RQ3); (3) the response content (valence; RQ2) was predictive, across conditions, of subsequent screening (at same time-points; RQ3); (4) in a multivariate analysis, the response valence was predictive of subsequent screening, while experimental conditions were not (ie, respondents across conditions displayed similar screening rates; RQ3); (5) response to the SMS reminders positively and significantly mediated between the experimental condition and CRC screening; and (6) age was related to the response content and subsequent FOBT screening; previous FOBT performance was related to repeating this behavior (RQ1).

The interrogative conditions in this study yielded 10-fold more responses than the declarative conditions, and the positive responses were associated with the target behavior. Furthermore, responding to the message mediated the effect of the

experimental condition on the FOBT performance. This joins previous findings in the health domain [27,28] and further attests to the motivating power of asking questions about the intention to enact health behaviors. Moreover, the linguistic form (interrogative vs declarative) and the responses it generated afforded a rare glimpse into the "introspective self-talk" [29], which theorists posited as enhancing intrinsic motivation. As opposed to questions, statements rarely elicited an internal dialogue, as manifested in more intentions expressed by interrogative SMS receivers than by declarative SMS receivers.

The content of the responses covered the entire range from "yes" and "send me the kit" to explanations why one would not perform the targeted behavior, including "not interested" or "no." The longer responses (ie, extended field length) comprised explanations why respondents would not perform the behavior; this was participants' way to share ideas from their introspective self-talk. The circumstances described ranged from a cancer diagnosis (ineligible to test for the early detection of this disease) to having already screened under a different modality, timing, or health provider.

What is more important for FOBT screening in an SMS reminder intervention—the act of responding or the content of the response? This study could not address this question directly, as there were no content data for nonrespondents. The multivariate analysis provided indirect indications that, among respondents, content (attitudes and intentions) was predictive of screening, while the experimental condition was not. While the content of the "introspective dialogue" matters, it is activated by questions, suggesting a possible mechanism behind the advantage of interrogative reminders.

The respondents in the younger category (≤60 years) were more positive toward CRC screening, expressing an intention to screen more frequently than respondents in the older age category (60+). The intention to conduct a recommended health behavior was a strong predictor of carrying out the behavior [30]. Nevertheless, the older age group screened significantly more than the younger age group within the next 6 months, possibly affected by their previous higher screening rates. Indeed, a gap is apparent between the attitude and intention, on the one hand, and the behavior, on the other; more work is needed to promote screening among younger individuals.

To date, studies have documented adherent individuals to CRC screening with FOBT—as consistently being women and older individuals worldwide [31-33] as well as in Israel [4,34,35]. Increasing FOBT uptake among men and younger age groups (50-60 years) of the target population has been the central aim of screening program organizers for some time. The current findings regarding the efficacy of the interrogative SMS text messages reaching the younger age group, both women and men as well as their expression of positive attitudes and intentions, are evidence that a nonadherent sector of the target population for CRC screening has been reached by the interrogative SMS text messages. The use of SMS "filtered" respondents, inspiring feedback from those who, thus far, have not (regularly, or at all) been involved in CRC screening. This has not yet materialized to a screening behavior among individuals in the younger age group, epitomizing the intention–behavior gap [36].

Undergoing FOBT once is a predictor of repeating this annually recommended health behavior [37]. Including cycle screenees who have never tested and those who have undergone the test irregularly in the screening may contribute to their future routine screening. The more individuals repeat screening, the more this health behavior becomes part of their lifestyle [38]. The 2-year CRC screening follow-up of an SMS reminder, not reported previously, may be a chain-reaction triggered by the SMS, in which participants entered the screening cycle following the reminder, remaining "in the loop" for years to come.

To date, few studies have addressed the unique characteristic of SMS immediacy combined with social contact [39], particularly the space for dialogue carved out by the interrogative wording. Such a dialogue is central to health care and supporting patients in taking recommended action to enhance their health. The technique is scalable to population-level health interventions. Response content and respondents' characteristics and screening patterns highlight a complex, dynamic aspect of "nonadherence" to CRC screening, which program administrators could address; for example, by preprogramming responses sent as a reply to frequently used comments and sharing patient concerns/questions (a mere 1.4% of the responses) with the attending physician.

Strengths and Limitations

The strengths of this field experiment are as follows [4]: an objective outcome measure, a large sample size, and the mHealth method: simple, inexpensive, and parsimonious. The additional 2-year follow-up of a single, short, interrogative SMS reminder, to engaging in health behavior (CRC screening) years later, attests to the impact of the technique's bidirectionality. The comparison of the response content, in receivers of question- vs statement-mode reminders, also pointed, in addition to the higher yield of responses to questions, to a possible explanation. Potentially wide, scalable [40] applications to enhance health behaviors are implied here, which could be used in everyday practice, replacing declarative recommendations—asking patients to predict what they would do, "Do you intend to...?" [41] activates the introspective self-talk [29], which is more effective than "you need to do this." The interrogative wording has rarely been used in SMS text messages. Even though SMS text messages are used abundantly, an examination of alternative wordings has not yet been published. Finally, the mediating effect of the SMS response was indicated using a state-of-the-art statistical technique of mediation analysis.

Study limitations include the lack of evidence that participants read the message. Second, the organizational signature concluding the message may have been less effective than if the attending physician had signed it. In addition, this study did not directly address mechanisms underlying QBE, which may be the goal of further work. Furthermore, implementing the recommendation to *ask* rather than *tell* in interpersonal encounters in the health care setting may seem challenging for established professionals. Finally, the analysis is limited by the lack of data on potentially important confounders such as digital literacy and health status.

Future studies may examine the routine use of SMS interrogative reminders to encourage FOBT kit holders who procrastinate in undergoing testing or supporting other behavioral modifications such as appointment attendance or medication adherence. The strength of posing questions stems from the internal dialogue which follows; interviews with respondents to SMS reminders may shed light on this phenomenon and possibly shape the design of future studies that will attempt to tease the effect of responding apart from the effect of the response content.

Conclusions

SMS interrogative reminders to undergo CRC screening with FOBT have had a long-term effect on sectors in the target population who rarely tested previously, reaching men and younger adults, who expressed positive attitudes toward screening and intentions to test. Medical recommendations, phrased interrogatively, may be more effective than statements. This work provides evidence for this also in the mHealth arena; asking patients may promote behavior change in face-to-face encounters in the clinic and other patient communications.

Acknowledgments

The authors appreciate participants' contribution to the study in responding to the SMS and thank an anonymous reviewer (of the first report of this study) for an interesting comment that initiated the analysis reported here. Finally, the authors appreciate the current reviewers' work and insightful comments to this manuscript. This work was supported by the Israeli Cancer Association (Grant No. 20132129).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Text messages.

[PDF File (Adobe PDF File), 22KB - mhealth_v7i1e9351_app1.pdf]

Multimedia Appendix 2

Mediation analysis.

[PDF File (Adobe PDF File), 71KB - mhealth_v7i1e9351_app2.pdf]

References

1. Hiatt R, Wardle J, Vernon S, Austoker J, Bistanti L, Fox S, et al. Workgroup IV: public education. UICC International Workshop on Facilitating Screening for Colorectal Cancer, Oslo, Norway (29 and 30 June 2002). Ann Oncol 2005 Jan;16(1):38-41. [doi: 10.1093/annonc/mdi033] [Medline: 15598935]
2. Baker DW, Brown T, Buchanan DR, Weil J, Balsley K, Ranalli L, et al. Comparative effectiveness of a multifaceted intervention to improve adherence to annual colorectal cancer screening in community health centers: a randomized clinical trial. JAMA Intern Med 2014 Aug;174(8):1235-1241. [doi: 10.1001/jamainternmed.2014.2352] [Medline: 24934845]
3. Calderwood AH, Roy HK. Increasing colorectal cancer screening adherence: comment on "A randomized comparison of print and web communication on colorectal cancer screening". JAMA Intern Med 2013 Jan 28;173(2):129-131 [FREE Full text] [doi: 10.1001/jamainternmed.2013.2527] [Medline: 23247843]
4. Hagoel L, Neter E, Stein N, Rennert G. Harnessing the question-behavior effect to enhance colorectal cancer screening in an mHealth experiment. AJPH 2016 Nov;106(11):1998-2004 [FREE Full text] [Medline: 27631750]
5. Hillebrand F. Who invented SMS? In: Short Message Service (SMS): The Creation of Personal Global Text Messaging. Chichester, UK: Wiley, John & Sons, Incorporated; 2010:15-21.
6. Hillebrand F. Conclusions. In: Short Message Service (SMS): The Creation of Personal Global Text Messaging. Chichester, UK: Wiley, John & Sons, Incorporated; 2010:131-134.
7. Nielsen. 2013 Jun 6. Mobile Majority: U.S. Smartphone Ownership Tops 60% URL: https://www.nielsen.com/us/en/insights/news/2013/mobile-majority--u-s--smartphone-ownership-tops-60-.html [accessed 2018-11-14] [WebCite Cache ID 73v6dGcos]
8. Epley R. How Different Generations View the Smartphone URL: http://www.tech-stress.com/1-5-how-the-smartphone-is-viewed-by-different-generations/ [accessed 2018-11-14] [WebCite Cache ID 73v6mxSQA]
9. Starr K, McPherson G, Forrest M, Cotton SC. SMS text pre-notification and delivery of reminder e-mails to increase response rates to postal questionnaires in the SUSPEND trial: a factorial design, randomised controlled trial. Trials 2015 Jul 08;16:295 [FREE Full text] [doi: 10.1186/s13063-015-0808-9] [Medline: 26152519]
10. Boksmati N, Butler-Henderson K, Anderson K, Sahama T. The Effectiveness of SMS Reminders on Appointment Attendance: a Meta-Analysis. J Med Syst 2016 Apr;40(4):90. [doi: 10.1007/s10916-016-0452-2] [Medline: 26852337]

11. Guy R, Hocking J, Wand H, Stott S, Ali H, Kaldor J. How effective are short message service reminders at increasing clinic attendance? A meta-analysis and systematic review. Health Serv Res 2012 Apr;47(2):614-632 [FREE Full text] [doi: 10.1111/j.1475-6773.2011.01342.x] [Medline: 22091980]
12. Kampman C, Koedijk F, Driessen-Hulshof H, Hautvast J, van den Broek I. Retesting young STI clinic visitors with urogenital Chlamydia trachomatis infection in the Netherlands; response to a text message reminder and reinfection rates: a prospective study with historical controls. Sex Transm Infect 2016 Mar;92(2):124-129. [doi: 10.1136/sextrans-2015-052115] [Medline: 26404946]
13. Park LG, Howie-Esquivel J, Dracup K. A quantitative systematic review of the efficacy of mobile phone interventions to improve medication adherence. J Adv Nurs 2014 Sep;70(9):1932-1953. [doi: 10.1111/jan.12400] [Medline: 24689978]
14. Park LG, Beatty A, Stafford Z, Whooley MA. Mobile Phone Interventions for the Secondary Prevention of Cardiovascular Disease. Prog Cardiovasc Dis 2016;58(6):639-650. [doi: 10.1016/j.pcad.2016.03.002] [Medline: 27001245]
15. Chow C, Redfern J, Hillis G, Thakkar J, Santo K, Hackett M, 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]
16. Car NJ, Christen EW, Hornbuckle JW, Moore GA. Using a mobile phone Short Messaging Service (SMS) for irrigation scheduling in Australia – Farmers' participation and utility evaluation. Computers and Electronics in Agriculture 2012 Jun;84:132-143. [doi: 10.1016/j.compag.2012.03.003]
17. Revere D, Calhoun R, Baseman J, Oberle M. Exploring bi-directional and SMS messaging for communications between Public Health Agencies and their stakeholders: a qualitative study. BMC Public Health 2015 Jul 08;15(1):621 [FREE Full text] [doi: 10.1186/s12889-015-1980-2] [Medline: 26152142]
18. Walsh E, Brinker J. Short and Sweet? Length and Informative Content of Open-Ended Responses using SMS as a Research Mode. J Comput Commun? 2016;21:100.
19. Sequist TD, Zaslavsky AM, Colditz GA, Ayanian JZ. Electronic patient messages to promote colorectal cancer screening: a randomized controlled trial. Arch Intern Med 2011 Apr 11;171(7):636-641 [FREE Full text] [doi: 10.1001/archinternmed.2010.467] [Medline: 21149743]
20. Rodrigues AM, O'Brien N, French DP, Glidewell L, Sniehotta FF. The question-behavior effect: genuine effect or spurious phenomenon? A systematic review of randomized controlled trials with meta-analyses. Health Psychol 2015 Jan;34(1):61-78. [doi: 10.1037/hea0000104] [Medline: 25133835]
21. Mahler HIM, Kulik JA, Gerrard M, Gibbons FX. Effects of upward and downward social comparison information on the efficacy of an appearance-based sun protection intervention: a randomized, controlled experiment. J Behav Med 2010 Dec;33(6):496-507 [FREE Full text] [doi: 10.1007/s10865-010-9279-3] [Medline: 20652391]
22. Lipkus I, Klein W. Effects of communicating social comparison information on risk perceptions for colorectal cancer. J Health Commun 2006 Jun;11(4):391-407. [doi: 10.1080/10810730600671870] [Medline: 16720537]
23. Strauss AL. Qualitative Analysis Soc Scientists. Cambridge: Cambridge University Press; 1987:55-81.
24. Muthén B, Muthén L, Asparouhov T. Regression and mediation analysis using Mplus. Los Angeles: Muthén & Muthén; 2017.
25. Muthén L, Muthén BO. Mplus Version 8 User's Guide. Los Angeles: Muthén & Muthén; 2017.
26. Mackinnon DP, Lockwood CM, Williams J. Confidence Limits for the Indirect Effect: Distribution of the Product and Resampling Methods. Multivariate Behav Res 2004 Jan 01;39(1):99 [FREE Full text] [doi: 10.1207/s15327906mbr3901_4] [Medline: 20157642]
27. Godin G, Sheeran P, Conner M, Delage G, Germain M, Bélanger-Gravel A, et al. Which survey questions change behavior? Randomized controlled trial of mere measurement interventions. Health Psychol 2010 Nov;29(6):636-644. [doi: 10.1037/a0021131] [Medline: 20939639]
28. Godin G, Bélanger-Gravel A, Vézina-Im L, Amireault S, Bilodeau A. Question-behaviour effect: a randomised controlled trial of asking intention in the interrogative or declarative form. Psychol Health 2012 Sep;27(9):1086-1099. [doi: 10.1080/08870446.2012.671617] [Medline: 22471288]
29. Senay I, Albarracín D, Noguchi K. Motivating goal-directed behavior through introspective self-talk: the role of the interrogative form of simple future tense. Psychol Sci 2010 Apr;21(4):499-504 [FREE Full text] [doi: 10.1177/0956797610364751] [Medline: 20424090]
30. Sheeran P, Maki A, Montanaro E, Avishai-Yitshak A, Bryan A, Klein WMP, et al. The impact of changing attitudes, norms, and self-efficacy on health-related intentions and behavior: A meta-analysis. Health Psychol 2016 Nov;35(11):1178-1188. [doi: 10.1037/hea0000387] [Medline: 27280365]
31. Meissner HI, Breen N, Klabunde CN, Vernon SW. Patterns of colorectal cancer screening uptake among men and women in the United States. Neuropharmacology 2006 Feb;15(2):389-394 [FREE Full text] [doi: 10.1158/1055-9965.EPI-05-0678] [Medline: 16492934]
32. Javanparast S, Ward P, Young G, Wilson C, Carter S, Misan G, et al. How equitable are colorectal cancer screening programs which include FOBTs? A review of qualitative and quantitative studies. Prev Med 2010 Apr;50(4):165-172. [doi: 10.1016/j.ypmed.2010.02.003] [Medline: 20153359]

33. Klabunde C, Blom J, Bulliard J, Garcia M, Hagoel L, Mai V, et al. Participation rates for organized colorectal cancer screening programmes: An international comparison. J Med Screen 2015 Sep;22(3):119-126. [doi: 10.1177/0969141315584694] [Medline: 25967088]
34. Ore L, Hagoel L, Lavi I, Rennert G. Screening with faecal occult blood test (FOBT) for colorectal cancer: assessment of two methods that attempt to improve compliance. Eur J Cancer Prev 2001 Jun;10(3):251-256 [FREE Full text] [Medline: 11432712]
35. Neter E, Stein N, Barnett-Griness O, Rennert G, Hagoel L. From the bench to public health: population-level implementation intentions in colorectal cancer screening. Am J Prev Med 2014 Mar;46(3):273-280 [FREE Full text] [doi: 10.1016/j.amepre.2013.11.008] [Medline: 24512866]
36. McEachan RR, Conner M, Taylor NJ, Lawton RJ. Prospective prediction of health-related behaviours with the theory of planned behaviour: A meta-analysis. Health Psychology Review 2011 Sep;5(2):97-144 [FREE Full text] [doi: 10.1080/17437199.2010.521684]
37. Conner M, Norman P, Bell R. The theory of planned behavior and healthy eating. Health Psychol 2002 Mar;21(2):194-201. [Medline: 11950110]
38. Hagoel L, Ore L, Neter E, Shifroni G, Rennert G. The gradient in mammography screening behavior: a lifestyle marker. Soc Sci Med 1999 May;48(9):1281-1290. [Medline: 10220026]
39. Riley WT, Rivera DE, Atienza AA, Nilsen W, Allison SM, Mermelstein R. Health behavior models in the age of mobile interventions: are our theories up to the task? Transl Behav Med 2011 Mar;1(1):53-71 [FREE Full text] [doi: 10.1007/s13142-011-0021-7] [Medline: 21796270]
40. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A. Mobile Telephone Text Messaging for Medication Adherence in Chronic Disease. JAMA Intern Med 2016 Mar;176(3):340-349 [FREE Full text] [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]
41. Wilding S, Conner M, Sandberg T, Prestwich A, Lawton R, Wood C. The question-behaviour effect: A theoretical and methodological review and meta-analysis. Eur Rev Soc Psychol 2016;27(1):196-230 [FREE Full text] [doi: 10.1080/10463283.2016.1245940]

Abbreviations

CRC: colorectal cancer
FOBT: Fecal Occult Blood Test
mHealth: mobile health
OR: odds ratio
QBE: question-behavior-effect
RQ: research question
SES: socioeconomic status
SMS: short message service

Edited by G Eysenbach; submitted 07.11.17; peer-reviewed by P Sheeran, A Beatty, A Knudsen; comments to author 15.03.18; revised version received 05.05.18; accepted 10.09.18; published 21.01.19.

Please cite as:
Hagoel L, Stein N, Rennert G, Neter E
Better Ask Than Tell: Responses to mHealth Interrogative Reminders and Associations With Colorectal Cancer Screening Subsequent Uptake in a Prospective Cohort Intervention
JMIR Mhealth Uhealth 2019;7(1):e9351
URL: https://mhealth.jmir.org/2019/1/e9351/
doi:10.2196/mhealth.9351
PMID:30664486

Original Paper

Design, Development, and Evaluation of an Injury Surveillance App for Cricket: Protocol and Qualitative Study

Najeebullah Soomro[1,2,3], MIPH, MBBS; Meraj Chhaya[4], MBA, MSci; Mariam Soomro[5], BDS, MPH; Naukhez Asif[2]; Emily Saurman[1], PhD; David Lyle[1], MBBS, PhD; Ross Sanders[3], PhD

[1]Broken Hill University Department of Rural Health, University of Sydney, Broken Hill, Australia

[2]Rural Clinical School, The University of Adelaide, Adelaide, Australia

[3]Discipline of Exercise and Sport Science, Faculty of Health Sciences, The University of Sydney, Lidcombe, Australia

[4]Academy of Computer Science and Software Engineering, University of Johannesburg, Johannesburg, South Africa

[5]School of Public Health and Community Medicine, University of New South Wales, Sydney, Australia

Corresponding Author:
Najeebullah Soomro, MIPH, MBBS
Broken Hill University Department of Rural Health
University of Sydney
Corrindah Court
Broken Hill, 2880
Australia
Phone: 61 880801282
Email: naj.soomro@sydney.edu.au

Abstract

Background: Injury surveillance and workload monitoring are important aspects of professional sports, including cricket. However, at the community level, there is a dearth of accessible and intelligent surveillance tools. Mobile apps are an accessible tool for monitoring cricket-related injuries at all levels.

Objective: The objective of this paper is to share the novel methods associated with the development of the free TeamDoc app and provide evidence from an evaluation of the user experience and perception of the app regarding its functionality, utility, and design.

Methods: TeamDoc mobile app for Android and Apple smartphones was developed using 3 languages: C++, Qt Modeling Language, and JavaScript. For the server-side connectivity, Hypertext Preprocessor (PHP) was used as it is a commonly used cross-platform language. PHP includes components that interact with popular database management systems, allowing for secure interaction with databases on a server level. The app was evaluated by administrating a modified user version of the Mobile App Rating Scale (uMARS; maximum score: 5).

Results: TeamDoc is the first complementary, standalone mobile app that records cricket injuries through a smartphone. It can also record cricketing workloads, which is a known risk factor for injury. The app can be used without the need for supplementary computer devices for synchronization. The uMARS scores showed user satisfaction (overall mean score 3.6 [SD 0.5]), which demonstrates its acceptability by cricketers.

Conclusions: Electronic injury surveillance systems have been shown to improve data collection during competitive sports. Therefore, TeamDoc may assist in improving injury reporting and may also act as a monitoring system for coaching staff to adjust individual training workloads. The methods described in this paper provide a template for researchers to develop similar apps for other sports.

(JMIR Mhealth Uhealth 2019;7(1):e10978) doi:10.2196/10978

KEYWORDS

cricket; injury surveillance; mobile app; mobile phone; TeamDoc; mHealth

Introduction

Emerging technologies are enabling new opportunities for science and medicine within high-performance sports [1]. High-performance sports benefit from applications of science, in the form of physiology, psychology and use of technology to monitor parameters of performance and injury [2,3]. In terms of technology, an emerging domain is mobile health (mHealth), which involves mobile computing by the use of apps on smartphones to improve health [4]. mHealth to monitor an athlete's health has also been identified as an area that can revolutionize sports medicine [5]. Electronic injury surveillance and monitoring tools (including mobile apps) are being used to monitor and predict injuries for sports including athletics, football, and handball [6-8]. However, their use in cricket is limited to elite players, with limited or no availability at the community level.

Cricket Australia (CA) uses the Athlete Management System for workload monitoring and injury reporting of their contracted players [9,10]. This system allows coaches and medical staff to monitor individual workloads, which may reduce the occurrence of overuse injuries. CA's 10-year injury report (2005-2014) indicated that the 2013-2014 season had the lowest prevalence of injury (10.8%) compared with the 10-year average of 11.9% [10]. One factor contributing to the lower injury prevalence during the 2013-2014 season was the introduction of mandatory use of the Athlete Management System [10]. Mandatory reporting by players on variables such as sleep and workload may have helped them adhere to the recommendations of the team's medical staff and, thereby, minimize reportable injuries.

However, the statistics for cricket injuries at the junior level tell a different story; a 5-year investigation of injuries in elite junior cricketers in South Africa indicated that 27% of the cricketers sustained injuries when only time-loss injuries were considered [11]. In Australia, however, the injury incidence in under-14 and under-16 players at the club level was 14.2%, despite the inclusion of both time-loss and nontime loss injuries [12]. Other studies have reported the injury incidence to range between 24% and 34%, and cricket-related musculoskeletal pain has been reported by 80% of school-aged cricketers in a season [13-17]. Yet, the actual injury burden may be higher than what is currently reported because most cricket injury reports discount the burden of nontime loss injuries, where the players continues to play despite the injury. Recording nontime loss injuries is now considered essential according to injury epidemiologists in other sports [7]. Therefore, in 2016, the new consensus statement for injury surveillance in cricket included the reporting of nontime loss injuries [18].

Adolescent athletes are susceptible to injuries due to rapid bone growth and musculoskeletal immaturity [19]. Evidence shows that increasing workload increases injury risk [15,20]. To tackle this problem, international cricket associations have proposed workload guidelines [21,22]. However, these are not being extensively followed at the junior level and may be attributed to the lack of support staff to keep track of the bowling and batting workloads and injuries [23].

Monitoring the training workload by using subjective measures from an athlete is an effective way to address the issue of training loads in sports such as cycling, athletics, and football [24-27]. Similarly, if cricketers record their workloads with a user-friendly mobile app, the increased surveillance may allow coaches to devise injury prevention strategies. Currently, no free-to-download mHealth apps are available that can record cricket-related injuries and monitor workload. Given that elite cricketers emerge from junior cricket, it would be logical to implement such a system at the junior or amateur level. This would have several benefits: first, it could reduce the possibility of talented cricketers being "lost" from the player pool because of injury. Second, it could provide exposure and experience with reporting injuries and workload for those cricketers who progress to the elite levels, where reporting is mandatory. Finally, reporting injuries may enable players to seek timely medical advice and minimize injury effect.

The primary aim of this paper is to outline the methods for app development used to design TeamDoc, a free mHealth app providing paperless, user-friendly solution for monitoring injuries and workloads in junior cricket. The sharing of novel methods associated with the development of TeamDoc will act as a foundation for future app developments in the area of injury surveillance and workload monitoring. The secondary aim of the paper is to provide evidence from a pilot evaluation of the user experience, functionality, utility, and design of the app. As end-user perceptions have been shown to be an important aspect for the long-term uptake of new interventions [28], behavior change was also appraised. The results of the evaluation will assist in improvement of future apps in this domain.

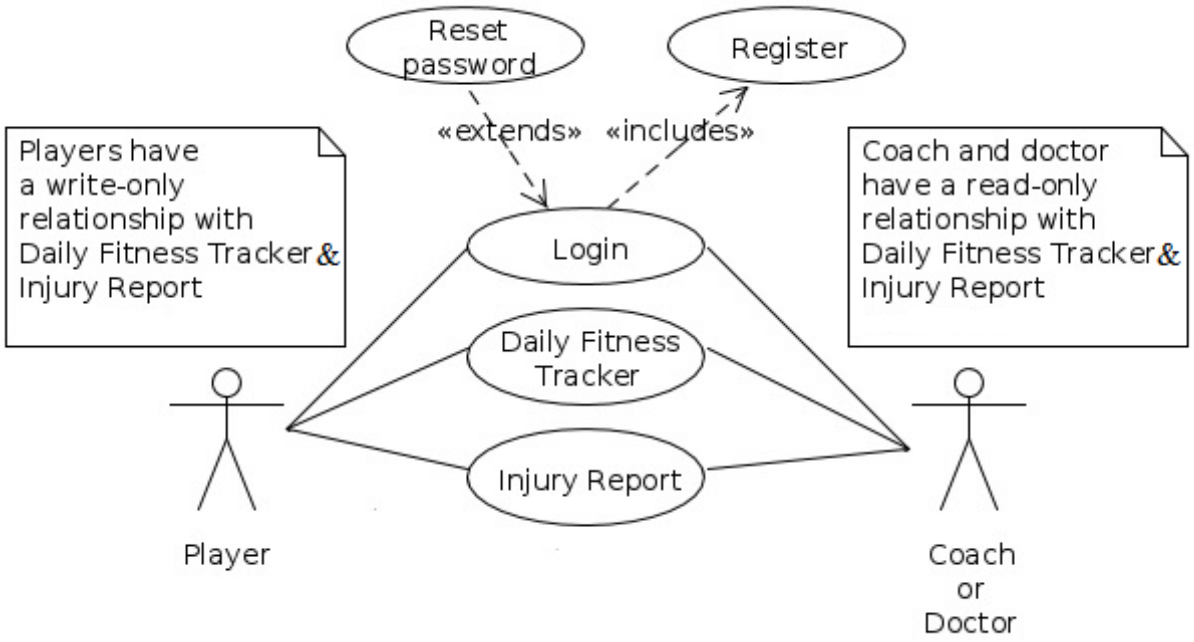
Methods

Software System Development

When designing the app, we took into account several important considerations. First, the app should ensure confidentiality of the data provided by the players. Second, the system needs to be user friendly with ease for quick data entry (not exceeding >2 minutes). Third, there needs to be a back-end server that stores the data for future analysis. Fourth, it should be usable and adaptable for common operating systems. Finally, the injury and workload data must be presented in a way that is easily read and interpreted.

We divided the TeamDoc software system design into 3 components: player interface, coach interface, and a back-end system to securely store the data. The player and coach interfaces are completely separate. This design protects player privacy, as the player interface only permits authorized players to log the data. Similarly, the coach interface only allows authorized coaching staff to access the data. The design of the software relies on client-server architecture, with the player and coach interfaces operating as clients (resource and service requesters) and the back-end system operating as the server (resource and service provider; Figure 1).

Figure 1. A Unified Modeling Language use-case diagram details the functionality offered to players and coaches.



Tools and Languages

Client-Side Software

The client-side software used was an open-source software development platform, Qt 5.3 (Qt Company Ltd, Finland, 2014). This platform was chosen because of its cross-platform compatibility (ability to work on multiple operating systems; eg, it supports Windows, Mac OS X, and Linux for desktops and Android, iOS, and Windows Phone for mobile phones) [29]. This meant that even though the initial release was compatible only with Android, the source code can be ported to 15 other operating systems with relative ease [30]. In addition, it allows developers to program software using a range of different programming languages. We used 3 languages—C++, Qt Modeling Language (QML), and JavaScript—because they are well documented and supported by Qt 5.3.

The chosen platform and languages simplify the construction of custom user interfaces (UIs) and provide the opportunity to augment UI components with high-level logic. The Qt software development kit was used to develop the client-side software. This integrated development environment had a compiler for the C++ language and a graphical user interface designer, allowing for rapid prototyping of UIs.

Server-Side Software

The most important considerations for designing the server-side software were data security and cross-platform connectivity. Hypertext Preprocessor (PHP) scripting language was used as it allows cross-platform server-side connectivity, and it is also commonly used in the website development. PHP includes components that interact with popular database management systems and provides protection against certain malicious attacks, such as structured query language (SQL) injections, which is a technique used to exploit data from servers (although this must be explicitly instructed in source code), allowing for secure interaction with databases on a server level.

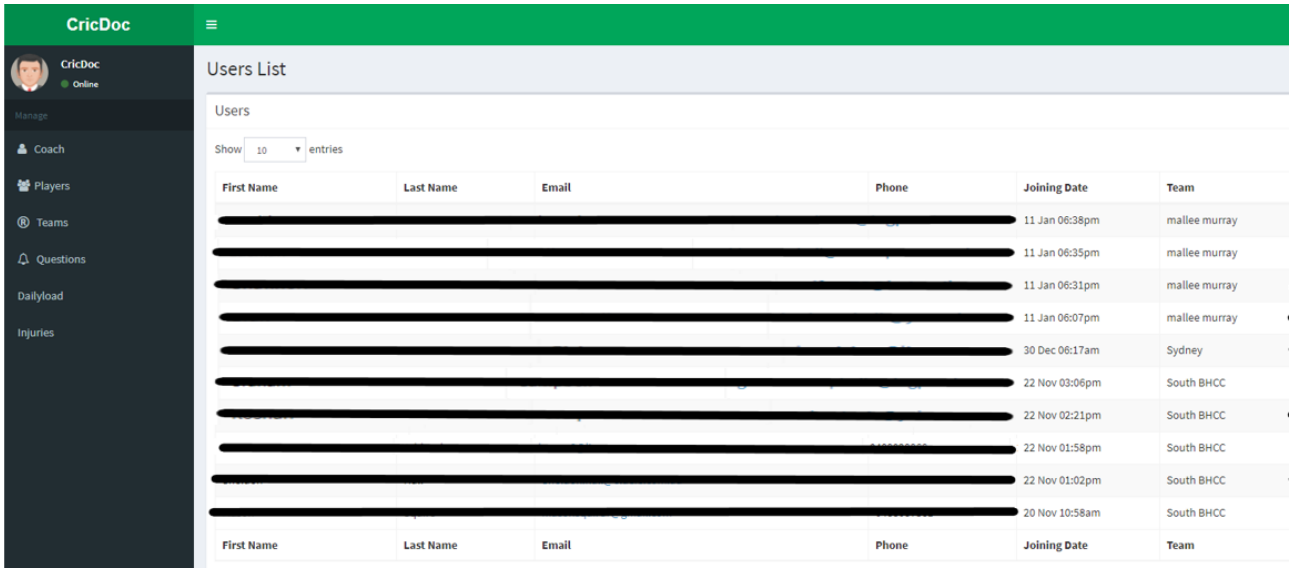
The data entry required to compile injury reports was quite simple, and this led to TeamDoc being built with a thin-server architecture, which is a type of architecture suitable for systems where the majority of computation occurs on the client side, as opposed to fat-server architecture, which requires higher computational resources on the server side. PHP was chosen instead of native executable code, as it is more suitable given this thin-server architecture. Alternatively, with native executable code, Transmission Control Protocol server sockets would have to be implemented and socket communication handled, which would have increased the complexity of this system.

Software Architecture

The 3 components of the system were as follows:

1. *Player App*: A write-only function permits only data entry and restricts access to individual output data, thereby ensuring data privacy. Players can submit information that will be compiled into the Daily Fitness Tracker and Injury Report System.
2. *Coach App*: A read-only function limits data access so that no amendment can be made by the coaching staff after the data have been recorded by the player. This ensures data security and authenticity as multiple people may be involved in a team's coaching staff. Only coaching staff (including doctors) can visualize the collected information, both graphically within the app and in the tabular form (via a Comma-Separated Values file compatible with Microsoft Excel and other spreadsheet apps; Figure 2).
3. *Server-side operations*: Both players and coaches or doctors can log-in and register, and if necessary, reset their password.

Figure 2. Screenshot from server interface showing data view available to the coaching staff.



The interaction of a design-focused language, QML, and 2 logic-focused programming languages, C++ and JavaScript, allowed for the use of the Model-View-Controller (MVC) software architectural pattern. MVC supports the UI, internal data representation, and logic of the project to be independent so that a change in one component does not directly affect the next one [31]. This architectural pattern also ensured that team members involved in the project design were able to engage with the project from their preferred aspect—design, logic, or database management—and minimized delay with parallel, rather than sequential development and implementation.

Database Implementation

To improve cross-platform connectivity in the future, we used a popular and robust open-source database management system called MySQL to capture, query, and administer the data collected by the TeamDoc app. The database comprises 5 tables that store raw and processed data:

- The *Users table* contains information about every user, including user-identification and password.
- The *ResetPasswords* table is an administrative table that stores computer-generated temporary passwords for users who have forgotten their passwords.
- The *DailyFitnessTracker* table stores processed final scores as well as the raw values input by players.
- The *InjuryReport* table collects the players' responses and can be accessed by coaches and affiliated medial staff.
- The *Attendance* table stores player attendance and injury details for a 40-week season, each week includes 3 practice sessions and 1 match. These variables may be adjusted as needed.

The information for each table is linked to each player through a user ID and time-stamp of submission. The values in the tables can be queried manually through SQL or through the interfaces available to coaches and doctors.

User Interface Design

The UI of TeamDoc was designed to be user friendly with little-to-no training time. Figure 3 illustrates the log-in and player and coach interfaces. The player interface provides forms with text-input-boxes, numeric sliders, check-boxes, and radio buttons to make data input quick and simple (Multimedia Appendix 1).

Injury Reporting

The injury reporting in the UI was based on the standard injury reporting form developed by Finch et al and used by the Sports Medicine Australia [32,33]. The form had questions on the activity at the time of injury, reason for presentation, site of injury, nature and mechanism of injury, initial treatment given, action taken after the injury, referral, etc. The site of injury function uses a "branching" logic, which means that when a specific body region was indicated as injured, only questions related to that region would appear. For example, if *Knee* or *Lower leg* was selected, then options such as calf muscles, knee joint, etc, came up. Figure 4 shows injury reporting forms in the player's interface; Multimedia Appendix 1 shows all UIs in the players' app, and Multimedia Appendix 2 shows a video run-down of the app.

Workload Reporting

Workload monitoring was designed for batting and bowling. For batting, the number of balls batted was the primary input and the number of balls bowled for bowling. CA's fast bowler workload guidelines were used to determine if a fast bowler overbowls or underbowls [21]. These guidelines are part of the coach training programs and are standard to monitor the training of fast bowlers. All input from the players is stored on the server and is accessible for the doctor and coach at any time.

Figure 3. User interface of the TeamDoc app. Top left: log-in view; top right: main tab of the player app; bottom left: main tab of the coach app; bottom right: injury report tab in the coach app.

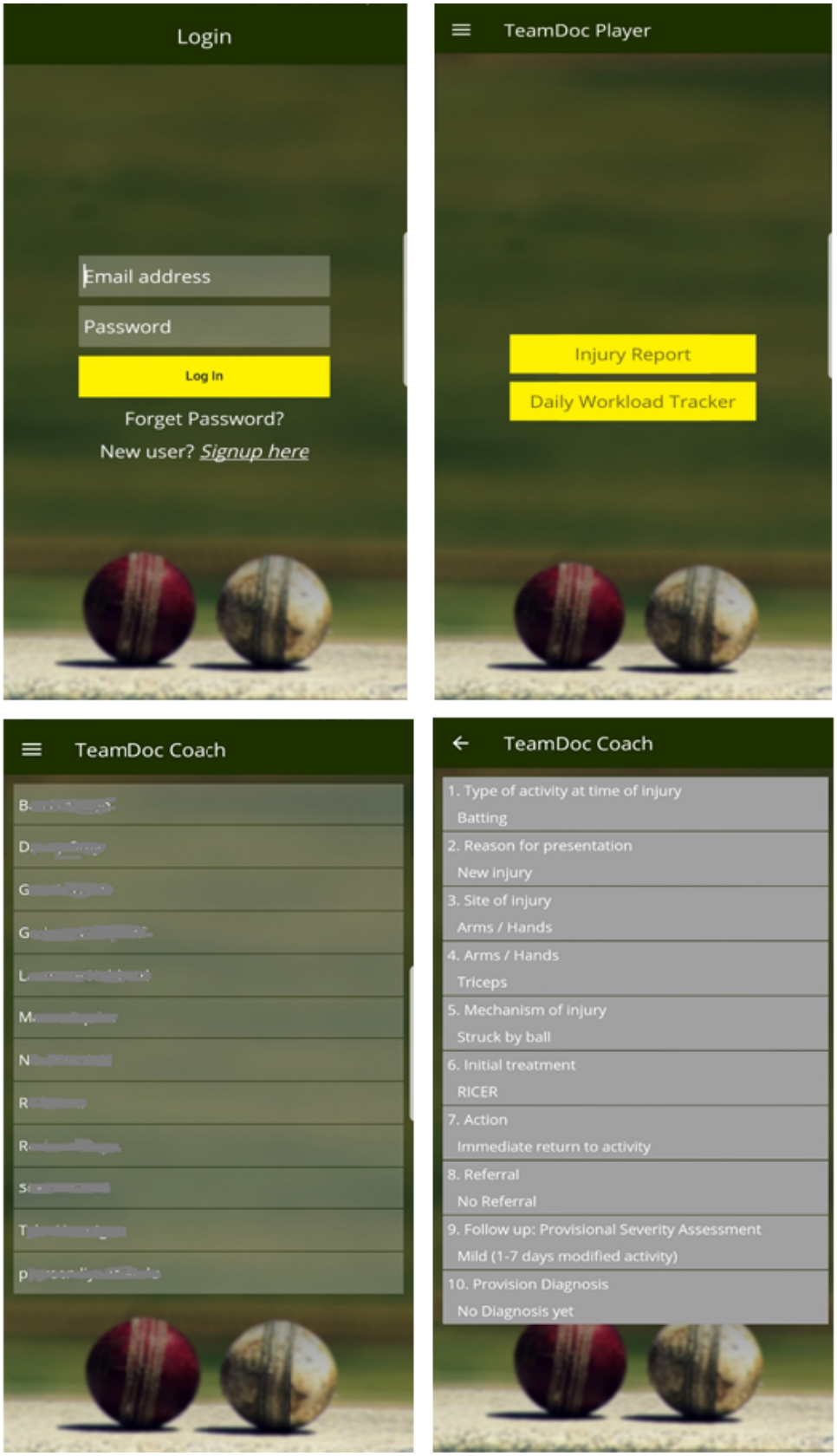
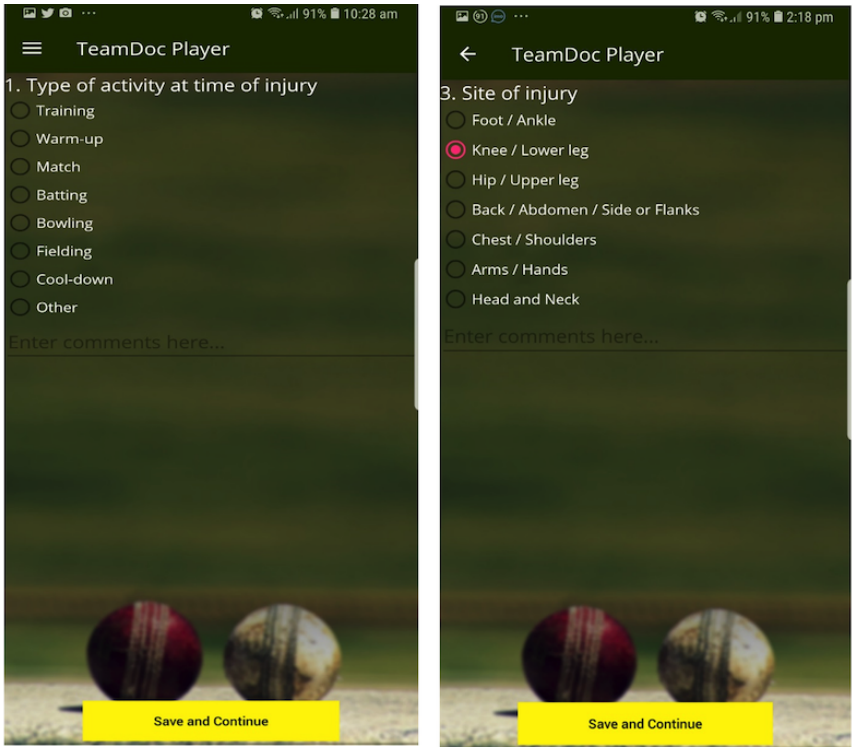


Figure 4. User interfaces for the injury reporting tab in the TeamDoc player app.



Mode of Availability of Software

We built the client side to simplify porting it to diverse smartphone operating systems, the initial release for the Android platform. In addition, we implemented the server side of TeamDoc to be compatible with various server operating systems and avoid dependence on a single technology. This consideration made the release to Apple's iOS platform.

App Evaluation

We used a modified user version of the Mobile App Rating Scale (uMARS) to critically appraise the app [34]. uMARS is derived from MARS, which is a validated mobile app evaluation tool and has been used extensively to rate the quality of medical-based apps [35-37]. It comprises 31 questions, mostly using a Likert-type rating scale, to evaluate an app on the following 3 domains: (1) Quality score: it examines engagement, functionality, aesthetics, and content information provided in the app; (2) Subjective quality: it questions likelihood of recommending the app to others, use in future, overall rating, etc; and (3) Behavior change: it assesses the perceived impacts on knowledge, attitude, awareness, and behavior. The internal consistency (alpha=.90) and interrater reliability (.79) for MARS is acceptable [35].

The MARS was modified by excluding 7 questions on the information content of the app. These questions are only relevant for apps that provide content to users. The 24 questions relevant to assess the TeamDoc app were used (see Multimedia Appendix 3). Field testing and validation of the app was conducted during the initial phase of development by collecting informal user reviews from 20 players on the University of Sydney Cricket team. After the final launch of TeamDoc on the Android and iOS app stores, 8 registered cricket clubs in New South Wales and Victoria, Australia, were invited to use the app during the 2017-2018 season using convenience sampling.

Procedure for User Version of the Mobile App Rating Scale Administration

We administered the modified uMARS at the end of the season through a Web-based survey using the Research Electronic Data Capture (REDCap) Survey instrument. REDCap is a secure, Web-based app for building, disseminating, and managing Web-based surveys and complies with the Health Insurance Portability and Accountability Act regulations. Participant information was secure and only available to the authors through the use of both server authentication and data encryption using secure Web authentication, data logging, and Secure Sockets Layer. The survey was available on Web, and an invitation with the survey link was sent to the registered users.

Data Analysis

We exported the survey results to Microsoft Excel v2013, and performed basic calculations to report standard descriptive statistics. For qualitative data, we performed the content analysis by categorizing the content into themes. Top 3 themes from each category were reported.

Results

Participants

In this study, 3 of the 8 club teams agreed to participate in the app testing and evaluation. Each club team had an average squad size of 14 players. In total, 42 club cricketers (14×3) registered and used the app. The data collected by the app were verified with the data stored on the server by the developers, and the results indicated 100% data accuracy.

App review using the modified uMARS was completed by 16 of 42 cricketers (38% of the app user base). All the respondents were current club cricketers with a mean age of 27.4 (range 16-42) years. Of all, 9 users were running the app on Android smartphones and 7 on Apple iOS (6 used an iPhone and 1 used an iPad). No coaches or doctors responded to the survey.

User Version of the Mobile App Rating Scale Ratings

The mean app quality score (maximum score=5) was 3.6 (SD 0.6); this was compiled from the mean scores on app functionality, engagement, and aesthetics. The mean subjective quality score was 3.1 (SD 0.7). Behavioral change, which included an assessment of the perceived impacts on knowledge, attitude, awareness, and behavior, had a mean score of 3.8 (SD 0.5). The overall mean uMARS score (maximum score=5) was 3.6 (SD 0.5), and the scores ranged from 2.9 to 4.8 (Table 1 and Figure 5).

- *Engagement.* This score ranged from 2 to 4.4 (mean 3.3 [SD 0.7]). Engagement scores were compiled from 5 questions on entertainment, interest, customization, interactivity, and appropriateness for target audience. Appropriateness for target group was rated highly among the "engagement" questions, with an average score of 3.8 [SD 0.9]. However, customization received the lowest score (mean 2.94 [SD 1.1]).
- *Functionality* score ranged from 2.3 to 5 (mean 3.9 [SD 0.7]). Functionality scores were compiled from 4 questions on performance, ease of use, navigation, and gestural interactivity. Ease of use scored highly within the category (mean 4.1 [SD 0.6]), while gestural interactivity was the lowest-rated category with a mean score of 3.8 (SD 1.0).
- *Aesthetics* scores ranged from 2.3 to 4.3 (mean 3.5 [SD 0.6]) and were for questions on the app's layout and graphics to the visual appeal. The layout of the app had the highest score (mean 4.2 [SD 1.3]), and visual appeal had the lowest score (mean 3.2 [SD 0.8]).
- *Subjective Quality* or *Satisfaction.* This score ranged from 2.25 to 4.75 (mean 3.14 [SD 0.7]). These scores were for questions on a recommendation to others, use in the next 12 months, overall star rating, and paying for the app. Recommendation to others received the highest score (mean 3.75 [SD 1.0]), while paying for the app received the lowest score with a mean of 2.75 (SD 1.2).
- *Behavior Change* scores ranged from 2.7 to 5 (mean 3.6 [SD 0.5]). These scores were from 6 questions about awareness, knowledge, attitudes, intention, behavior to change, and help-seeking. The question on behavior change describing the likelihood of the app in improving the understanding of injury and seeking help for it received the highest score (mean 3.9 [SD 0.7]). Conversely, the question on the role of the app to improve the knowledge about injuries received the lowest score (mean 3.6 [SD 0.9]).

User Perceptions

User perceptions were collected with 2 open-ended questions: (1) If you decide or decided not to use this app, what will be the possible reasons for it? (2) What improvements do you want to see in the future versions of the app?

Majority of the respondents (n=10) were not currently using the app (nonusers). The main reasons for not using the app were UI, time consumption, and forgetfulness. Users expressed the importance of functional design improvement of the app, which may have made them feel they were spending too much time on filling out information. Several users expressed that the app lacked the graphic interface and breadth of content to engage them for regular use.

> *I like using this app, however, I need more interactive options in it such as scores and health tips etc.*
>
> *Lack of feedback, unable to enter data in days after activity*
>
> *Time consuming*

Most current users (n=6) mentioned that the reasons for future disuse will be if they did not get injured, stopped playing cricket, or forgot using it.

> *Due to no injuries*
>
> *If I don't play in the future*
>
> *The only reason I wouldn't (use the app) would be forgetting to.*

On the question regarding future improvements in the app, the main reasons cited by current nonusers (n=10) were linked to lack of feedback, UI, and user experience.

> *Being able to see how the data is collated and be able to refer back to this data would help with the information entered. Entering data two days after activities by putting in date (not rely on entering data immediately after activity) would allow more entries to be input. Workload app would consider multiple activity types for example, running while not playing cricket, gym time etc.*
>
> *No graphics or engaging content*

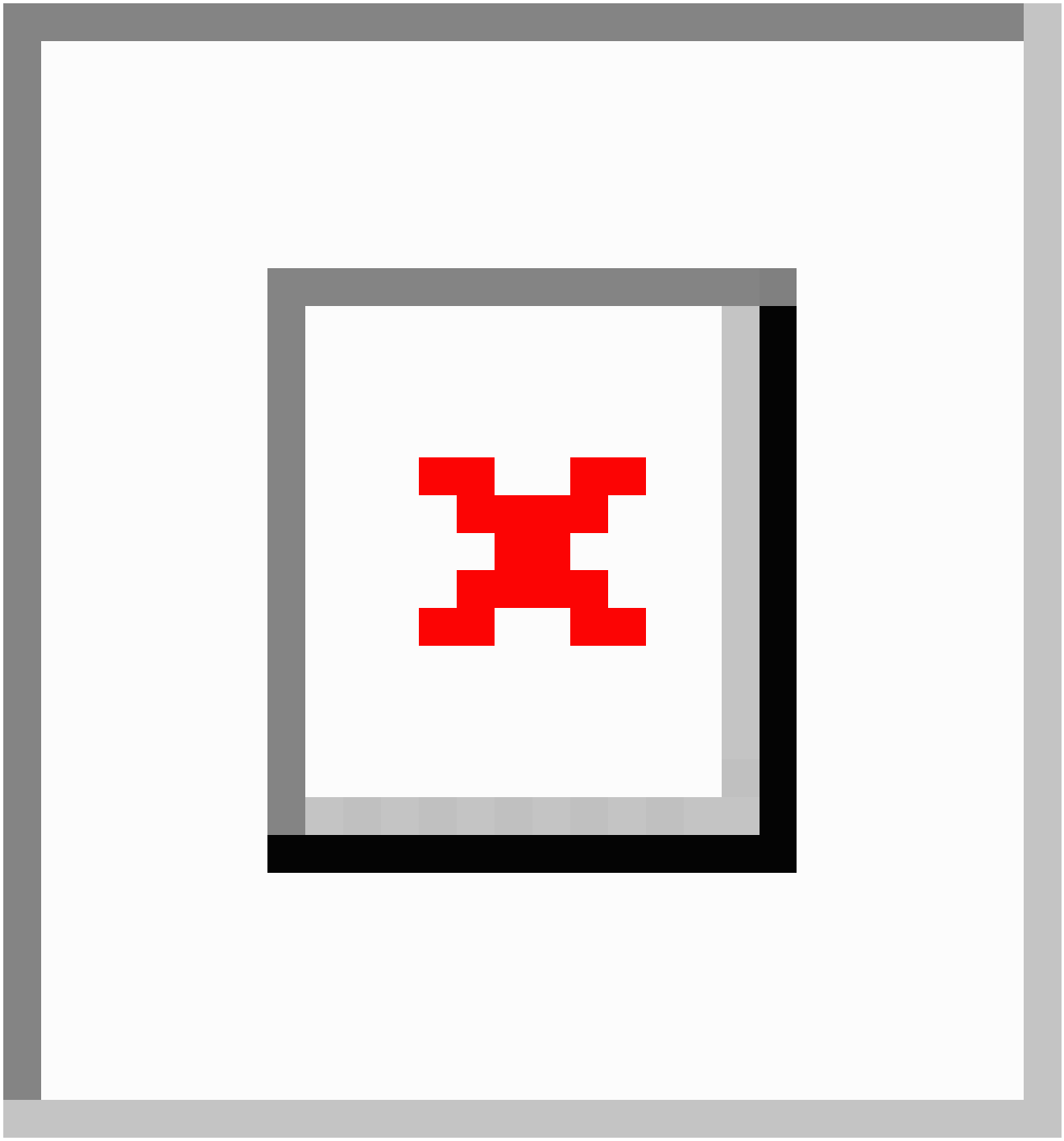
Reviews of current app users (n=6) identified two main themes where improvements in the future versions could be made, that is, improvement in the interactivity and content and improvement of UI.

> *Injury reports should be graphically displayed rather than plain text.*
>
> *User-design can be improved my making the content more interactive.*
>
> *I would like to see more information diet, such as calorie tracker, dietary recommendations before and after the game etc.*

Table 1. Mobile App Rating Scale ratings for the TeamDoc app.

Subject	Engagement	Functionality	Aesthetics	Subjective quality or satisfaction	Behavior change	Overall mean
1	3.6	4	3.3	4	3.7	3.8
2	4.2	5	4.3	2.8	4.2	3.8
3	4.4	4	4.3	3.8	4	4.0
4	2.6	4	3.3	2.8	2.7	2.9
5	3	2.3	2.3	2.8	3.5	2.9
6	2.8	3.5	3.7	2.3	3	2.9
7	2.6	2.8	2.7	4	4	3.6
8	4.4	4.8	4.3	4.8	5	4.8
9	3.4	4.8	3.7	3.3	4	3.8
10	3.8	3.5	3.3	3	4	3.5
11	3.6	4	4	3.3	3.8	3.7
12	2.6	4	3.7	2	3.7	3.0
13	2.8	4.5	2.7	3.8	3.8	3.6
14	2	4	3.3	2.3	3.3	2.9
15	3.6	4.3	3	2.8	3.5	3.3
16	3.2	3.5	3.3	3	3.8	3.4
Mean (SD)	3.3 (0.7)	3.9 (0.7)	3.5 (0.6)	3.2 (0.7)	3.8 (0.5)	3.6 (0.5)

Figure 5. Mean (SD) domain scores for the Mobile App Rating Scale (MARS).



Validation of Design Considerations

First, data confidentiality was assured because none of the test users were able to access the other player's records without validated authentication details. Second, the system's user-friendliness was validated with high mean "functionality" and "ease of use" scores of 3.9 and 4.1, respectively. However, on average, users took 3 minutes to fill out the injury and workload entries, so the time efficiency for data entry did not meet the aim of 2 minutes per entry. Future versions should reduce the number of data entry fields and make the entry more engaging to enhance user compliance and entry efficiency. The third consideration was data storage for future analysis. The testing and validation showed that data output through the server was 100% accurate and could be retrieved instantaneously; this was tested by the investigators by asking the players to enter data on the player app and then cross-verifying the data with the player after downloading the information from the server. The fourth consideration was app's cross-platform availability, and the app was made available for both Android and iOS operating systems. Finally, the UI for injury and workload data entry was to be presented in a way that was easy to input and interpret. This was validated during user-rating functionality and had an overall mean uMARS score of 3.8. Most features of the app scored >3 out of 5 on uMARS domains, showing an overall end-user satisfaction.

Discussion

Principal Findings

It is a common perception that cricket is a nonimpact sport associated with fewer injuries than other sports. However, the literature shows that when injuries are measured in terms of injury rates (ie, per hour of athletic exposure), junior and amateur cricketers have higher rates of injury than professional cricketers [38], and injury rates are comparable to other noncontact or quasi-contact team sports such as soccer, basketball, and tennis [39]. In recent times, the use of mHealth and Web-based technology to monitor an athlete's health is becoming an important component of sports medicine [5]. The development of the TeamDoc app was inspired by this concept and is the first standalone mobile app that can record injuries in cricket through a smartphone without the need for connectivity from parent software on computers. The injury questionnaire was tailored to effectively cater cricket-related injuries, for example, a finger injury while catching the ball.

The benefits of using electronic injury surveillance systems have been extensively documented by Karlsson [6]. It identifies the advantages of using such systems compared with a paper-based system as having less risk of error while transcribing and minimal to no logistic issues. TeamDoc provided the functionality to store the injury data on the server, thereby delivering a paperless solution for tracking injuries and providing ease of access for the team's coaching staff to track injury profiles of the players. Yet, of 8 club teams, 5 (63%) declined to participate when invited to use and evaluate the app. The two main reasons cited by the team coaches or captains for not participating were "time commitment" and "injury not a major issue for the team." To change these perceptions, educating the players and coaches about cricket-related injuries, injury prevention strategies, and the role of technology to prevent future injuries is important.

The overall mean uMARS score was 3.6 out of 5. This is comparable to the mean score for other health-related surveillance apps; for instance, a review of 7 apps for prostate cancer risk calculation had a mean quality score of 3.75 [36]. Similarly, a review of 20 epilepsy self-management apps found a mean quality score of 3.25 [37]. The results indicated that of the 16, only 6 (38%) respondents were current users. The high rate of attrition may be linked to the low scores on satisfaction on the subjective quality of the app (3.14/5) and engagement (3.3/5). The main reasons for the low score on engagement were that half of the users rated the app having no or very basic interactive features and a quarter rated the app as boring. Previous research has shown that providing feedback to users and considering their preferences are important aspects when introducing new injury prevention strategies [28,40]. Therefore, an understanding of perceptions and behaviors when adopting new technology for injury prevention is important. Escoffery et al advocated that for apps targeting behavioral change, developers should work with behavioral scientists to improve the engagement features within the app and encouraged the use of theoretical strategies for behavior change during conceptualization and design phases of app development [37].

The mandatory reporting of injuries and workload by the players may be another reason for high attrition rate from regular use of the app. Medical professionals often use the terms "compliance" and "adherence" to describe the rate at which patients follow their "requests, commands, orders, or rules" [41,42]. These rules and orders can range from following the advice on talking medications and performing investigations to engaging in physical activity, etc. When patients fail to perform the required tasks, they are deemed to have poor compliance or adherence. More recently, "concordance" rather than compliance or adherence has been proposed to be a better alternative when dealing with certain populations [41]. Concordance in medicine is defined as "a state of agreement" between the patient and the physician [43]. Similarly, in sports, concordance can be inferred as a state of agreement between the player and the coach. In medicine, low concordance has been shown to have poor outcomes in patient satisfaction and perception of care [43]. Therefore, before introducing mandatory reporting or asking players to report their injuries and workload, educating them on the benefits of reporting may improve concordance and improve the uptake of the app in the future.

The reason for the low mean subjective quality score (3.18/5) can be attributed to a low score (2.8/5) on "would you pay for this app?" Only 2 users indicated that they would be willing to pay for the app in the future. The app was not designed for commercial use, but user inclination to pay for it may be a surrogate for their perception about the value of the app. "Lack of feedback" to users may be another issue for the low ratings. This is associated with the design constraints of the app, which only allows the coaches to view the data entered by players. Previous research has shown that providing feedback to users is important to maintain their adherence while using Web-based injury surveillance systems [44]. In future versions, it may be useful to allow players a view of their own data so that they can track their activity levels and set up goals. Another feature that may be useful to improve user experience is the inclusion of "gamification" and "social media plugins." Gamification features may include features such as players setting up weekly targets for their activity and getting rewards if they achieve their target. Social media and sports news plugins may improve user experience and encourage regular use of the app.

There were multiple limitations within the current version of the app. For example, there was no mechanism for alerting fast bowlers or coaches if a player exceeded age-related bowling workload recommendations nor mechanisms for delivering reminder alerts if players forgot to key in their workloads. However, important design considerations, such as security and confidentiality of data, were ensured by designing the app on PHP, which provided protection against malicious attacks by hackers, and by designing separate app interfaces for players and coaches. Another consideration during the development was cross-platform connectivity with other eHealth platforms on the client side of the app to simplify porting it to diverse operating systems.

Conclusions

The use of mHealth in sports medicine can assist in wireless data capture that may be used to make informed, evidence-based

decisions. TeamDoc follows this concept by allowing the coaching staff and the players to record data on injury and workload on the go. The app may assist coaches to make informed decisions in real time during match conditions. TeamDoc is available for free, which means that community-based clubs can access and use it. This study provides a guide to the architecture and framework for developing an injury surveillance and workload monitoring mobile app, which can be applied to design similar systems for other sports. The results from the user survey indicate that future versions of the app should have improved UI and interactivity features.

Practical Implications

The following are the practical implications of the study:

- The ease to use the app "on the go" may mean better reporting of injuries at the junior level.
- The app can act as a monitoring tool for the coaching staff to adjust individual training loads for players, which may assist in reducing injuries.
- The methods of development used for this app can be applied by researchers and developers to introduce similar apps for other adolescent team sports.
- In the future, surveillance apps should focus on improved UI and interactivity to attract and retain users.

Acknowledgments

The authors are grateful to Dr René Ferdinands, Mr Gary Whittaker, and players from the University of Sydney Cricket team for their assistance in the testing of the app. We also thank James Alexander and Shane Herft from the Incubate Hub University of Sydney Union for their assistance in the development and promotion of the app.

Conflicts of Interest

None declared.

Multimedia Appendix 1

User Interfaces for the TeamDoc player App.

[PDF File (Adobe PDF File), 1MB - mhealth_v7i1e10978_app1.pdf]

Multimedia Appendix 2

A video showing features of the player app.

[MP4 File (MP4 Video), 12MB - mhealth_v7i1e10978_app2.mp4]

Multimedia Appendix 3

User version of the Mobile App Rating Scale (uMARS) survey tool used for evaluation.

[PDF File (Adobe PDF File), 288KB - mhealth_v7i1e10978_app3.pdf]

References

1. Speed CA, Roberts WO. Innovation in high-performance sports medicine. Br J Sports Med 2011 Sep;45(12):949-951. [doi: 10.1136/bjsm.2010.075374] [Medline: 20870810]
2. Qaisar S, Imtiaz S, Glazier P, Farooq F, Jamal A, Iqbal W, et al. A Method for Cricket Bowling Action ClassificationAnalysis Using a System of Inertial Sensors. In: Murgante B, Misra S, Carlini M, Torre C, Nguyen HQ, Taniar D, et al, editors. Dynamic Context for Document Search and Recovery. Berlin Heidelberg: Springer; Jun 2013:396-412.
3. Petersen CJ, Pyne DB, Portus MR, Dawson BT. Comparison of player movement patterns between 1-day and test cricket. J Strength Cond Res 2011 May;25(5):1368-1373. [doi: 10.1519/JSC.0b013e3181da7899] [Medline: 21273907]
4. Free C, Phillips G, Felix L, Galli L, Patel V, Edwards P. The effectiveness of M-health technologies for improving health and health services: a systematic review protocol. BMC Res Notes 2010 Oct 06;3:250 [FREE Full text] [doi: 10.1186/1756-0500-3-250] [Medline: 20925916]
5. Verhagen E, Bolling C. Protecting the health of the @hlete: how online technology may aid our common goal to prevent injury and illness in sport. Br J Sports Med 2015 Sep;49(18):1174-1178. [doi: 10.1136/bjsports-2014-094322] [Medline: 25614537]
6. Karlsson D. DiVA Portal.: Department of Computer and Information Science, Linköping University; 2013 Aug 25. Electronic Data Capture for Injury and Illness Surveillance: A usability study (Master thesis) URL: http://www.diva-portal.org/smash/get/diva2:681327/FULLTEXT01.pdf2013 [accessed 2018-12-03] [WebCite Cache ID 74NuCi5KR]

7. Clarsen B, Myklebust G, Bahr R. Development and validation of a new method for the registration of overuse injuries in sports injury epidemiology: the Oslo Sports Trauma Research Centre (OSTRC) overuse injury questionnaire. Br J Sports Med 2013 May;47(8):495-502. [doi: 10.1136/bjsports-2012-091524] [Medline: 23038786]
8. Nilstad A, Bahr R, Andersen T. Text messaging as a new method for injury registration in sports: a methodological study in elite female football. Scand J Med Sci Sports 2014 Feb;24(1):243-249. [doi: 10.1111/j.1600-0838.2012.01471.x] [Medline: 22537065]
9. Barwick H. Cricket Australia sends technology in to bat. 2012 Dec 06. URL: https://www.computerworld.com.au/article/443901/cricket_australia_sends_technology_bat/ [accessed 2018-12-03] [WebCite Cache ID 74NuXZgTd]
10. Orchard J, James T, Kountouris A, Blanch P, Beakley D, Orchard J. Injury Report 2014: Cricket Australia. John Orchard's Sports Injury Site. 2014 Jan 04. URL: https://sma.org.au/sma-site-content/uploads/2017/08/Sport-Health-Volume-32-Issue-4-2014-15.pdf [WebCite Cache ID 74NuXZgTd]
11. Stretch RA. Junior cricketers are not a smaller version of adult cricketers: A 5-year investigation of injuries in elite junior cricketers. South African Journal of Sports Medicine 2014;26(4):123.
12. Finch C, White P, Dennis R, Twomey D, Hayen A. Fielders and batters are injured too: a prospective cohort study of injuries in junior club cricket. J Sci Med Sport 2010 Sep;13(5):489-495. [doi: 10.1016/j.jsams.2009.10.489] [Medline: 20031485]
13. Stretch RA. The seasonal incidencenature of injuries in schoolboy cricketers. S Afr Med J 1995;85(11):1182-1184.
14. Stretch R. A five year investigation into the incidence and nature of cricket injuries in elite South African schoolboy cricketers. Br J Sports Med 2014 Mar 11;48(7):663.1-66663. [doi: 10.1136/bjsports-2014-093494.275]
15. Dennis R, Finch C, Farhart P. Is bowling workload a risk factor for injury to Australian junior cricket fast bowlers? Br J Sports Med 2005 Nov;39(11):843-6; discussion 843 [FREE Full text] [doi: 10.1136/bjsm.2005.018515] [Medline: 16244195]
16. Noorbhai M, Essack F, Thwala S, Ellapen T, Van Heerden J. Prevalence of cricket-related musculoskeletal pain among adolescent cricketers in KwaZulu-Natal. S Afr J Sports Med 2012 Mar 30;24(1). [doi: 10.17159/2078-516X/2012/v24i1a352]
17. Soomro N, Redrup D, Evens C, Strasiotto L, Singh S, Lyle D, et al. Injury rate and patterns of Sydney grade cricketers: a prospective study of injuries in 408 cricketers. Postgrad Med J 2018 Aug;94(1114):425-431. [doi: 10.1136/postgradmedj-2018-135861] [Medline: 30049730]
18. Orchard J, Ranson C, Olivier B, Dhillon M, Gray J, Langley B, et al. International consensus statement on injury surveillance in cricket: a 2016 update. Br J Sports Med 2016 Oct;50(20):1245-1251. [doi: 10.1136/bjsports-2016-096125] [Medline: 27281775]
19. Caine D, Maffulli N, Caine C. Epidemiology of injury in child and adolescent sports: injury rates, risk factors, and prevention. Clin Sports Med 2008 Jan;27(1):19-50, vii. [doi: 10.1016/j.csm.2007.10.008] [Medline: 18206567]
20. Schwellnus M, Soligard T, Alonso J, Bahr R, Clarsen B, Dijkstra H, et al. How much is too much? (Part 2) International Olympic Committee consensus statement on load in sport and risk of illness. Br J Sports Med 2016 Sep;50(17):1043-1052 [FREE Full text] [doi: 10.1136/bjsports-2016-096572] [Medline: 27535991]
21. Cricket Australia. Manning River District Cricket Association. Well Played: Australian Cricket's Playing Policies and Guidelines URL: http://mrdca.nsw.cricket.com.au/files/2955/files/Australian_Crickets_Playing_Policy_and_Guidelines.pdf [accessed 2018-12-03] [WebCite Cache ID 74NxXAqBy]
22. England Wales Cricket Board. The Liverpool and District Cricket Competition. 2003 Feb. ECB Fast Bowling Match Directives URL: http://www.ecb.co.uk/sites/default/files/3485_nfc_fastbowling_2015_p329-330_lr.pdf [accessed 2018-12-03] [WebCite Cache ID 74Nxips5p]
23. Soomro N, Hackett D, Freeston J, Blanch P, Kountouris A, Dipnall J, et al. How do Australian coaches train fast bowlers? A survey on physical conditioning and workload management practices for training fast bowlers. International Journal of Sports Science & Coaching 2018 Jul 19;13(5):761-770 [FREE Full text] [doi: 10.1177/1747954118790128] [Medline: 25877007]
24. Saw A, Main L, Gastin P. Monitoring the athlete training response: subjective self-reported measures trump commonly used objective measures: a systematic review. Br J Sports Med 2016 Mar;50(5):281-291 [FREE Full text] [doi: 10.1136/bjsports-2015-094758] [Medline: 26423706]
25. Halson S, Bridge M, Meeusen R, Busschaert B, Gleeson M, Jones D, et al. Time course of performance changes and fatigue markers during intensified training in trained cyclists. J Appl Physiol (1985) 2002 Sep;93(3):947-956 [FREE Full text] [doi: 10.1152/japplphysiol.01164.2001] [Medline: 12183490]
26. Chennaoui M, Desgorces F, Drogou C, Boudjemaa B, Tomaszewski A, Depiesse F, et al. Effects of Ramadan fasting on physical performance and metabolic, hormonal, and inflammatory parameters in middle-distance runners. Appl Physiol Nutr Metab 2009 Aug;34(4):587-594. [doi: 10.1139/H09-014] [Medline: 19767792]
27. Borresen J, Lambert M. Quantifying training load: a comparison of subjective and objective methods. Int J Sports Physiol Perform 2008 Mar;3(1):16-30. [Medline: 19193951]
28. Donaldson A, Lloyd D, Gabbe B, Cook J, Young W, White P, et al. Scientific evidence is just the starting point: A generalizable process for developing sports injury prevention interventions. J Sport Health Sci 2016 Sep;5(3):334-341 [FREE Full text] [doi: 10.1016/j.jshs.2016.08.003] [Medline: 30356506]

29. Andersson T, Johansson E. DiVA Portal. 2018 Jan 11. A closer look and comparison of cross-platform development environment for smartphones (Dissertation) URL: http://mdh.diva-portal.org/smash/record.jsf?pid=diva2%3A726850&dswid=-4332 [accessed 2018-12-03] [WebCite Cache ID 74NyF38Kl]
30. Maduka O, Tobin-West C. Barriers to HIV treatment adherence: Perspectives from the nonadherent at a treatment center in South-South, Nigeria. Ann Afr Med 2015;14(3):159-160 [FREE Full text] [doi: 10.4103/1596-3519.149879] [Medline: 26021398]
31. Reenskaug T, Coplien J. Artima. 2009 Mar 20. The DCI Architecture: A New Vision of Object-Oriented Programming URL: https://www.artima.com/forums/flat.jsp?forum=226&thread=253008
32. Sports Medicine Australia. 2015. Sports Injury Reporting Form URL: http://sma.org.au/sma-site-content/uploads/2018/10/Injury-Reporting-Form_General.pdf [accessed 2018-12-03] [WebCite Cache ID 74NyzKaAG]
33. Finch C, Valuri G, Ozanne-Smith J. Injury surveillance during medical coverage of sporting events--development and testing of a standardised data collection form. J Sci Med Sport 1999 Mar;2(1):42-56. [Medline: 10331475]
34. Stoyanov S, Hides L, Kavanagh D, Wilson H. Development and Validation of the User Version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth 2016 Jun 10;4(2):e72 [FREE Full text] [doi: 10.2196/mhealth.5849] [Medline: 27287964]
35. Stoyanov S, Hides L, Kavanagh D, 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]
36. Adam A, Hellig J, Perera M, Bolton D, Lawrentschuk N. 'Prostate Cancer Risk Calculator' mobile applications (Apps): a systematic review and scoring using the validated user version of the Mobile Application Rating Scale (uMARS). World J Urol 2018 Apr;36(4):565-573. [doi: 10.1007/s00345-017-2150-1] [Medline: 29222595]
37. Escoffery C, McGee R, Bidwell J, Sims C, Thropp E, Frazier C, et al. A review of mobile apps for epilepsy self-management. Epilepsy Behav 2018 Apr;81:62-69. [doi: 10.1016/j.yebeh.2017.12.010] [Medline: 29494935]
38. Soomro N, Chua N, Freeston J, Ferdinands R, Sanders R. Cluster randomised control trial for cricket injury prevention programme (CIPP): a protocol paper. Inj Prev 2017 Sep 28. [doi: 10.1136/injuryprev-2017-042518] [Medline: 28971855]
39. Soomro N, Strasiotto L, Sawdagar T, Lyle D, Mills D, Ferdinands R, et al. Cricket Injury Epidemiology in the Twenty-First Century: What is the Burden? Sports Med 2018 Oct;48(10):2301-2316. [doi: 10.1007/s40279-018-0960-y] [Medline: 30019111]
40. Finch C, Bahr R, Drezner J, Dvorak J, Engebretsen L, Hewett T, et al. Towards the reduction of injury and illness in athletes: defining our research priorities. Br J Sports Med 2017 Aug;51(16):1178-1182. [doi: 10.1136/bjsports-2016-097042] [Medline: 28003237]
41. Benson J. Concordance--An alternative term to 'compliance' in the Aboriginal population. Aust Fam Physician 2005 Oct;34(10):831-835 [FREE Full text] [Medline: 16217567]
42. Smith V, DeVellis B, Kalet A, Roberts J, DeVellis R. Encouraging patient adherence: primary care physicians' use of verbal compliance-gaining strategies in medical interviews. Patient Educ Couns 2005 Apr;57(1):62-76. [doi: 10.1016/j.pec.2004.03.017] [Medline: 15797154]
43. Thornton R, Powe N, Roter D, Cooper L. Patient-physician social concordance, medical visit communication and patients' perceptions of health care quality. Patient Educ Couns 2011 Dec;85(3):e201-e208 [FREE Full text] [doi: 10.1016/j.pec.2011.07.015] [Medline: 21840150]
44. Barboza S, Bolling C, Nauta J, van Mechelen W, Verhagen E. Acceptability and perceptions of end-users towards an online sports-health surveillance system. BMJ Open Sport Exerc Med 2017;3(1):e000275 [FREE Full text] [doi: 10.1136/bmjsem-2017-000275] [Medline: 29071115]

Abbreviations

CA: Cricket Australia
uMARS: user version of the Mobile App Rating Scale
mHealth: mobile health
MVC: Model-View-Controller
PHP: Hypertext Preprocessor
QML: Qt Modeling Language
SQL: structured query language
UI: user interface

Edited by G Eysenbach; submitted 11.05.18; peer-reviewed by T Esgin, C Bolling, J Thomas; comments to author 03.09.18; revised version received 07.09.18; accepted 25.09.18; published 22.01.19.

Please cite as:
Soomro N, Chhaya M, Soomro M, Asif N, Saurman E, Lyle D, Sanders R
Design, Development, and Evaluation of an Injury Surveillance App for Cricket: Protocol and Qualitative Study
JMIR Mhealth Uhealth 2019;7(1):e10978
URL: http://mhealth.jmir.org/2019/1/e10978/
doi:10.2196/10978
PMID:30668516

Original Paper

Public Views on Using Mobile Phone Call Detail Records in Health Research: Qualitative Study

Kerina Helen Jones[1], BSc (Hons), PhD; Helen Daniels[1], BSc (Hons), PhD; Sharon Heys[1], BSc (Hons); David Vincent Ford[1], MBA

Population Data Science, School of Medicine, Swansea University, Swansea, United Kingdom

Corresponding Author:
Kerina Helen Jones, BSc (Hons), PhD
Population Data Science
School of Medicine
Swansea University
Swansea,
United Kingdom
Phone: 44 1792602764
Email: k.h.jones@swansea.ac.uk

Abstract

Background: Mobile phone call detail records (CDRs) are increasingly being used in health research. The location element in CDRs is used in various health geographic studies, for example, to track population movement and infectious disease transmission. Vast volumes of CDRs are held by multinational organizations, which may make them available for research under various data governance regimes. However, there is an identified lack of public engagement on using CDRs for health research to contribute to an ethically founded framework.

Objective: This study aimed to explore public views on the use of call detail records in health research.

Methods: Views on using CDRs in health research were gained via a series of three public workshops (N=61) informed by a pilot workshop of 25 people. The workshops included an initial questionnaire to gauge participants' prior views, discussion on health research using CDRs, and a final questionnaire to record workshop outcome views. The resulting data were analyzed for frequencies and emerging themes.

Results: At the outset, most participants (66%, 40/61) knew that location data were collected by operators, but only 3% (2/61) knew they were being used for health research. Initially, the majority of the participants (62%, 38/61) was content for their anonymous CDRs to be used, and this increased (80%, 49/61) after the discussion explained that safeguards were in place. Participants highlighted that terms and conditions should be clearer, as should information to phone users on data collection, privacy safeguards, sharing, and uses in research.

Conclusions: This is the first known study exploring public views of using mobile phone CDRs in health research. It revealed a lack of knowledge among the public on uses of CDRs and indicated that people are generally amenable to the use of anonymized data for research, but they want to be properly informed and safeguarded. We recommend that public views be incorporated into an ethically founded framework for the use of CDRs in health research to promote awareness and social acceptability in data use.

(JMIR Mhealth Uhealth 2019;7(1):e11730) doi:10.2196/11730

KEYWORDS

qualitative research; mobile phone use

Introduction

Background

Mobile phone penetration is constantly rising and is predicted to exceed 5 billion users by 2019; the number of mobile connections already exceeds the world population at over 8 billion [1]. Call detail records (CDRs) are collected passively each time a mobile phone user connects to a mobile network, by either voice call or short message service (SMS) text message. The record generated includes the starting time of the call (or SMS text message), its duration, the caller's and receiver's phone numbers, and the locations of the activated

towers. Locations can be made more precise via tower triangulation and Wi-Fi connections [2]. Billions of CDRs are collected by mobile network operators (MNOs) such as Orange, O_2, and EE: they are essential for service operation and are used for billing, monitoring data usage, and targeting customers according to their cell phone use [2]. MNOs may make subsets of CDRs available for research under various data governance regimes [3], enabling the location element in CDRs to be used in a variety of health geographic studies, such as tracking population movement and infectious disease transmission, as shown in a recent review [4].

CDRs are not health data *per se* but can be used alone or in conjunction with other datasets for health research. Just using CDRs alone, researchers have been able to model how to disseminate emergency information during an epidemic in the Ivory Coast [5], work on ways to arrest contagious diseases at an early stage in Belgium [6], and map human mobility after a natural disaster to inform humanitarian response in Nepal [7]. Value is added when CDRs are combined with other datasets for research. In general, these are not linked at the individual level but are overlaid in aggregated form (or at least anonymized) so that individual identities are not exposed. Some examples included in our review are given as illustrations of this kind of study. CDRs with population data and incidence data of Dengue fever were used to predict the timing and spatial extent of disease outbreak in Pakistan [8]. CDRs were used with data on confirmed malaria cases in Namibia to identify areas where malaria surveillance should be increased [9]. CDRs have been used to model likely malaria importation into Zanzibar by combining them with ferry traffic data and malaria surveys [10].

Public perceptions on the use of person-based health and administrative data for research have been, and still are, the subject of extensive work. We define health data as information relating to the health status of individuals, typically, as collected in the course of care provision. We use the term administrative data to include broader public service information such as records on education, housing, and social services. Understanding Patient Data is a prime example of an initiative to support public engagement in the reuse of health data for research [11], along with the Data Saves Lives campaign [12]. In 2014, the Economic and Social Research Council (ESRC) and the Office for National Statistics commissioned a report to explore public understanding and views of administrative data and data linkage [13]. During 7 workshops held with members of the public, it came to light that the public had only limited knowledge of data use in social research, despite acknowledging the importance of data in digital societies. Participants questioned why such research had to take place; their main concerns centered on the risk of reidentification and the level and limits of security protecting the databases where information was held.

The Welcome trust published a report in 2015 on public opinion about the use of health service data for research by commercial companies. Overall, 6 workshops were conducted and researchers found in general that members of the public felt that their identifiable health data should not be shared without their explicit consent and that the risk of reidentification from anonymized datasets was a concern. Participants expressed feelings of mistrust over the use of their health data by commercial companies and questioned the possible motivation for their use [14]. As a view from the opposite perspective, that is, the use of commercial data for research by the public sector, the ESRC commissioned a study in 2015 of public views on private sector data being used for social research. A series of 3 workshops was held, and the findings showed broad support for the reuse of data. The dialogue with the public alleviated many of their concerns about privacy and security and about the role commercial companies can play in research for public benefit [15].

Although there have been surveys of public views on other aspects of mobile phone usage [16], a 2015 seminal work observed that there was no known literature on public perceptions of using CDRs for health research [2]. Despite updated searches, and a review of CDRs in health research [4], no published work on public perceptions of using CDRs for health research was identified. There are studies on public views of using mobile phone apps for monitoring health conditions, but these are distinct and outside our area of interest.

Objective

Due to the importance of public engagement on the reuse of data for research and the dearth of published work, the aim of this study was to gain public views via a series of workshops and to use the information gained to contribute to an ethically founded framework for the socially acceptable use of CDRs for health research.

Methods

Ethical Approval and Study Documentation

Public knowledge on the extent of passive and active data collection via mobile phones, the uses the data are put to, and views on acceptability were gained via a series of 3 public workshops. Ethical approval for research with public participants was obtained from the Swansea University Medical School Research Ethics and Governance Committee. Participants were invited to attend a workshop and did so voluntarily. Information sheets and consent forms were provided, and these set out what participants could expect while taking part in the study and assured them that no identifiable information would be collected, and any comments made would not be attributed to them.

Pilot Workshop

To inform the structure and content of the public workshops, a pilot workshop was held on March 7, 2017, for members of the Population Data Science department, Swansea University Medical School, and 25 participants attended voluntarily, giving their consent to take part. This department was chosen because it includes experienced data analysts, computer scientists, data managers and data-facing researchers, as well as mobile phone users. It was anticipated that their knowledge would be valuable and insightful in shaping the workshops for more general groups. All the workshops (pilot and public series) were informed by a literature review of studies using mobile phone data for health research [4], and participants were provided with information on examples with their respective benefits and limitations. The

types of data that are collected by mobile phone operators and the uses that these data are put to (including health-related research) were introduced. Examples of research studies using CDRs alone and in conjunction with other datasets were given, as described in the introduction. This was intended to enable the participants to consider the types of research that could be conducted, along with the respective risks and benefits.

Participants were introduced to the purpose of the pilot workshop, and to gauge prior knowledge, they were asked for a show of hands in answer to the following questions:

1. Did you know that mobile phone network operators routinely collect data on your location? (The distinction between MNOs and phone handset manufacturers is made for clarity).
2. Did you know that mobile phone network operators and third parties are using the data for research?
3. Have you read the terms and conditions of your mobile phone network operator?

This was followed by a short presentation with examples of research studies using mobile phone CDRs. For example, characterizing the patterns of malaria transmission in Namibia [9], the location of hospitals in relation to travel time following a myocardial infarction or stroke for public health planning in Senegal [17], and monitoring exposure to air pollution in Belgium [18].

Participants were then asked to consider the following in an informal discussion:

1. How aware do you feel about data being collected via mobile phones and likely potential of data for health research?
2. How do you see the pros and cons of using mobile phone data in health research?
3. Any other comments or suggestions regarding workshops with the public?

Finally, the participants were asked to write down their responses to the following:

1. What data do you believe are collected by your phone?
2. How does the operator use the data or make them available to others?
3. What in your view are the data governance issues and risks?
4. How do you feel about these data collection and use?
5. What do you think should be included in the terms and conditions?
6. Which data users or sectors are more acceptable or less acceptable?
7. What do you think a general public group would know?
8. What do you believe the general public would think?

Public Workshops

The findings of the pilot workshop were used to shape the public workshops. The first of these took place on June 14, 2017, with a convenient workforce group as part of a Swansea University seminar series. Participants were employees of the University from various departments and disciplines; they included academics, researchers, students, and administrators. Overall, 21 people attended the workshop (5 men and 16 women). The second workshop took place on June 28, 2017, with the Consumer Panel for Data Linkage who provide a public perspective on information governance issues in connection with big data and data linkage research in Swansea University–based data initiatives [19]. Overall, 14 people took part in the workshop (7 men and 7 women). The final workshop was held at Pembrokeshire College of Further Education on September 6, 2017, with an adult group attending level 3 health and social care. In total, 26 people (6 men and 20 women) attended the workshop on this occasion. The total number of people who attended the public workshops was 61, and this number is used as the denominator in presenting the results. The age breakdown is shown in Table 1. As the data were collected in age bands, mean age and SD are not shown.

To gauge the representativeness of the sample compared with the UK population, it was compared with the 2011 census figures [20]. The age bands are slightly different in the census, but are close enough to provide an indicative measure. Moreover, we used 18 to 25 years range, as we did not include anyone less than 18 years of age, whereas the census category is 15 to 24 years. The census covers a broader age range than our sample, and so the percentages in the age bands have been adjusted to mirror our range being 100%. Having done this, we have for the census the following: 15 to 24 years: 17%, 25 to 34 years: 18%, 35 to 44 years: 18%, 45 to 54 years: 18%, 55 to 64 years: 17%, and 65 to 74 years: 12%. This indicates that our sample is heavier in the younger (18-45 years) age bands and lighter in the older (46-75 years) age bands.

All the public workshops followed the same format. At the beginning of the workshop, participants were asked to complete a questionnaire (Multimedia Appendix 1) based on their prior knowledge. This questionnaire covered mobile phone use, knowledge of the collection of mobile phone data, their use in research, and the participants' willingness for their mobile phone data to be used for health research purposes.

Table 1. Numbers of public workshop participants in age bands.

Workshop	Age band (years), n (%)						Participants, n (%)
	18-25	26-35	36-45	46-55	56-65	66-75	
1	1 (5)	7 (33)	9 (43)	4 (19)	0 (0)	0 (0)	21 (100)
2	0 (0)	3 (21)	3 (21)	1 (7)	2 (14)	5 (36)	14 (100)
3	14 (52)	8 (31)	3 (12)	0 (0)	1 (4)	0 (0)	26 (100)
All	15 (25)	18 (30)	15 (25)	5 (8)	3 (5)	5 (8)	61 (100)

They were also asked if they had read the terms and conditions of their mobile phone operator, as a show-of-hands. Participants were presented with examples of the following:

- Mobile phone contract terms and conditions highlighting the types of data that are collected and how they are used
- Other types of day-to-day Big Data collection (eg, supermarket loyalty cards)
- Commercial uses of mobile phone data (retail, council planning, and transport models)
- Health research using phone data (as for the pilot workshop)

A general discussion followed using the following questions as prompts:

- What data do you believe are collected by your phone?
- What in your view are the data governance issues and risks?
- How do you feel about this data collection and use?
- What do you think should be included in terms and conditions?
- Which users or sectors are more acceptable or less acceptable?

Further explorations focused on how the public should or could be involved and informed about research using mobile phone data considering the formats in which the data are collected and used, the practicalities of seeking meaningful consent, the acceptability of agreement via MNO terms and conditions, and the implications for individuals and society. Finally, participants were asked to complete a second questionnaire (Multimedia Appendix 2) covering some of the same topics as initial to assess if finding out more via the workshop had altered their opinions on the collection and use of mobile phone data for health research. Questionnaire responses were collected in an anonymous format, but a unique number was applied to each set so that before and after responses for each individual could be compared. Quantitative responses were analyzed as frequencies in IBM SPSS (v.22), and free-text qualitative responses were analyzed thematically by manual assessment and comparison between members of the research team for consensus on theme identification and data convergence.

Results

Pilot Workshop

Beginning with the initial 3 show-of-hands questions on prior knowledge, all 25 participants knew that MNOs collect data about their customers, and 24% (6/25) knew that MNOs and third parties are using the data for research, and no one had read their MNO terms and conditions fully, if at all.

The informal discussion based on questions 4 to 6 yielded interesting points. Participants felt reasonably aware of the types of data collected by MNOs. Examples they gave included the location of the user when making phone calls, which mobile apps were on users' phones, and the data that users had downloaded onto their phones. The general consensus was that there is potential for these type of data to benefit health research, particularly in population health, and participants felt that using their data for this purpose was acceptable as long as their data were anonymized.

However, some concerns were also expressed such as MNOs might sell location data to (potential) employers or to companies such as insurers for profit. Moreover, there was some concern about the risk of disclosure for individuals who live in remote areas. Participants were in consensus in believing that young people would be more likely to be accepting the use of their mobile phone data in health research because of their high usage of mobile phones. Some felt that they would need certain questions answered before being able to decide whether or not they would be happy with this, for example, how secure the identifiable data are before anonymization, whether real-time data are used, and the levels of aggregation applied.

The written responses (questions 7-14) provided the following collated information. Participants listed the data types they believed to be collected by mobile phones as call data: date, time, start or finish, who called or SMS text messaged; data usage; demographic data; location; financial data; online purchasing history; internet search history; app data (eg, about health); and emails (question 7). Participants believed that mobile phone operators may use data to inform advertising strategies or for improving network and data coverage and services. Some also thought that they may make data available by sharing or selling them to third parties such as insurance companies. It was also noted that data would likely be shared with the government if requested (question 8). In terms of data governance issues, the main points were whether data were anonymized and aggregated to a sufficient standard and whether informed consent had occurred for the identifiable data to be collected and used in the first instance. Risks identified included disclosure and data misuse and the increased possibility of reidentification from the use of multiple datasets (question 9).

In total, 11 participants felt happy for their mobile phone data to be used for research purposes as long as they were not identifiable. Several stipulated that they would prefer this to contribute to an improvement for the general population, not just be used for commercial gain. Others wanted to be able to give fully informed consent as way of guaranteeing that they knew exactly what data were being collected, for what purpose, and to be used by whom. Some felt uncomfortable about their data being used in this way and had continuing questions over the identifiability of the data and the corresponding risks of fraud and malicious use. Some participants were concerned that individual-level data (rather than aggregated) could be released to third parties (question 10).

Participants were in agreement that MNO terms and conditions should be written in basic language and be more concise, but also more explicit, and should include information about how, and with whom, the data would be shared. Opting out to certain uses should be available, rather than an *all or nothing* approach (question 11). Participants suggested that use by all sectors could be made possible, but phone users should have the option of opting out of some or all of them. Participants were not in favor of their data being sold to large commercial companies; those that would use their data to improve health were felt to be more acceptable (question 12).

The final 2 questions invited the group to provide their opinions on what the general public would know and think (questions

13 and 14). These questions were asked because the participants were suspected to be more tech-savvy than the public at large by reason of their data-focused work roles, and their views would help in guiding the public workshops. Potential knowledge among the public was hypothesized to be variable and likely to vary with age. Participants believed that a public group would know less than their group because of their experience with big data and analysis. Several participants suggested that members of the public would be less likely to understand what types of data were being collected and may confuse them with data collected via mobile phone apps. It was thought that some people might be alarmed to find out about data being collected via their mobile phones. Participants considered young people to be more familiar with their mobile device and, therefore, more likely to be aware of and accepting the amount of data that is collected.

Even among this group of people working in a data-intensive field of work, no one had engaged with the information provided in the terms and conditions. The group was reasonably aware of the types of data being collected by MNOs and believed the use of CDR data was beneficial, provided that safeguards were in place. A number of concerns were raised, including consent to collect identifiable data in the first place, the effectiveness of anonymization applied, the security of data systems, the potential for data misuse, and the possibility of data being sold to insurers or employers. The discussions and points raised gave us an insight into the types of issues that may arise in the public workshops.

There were some key learning points from the pilot workshop that helped shape the series of public workshops. First, to avoid any confusion, the types of data that are collected by MNOs were described in sufficient detail so that they could be clearly distinguished from mobile app data. Second, to put the topic in context, other types of big data that are collected day-to-day were presented, for example, the data that are collected by supermarkets via their loyalty cards. Finally, it was noted that although pilot participants admitted to not fully reading the terms and conditions of their mobile phone use, all believed that the terms and conditions needed to be changed in some way. Therefore, it was decided that examples of MNO terms and conditions should be presented during the public workshops, so people could give a balanced and informed opinion on whether they thought the current content was sufficient to constitute informed consent.

Public Workshops

All the public workshops followed the same format, as described above. To gain an understanding of the participants' level of familiarity with mobile phones, they were asked some questions about their mobile phone use (Multimedia Appendix 1, questions 3-6). We chose to mention smartphones (a smartphone is taken as a more advanced mobile phone that functions as a small personal computer with full internet access, social media connectivity, apps, and games) and *standard* mobile phones in the questionnaire for clarity and to ensure that both types of mobile phones were included but noting that our particular interest was in CDRs, which are collected by all mobile phones.

Questions 3 to 6 were not intended for detailed analysis and so a summary is given here. The majority of public participants (92%, 56/61) owned a mobile phone at the time of the workshop (smartphone or standard mobile phone), with the other 5 making use of a family member's mobile as needed. All but 2 specified that they used a mobile several times a day, 1 used it several times a week, and for the remaining participants, mobile phone use was seldom. All participants said they used a mobile for making phone calls and SMS text messaging and for accessing the internet and emailing. Those with a smartphone noted that they used it for apps, watching videos, playing games, finding places (global positioning system), and several used their phones for reading and making diary appointments. None of the participants had read the terms and conditions.

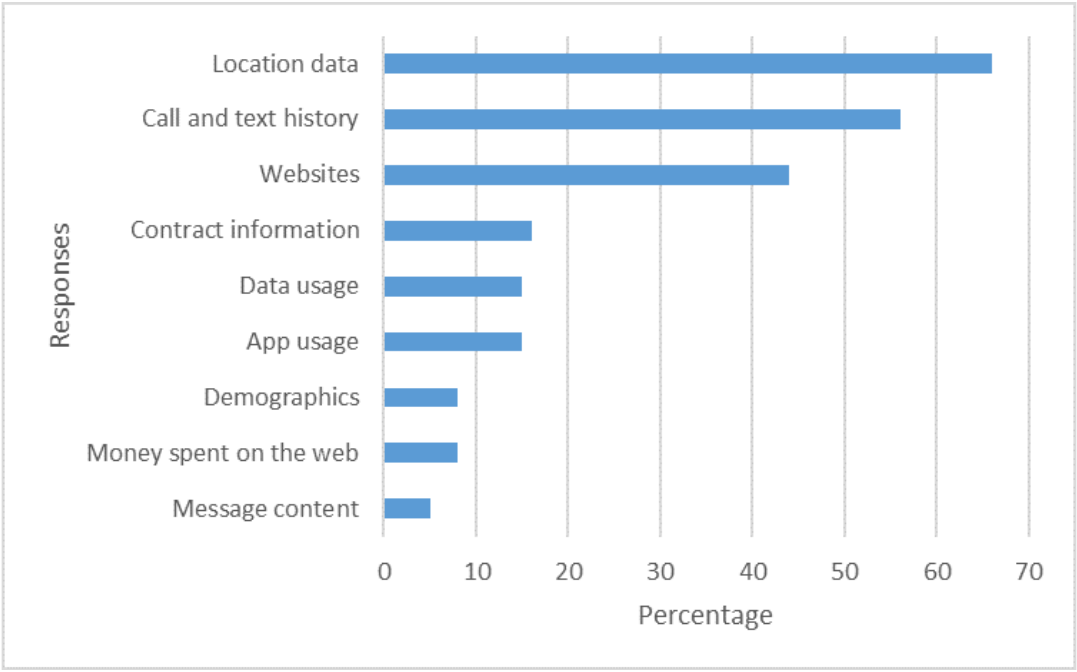
On the basis of the responses to questionnaire 1 (Multimedia Appendix 1), at the outset of the workshops, a majority (61%, 37/61) of the participants knew mobile phone operators collect data about the mobile phone user, 11 (18%, 11/61) were not aware of this, and 13 (21%, 13/61) were unsure. When asked to list which types of data they thought that mobile phone operators were collecting (Figure 1), without prompt, the most popular response was mobile phone user location (66%, 40/61).

In terms of how they believe data are used, a variety of uses were listed in the responses, with market research and targeted advertising being the most frequent (43%, 26/61). However, only 2 (3%, 2/61) participants were already aware that the data were being used in health research (Figure 2).

Despite having little knowledge of health research using mobile phone data, most of the participants (62%, 38/61) agreed that they would be happy for their mobile phone data to be used for this purpose. In total, 4 (7%, 4/61) recorded that they would not be happy for this to occur, and 19 (31%, 19/61) were unsure. Participants were also asked to comment on their response to this question. Overall, 6 people stated that they were happy for their data to be used in this manner as long as their data were safeguarded and anonymized, and 4 participants stated that their data could be used in this manner only if their consent was sought. Other participants explained that they would also be happy with this, on the condition that they were given more information. For example, participants felt that they would want to know what kind of research their data were being used for; one stipulated that they would want to be given the ability to exclude him or her on a project-by-project basis, and another wanted reassurance that the data would not be sold for profit. Overall, 2 participants were keen to be given the opportunity to take part in health research in this way.

Following a presentation of examples and a general discussion (as outlined above), participants were asked to complete a second questionnaire (Multimedia Appendix 2) before leaving the workshop. Having received information regarding mobile phone data and health research, participants were asked again whether they were happy for their data to be used. More participants (80%, 49/61) compared with 62% (38/61) were happy for the data collected via their mobile phone to be used in health research after the workshop than before.

Figure 1. Public views on the types of data collected by mobile network operators in the course of phone use.



Although the majority trend among those who changed their minds was to a more positive viewpoint, there was a small degree of crossover as some became less happy for their data to be used. In total, 13 participants (21%, 13/61) explained that they had changed their minds as they had not been aware that their data was being used for health research in a positive way. Moving in the other direction, 1 participant said they was more concerned as they had been unaware so much data were being collected, and another stated that they felt more concerned but did not give a reason. The comparison between the outset and exit responses is shown in Figure 3.

The public identified a variety of benefits and concerns in using mobile phone data for health research. Benefits fell into 2 broad categories: (1) improvements in health, for example, *optimize hospital locations*, *public health*, *track disease spread* and *cures, treatments, and extended health expectancy* and (2) advancements in big data, for example, *easy access to large datasets*, *a cheap and easy method of data collection*, *detailed profiling of a cohort to understand people*, and to *link demographics to health outcomes*. The views about concerns were varied; however, the most frequent was the risk of breaching their anonymity (33%, 20/61), followed by data being sold for commercial gain (25%, 15/61), and unknown third-party use (21%, 13/61). Views on how to address these concerns are illustrated in Figure 4 and include greater transparency, clearer information and options for users, and better data governance.

The majority (84%, 51/61) of participants said that information on the use of anonymized mobile phone data for health research should be included in the terms and conditions. They should be written in simple language to include details on who their data were being shared with, which data were being shared, and why.

An opt-out system was also suggested where mobile phone users could choose specifically if and when their data could be shared. The majority was content for MNOs to share their phone data with academia (59%, 36/61), whereas government (34%, 21/61) and charities (26%, 16/61) were less popular options. Only 2 participants were happy to have their data shared with insurance companies and 5 with the pharmaceutical industry. Overall, 6 participants said they would like to be involved with research using mobile phone and health data. They indicated they would have an appetite for influencing topics for research, taking part in further research activities, and advising on public engagement and dissemination strategies.

Key Learning Points

At the outset of the public workshops, the majority was aware that a variety of data items were collected routinely, but a sizeable proportion either did not know or were unsure. This suggests that although there is a level of awareness among the general public about data collection by MNOs, there are many with limited knowledge. Of those who were aware, there was a reasonable grasp of the types of data collected, but very few people knew the data were being used for health research. Although, a majority of participants were happy for their mobile phone data to be used for this purpose, quite a proportion was unsure or unhappy. When asked what would make people more comfortable with their data being used for health research, the responses included:

- Data being safeguarded and anonymized
- Consent being sought for usage in research
- More information on types of research
- The option to opt-in and out on a project-by-project basis
- Reassurance that the data would not be sold for profit

Figure 2. Public responses on how mobile phone data are being used.

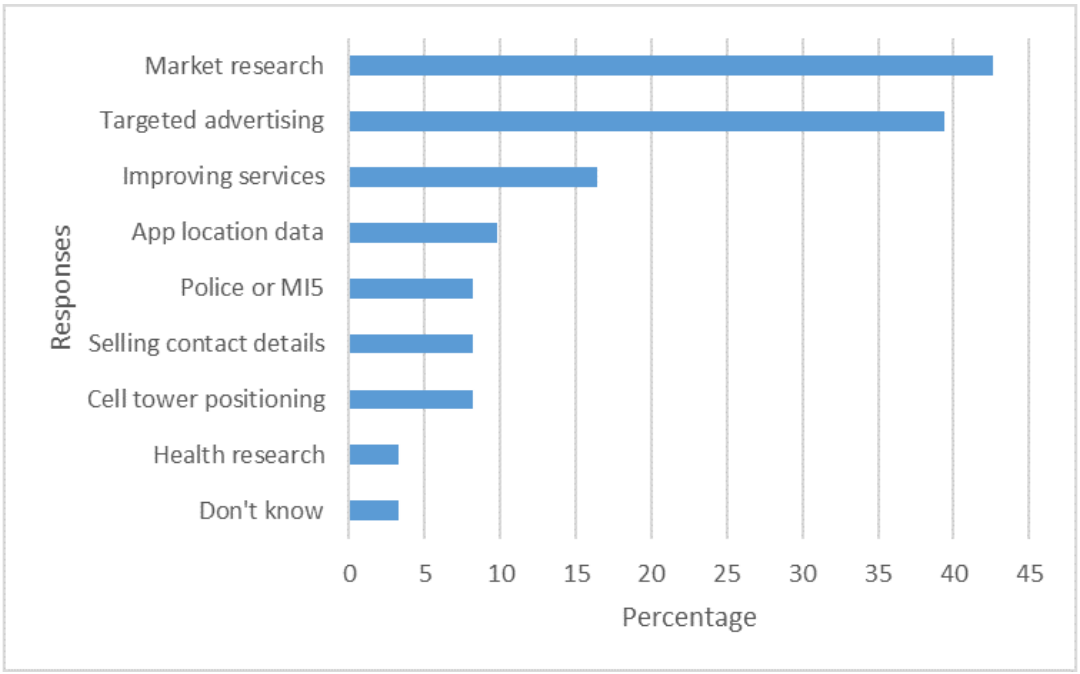
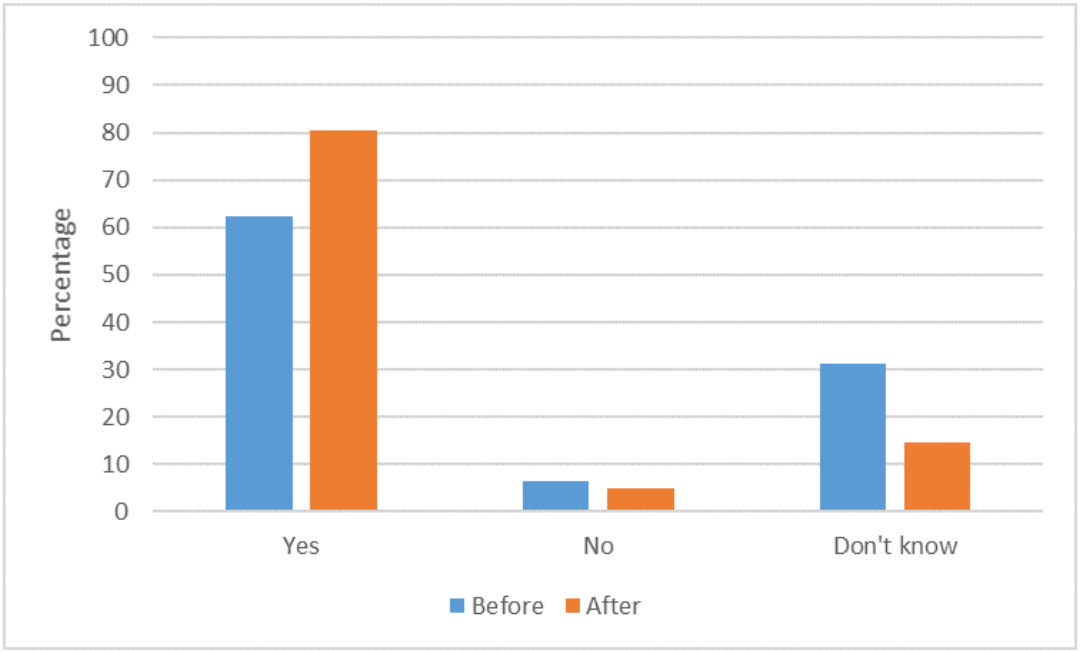


Figure 3. Public views on willingness to use mobile phone data for research before and after the workshop.



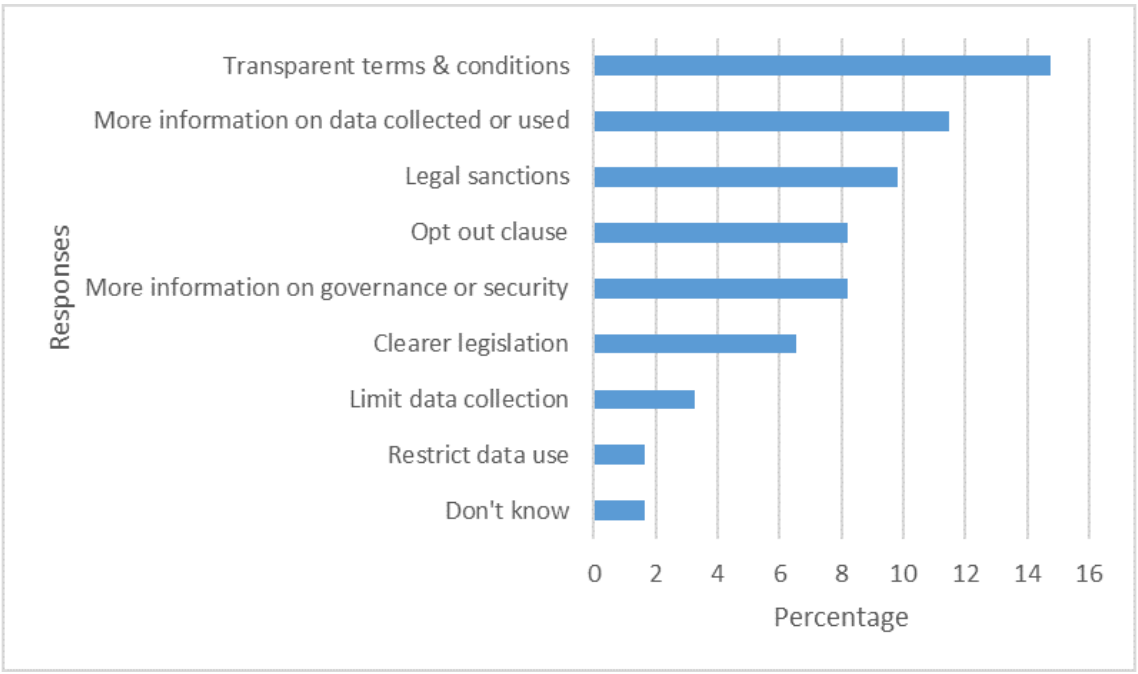
At the end of the workshops, a greater number of participants stated they were happy for their phone data to be used for health research than at the outset, with the principal reason for the changed viewpoint being that they had not realized the data could be put to beneficial health uses. Participants listed some notable benefits of using phone data in this way but also some important concerns, notably:

- Risk of reidentification
- Data being sold for commercial gain
- Unknown third-party use

A number of suggestions on how to improve the terms and conditions were made by the workshop participants and these were:

- Use of more basic language
- Wording to be more concise but more explicit
- Inclusion of information about how, and with whom, data would be shared
- Allowing opt out to different data sharing options, rather than an *all or nothing* approach
- Stronger, more transparent, information governance

Figure 4. Public views on how to address concerns about mobile phone data use.



Participants were also in favor of information about the use of anonymized mobile phone data for health research to be included in the terms and conditions. Although the majority was happy for their data to be shared with academic institutions, few people thought it acceptable for data to be shared with pharma and insurance companies. In summary, the public workshops indicated that people want clearer and more information; to be informed with whom and for what purpose data are shared; their views to be taken into account; more assurance of good data governance; and greater transparency and overt accountability from the MNO.

Discussion

Principal Findings

This is the first known study to explore public views on the use of mobile phone data, specifically CDRs, for health research. The workshops revealed many relevant and potentially valuable findings.

Knowledge and Engagement

As anticipated, the pilot group members were more knowledgeable than the general public groups at the outset. All members of the pilot group knew that MNOs collect data about phone users, compared with less than two-thirds among the general groups. Similarly, there was considerably greater awareness of data being used for research. This supported the value of running the pilot group with data-focused staff and of their insights to inform the public workshops. None of the total participants (25+61) said they had read the terms and conditions.

Concerns

The main concerns raised across the groups were in relation to data privacy and perceived inappropriate use. These centered on the nature of consent to collect identifiable data, the risk of reidentification in data purported to be anonymized, the potential for unknown use and misuse, and data being sold to potentially discriminatory parties.

Solutions

The solutions suggested by the groups largely converged on being provided with more information, more choice, and greater assurance of proper data governance. Participants wanted to know that their data were safeguarded and wanted reassurance that their data would not be sold for profit without their engagement. They wanted to know about the types of research that might take place, by whom and for what purpose, with public benefit being an important factor. Depending on the form of data being used, participants wanted to be asked for informed consent or to be able to opt-out of certain data uses. Across the groups, participants wanted clearer and more transparent terms and conditions.

An interesting and somewhat ironic finding was that having clearer terms and conditions and more information on data collection and use were strongly recommended although no one read the information already provided. This is an apparent contradiction and one that raises questions as to why, if people have concerns and want to know more, they do not read the information given to them. This is a common problem, not limited to mobile phone contracts, and one for which there have been a number of social experiments. The complexity of the wording used in various social media terms and conditions was recently highlighted as requiring a university degree to understand them fully [21]. Another study concerned a group of over 500 students signing up to a fictitious social media channel. None of the students read the terms and conditions well enough to notice that they had agreed to hand over their first-born child [22]. A further example involved over 7500 people agreeing to forgo the rights to their immortal soul by failing to read the terms and conditions for a gaming site download [23]. This phenomenon is also observed in other domains. In a survey completed by 550 direct-to-consumer

genetic testing customers, most respondents considered themselves aware of privacy issues and the risk of troubling repercussions of data donation to be negligible. However, over 50% men and almost 30% women also said they had not read the terms and conditions [24].

These examples, as well as the situation with mobile phone terms and conditions, leave us with doubts about the adequacy of informatory processes in some spheres. However, they also highlight the dilemma of where the responsibility lies, as to whether the onus should be more on the individual or the data collector [25]. Clearly, it is important that people read terms and conditions, but there are issues with the current formats of these documents that discourage them from doing so. It is possible that this is because of requiring to agree to the terms and conditions to obtain the phone and also the user-unfriendliness of the layout, with extensive small print in often opaque language. As it appears that the current terms and conditions are not hitting the spot, we propose that the format of information, and the way it is provided to the public, needs to be revised.

The public engagement workshops were particularly revealing about people's awareness and viewpoints on the use of mobile phone CDR data. Although MNOs are clearly profit-making organizations, participants did not want their data used for profit. This could be considered to be a naïve position as commercial gain via provision of a service is the raison d'être of MNOs, or perhaps it is because of the separation in our perceptions between the use of our phones and the systems that operate behind the scenes [26]. It should be acknowledged that there is sometimes a lack of understanding among the general public about identity disclosure risks in the use of anonymized or strongly pseudonymized data and aggregated data. Similarly, there may be some misunderstandings about the regulatory and legal requirements of using such data. However, the concerns raised are valid, as although something is lawful, it might not be socially acceptable [27]. Some argue that we have entered a state of surveillance realism characterized by a combination of unease and resignation to the use of our data [28]; however, this does not negate the need for more meaningful engagement with the public and the onus on us all to engage with our social responsibilities [25].

In common with a study on other private sector data being used for social research, dialogue with the public alleviated many of their concerns and clarified the role commercial companies can play in research for public benefit [15]. Furthermore, it had already been identified that there is an absence of a clear, holistic, ethical, and regulatory framework to guide research using CDRs [2]. The findings of this study support the need for such a framework, incorporating public views to guide its development.

Limitations

The main limitations of this study were the number of participants and the time available for the workshops (1 hour each). A longer time might have allowed more deliberative engagement and drawn out additional points. Although the findings of the workshops cannot claim to be fully representative of the public at large, as they are weighted toward younger or middle-aged adults, they should provide food for thought for MNOs and other parties with an interest in using mobile phone CDRs for research. It is possible that bias was introduced in drawing out the themes, but we aimed to avoid this by coming together to discuss data reduction and approach consensus.

Conclusions

Mobile phone CDRs are increasingly being used for health research with a geographical element. This novel study engaged with a cross-section of the public to gain their perspectives on the use of mobile phone data for health research. It was evident that this is a topic that has lacked public engagement in the past, and it showed that while the public recognizes the value of CDRs for health research, important concerns were raised and solutions suggested. The findings of this study lead us to recommend future work on gathering the views of a wider spectrum of participants to verify our findings, ascertain why people do not engage with the information provided in the terms and conditions and what a better information vehicle would look like, and further explore perceptions around profit-making from mobile phone and other networked device data. Crucially, our findings support the inclusion of public views into the development of an ethically founded framework for the socially acceptable use of CDRs in health research. We suggest that the design of this study might be of value to others seeking to work with the public on this important topic and in relation to data from other networked devices.

Acknowledgments

The authors acknowledge funding for the study from the Natural Environmental Research Council (NE/N007255/1).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaire 1: Public workshop prior knowledge questionnaire.

[PDF File (Adobe PDF File), 26KB - mhealth_v7i1e11730_app1.pdf]

Multimedia Appendix 2

Questionnaire 2: Public workshop exit questionnaire.

[PDF File (Adobe PDF File), 21KB - mhealth_v7i1e11730_app2.pdf]

References

1. GSMA Intelligence. 2018. Definitive data and analysis for the mobile industry URL: https://www.gsmaintelligence.com/ [accessed 2018-11-12] [WebCite Cache ID 73s3rfOsr]
2. Letouze E, Vinck P, Kammourieh L. Data Pop Alliance. 2015. The law, politics and ethics or cell phone data analytics URL: http://datapopalliance.org/wpcontent/uploads/2015/04/WPS_LawPoliticsEthicsCellPhoneDataAnalytics.pdf [accessed 2018-11-12] [WebCite Cache ID 73s47yn5U]
3. United Global Nations Pulse. 2013. Winning Research From The Data 4 Development Challenge URL: https://www.unglobalpulse.org/D4D-Winning-Research [accessed 2018-11-12] [WebCite Cache ID 73s4L7M4n]
4. Jones KH, Daniels H, Heys SM, Ford DV. Challenges and potential opportunities of mobile phone call detail records in health research: review. JMIR Mhealth Uhealth 2018;6(7):e161 [FREE Full text] [doi: 10.2196/mhealth.9974] [Medline: 30026176]
5. Agarwal R. Institut National des Télécommunications. 2013. Towards enhancing information dissemination in wireless networks URL: https://tel.archives-ouvertes.fr/tel-00919417/file/AGARWAL_Rachit-2.pdf [accessed 2018-12-18] [WebCite Cache ID 74lPJl0tc]
6. Farrahi K, Emonet R, Cebrian M. Computer-Supported Cooperative Work and Social Computing. 2015. Predicting a Community's Flu Dynamics with MobilePhone Data URL: https://hal.archives-ouvertes.fr/hal-01146198/document [accessed 2018-11-12] [WebCite Cache ID 73s5inLVD]
7. Wilson R, Zu EE, Albert M, Power D, Tudge S, Gonzalez M, et al. Rapid and near real-time assessments of population displacement using mobile phone data following disasters: the Nepal earthquake. PLoS Curr 2016 Feb 24;8 [FREE Full text] [doi: 10.1371/currents.dis.d073fbece328e4c39087bc086d694b5c] [Medline: 26981327]
8. Wesolowski A, Qureshi T, Boni MF, Sundsøy PR, Johansson MA, Rasheed SB, et al. Impact of human mobility on the emergence of dengue epidemics in Pakistan. Proc Natl Acad Sci U S A 2015 Sep 22;112(38):11887-11892 [FREE Full text] [doi: 10.1073/pnas.1504964112] [Medline: 26351662]
9. Tatem A, Huang Z, Narib C, Kumar U, Kandula D, Pindolia D, et al. Integrating rapid risk mapping and mobile phone call record data for strategic malaria elimination planning. Malar J 2014 Feb 10;13:52 [FREE Full text] [doi: 10.1186/1475-2875-13-52] [Medline: 24512144]
10. Le Menach A, Tatem AJ, Cohen JM, Hay SI, Randell H, Patil AP, et al. Travel risk, malaria importation and malaria transmission in Zanzibar. Sci Rep 2011 Sep;1:93 [FREE Full text] [doi: 10.1038/srep00093] [Medline: 22355611]
11. Understanding Patient Data. 2018. URL: http://understandingpatientdata.org.uk/ [accessed 2018-11-12] [WebCite Cache ID 73s6JnTmr]
12. Data Saves Lives. 2018. What is patient data? How and why do UK universities use this data? Can data really save lives? URL: https://www.datasaveslives.info/ [accessed 2018-11-12] [WebCite Cache ID 73s6SCkAM]
13. Cameron D, Pope S, Clemence M, Ipsos MORI Research Institute. Economic & Social Research Council. 2014. Dialogue on data: Exploring the public's views on using administrative data for research purposes URL: https://tinyurl.com/y87k9kyx [accessed 2018-11-12] [WebCite Cache ID 73s6knxwr]
14. Ipsos Mori. Wellcome Trust. 2016. The One-Way Mirror: Public attitudes to commercial access to health data URL: https://wellcome.ac.uk/sites/default/files/public-attitudes-to-commercial-access-to-health-data-wellcome-mar16.pdf [accessed 2018-11-12] [WebCite Cache ID 73s6rBcqD]
15. van Mil A, Hopkins A. Economic and Social Research Council. 2015. Big Data: Public views on the use of private sector data for social research URL: https://tinyurl.com/ya7yfn9x [accessed 2018-11-12] [WebCite Cache ID 73s7Qexp6]
16. Statista. 2018. Mobile phone user penetration as percentage of the population worldwide from 2013 to 2019 URL: https://www.statista.com/statistics/470018/mobile-phone-user-penetration-worldwide/ [accessed 2018-11-12] [WebCite Cache ID 73s8Bvle6]
17. Mutafungwa E, Thiessard F, Pathe D, Gore R, Jouhet V, Karray C, et al. Aalto University. 2015. Mobile Data as Public Health Decision Enabler: A Case Study of Cardiac and Neurological Emergencies URL: https://tinyurl.com/ydcq6pfw [accessed 2018-11-12] [WebCite Cache ID 73s8YOm01]
18. Dewulf B, Neutens T, Lefebvre W, Seynaeve G, Vanpoucke C, Beckx C, et al. Dynamic assessment of exposure to air pollution using mobile phone data. Int J Health Geogr 2016 Apr 21;15(1):14 [FREE Full text] [doi: 10.1186/s12942-016-0042-z] [Medline: 27097526]
19. Jones KH, McNerney CL, Ford DV. Involving consumers in the work of a data linkage research unit. Int J Consum Stud 2013 Oct 07;38(1):45-51. [doi: 10.1111/ijcs.12062]
20. Office for National Statistics. 2011. 2011 Census: Population and Household Estimates for the United Kingdom URL: https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/datasets/2011censuspopulationandhouseholdestimatesfortheunitedkingdom [accessed 2018-11-12] [WebCite Cache ID 73s8qh3ja]

21. BBC News. 2018. Social site terms tougher than Dickens URL: https://www.bbc.co.uk/news/business-44599968 [accessed 2018-11-12] [WebCite Cache ID 73s91azuH]
22. Kravets D. Ars Technica. 2016. TOS agreements require giving up first born—and users gladly consent URL: https://arstechnica.com/tech-policy/2016/07/nobody-reads-tos-agreements-even-ones-that-demand-first-born-as-payment/ [accessed 2018-11-12] [WebCite Cache ID 73s9Dtlrl]
23. Out-law.com. 2010. Special Reports Sign in to my account Topics TMT & Sourcing Nobody reads terms and conditions: it's official URL: https://www.out-law.com/page-10929 [accessed 2018-11-12] [WebCite Cache ID 73s9zQFBX]
24. Haeusermann T, Greshake B, Blasimme A, Irdam D, Richards M, Vayena E. Open sharing of genomic data: who does it and why? PLoS One 2017;12(5):e0177158 [FREE Full text] [doi: 10.1371/journal.pone.0177158] [Medline: 28486511]
25. Jones KH. Perspectives on the ethics of data donation: juggling our 1s and 0s. In: Springer Philosophy Studies Series. New York: Springer; 2019.
26. Greenfield A. Smartphone? The networking of the self. In: Radical Technologies: The Design of Everyday Life. London: Verso; 2018.
27. Carter P, Laurie GT, Dixon-Woods M. The social licence for research: why care.data ran into trouble. J Med Ethics 2015 May;41(5):404-409 [FREE Full text] [doi: 10.1136/medethics-2014-102374] [Medline: 25617016]
28. Dencik L, Cable J. The advent of surveillance realism: public opinion and activist responses to the Snowden leaks. Int J Commun 2017;11 [FREE Full text]

Abbreviations

CDRs: call detail records
ESRC: Economic and Social Research Council
MNO: mobile network operators
SMS: short message service

Edited by G Eysenbach; submitted 30.07.18; peer-reviewed by N Lea, E Ford, A Beukenhorst; comments to author 13.09.18; revised version received 19.09.18; accepted 03.10.18; published 16.01.19.

Please cite as:
Jones KH, Daniels H, Heys S, Ford DV
Public Views on Using Mobile Phone Call Detail Records in Health Research: Qualitative Study
JMIR Mhealth Uhealth 2019;7(1):e11730
URL: https://mhealth.jmir.org/2019/1/e11730/
doi:10.2196/11730
PMID:30664467

Viewpoint

Examining Diabetes Management Apps Recommended From a Google Search: Content Analysis

Geronimo Jimenez[1], MA; Elaine Lum[1,2,3], PhD; Josip Car[1,4], MD, PhD

[1]Centre for Population Health Sciences, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore, Singapore

[2]Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Australia

[3]School of Clinical Sciences, Faculty of Health, Queensland University of Technology, Brisbane, Australia

[4]Global eHealth Unit, Department of Primary Care and Public Health, School of Public Health, Imperial College London, London, United Kingdom

Corresponding Author:
Geronimo Jimenez, MA
Centre for Population Health Sciences
Lee Kong Chian School of Medicine
Nanyang Technological University
11 Mandalay Road, Level 18
Clinical Sciences Building, LKC Medicine
Singapore, 308232
Singapore
Phone: 65 69047027
Email: geronimo.jimenez@ntu.edu.sg

Abstract

Background: The availability of smartphone health apps empowers people to manage their own health. Currently, there are over 300,000 health apps available in the market targeting a variety of user needs from weight loss to management of chronic conditions, with diabetes being the most commonly targeted condition. To date, health apps largely fall outside government regulation, and there are no official guidelines to help clinicians and patients in app selection. Patients commonly resort to the internet for suggestions on which diabetes app to use.

Objective: The objective of this study was to investigate apps identified through a Google search and characterize these apps in terms of features that support diabetes management.

Methods: We performed a Google search for the "best diabetes apps 2017" and explored the first 4 search results. We identified and compiled a list of the apps recommended in the returned search results, which were Web articles. Information about each app was extracted from the papers and corresponding app store descriptions. We examined the apps for the following diabetes management features: medication management, blood glucose self-management, physical activity, diet and nutrition, and weight management.

Results: Overall, 26 apps were recommended in 4 papers. One app was listed in all 4 papers, and 3 apps appeared on 3 of the 4 lists. Apart from one paper, there were no explicit criteria to justify or explain the selection of apps. We found a wide variation in the type and the number of diabetes management features in the recommended apps. Five apps required payment to be used. Two-thirds of the apps had blood glucose management features, and less than half had medication management features. The most prevalent app features were nutrition or diet-related (19/24, 79%) and physical activity tracking (14/24, 58%).

Conclusions: The ambiguity of app selection and the wide variability in key features of the apps recommended for diabetes management may pose difficulties for patients when selecting the most appropriate app. It is critical to involve patients, clinicians, relevant professional bodies, and policy makers to define the key features an app should have for it to be classified as a "diabetes management" app. The lessons learned here may be extrapolated for the development and recommendation of apps for the management of other chronic conditions.

(JMIR Mhealth Uhealth 2019;7(1):e11848) doi:10.2196/11848

KEYWORDS

chronic diseases; diabetes; Google; health apps; mobile phone

Introduction

We live in a digital era, where people turn first to the internet rather than a health provider to learn about a health condition or to look up health-related information [1]. Additionally, the use of health apps is increasingly widespread with over 318,000 such apps available in the market [2]. Notably, of all medical conditions, diabetes is the most commonly targeted condition by health app developers [3].

The availability of smartphone health apps may empower people even further, as these provide an opportunity to assist or support patients to better self-manage long-term conditions, like diabetes, and influence healthier lifestyles [2]. Evidence-based diabetes guidelines emphasize lifestyle management like healthy eating and physical activity to manage diabetes [4,5]. Patients who actively engage in their own care between clinic visits are more successful in managing their diabetes [6,7], by, for example, using apps with blood glucose diary features and insulin calculators [8,9].

Most diabetes apps are neither regulated nor accredited by relevant governmental bodies [10,11]. To date, no clinical guidelines for diabetes management explicitly recommend the use of specific diabetes apps, nor list mandatory or desirable features of such apps [11]. Hence, clinicians have largely been concerned about app safety and reticent about recommending diabetes management apps to patients to support self-management between clinic visits [12,13].

In the absence of any official guidance, patients exploring the use of health apps to manage diabetes often resort to the internet for suggestions. Given that in terms of search engines, Google currently holds the majority market share of >86% [14], it is likely that our hypothetical patient would "google it." Google Trends for the search term "diabetes apps" showed increasing interest from August 2008, which peaked in August 2012 and remains at relatively high levels today [15]. We wondered about the results that such a Google search would yield and the features that the recommended apps will have.

Therefore, we aimed to investigate the characteristics (main features and type of support provided) of the apps recommended in the most popular results returned through a Google search for diabetes apps.

Methods

Google Search

We performed a Google search on April 18, 2018, with the following phrase: "best diabetes apps 2017." From this search, we identified and accessed the weblinks of the first 4 search results, in line with digital market research that showed a drop in user click-through after the fifth search result [16].

App Identification, Data Extraction, and Assessment

We identified the listed apps from each paper and extracted the following information from the papers and the Apple App store or Google Play store descriptions (respective to the listed app): App name, developer, user rating, number of ratings, whether apps were free or paid, requirements related to external blood glucose monitoring devices, diabetes management features, and general characteristics derived from paper and store descriptions.

Health interventions important for the successful management and treatment of diabetes include medications, self-management of blood glucose levels, physical activity, healthy eating, and weight management [4,5]. The presence or absence of the following app functions was noted. They were as follows: medication management—the capability to log or track medications taken or used, including insulin, other insulin-related features such as bolus calculation; self-management of blood glucose—the capability to log or track; physical activity—capability to log or track; diet and nutrition—the capability to log or track food consumption, carbohydrate, or calorie counting; and weight management—the capability to log or track weight, body mass index calculation.

Other functions were extracted as "additional functions," which included the following: blood pressure management; glycated hemoglobin (HbA_{1c}) tracking, prediction, or calculation; cholesterol management; provision of statistics and data visualization; data sharing; and the capability to connect to family members or friends, community, or health providers.

Data Reporting and Analysis

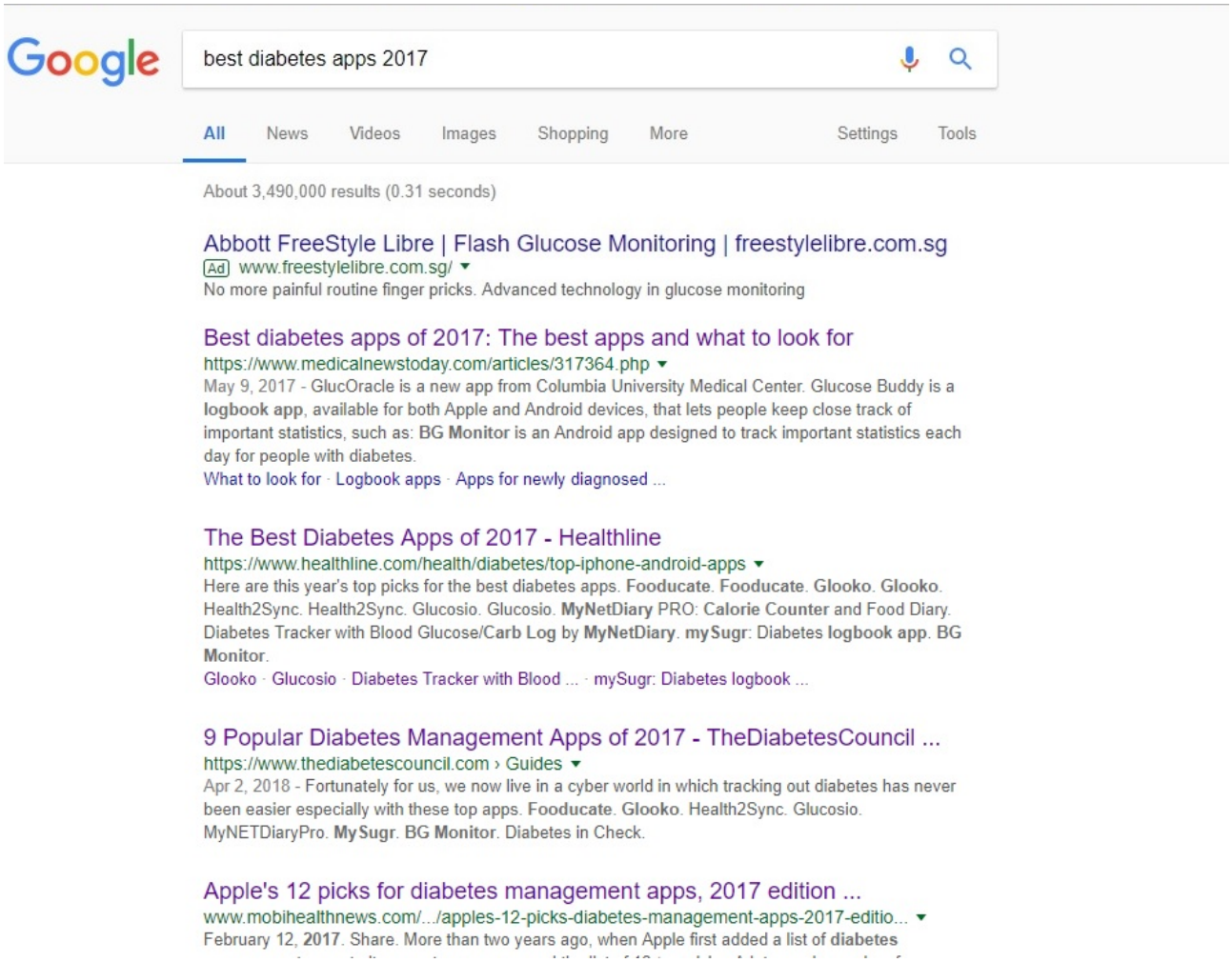
A narrative synthesis was used to describe the papers and app store descriptions of the apps. In terms of the papers that provided the lists of recommended apps, a descriptive summary of the paper content, recommended apps, and readership of the respective website are presented. Data about the apps are presented as quantitative aggregations and tallies in terms of which and how many diabetes management features were found.

Results

Content Summary of Identified Papers

The first 4 search results were all papers providing lists with the "best" or "most popular" diabetes or diabetes management apps (Figure 1).

Figure 1. A screenshot of Google search for "best diabetes apps 2017." Source: Google.com; screenshot taken by Geronimo Jimenez on April 20, 2018.



The content of each paper is described below:

1. Medical News Today—"Best diabetes apps of 2017: The best apps and what to look for" [17]. This paper guides the reader on what to look for in a diabetes management app, mentioning that it should be able to monitor blood sugar levels, carbohydrate intake, and weight management. It lists the different categories of apps (logbook, calorie counters, diet apps, carbohydrate counting apps, and general diabetes management) and reviews some of the "best diabetes apps of the year." The website claims to have "more than 15 million monthly visits, 13 million monthly unique visitors, and 20 million monthly page views" [18].
2. Healthline.com—"The Best Diabetes Apps of 2017" [19]. This paper provides a description of the complications of diabetes and mentions the aspects that help a person live better with diabetes, such as healthy eating, exercising, taking medicines, and sticking to treatment plans. The paper then goes on to mention that several apps have emerged to help patients "keep all the pieces of their care together," followed by a list of "this year's top picks." This is the only paper that explicitly mentioned app selection criteria: "We've selected these apps based on their quality, user reviews, and overall reliability as a source of support for people living with diabetes." The website claims that over 85 million people access it every month [20].
3. The Diabetes Council—"9 Popular Diabetes Apps of 2017" [21]. This paper describes the complications related to diabetes and then mentions that "we live now in a cyber-world in which tracking diabetes has never been easier especially with these top apps." The site claims to be the "#1 diabetes blog reaching millions of readers" [22]. We contacted the organization for more accurate readership figures but received no reply.
4. MobiHealthNews—"Apple's 12 picks for diabetes management apps, 2017 edition" [23]. This paper reports on Apple's list of 12 diabetes management apps for 2017. To ascertain Apple's selection criteria for this list, we contacted the author of the paper directly. The author mentioned that at times, the Apple Store would publish "best apps" lists for different apps categories or purposes but, unfortunately, was not able to shed any light on the criteria Apple used to come up with the list. The site claims to reach ">over 150,000 readers every month" and that its readers are "senior-level decision makers and influencers in healthcare and technology companies" [24].

Table 1. Diabetes management-related functions included in each app (n=24).

App Name[a]	Medication management	Blood glucose management	Physical activity features	Nutrition or diet features	Weight management	Additional useful features
Accu-Check Connect Diabetes Management App	Yes	Yes	Yes	Yes	Yes	Data sharing and blood pressure monitoring
BeatO	N/A[b]	Yes	Yes	Yes	N/A	Diabetes education
BG Monitor	Yes	Yes	Yes	Yes	N/A	Statistics and data visualization
Blip Notes, aka Tidepool Mobile	N/A	Yes	Yes	Yes	N/A	Aggregates data from different devices
Calorie Counter & Diet Tracker by MyFitnessPal	N/A	N/A	Yes	Yes	N/A	Connects to communities
Carb counting with Lenny	N/A	N/A	N/A	Yes	N/A	Food guide
Dexcom Follow	N/A	N/A	N/A	N/A	N/A	App for recipient of shared data
Dexcom G5 Mobile	N/A	Yes	N/A	N/A	N/A	For continuous glucose monitoring
Diabetes Kit Blood Glucose Logbook	Yes	Yes	Yes	Yes	Yes	HbA_{1c}[c] monitoring, data sharing, and blood pressure monitoring
Diabetes Tracker with Blood Glucose or Carb Log by MyNetDiary	Yes	Yes	Yes	Yes	N/A	Analysis and charts
Diabetes:M	Yes	Yes	N/A	Yes	N/A	Data export options
DiabetesConnect	Yes	Yes	N/A	Yes	N/A	Data export options, features can be turned off
Fooducate	N/A	N/A	Yes	Yes	Yes	Additionally tracks sleep schedule and moods
Glooko	Yes	Yes	Yes	Yes	N/A	Statistics and data visualization as well as data export options
GlucOracle	N/A	N/A	N/A	Yes	N/A	Projects blood glucose levels based on meal
Glucose Buddy	Yes	Yes	Yes	Yes	Yes	HbA_{1c} and blood pressure functions, data export options
Glucosio	N/A	Yes	N/A	N/A	Yes	HbA_{1c}, cholesterol, blood pressure, and ketones monitoring functions
Health2Sync	N/A	Yes	N/A	N/A	Yes	Blood pressure monitoring, connects with family and friends, and data exporting options
Helparound - diabetes dialysis	N/A	N/A	N/A	N/A	N/A	Information and resources for conditions and connects to community for help and support
Lose It!–Calorie Counter	N/A	N/A	Yes	Yes	Yes	Allows input of fitness goals
MyNETDiaryPro	N/A	N/A	Yes	Yes	Yes	Can sync with devices for other features and connects with friends and dietitian
mySugr: Diabetes Tracker Log	Yes	Yes	N/A	Yes	N/A	Coaching services with the upgrade
One Drop for Diabetes Management	Yes	Yes	Yes	Yes	N/A	Can connect to experts
OneTouch Reveal	Yes	Yes	Yes	Yes	N/A	Data visualization and data export options
Total functions, n (%)	11 (46)	16 (67)	14 (58)	19 (79)	8 (33)	N/A

[a]Two apps were excluded from the analysis (Diabetes in Check and GlucoSuccess) because these were no longer available in the app store.

[b]N/A: not available.

[c]HbA_{1c}: glycated hemoglobin.

Identified Apps and Corresponding Characteristics

A total of 26 apps (after removing duplicates) were identified (Multimedia Appendix 1). Of these, 2 apps were no longer available or were removed from the app stores (GlucoSuccess and Diabetes in Check) and were hence excluded from the analysis. One app was listed in all 4 papers (mySugr), and 3 apps appeared on 3 of the 4 lists (BG Monitor, Glooko, and myNetDiaryPro). In addition, 20 apps had a user rating, which varied widely among apps; 14 had user ratings of ≥4.5 (out of 5), 5 apps scored ratings <3.0, and the remaining app rated 3.3. Likewise, the number of ratings per app varied greatly; 2 apps received 384,000 and 88,000 ratings, respectively, whereas 4 apps had >4500 ratings. On the other end, 8 apps received between 100 and 700 ratings, 5 apps received 10-99 ratings, and 1 app received fewer than 10 ratings.

While 10 apps were completely free to use, 11 were either free to access but offered in-app purchases to unlock additional features, or had both a free and a paid version with more features. Three apps required payment or a subscription to access the app, although one of them (Glooko) would be free if sponsored by a doctor or covered by health insurance.

Three of the apps were part of a package that entailed purchasing a blood glucose monitoring device to use the app, that is, the Dexcom G5 Mobile Continuous Monitoring System and Roche's Accu-Chek glucose meter. Therefore, although the app was free, users do have a financial outlay upfront. In addition, another app (One Touch Reveal) was part of a package with a glucometer, but the app can still be used without the device. Similarly, 8 other apps can be connected to blood glucose monitoring devices but would still function without a specific device. Overall, 5 apps required upfront payment either as a subscription to the app or the purchase of a required blood glucose monitoring device.

We observed wide variability among the analyzed apps in terms of features and functions for diabetes management (Table 1). Two-thirds (16/24, 67%) had blood glucose management features, and less than half (11/24, 46%) had medication management features. The most prevalent feature was a nutrition or diet function present in 79% of the apps (19/24). In addition, 58% (14/24) included functions for physical activity tracking, and 33% (8/24) had weight tracking functions. Additional useful functionalities available in the apps were statistics and data visualization (eg, charts and graphs), data sharing options, the capability to connect to a community, friends and family, or health provider or expert or dietician. Some apps included additional monitoring functions for blood pressure, HbA_{1c}, and cholesterol.

Discussion

Principal Findings

This study explored the features of the "best/top diabetes management apps of 2017" recommended by the first 4 papers returned through a Google search. We found wide variation in the type of apps being recommended and in the number of diabetes management features included, reflecting ambiguous interpretation of the "diabetes management" concept.

The wide variability among selected apps signals no consensus regarding what diabetes management means. For instance, although there is no doubt that physical activity and healthy eating habits are important for managing diabetes, considering apps that *only* provide these features, "diabetes management" apps remains unclear and may be misleading. Several of these recommended apps only focused on diet and nutrition tracking and physical exercise which, while important, may not in itself be sufficient to manage diabetes. In addition, the selection of apps for these lists suggests little or no input from diabetes specialists (eg, diabetologists, diabetes nurse educators, and so forth) who would be mindful of *all* aspects required to manage this condition.

Are these apps capable of supporting the management of diabetes as the papers claimed? It depends. If we consider the capability of an app to log and track blood glucose levels to be an essential element for the management of diabetes, as defined by evidence-based diabetes guidelines [4,5], then 16 of the 24 available apps could be recommended for the management of diabetes based simply on the availability of this function (without exploring it in further detail). However, if we determine that the app must also support medication management, the number of apps deemed suitable decreases to 11. These results correlate with studies that found that, although diabetes management apps may have relatively high scoring functionality and aesthetics aspects, the majority do not integrate the most important diabetes management tasks [25]. Should apps that target one or two aspects of diabetes management be considered "diabetes management" apps? If so, which aspect or aspects should be deemed essential? The fact that these articles recommended some single-feature apps under the banner of "diabetes management" apps may mislead users about what to expect regarding the health consequences of living with diabetes and its successful management.

Taking all of the above into consideration, we may need to take a collective step back to examine the following questions: (1) what is diabetes management and (2) what does it really entail for each part of the patient journey? Conceptual segmentation can be helpful in designing or guiding the selection of diabetes apps—for example, in terms of the stage of diagnosis (newly diagnosed and delayed diagnosis), which health goals to emphasize at various stages, patient education or information about diabetes and health consequences, and types of support required (clinical, social or community, practical help).

For clinicians, these findings mean that patients with diabetes may be or become confused about what it means to have diabetes and how to manage it. Apps focusing solely on healthy eating and physical activity recommended under the "banner" of diabetes apps may mislead patients into thinking these strategies alone are adequate to keep diabetes in check. As such, there is a key role for clinicians to guide not only their patients but also the people creating these potentially very influential lists. Similarly, for policy makers involved with enacting

health-related policies and legislation, these findings emphasize a need to work with professional organizations (eg, American Diabetes Association and American Association of Diabetes Educators) to establish and clearly define what diabetes management should "look like" in a health app based on the existing diabetes management guidelines. Jointly endorsed guidance from policy makers and professional organizations on essential features of a diabetes management app would raise not only the quality bar for app development but also promote patient safety.

More transparency of app selection for recommendations is needed, as there were no explicit criteria to justify or explain the selection of apps, apart from the Healthline.com paper [19]. Other studies have also pointed out that search capabilities in app stores are limited and algorithms are not transparent, which increases the difficulty for patients to find the most suitable app to manage their condition [10]. The criteria should be clearly stated so that users know why an app made the cut for a certain list. Given the high potential of such lists in influencing patient selection of apps, nontransparency of app selection is concerning; this will also assure users that apps were not listed because of industry influences or other unclear reasons.

For app recommendations to be user-friendly, we need a more nuanced provision of recommendations tailored to an individual patient's journey with diabetes. For example, apps with physical activity and healthy nutrition functions may be sufficient for patients recently diagnosed with type 2 diabetes, whereas those being treated with medications and insulin may require blood glucose and medication management functions. Different types of apps may be more effective for type 1 or type 2 diabetes, younger patients, older adults, and so forth. The MedicalNewsToday article [17], for example, provided the features of their recommended apps and categorized the list into logbook apps, calorie or carbohydrate counters, diet apps, and general diabetes management apps, which may provide better guidance for users.

Limitations

This study has several limitations. First, Google searches vary depending on the country location in which they were performed. Therefore, this same search performed elsewhere may result in different retrieved websites. However, we investigated the reach and readership of each of the websites we accessed, in terms of the number of visitors and readers, which was substantial. Second, given the rapid changes in this field, some apps may no longer be available in the app stores. For example, at the time of analysis, 2 out of the 26 recommended apps were no longer available. Nonetheless, to the best of our knowledge, the rest of the apps included in this study are still widely used and available.

Although the focal point of this study was on diabetes apps, these findings and implications also apply to health apps for the management of other chronic conditions. It is important to clearly define what the management of a chronic condition entails to inform the development of safe and efficient apps that may help in managing such conditions. For this to happen, it is essential for patients, clinicians, professional bodies, and policy makers to interact and be involved with app developers, so that apps are appropriately designed to tackle key aspects of the target condition. Subsequently, these requirements should be made known to the wider public, so that organizations and websites promoting apps for the management of a chronic condition are aware of the required app features and functions.

Conclusions

The ambiguity of app selection and the wide variability in key features of the apps recommended are largely not helpful in guiding patients to select the most appropriate diabetes management apps. It is imperative for websites and organizations to be as transparent as possible when recommending apps, by clearly stating the selection criteria and justifications for the recommendations. There is a clear need to involve patients, clinicians, relevant professional bodies, and policy makers to define what makes an app a "diabetes management" app, including specifying the minimum features an app should have to be classified in this category. The lessons highlighted in this study could be extrapolated to other chronic diseases to inform the development and recommendation of safe and reliable apps.

Acknowledgments

This work was supported by the Ageing Research Institute for the Society and Education, Nanyang Technological University, Singapore (Grant no: ARISE/2017/21) and by the Centre for Population Health Sciences, Lee Kong Chian School of Medicine, Nanyang Technological University, Singapore.

The funder did not have direct involvement in the research, including manuscript preparation, apart from employing the authors. The research presented in this paper is solely the responsibility of the authors and does not reflect the views of the funder.

Conflicts of Interest

None declared.

Multimedia Appendix 1

List of recommended apps and characteristics as described in articles and app store descriptions.

[PDF File (Adobe PDF File), 42KB - mhealth_v7i1e11848_app1.pdf]

References

1. Jacobs W, Amuta A, Jeon K. Health information seeking in the digital age: An analysis of health information seeking behavior among US adults. Cogent Social Sciences 2017 Mar 13;3(1). [doi: 10.1080/23311886.2017.1302785]
2. Aitken M, Clancey B, Nass D. The Growing Value of Digital Health. New Jersey, USA: IQVIA Institute for Human Data Science; 2017. Evidence and Impact on Human Health and the Healthcare System URL: https://www.iqvia.com/institute/reports/the-growing-value-of-digital-health [accessed 2018-11-23] [WebCite Cache ID 748VhvmRW]
3. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M. Mobile health applications for the most prevalent conditions by the World Health Organization: review and analysis. J Med Internet Res 2013 Jun 14;15(6):e120 [FREE Full text] [doi: 10.2196/jmir.2600] [Medline: 23770578]
4. American Diabetes Association. Standards of Medical Care in Diabetes. Diabetes Care 2018;41(Suppl. 1):S1-159.
5. National IFHE. NICE Guideline. 2015. Type 2 diabetes in adults: Management URL: https://www.nice.org.uk/guidance/ng28 [WebCite Cache ID 71SvpE7iJ]
6. McAndrew L, Schneider S, Burns E, Leventhal H. Does patient blood glucose monitoring improve diabetes control? A systematic review of the literature. Diabetes Educ 2007;33(6):991-1011; discussion 1012. [doi: 10.1177/0145721707309807] [Medline: 18057267]
7. Hou C, Carter B, Hewitt J, Francisa T, Mayor S. Do Mobile Phone Applications Improve Glycemic Control (HbA1c) in the Self-management of Diabetes? A Systematic Review, Meta-analysis, and GRADE of 14 Randomized Trials. Diabetes Care 2016 Dec;39(11):2089-2095. [doi: 10.2337/dc16-0346] [Medline: 27926892]
8. Boyle L, Grainger R, Hall RM, Krebs JD. Use of and Beliefs About Mobile Phone Apps for Diabetes Self-Management: Surveys of People in a Hospital Diabetes Clinic and Diabetes Health Professionals in New Zealand. JMIR Mhealth Uhealth 2017 Jun 30;5(6):e85 [FREE Full text] [doi: 10.2196/mhealth.7263] [Medline: 28666975]
9. El-Gayar O, Timsina P, Nawar N, Eid W. Mobile applications for diabetes self-management: status and potential. J Diabetes Sci Technol 2013 Jan 01;7(1):247-262 [FREE Full text] [doi: 10.1177/193229681300700130] [Medline: 23439183]
10. Eng DS, Lee JM. The promise and peril of mobile health applications for diabetes and endocrinology. Pediatr Diabetes 2013 Jun;14(4):231-238 [FREE Full text] [doi: 10.1111/pedi.12034] [Medline: 23627878]
11. Huang Z, Soljak M, Boehm B, Car J. Clinical relevance of smartphone apps for diabetes management: A global overview. Diabetes Metab Res Rev 2018 Dec;34(4):e2990. [doi: 10.1002/dmrr.2990] [Medline: 29431916]
12. Wyatt JC, Thimbleby H, Rastall P, Hoogewerf J, Wooldridge D, Williams J. What makes a good clinical app? Introducing the RCP Health Informatics Unit checklist. Clin Med (Lond) 2015 Dec;15(6):519-521. [doi: 10.7861/clinmedicine.15-6-519] [Medline: 26621937]
13. Huckvale K, Adomaviciute S, Prieto J, Leow M, Car J. Smartphone apps for calculating insulin dose: a systematic assessment. BMC Med 2015 May 06;13:106 [FREE Full text] [doi: 10.1186/s12916-015-0314-7] [Medline: 25943590]
14. Statista. Worldwide desktop market share of leading search engines from January 2010 to April 2018. 2018. URL: https://www.statista.com/statistics/216573/worldwide-market-share-of-search-engines/ [accessed 2018-08-06] [WebCite Cache ID 71SvTVrDi]
15. Google LLC. Google Trends. 2018. Diabetes apps URL: https://trends.google.com/trends/explore?date=all&q=Diabetes%20apps [accessed 2018-05-03]
16. Lincoln J. Ignite Visibility. 2018. Google click-through rates in 2017 by ranking position URL: https://ignitevisibility.com/ctr-google-2017/ [accessed 2018-11-23] [WebCite Cache ID 748WIBC3e]
17. Johnson J. MedicalNewsToday. 2017 May 09. What are the best diabetes apps of 2017? URL: https://www.medicalnewstoday.com/articles/317364.php [accessed 2018-04-18]
18. MedicalNewsToday. 2016. About Medical News Today URL: https://www.medicalnewstoday.com/about [accessed 2018-04-28] [WebCite Cache ID 71SvsHtKh]
19. Cronenberg C. Healthline. 2017. The Best Diabetes Apps of the Year URL: https://www.healthline.com/health/diabetes/top-iphone-android-apps [accessed 2018-04-18]
20. Healthline. 2018. About Us - Healthline URL: https://www.healthline.com/about [accessed 2018-04-25]
21. Montgomery B. The Diabetes Council. 2017. 9 Popular Diabetes Management Apps of 2017 URL: https://www.thediabetescouncil.com/popular-diabetes-management-apps-review/ [accessed 2018-04-18]
22. The Diabetes Council Facebook Page. 2017 Dec 29. Our Story URL: https://www.facebook.com/pg/thediabetescouncil/about/ [accessed 2018-04-25]
23. Comstock J. MobiHealthNews. 2017 Feb 12. Apple's 12 picks for diabetes management apps, 2017 edition URL: https://www.mobihealthnews.com/content/apples-12-picks-diabetes-management-apps-2017-edition [accessed 2018-04-18]
24. HIMSS Media. 2018. Media: MobiHealthNews URL: https://www.himssmedia.com/media/mobihealthnews [accessed 2018-04-25]
25. Chavez S, Fedele D, Guo Y, Bernier A, Smith M, Warnick J, et al. Mobile Apps for the Management of Diabetes. Diabetes Care 2017 Dec;40(10):e145-e146 [FREE Full text] [doi: 10.2337/dc17-0853] [Medline: 28774944]

Abbreviations

HbA $_{1c}$: glycated hemoglobin

Edited by G Eysenbach; submitted 07.08.18; peer-reviewed by M Adu, U Öberg; comments to author 13.09.18; revised version received 20.09.18; accepted 29.09.18; published 16.01.19.

Please cite as:
Jimenez G, Lum E, Car J
Examining Diabetes Management Apps Recommended From a Google Search: Content Analysis
JMIR Mhealth Uhealth 2019;7(1):e11848
URL: http://mhealth.jmir.org/2019/1/e11848/
doi:10.2196/11848
PMID:

Original Paper

Mobile Apps for Management of Tinnitus: Users' Survey, Quality Assessment, and Content Analysis

Magdalena Sereda[1,2], MSc, PhD; Sandra Smith[1,2], BSc (Hons); Kiri Newton[3], BSc (Hons); David Stockdale[4], BA (Hons)

[1]National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre, Nottingham, United Kingdom

[2]Hearing Sciences, Division of Clinical Neuroscience, School of Medicine, University of Nottingham, Nottingham, United Kingdom

[3]Division of Audiology, Leicester School of Allied Health Sciences, De Montfort University, Leicester, United Kingdom

[4]British Tinnitus Association, Sheffield, United Kingdom

Corresponding Author:
Magdalena Sereda, MSc, PhD
National Institute for Health Research Nottingham Biomedical Research Centre
Ropewalk House
113 The Ropewalk
Nottingham, NG1 5DU
United Kingdom
Phone: 44 1158232600
Fax: 44 1158232615
Email: Magdalena.Sereda@nottingham.ac.uk

Abstract

Background: Tinnitus is the perception of a sound without any outside source. It affects 6 million people in the United Kingdom. Sound therapy is a core component of many tinnitus management programs. Potential mechanisms of benefit include making tinnitus less noticeable, habituation, distracting attention from tinnitus, relaxation, and promoting neuroplastic changes within the brain. In recent years, there has been a substantial increase in the use of mobile technology. This provided an additional medium through which people with tinnitus can access different tinnitus management options, including sound therapy.

Objective: The aim of this study was to (1) generate the list of apps that people use for management of their tinnitus, (2) explore reasons for app use and nonuse, (3) perform quality assessment of the most cited apps, and (4) perform content analysis to explore and describe options and management techniques available in the most cited apps.

Methods: A Web-based survey consisting of 33 open and closed questions captured (1) demographic information, information about tinnitus, and hearing loss and (2) mobile app–specific information about the motivation to use an app, the apps which respondents used for tinnitus, important factors when choosing an app, devices used to access apps, and reasons for not using apps. The quality of the most cited apps was assessed using the Mobile Apps Rating Scale (MARS). Content and features of the most cited apps were analyzed.

Results: Data from 643 respondents were analyzed. The majority of respondents (482/643, 75.0%) had never used an app for management of tinnitus mainly because of lack of awareness (381/643, 59.3%). The list of the 55 apps that people use for their tinnitus was generated. These included apps that were developed specifically for the management of tinnitus; however, the majority of cited apps were developed for other problems (eg, sleep, depression or anxiety, and relaxation). Quality assessment of the 18 most popular apps using MARS resulted in a range of mean scores from 1.6 to 4.2 (out of 5). In line with the current model of tinnitus management, sound was the main focus of the majority of the apps. Other components included relaxation exercises, elements of cognitive behavioral therapy, information and education, and hypnosis.

Conclusions: People used apps for the management of their tinnitus; however, this was done mostly as a self-help option, without conjunction with management provided by hearing health care professionals. Further research should consider the place for apps in tinnitus management (stand-alone self-management intervention vs part of the management by a hearing professional). As the content of the apps varies with respect to sound options, information, and management strategies, it seems that the choice of the best management app should be guided by individual patient's needs and preferences.

(JMIR Mhealth Uhealth 2019;7(1):e10353) doi:10.2196/10353

KEYWORDS

tinnitus; mobile apps; disease management; surveys and questionnaires; Mobile Apps Rating Scale

Introduction

Background

Tinnitus is the perception of a sound without any outside source. It affects 6 million people in the United Kingdom. Sound therapy, in the form of hearing aids or sound generators, is a core component of many tinnitus management programs. Potential mechanisms of benefit include making tinnitus less noticeable, promoting habituation, distracting attention from tinnitus, relaxation, and promoting neuroplastic changes within the auditory system. Sound therapy can be provided by a range of media, including hearing aids, wearable sound generators, combination hearing aids, or bedside or tabletop sound generators [1].

Mobile technology, including mobile phone provides an additional medium through which people with tinnitus can access different tinnitus management options, including sound therapy. Recent years have seen a substantial increase in the use of mobile technology. According to the recent Global Mobile Consumer Survey by Deloitte, 85% of adults in the United Kingdom own a mobile phone, and this number is expected to increase to 90% by 2020. More than 37 million people aged 16 to 75 years use their device every day, and 34% look at their device within 5 min of waking [2].

The use of mobile apps to deliver health care (mobile health, mHealth) has several advantages, including (1) improved access to health care, (2) improved quality of health care, and (3) lowering the cost of health care [3]. There are also potential issues associated with mHealth, and these include safety or misuse [4], quality and effectiveness [5,6], responsibility, and risk [7]. The attitudes of patients and health care professionals toward these new developments also need to be assessed and addressed [8].

The quality and functionality of health care apps, including tinnitus apps can vary greatly. The IMS Institute for Healthcare Informatics [9] assessed the functionality of 16,275 health care apps according to 25 individual criteria, including the type and quantity of information provided by the app, how the app tracks or captures user data, the communication processes utilized by the app, and the quantity of device capabilities included in the app. More than 90% of the apps tested received a score of 40 or less out of a possible 100, which indicates the general low quality of the apps tested.

In 2013, the National Health Services (NHS) Commissioning Board created a digital apps library for health care apps. Currently, in the early version of the library, there are 42 apps listed that “meet the high standard of quality, safety, and effectiveness” [10]. Some apps were tested further to assure they meet NHS standards for clinical effectiveness, safety, usability, and accessibility. Although the library includes apps developed for variety of health care conditions as well as healthy living in general, it does not currently list apps for management of tinnitus. However, people with tinnitus might find some of the apps helpful, for example, those developed for the management of stress and anxiety.

To date, no research has looked specifically at the use of mobile apps for tinnitus management. A study by Paglialonga et al [11] identified and assessed apps for hearing science and care in general, which were available on the leading platforms (iOS, Android, and Windows phone stores). Tinnitus apps identified by the authors were mentioned in 2 categories: (1) screening and assessment (estimation of tinnitus pitch and loudness) and (2) intervention and rehabilitation (tinnitus management tools such as maskers and sound stimulation). The identified apps were intended to be used by hearing health care professionals, people with tinnitus, or both.

Objectives

Despite the increasing popularity of apps in general, it is unclear what proportion of people use apps for tinnitus management and which apps are the most popular. The purpose of this study was to (1) generate the list of apps that people use for management of their tinnitus, (2) explore reasons for apps use and nonuse, (3) perform quality assessment of the most cited apps, and (4) perform content analysis to explore and describe options and management techniques available in these most cited apps.

Methods

Web-Based Survey

The Checklist for Reporting Results of Internet e-Surveys was used to report the methods and results of the survey ([12] Multimedia Appendix 1). Ethics approval for the study was granted by the University of Nottingham Faculty of Medicine & Health Sciences Research Ethics Committee reference number LT18082016. As this was an anonymous Web-based survey, completion of the survey was taken as informed consent. No identifiable data were collected.

Survey Development

Items for the survey were decided through an iterative process. A list of questions was generated to capture (1) demographic information about respondents (gender, age group, and country of residence); information about tinnitus (presence, duration, and severity); and hearing loss (presence, severity, and use of devices to address hearing loss) and (2) mobile app–specific questions asked about motivation to use an app to manage tinnitus; a list of apps respondents used for managing tinnitus; important factors when choosing an app; devices used to access apps; and reasons for not using apps to manage tinnitus. First, questions were generated in collaboration with the British Tinnitus Association (BTA) and based on the information about the apps that patients were seeking when contacting the BTA. Second, questions were generated to capture information missing from the general tinnitus literature (eg, factors that drive the decision to try apps or factors important when choosing apps for tinnitus management). Questions were first drafted by one of the authors (MS) and then appraised and reduced by other

coauthors toward strong face validity and relative merit of the included items. The final questionnaire included 33 items presented on 15 pages. The final survey comprised a mix of open and closed questions and took between 15 and 30 min to complete. The survey used skip logic depending on if a participant had used or not used apps for tinnitus management before or had or did not have tinnitus. No randomization of items was used. All questions, with exception of questions asking about additional comments, were mandatory. Respondents were unable to change their responses once submitted.

Administration

Over a 2-month period, people were invited to take part in an anonymous Web-based survey, which was hosted on Survey Monkey (Survey Monkey Inc., San Mateo, California, USA). Responses were collected between August 15, 2016, and November 15, 2016. The survey was open to anyone who wanted to take part, and both app users and nonusers were invited. The survey was advertised via email to current BTA members and National Institute for Health Research (NIHR) Nottingham Biomedical Research Centre (BRC) participants' database members. The link to the questionnaire was sent out using social media to people following the BTA and BRC via Facebook and Twitter. Only 1 submission from each internet protocol address was permitted by the survey software.

Analysis

Closed questions were analyzed in IBM's SPSS Statistics 24 using descriptive statistics, including frequencies, means, and SDs. Patterns of use depending on age, tinnitus severity and duration, hearing loss, and gender were analyzed using chi-square statistics. Qualitative data from the open questions were analyzed separately using inductive thematic analysis.

Quality Assessment of the Apps

The quality of the most cited apps listed by respondents was assessed by 3 researchers using the Mobile Apps Rating Scale (MARS) [13]. To be included in the quality assessment, an app needed to be cited by 2 or more people. The MARS scale was developed to be a simple, objective, and reliable tool for assessing the quality of mHealth apps. It contains 23 items rated on a 5-point scale (1=inadequate, 2=poor, 3=acceptable, 4=good, and 5=excellent) or not applicable. A total of 19 questions form the objective quality section, which is divided into 4 scales: engagement, functionality, aesthetics, and information quality. In addition, 4 questions form the subjective quality section evaluating users' satisfaction. Each app was scored independently by 3 researchers (MS, SS, KN, or DS) using MARS. Apps were tested on Android and iPhone devices where the app was available on both devices. This was followed by a consensus meeting where the scores and reasons for them were discussed. Consensus on the final scoring was then reached by all 3 raters for the objective scales. For the subjective scale, an average rating was taken.

Content and Features Analysis

Content and features of the most cited apps were analyzed using a bottom-up approach. MS developed a coding manual based on the features listed in the Web-based description of the apps in the Apple App Store, Google Play, and the Amazon App Store, including descriptions and example quotes from the text. The coding manual was reviewed by MS and SS to assure clarity of definitions and examples. A small sample of the cited apps was then assessed by MS, and any missing codes generated were added to the coding manual. This coding manual was then used to identify the content and features of the most cited apps. MS and SS independently applied the coding manual to each mobile app to clarify ambiguous codes, remove duplicate codes, and identify data that did not fit the coding scheme. Coding was then compared and discussed between coders, and subsequent modifications made to the coding manual, resulting in final version of the manual (Multimedia Appendix 2).

Results

Participation Data

A total of 675 people responded to the survey. Responses were collected between August 15, 2016, and November 15, 2016. Of the 675 participants who read the welcome page and proceeded to consenting, 671 consented to take part in the survey, which translated to 99.4% (671/675) participation rate. The data were included in the analysis if the respondents provided a response to the question asking if they had ever used an app to manage their tinnitus, which left 643 responses for further analysis. Moreover, 32 people provided only initial demographic information and, therefore, were excluded from the analysis.

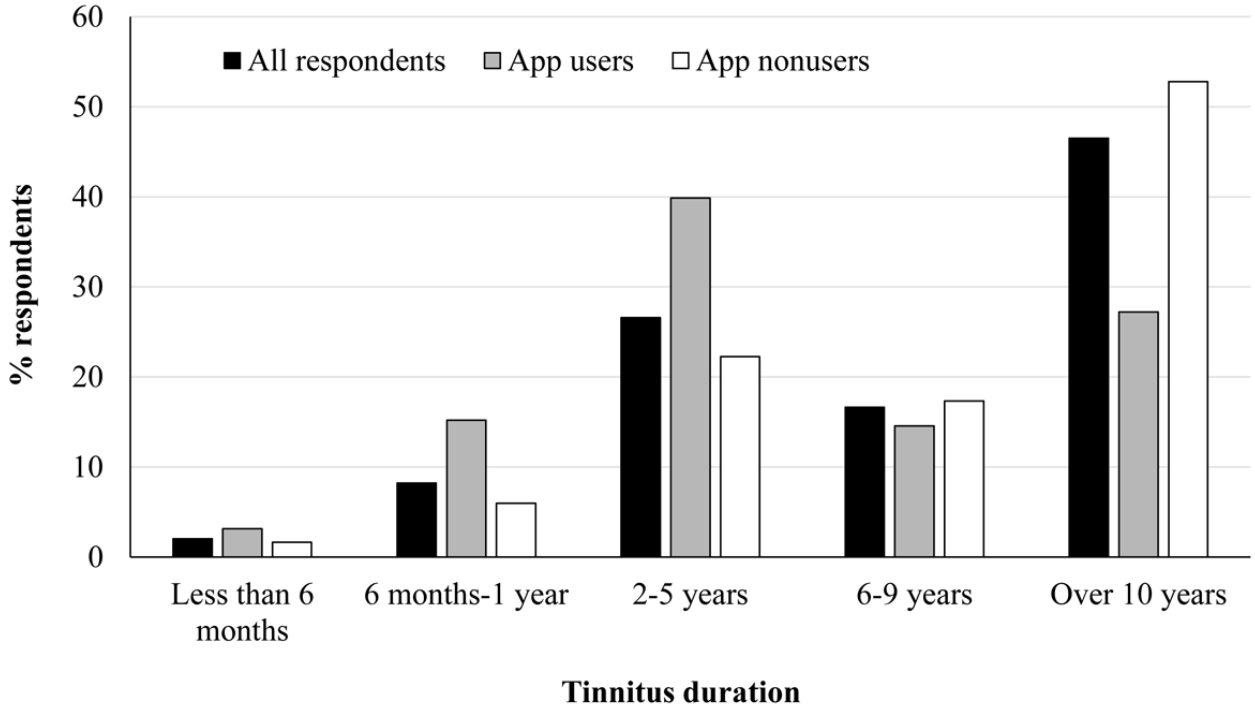
Of 643 respondents, 158 respondents had used an app, whereas 485 had never used an app to manage their tinnitus. The majority of participants were UK residents (627/643, 97.5%), with 16 residents from other countries, including Australia (n=5), Canada (n=4), Norway (n=2), Cyprus (n=1), Denmark (n=1), Egypt (n=1), Ireland (n=1), and Malaysia (n=1).

Demographic Data

The majority of respondents (n=637) had tinnitus at the point of completing the survey, whereas 6 had tinnitus in the past. The largest group of respondents were people who had tinnitus for more than 10 years (299/637, 46.9%; Figure 1). There was a significant association between tinnitus duration and use or nonuse of apps, χ^2_4=44.8, P<.001. Among the users, there were significantly more people in *6 months to 1 year* (z=3.0) and *2 to 5 years'* (z=3.2) groups and significantly less users in *over 10 years'* group (z=−3.6). Among the nonusers, there were significantly more people in the *over 10 years'* group (z=2.0).

The majority of respondents (278/673, 41.3%) reported their tinnitus to be moderate, whereas 33 respondents reported slight, 151 mild, and 181 reported severe tinnitus. For the chi-square analysis of tinnitus severity in app users and nonusers, we have combined *slight* and *mild* categories to achieve at least five observations in each category. There was a significant association between tinnitus severity distribution and use or nonuse of apps, χ^2_2=11.3; P=.004. Among users, there were significantly less people who reported slight or mild tinnitus (z=−2.8).

Figure 1. Tinnitus duration in all respondents (black bars), app users (grey bars), and nonusers (white bars).



The age of survey respondents ranged from less than 18 to 75 years and over, with the largest representation from people aged 55 to 64 years (216/675, 32.0%) and 65 to 74 years (216/675, 32.0%; Figure 2). The majority of survey respondents were people with tinnitus; therefore, such age distribution is in line with the data showing higher prevalence of tinnitus with age [14]. There was a significant association between the age distribution and use or nonuse of apps, χ^2_7=40.9; *P*<.001. Among the users, there were significantly more people in the 45 to 54 years group (*z*=3.2) and significantly less in the 65 to 74 years (−2.7) and more than 75 years (*z*=−2.8) groups.

Of 643 respondents, 289 (44.9%) were female, 350 (54.4%) were male, and 4 (0.6%) identified in another way. The proportion of males versus females was similar among app users (84/158, 53.1% vs 74/158, 46.9%) and nonusers (267/485, 55.1% vs 218/485, 44.9%; χ^2_1=0.43; *P*=.51).

Figure 2. Age distribution for all respondents (black bars), app users (grey bars), and nonusers (white bars).

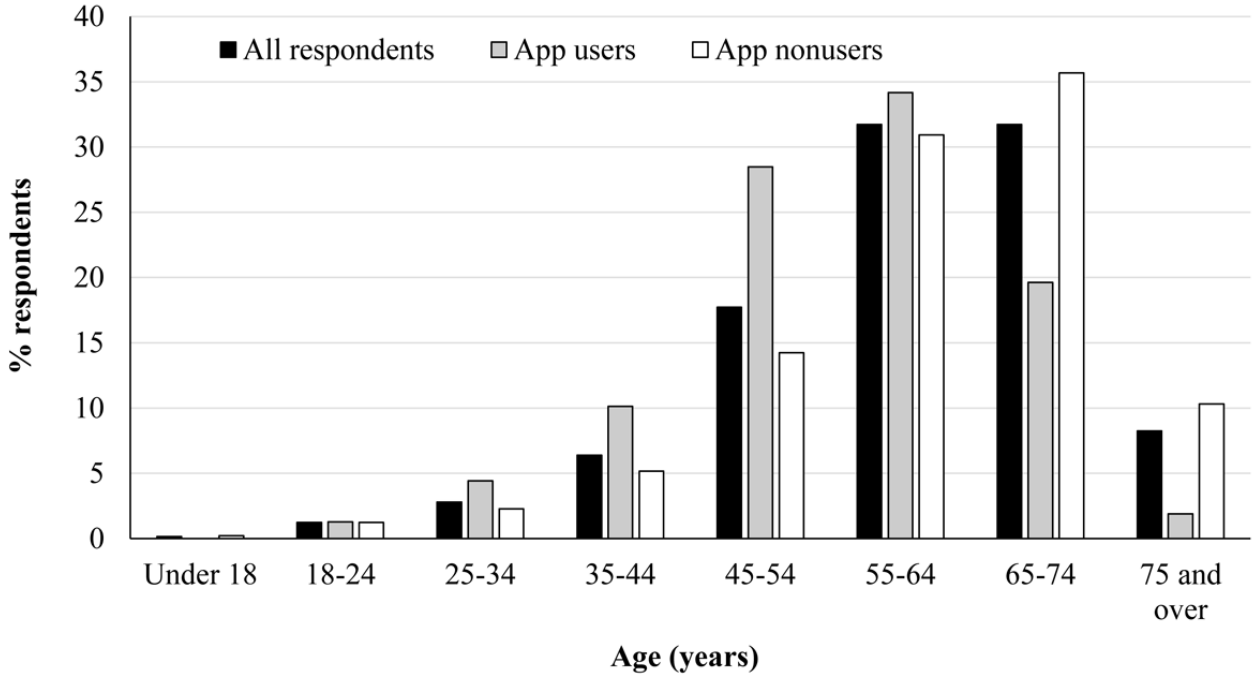
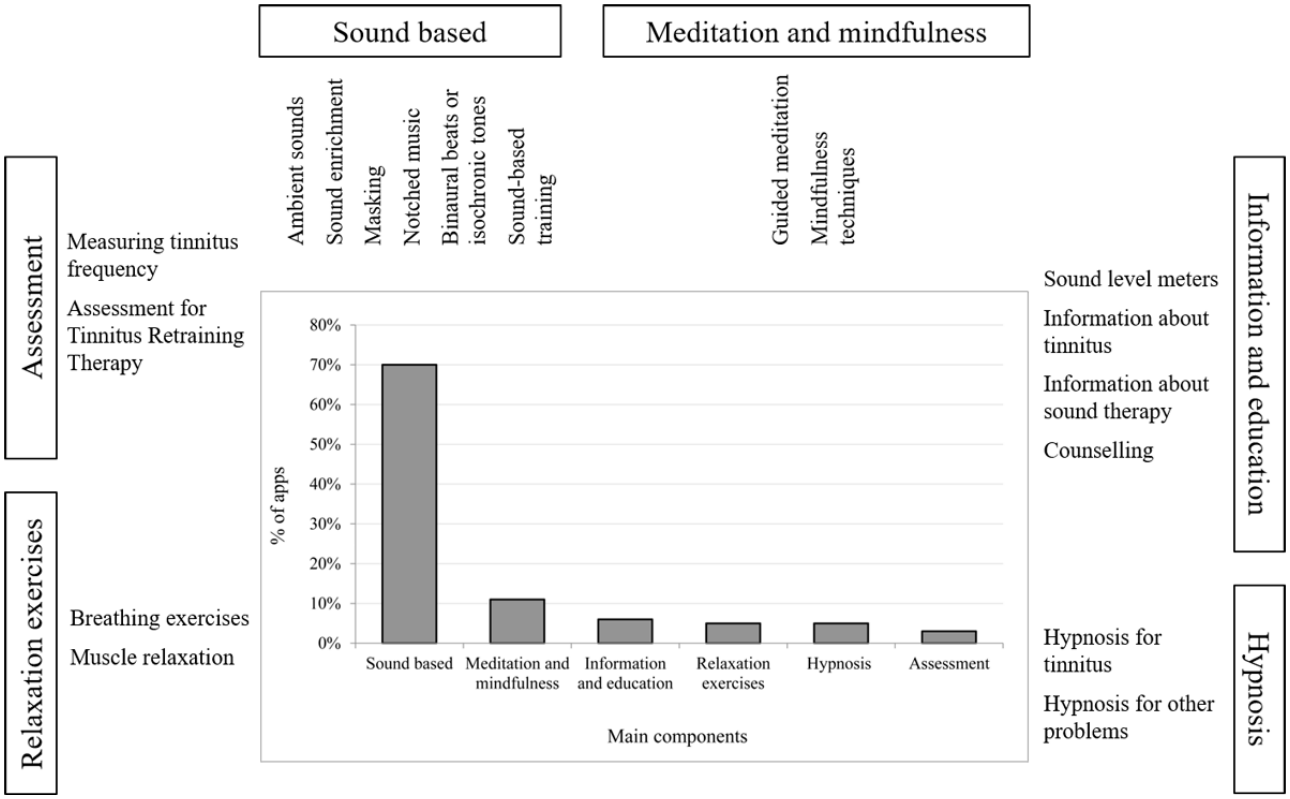


Figure 3. Components of the apps used by respondents for tinnitus management.



The majority of respondents reported some degree of hearing loss (494/643, 76.8%), consistent with the association between tinnitus and hearing loss [15,16]. The largest group of respondents reported mild hearing loss (261/643, 40.6%), with 149 respondents reporting no hearing loss (23.2%), 172 reporting moderate (/643, 26.7%), and 61 reporting severe or profound hearing loss (61/64, 9.5%). There was a significant association between the degree of hearing loss distribution and use or nonuse of apps, χ^2_3=17.5; *P*=.001. There were significantly less app users in the severe or profound hearing loss group (*z*=−2.3).

Of 494 respondents with hearing loss, 263 reported wearing hearing aids and 13 reported wearing cochlear implants. Moreover, 56 hearing aid users and only 2 cochlear implant users reported using apps to manage their tinnitus. In addition, 59 respondents reported using assistive listening devices, and 19 of those reported using apps.

Apps That People Tried for Tinnitus Management

Altogether, 120 respondents listed 55 apps that they have tried to manage their tinnitus. In addition, 15 people listed apps with a more general context such as radio, YouTube, podcast apps, and audiobook apps, without specifying the exact content that they are using to manage their tinnitus. As there was no way of verifying which contents have been used, those were excluded from further analysis. A full list of the 55 apps listed by respondents and their characteristics is available in Multimedia Appendix 2.

A total of 6 main components of the apps have been identified based on the description in the app stores (Apple, Google, and Amazon): (1) sound generation or therapy, (2) meditation and mindfulness, (3) information and education, (4) hypnosis, (5) relaxation exercises, and (6) assessment (Figure 3). In 70% (38/55) of listed apps, sound was the main focus of the app, including providing a selection of ambient sounds, sound enrichment, sounds for masking distracting sounds or tinnitus, notched music, binaural beats or isochronic tones, and sound-based training. Moreover, 11% (6/55) of apps included guided meditation and mindfulness techniques, and 5% (3/55) of apps had an extensive information and education component and included sound level meters, apps containing information about tinnitus, information about sound therapy, and counseling. Hypnosis for tinnitus or for other problems was a focus of 5% (3/55) of apps. Relaxation exercises such as breathing exercises and muscle relaxation were the components of 5% (3/55) of apps. In addition, 3% (2/55) of apps focused on assessment, including measuring tinnitus frequency and assessment for tinnitus retraining therapy (TRT).

Of 55 listed apps, 14 were developed specifically for tinnitus (Multimedia Appendix 1). In addition, 6 apps used sound or sound therapy to provide relief or distraction from tinnitus (eg, *Tinnitus Therapy Lite*, *Sound Relief*, and *Tinnitus Therapy Tunes*), 3 apps provided a combination of sound and relaxation exercises (*Beltone Tinnitus Calmer* and *ReSound Relief*), 3 apps implemented specific tinnitus management programs (*TRT* [*iTinnitus*], Progressive Tinnitus Management [*Tinnitus Balance*], and Zen Therapy [*Widex Zen*, *Tinnitus Management*]), 1 app indicated it was a combination of informational resource and sound therapy (*Starkey Relax*), 1 app used hypnosis (*Overcome Tinnitus*), and 1 app aimed to measure tinnitus pitch (*Tinnitus Measurer*). Moreover, 5 apps were not developed specifically for tinnitus but mentioned tinnitus as one of the possible applications either through masking or without

specifying a specific mechanism through which the app might be helpful for tinnitus. A total of 32 apps were developed for sleep, relaxation, concentration, meditation, stress, anxiety, and general well-being and have not mentioned tinnitus as a potential application of the app, and 2 apps were sound level meters.

Each app was listed by between 1 and 21 respondents. The apps that were the most often listed by respondents as the ones they have tried for managing their tinnitus were *White Noise Free* (n=21), *Oticon Tinnitus Sound* (n=13), *Relax Melodies: Sleep Sounds* (n=10), *myNoise* (n=7), and *Tinnitus Therapy Lite* (n=7).

Given that the majority of the apps (n=37) were only mentioned by 1 respondent, we performed a further analysis for those 18 apps that were listed by at least two people (Table 1; for the full list of apps, please see Multimedia Appendix 3). This included quality assessment using MARS [12] and detailed content analysis.

Reasons for Nonuse

The most commonly listed reason for not using apps for management of tinnitus was lack of awareness of apps (364/485, 75%). In addition, 20.0% (73/364) of respondents declared that they did not use apps as they were not good with technology, 13.2% (48/364) could not find an app that they thought would be helpful for their tinnitus, 11.5% (42/364) said that they did not have mobile phone or tablet, 4.7% (17/364) had only a basic phone that did not support apps or problems with their phone, 3.0% (11/364) did not need to use apps, 2.7% (10/364) did not think the apps would help with their tinnitus, and 2.5% (9/364) used other technologies such as a bedside sound generator or CD player. Other reasons (<1%) for app nonuse included hearing problems, wanting a cure rather than management option, lack of knowledge about which apps could help, not willing to rely on technology, not willing to pay attention to tinnitus, apps exacerbating tinnitus, hyperacusis, preference for a personal contact, too many apps to choose from, having tinnitus for a short period, or lack of interest in apps.

Table 1. Characteristics of the apps mentioned by at least three respondents (N=number of times app was cited).

Name	Developer	Category	Star rating[a]	Cost GBP[b] (£)	In-app purchases	Installs[c]	Platform and version	Last update
White Noise Free (N=21)	TMSOFT	Health and fitness	4.5	Free	No	1-5 million	Apple (7.0), Google (varies with device), and Amazon (7.2.3)	2016
Oticon Tinnitus Sound (N=13)	Oticon A/S	Medical	4.2	Free	No	50,000-100,000	Apple (1.0.2) and Google (1.0.1)	2015
Relax Melodies: Sleep Sounds (N=10)	Ipnos Software	Health and fitness	4.7	Free	Yes	5-10 million	Apple (6.2), Google (varies with device), and Amazon (6.1.2)	2017
myNoise (N=7)	myNoise BVBA	Health and fitness	4.5	Free	Yes	50,000-100,000	Apple (2.4.2) and Google (1.2)	2017
Tinnitus Therapy Lite (N=7)	Sound Oasis	Health and fitness	4.5	Free sample of 5 sounds and basic options	No Pro version available	10,000-50,000	Apple (1.1.6), Google (1.1.6), and Amazon (1.0.3)	2017
Headspace: Guided Meditation & Mindfulness (N=6)	Headspace, Inc	Health and fitness	3.9	Free sample of 10-day meditation	No Pro version available	1-5 million	Apple (3.4.0), Google (3.1.2), and Amazon (2.2.0)	2017
Sleep Bug: White Noise Soundscapes & Music Box (N=6)	Panzertax	Health and fitness	4.4	Free	Yes	100,000-500,000	Apple (3.4) and Google (1.6)	2017
Beltone Tinnitus Calmer (N=4)	Beltone	Medical	4.3	Free	No	10,000-50,000	Apple (3.4.2) and Google (3.1.4)	2017
Sleep Pillow (N=4)	FITNESS22 LTD	Health and fitness	4.8	Free	Yes	100,000-500,000	Apple (7.4) and Google (4.3)	2016
Soothing Sounds Lite (N=4)	Lost Ego Studios Limited	Apple: Medical and Google: Lifestyle	3.5	Free	No	1000-5000	Apple (1.22) and Google (1.0)	2017

[a]Average star rating across platforms.

[b]GBP: Great Britain Pound/Pound Sterling

[c]Data on number of installs were available only on Google Play.

Motivation to Try an App

Of the 158 respondents, 36 (22.8%) of respondents tried an app to address sleep problems, including getting to sleep and staying asleep, 33 (20.9%) hoped to achieve masking of their tinnitus, and 31 (19.6%) followed a recommendation from a hearing professional, family member, or people on the Web. Moreover, 29 (18.4%) tried an app to achieve more general goals such as tinnitus relief, find ways of managing their tinnitus, help with tinnitus, and coping with tinnitus, without specifying ways or mechanisms through which that could be achieved. For 15 (9.5%) respondents, the motivator to try an app was desperation and frustration because of tinnitus, and 14 (8.9%) were looking for a source of sound generation and sound enrichment. In addition, 12 (7.6%) respondents reported that they had tried an app because of the convenience (ie, when they travel because of portability), and 10 (6.3%) looked for an alternative to other technologies such as CDs, radio, pillow speakers, or combination aids. Other motivators for trying an app (<5%) included achieving tinnitus reduction or alleviation, distraction, relaxation, reducing stress or anxiety, having an option to stream via hearing aid, variety or choice of sounds in the app, curiosity, free trial of an app, and aiding habituation.

Important Factors When Choosing an App

Ease of use (87/120, 72.5%), followed by trustworthy source (53/120, 44.2%), reviews (47/120, 39.2%), and cost (47/120, 39.2%) were most commonly listed as important factors when choosing an app. Additional factors included recommendation by a medical professional (31/120, 25.9%), recommendation by another person with tinnitus (22/120, 18.3%), and recommendation by friend or family (7/120, 5.9%), followed by name of an app (4/120, 4.2%).

Mobile Apps Rating Scale App Quality Scores

Overall, 3 researchers rated the apps and reliability of objective scales calculated as Cronbach alpha before consensus was .76. Consensus was reached on all the ratings for all the rated apps. Table 2 presents final scores for the 4 subscales (engagement, functionality, aesthetics, and information), overall quality score (mean of 4 subscales), and subjective quality score (satisfaction) for the 18 apps that at least two respondents listed as those that they have tried for management of their tinnitus. Overall, the average MARS quality scores for 18 apps mentioned by at least two respondents varied from 1.5 to 4.2 (out of 5), with scores for individual subscales varying from 1 to 4.6 (Table 2). Subjective scores varied from 1 to 4.1. Of the 4 subscales, functionality had the highest median score (4.4) and aesthetics had the lowest median score (3.15). The *White Noise Free* app had the highest overall MARS score (4.2), followed by *Relax Melodies* (4.1), *Headspace* (4.1), *Oticon Tinnitus Sound* (3.9), and *Sleep Pillow* (3.9). All but 2 apps (*Soothing Sounds Lite* and *Sleep well Hypnosis)* met or exceeded the minimum acceptability score of 3.0.

Characteristics of the Most Popular Apps

The characteristics of the 18 most often mentioned apps (listed by at least three respondents) are summarized in Table 1. Those included 6 apps developed specifically for tinnitus (*Beltone Tinnitus Calmer, Oticon Tinnitus Sound, ReSound Relief, Tinnitus Aid, Tinnitus Balance,* and *Tinnitus Therapy Lite*), 4 apps that were developed for other problems but mentioned tinnitus as one of the possible apps (*myNoise, Relax Noise 3, Soothing Sounds,* and *White Noise Free*), and 8 apps that were developed for other problems and did not mention tinnitus (*Nature Sounds, Rain Rain Sleep Sounds, Relaxed Melodies, Headspace, Sleep Bug, Sleep Pillow, Soothing Sounds Lite,* and *Sleep Well Hypnosis*). *Beltone Tinnitus Calmer, Oticon Tinnitus Sound, ReSound Relief,* and *Tinnitus Balance* were developed by hearing aid manufacturers and included an information that they should be used as a part of a tinnitus management plan provided by hearing care professional.

The tinnitus-specific goals listed by the apps included masking tinnitus (*myNoise, Relax Noise 3, Tinnitus Therapy Lite,* and *White Noise Free*), decreasing the annoyance of tinnitus (*Oticon Tinnitus Sound*), providing temporary relief from tinnitus (*Oticon Tinnitus Sound*), shifting attention away or distracting from tinnitus (*Beltone Tinnitus Calmer, Oticon Tinnitus Sound,* and *ReSound Relief*), managing tinnitus using sound therapy (*Tinnitus Therapy Lite*), helping prevent problems associated with tinnitus (*Soothing Sounds Lite*), easing the problems associated with tinnitus (*Soothing Sounds*), and relieving tinnitus symptoms (*Tinnitus Aid*).

Apps that were not developed specifically for tinnitus but mainly for other problems usually listed multiple goals. A total of 9 apps (*myNoise, Nature Sounds, Rain Rain Sleep Sounds, Relax Melodies, Sleep Bug, Sleep Pillow, Sleep Well Hypnosis, Soothing Sounds,* and *White Noise Free*) addressed sleep problems, including falling and staying asleep, insomnia, and improving quality of sleep. Moreover, 5 apps listed relaxation (*Rain Rain Sleep Sounds, Relax Melodies, Sleep Well Hypnosis, Soothing Sounds,* and *White Noise Free*), and 5 apps (*myNoise, Relax Melodies, Sleep Well Hypnosis, Soothing Sounds,* and *White Noise Free*) listed reducing stress or anxiety as one of the goals. Overall, 5 apps included the aim to block distractions or background noises or mask interruptions or noises one disliked (*myNoise, Relax Noise 3, Sleep Bug, Soothing Sounds,* and *White Noise Free*). Three apps listed increasing focus or improving concentration (*myNoise, Sleep Bug,* and *White Noise Free*), and 2 apps aimed to enhance or increase privacy (*Sleep Bug* and *White Noise Free*). Other listed aims included pacifying fussy and crying babies (*White Noise Free*); soothing headaches and migraines (*White Noise Free*); living a healthier, happier, more enjoyable life (*Headspace*); and helping calm a busy mind (*Sleep Bug*).

All the 18 apps had a free version available; however, *Headspace* and *Tinnitus Therapy Lite* had only a limited demonstration of meditations or sounds available for free, with an option to purchase the pro version. The *Tinnitus Aid app* was a free version of the *Tinnitus HQ app*, with more sounds to choose from and more filters available in the pro version. Of the 18 apps, 7 apps had in-app purchases, which allowed the purchase of a larger selection of sounds and of unlocking more advanced options.

Table 2. The Mobile App Rating Scale mean scores for 18 most cited apps.

Apps	Engagement	Functionality	Aesthetics	Information	Mean	Subjective
White Noise Free	4.2	4.5	4.3	3.7	4.2	4.1
Oticon Tinnitus Sound	3.6	4.5	4	3.4	3.9	3.1
Relax Melodies	4.4	4.5	3.3	4.2	4.1	4
myNoise	3.8	3.25	3	3	3.3	2.6
Tinnitus Therapy Lite	2.2	5	3.3	2.6	3.3	2.6
Headspace	4.6	4.5	3.3	4.2	4.1	4.1
Sleep Bug	3.4	4	3.7	3.3	3.6	2.6
Beltone Tinnitus Calmer	4.6	3.8	3	3.6	3.8	3.3
Sleep Pillow	3.2	5	4	3.3	3.9	3
Soothing Sounds Lite	1.6	2.3	1.7	1	1.6	1
Tinnitus HQ	3.6	3.5	2.7	3.3	3.3	2.4
Tinnitus Balance	3.2	4	2.7	3	3.2	2.3
Rain Rain Sleep Sounds	2.6	4.3	3	3	3.2	2.25
Nature Sounds	3.2	5	3.3	2.7	3.5	2.5
Relax Noise 3	2.4	5	2.7	3.4	3.4	2
ReSound Relief	4.6	3.8	3	3.6	3.8	3.3
Sleep Well Hypnosis	2.6	4	2.3	2.6	2.9	1.9
Zenways	2.6	5	3.3	2.5	3.4	2.4

Overall, 16 apps were updated in 2016 or 2017, with *Zenways* last updated in 2013 and *Relax Noise 3* in 2015. The number of installs ranged from 1000 to 5000 (*Soothing Sounds Lite*) to 5 to 10 million (*Relax Melodies*). The majority of the apps were classified in the health and fitness category (n=11), with 5 apps (*Beltone Tinnitus Calmer, Oticon Tinnitus Sound, ReSound Relief, Tinnitus Aid*, and *Tinnitus Balance*) in the medical category, 1 app (*Zenways*) in Lifestyle, and 1 app classified as medical in the App store and as Lifestyle in Google store. *Oticon Tinnitus Sound* was the only app stating an age limit, which was more than 17 years.

Content and Features of the Apps

Detailed content analysis was conducted for those 18 apps that at least two respondents listed as those they have tried to manage their tinnitus. Full list of apps content and features is available in Multimedia Appendix 4.

Sound

All but 2 apps (n=16) featured sound generation or sound therapy. Sound generation features included a wide selection of sounds. Overall, 1 app offered the possibility to record and loop your own sounds (*White Noise Free*), 7 apps offered an option to import or download additional sounds for free (*Beltone Tinnitus Calmer, myNoise, ReSound Relief*, and *White Noise Free*) or purchase sounds (*Sleep Pillow*) from the app library or your own library (*Oticon Tinnitus Sound* and *Tinnitus Balance*). The *White Noise Free* app linked to the *White Noise Market* app for even more choice of sounds to download. All 16 apps featured volume control for the sounds. Of the 16 apps, 6 apps allowed adjustments beyond volume control such as adjusting sound balance (*Beltone Tinnitus Calmer, ReSound Relief*, and *White Noise Free*), pitch (*White Noise Free*), frequency shaping (*my Noise* and *Tinnitus Aid*), variance (*White Noise Free*), speed (*White Noise Free*), and intensity (eg, a small log fire to a roaring beach fire and *Soothing Sounds Lite*). *myNoise* had a range of frequency shaping options such as animating sounds (ie, zen, subtle, moderate, allegro, and wobbler), setting color of the sound (brown, grey, pink, and white), setting frequency bandwidth (eg, centered around specific frequency), and scene (eg, dark rain, fairy rain, and under the leaves for Rain Noise). Moreover, 15 apps offered endless sounds, with 3 apps using an option to loop sounds (*Relax Melodies, Tinnitus Balance*, and *White Noise Free*) and *Soothing Sounds*, claiming they were using an advanced soundscape generator, which does not loop sounds but generates them in a way that one would not hear the same 10 seconds of sound. In addition, *Relax Melodies* featured a loop correction option, which allowed the user to try other modes in case the pause could be heard in the looped sounds. The remaining 12 apps have not specified how they have achieved the endless sound. *Tinnitus Aid* offered *long high-quality recordings*.

Overall, 9 apps included the possibility to mix different sounds to create personalized *soundscapes*, with all those apps allowing to adjust the volume of the sounds individually and some of the apps allowing to adjust balance (*Beltone Tinnitus Calmer, ReSound Relief*, and *White Noise Free*) and pitch of individual sounds in the mix (*White Noise Free*). In total, 2 apps allowed users to add random sound effects to the sounds (*Nature Sounds* and *Sleep Bug*). *myNoise* had an option to mix sound available only on iPhone.

Overall, 5 apps allowed to rate or mark the favorite sounds and store them in the favorite folder. Of the 5 apps, 4 apps (*Beltone Tinnitus Calmer, Oticon Tinnitus Sound, ReSound Relief,* and *Tinnitus Balance*) allowed the user to create a personalized sound plan and organize the sounds according to sound type, for example, soothing, interesting, or background (*Beltone Tinnitus Calmer, Oticon Tinnitus Sound, Tinnitus Balance,* and *ReSound Relief*) or according to situations when a particular sounds are preferred (*Oticon Tinnitus Sound* and *Tinnitus Balance*).

In total, 4 apps included binaural beats or isochronic tones in the free versions (*myNoise, Relax Melodies, Soothing Sounds,* and *Zenways*). *Relax Melodies* offered 6 different frequencies, between 2.5 and 20 Hz, of binaural beats, which can affect the brain in different ways. For example, the description in the app suggests that 2.5-Hz delta wave *helps you reach the deepest portion of your sleep cycle*, whereas 10-Hz mid-alpha wave *helps to calm and relax your mind after you have been active*. *myNoise* includes *Binaural Beat Machine*, with 10 carriers between 1 and 32 Hz to induce a particular mental state (eg, deep sleep, relaxed, conscious, and focused). *Soothing Sounds Lite* contained binaural sounds to *help improve concentration as well as relaxation*, with a slider to adjust tones' frequency.

Overall, 12 apps had a feature to play sound in the background while using other apps (*Beltone Tinnitus Calmer, myNoise, Oticon Tinnitus Sound, Rain Rain Sleep Sounds, Relax Melodies, Relax Noise 3, ReSound Relief, Sleep Bug, Sleep Pillow, Tinnitus Aid, Tinnitus Balance,* and *Zenways*).

Tinnitus Balance app used sound in the context of specific management program (Progressive Tinnitus Management).

Meditation and Mindfulness

A total of 5 apps featured meditation and mindfulness (*Beltone Tinnitus Calmer, Headspace, Relax Melodies, ReSound Relief,* and *Zenways*), with all 5 featuring guided meditation and 2 using imagery (*Beltone Tinnitus Calmer* and *ReSound Relief*). In 2 apps, meditation and mindfulness was the main focus of the app (*Headspace* and *Zenways*), whereas, in 3 apps, it was one of the features alongside other components.

Relax Melodies offered guided meditation programs and single sessions to help sleep. *Beltone Tinnitus Calmer* and *ReSound Relief* offered 6 guided meditation sessions to practice techniques for managing stress and tension caused by tinnitus. *Headspace* offered a wide selection of themed meditations on a variety of topics (eg, depression, self-esteem, stress, cancer, sleep, pregnancy, and anxiety); however, it is worth noting that the only free option is a 10-day meditation program that *taught you the essentials of living a healthier, happier life*. *Zenways* offers mindfulness of the breath meditation to *help relax and find your Zen*.

Relaxation Exercises

A total of 3 apps included relaxation exercises (*Beltone Tinnitus Calmer, Oticon Tinnitus Sound,* and *ReSound Relief*) alongside other components. All 3 apps had breathing exercises, where the task was to breathe in sync with expanding and collapsing bauble in the screen, together with the voice asking you to breathe in and breathe out. There was also an option to set the tempo of the breathing as deep, slow, or normal in *Beltone Tinnitus Calmer* and *ReSound Relief* and number of breaths per minute in *Oticon Tinnitus Sound*. *Oticon Tinnitus Sound* also featured muscle relaxation, which asks to tense and relax certain group of muscles according to spoken instructions. As all 3 apps containing relaxation exercises were apps developed specifically for tinnitus, the aim of the relaxation was to counteract tension and stress caused by tinnitus and in return notice tinnitus less.

Elements of Cognitive Behavior Therapy

Two apps included elements of cognitive behavioral therapy (CBT; *Beltone Tinnitus Calmer* and *ReSound Relief*). One of those apps was changing unpleasant thoughts about tinnitus into something less upsetting, including lack of help, tinnitus ruining life, loud tinnitus or bad day, lack of support from partner, tinnitus getting worse, and lack of understanding. The second CBT element was pleasant activities, where the users are asked to nominate activities they would like to do such as meet a friend for tea, learn a new skill, or play music and receive weekly reminders to do them, on the basis that doing things they enjoy makes life with tinnitus easier.

Information and Education

Information and education within the apps included information about tinnitus, using sound for the management of tinnitus, including binaural beats, sleep hygiene, insomnia and its causes, and meditation and mindfulness. *Beltone Tinnitus Calmer* and *ReSound Relief* apps provided information about tinnitus. The apps included the separate section entitled "What is tinnitus?," covering such topics as how tinnitus is defined, prevalence of tinnitus, causes of tinnitus, what can be done, how to live with tinnitus, and common therapies. *Tinnitus Balance* app contains brief information regarding the prevalence of tinnitus. Overall, 5 apps contained information about using sound for management of tinnitus. *Beltone Tinnitus Calmer, Oticon Tinnitus Sound, ReSound Relief,* and *Tinnitus Balance* explain the different role that soothing, interesting, and background sounds can play in the management of tinnitus. *Tinnitus Therapy Lite* contained a description of their tinnitus relief sounds. *Beltone Tinnitus Calmer* and *ReSound Relief* offered information about sleep hygiene, including sections on eating and drinking, relaxing before bedtime, sleep behavior, sleeping environment, and timing. *Sleep Well Hypnosis* included a spoken introduction at the beginning of the hypnosis session, explaining what is insomnia and possible causes of insomnia. *Relax Melodies* app described the role of different frequencies of binaural beats from 2.5 to 20 Hz. *myNoise* included an explanation of what binaural beats do, but this was only available on the iPhone. The *Headspace* app had a link to a short video explaining what mediation and mind training is.

Overall, 7 apps had weblinks to more information or app help and troubleshooting. *Tinnitus Therapy Lite* included a link to the Sound Oasis website, including more information on tinnitus and how sound therapy can help. *Beltone Tinnitus Calmer* and *ReSound Relief* included a link to the ReSound GN website, with information about their hearing and tinnitus products, links to national tinnitus charities and associations, treatment centers, and information resources about tinnitus management options.

Sleep Well Hypnosis included a link to answers to *hypnosis questions* such as "how long will it take to notice changes," "how does hypnosis work," and "will I lose control while I am under hypnosis." *myNoise* had a link to *myNoise* on the Web, with detailed information about noise generators, their calibration, extensive sounds library, and using noise for different purposes (at the office, studying, tinnitus, hyperacusis, and relaxation). *Relax Noise 3* had a link to the *Relaxed Noise 3* website, with information about the 3 different types of noises used in the app: white, pink, and red, using those sounds to aid concentration, as tinnitus maskers or noisers, for meditation and as a sleeping aid. *Sleep Well Hypnosis* app features Sleep Booster with binaural beats *to induce your brainwave frequency into an optimal state for deep, restorative sleep*; however, that option is only available in the pro version of the app. Moreover, 6 apps included a help section or brief introduction to an app within the app (available offline).

Hypnosis

Sleep Well Hypnosis features a single 25-min hypnosis audio session read by a certified hypnotherapist, which aims to *help reduce anxious thoughts and prepare the mind for deeper, more restorative rest*. This can be combined with background music and sleep booster, with binaural beats (only in pro version).

Nonauditory Stimuli

Beltone Tinnitus Calmer and *ReSound Relief* contain some secondary stimuli, that is, colors. The role of those was described as " keeping your mind occupied." For each of the soundscapes, there was a possibility of choosing *color mood*, which would be displayed while playing the sound.

Nature Sounds, *Sleep Bug*, *Sleep Pillow*, and *Tinnitus Aid* pointed in the app descriptions to using high-quality graphics. However, some other apps that did not explicitly specify that feature also featured high-quality images (eg, *White Noise Free*).

Technical Features

All 18 apps did not require streaming, but instead the content was downloaded to the device and worked offline. Of the 18 apps, 7 apps (*Beltone Tinnitus Calmer, Oticon Tinnitus Sound, Relax Melodies, ReSound Relief, White Noise Free, Sleep Bug*, and *Sleep Well Hypnosis*) had remote controls allowing to adjust volume (*White Noise Free*) and/or pause or start or close the apps while on the screen lock.

Moreover, 5 apps had different options for sharing, with *White Noise Free* featuring the most advanced sharing options of all the apps. These included the possibility of the user sharing their own recordings and mixes and photos. Sharing recordings or mixes is possible via *White Noise Market app*, which connects you to an app community or via email. *Headspace* allows you to invite up to 5 *buddies* through an email message, sends the information about the app, a short video, and link to the website. The *buddies* system allows you to access your buddies' statistics and progress and motivate them if they fail to meet the goals. *Relax Melodies, Sleep Well Hypnosis, Tinnitus Aid*, and *Zenways* had an option to share the link to the app, for example, via email or messaging apps. *myNoise* and *White Noise Free* featured their own app communities where you can upload, download, and/or rate different sounds and/or post comments.

Overall, 13 apps were advert free (*Beltone Tinnitus Calmer, Headspace, myNoise, Nature Sounds, Oticon Tinnitus Sound, Rain Rain Sleep Sounds, Relax Noise 3, ReSound Relief, Sleep Bug, Sleep Well Hypnosis, Tinnitus Balance, Tinnitus Therapy Lite*, and *Zenways*), whereas 3 apps featured advertisements on the small stripe at the top or bottom of the screen, not interfering with the apps content (*Relax Melodies, Sleep Pillow*, and *White Noise Free*). *Soothing Sounds Lite* was the only app where the adverts took considerable space on the screen, making it difficult to navigate.

Beltone Tinnitus Calmer, *Headspace*, *Relax Melodies*, *ReSound Relief*, and *Tinnitus Balance* offered progress or usage tracking. *Headspace* captured the total time spent on meditation, number of sessions completed, and average duration. *Beltone Tinnitus Calmer* and *ReSound Relief* captured total hours used and separately time spent using sounds and exercises. *Tinnitus Balance* reported average usage per day, percentage time spent on sounds sorted according to sound type and sounds sorted according to situation down, to percentage of time spent listening to individual sounds. *Relax Melodies* had an option to track *mindful minutes* using the Apple Health app.

A total of 10 apps were available in multiple language options in addition to English (*Beltone Tinnitus Calmer, Nature Sounds, Oticon Tinnitus Sound, Rain Rain Sleep Sounds, Relax Melodies, ReSound Relief, Sleep Pillow, Soothing Sounds Lite, Tinnitus Aid*, and *Tinnitus Balance*) and *White Noise free* was available in English and Spanish.

Overall, 15 apps featured a timer for controlling the length of the sound or meditation sessions (*Zenways*). Some of them had an option to fade audio out (*Relax Noise 3*, *Tinnitus Aid*, and *White Noise Free*). The *Headspace* app did not feature a timer and had a predefined meditation sessions length. *Soothing Sounds Lite* did not have a timer meaning that sounds would play until they were turned off. The timer option was displayed but not accessible in the free version of *Sleep Well Hypnosis* app. Moreover, 3 apps had a clock (*White Noise Free, Sleep Bug*, and *Relax Melodies*— iPhone only), 3 had alarms (*White Noise Free, Soothing Sounds Lite*, and *Relax Melodies*— iPhone only), and 1 had date display (*Sleep Bug*). Two apps (*Rain Rain Sleep Sounds* and *Relax Melodies*) featured bedtime reminders allowing the user to set days and times for going to sleep.

In the description in the app store, *White Noise Free* claimed to feature swipe gesture support for navigating sound collection, it is not clear, however, how that differed from other tested apps. *Sleep Bug* claimed to use accessibility support but did not specify in what way. *Sleep Bug* claimed *great user support* but again did not clarify what it would feature.

Discussion

Principal Findings

This study generated the list of 55 apps that people used for the management of their tinnitus and explored reasons for app nonuse as well as motivators for using apps for tinnitus

management. The main reason for app nonuse was lack of awareness of their existence. Quality assessment of the 18 most popular apps using MARS resulted in a range of mean scores from 1.6 to 4.2 (out of 5), depending on an app. Sound was the main component of the majority of the apps chosen by people with tinnitus.

The data from the Office of National Statistics [17] showed an increasing use of the internet by those aged over 65 years despite that this group has been consistently the lowest users of internet over the years. A similar pattern was found for apps' use, with the app users group being slightly younger than the apps nonusers group. The users group showed a lower proportion of people aged 65 years and older and a higher proportion of people aged 45 to 54 years. However, there was still a considerable proportion of people aged over 65 years who were app users.

People listed a large number of apps used for tinnitus management. A majority of the 55 apps were mentioned by only 1 person. There might be several reasons for such variability. First, people were looking for apps for a range of different reasons, including helping with sleep, masking, sound enrichment, distraction, relaxation, and reducing stress or anxiety; therefore, they were choosing apps that were addressing those specific goals or problems. This also explains why the majority of apps used by people were developed for other problems rather than for tinnitus. Second, only 19% of respondents reported that they followed a recommendation when choosing an app, suggesting that the majority of respondents found apps through a search in the app stores. A quick search for tinnitus apps in the Google Play Store returned 248 apps available for download. The large number of apps that can potentially be useful for the management of tinnitus, although encouraging, also poses a challenge for people with tinnitus and hearing care professionals equally. The search results for apps in the app store can be overwhelming, with several hundred apps available when searching for *tinnitus*. Without a clear criteria or guidance on which apps to choose, it is not surprising that people tended to choose different apps based on the reviews or personal preference. Given that ease of use was listed as the most important factor when choosing an app, it would not be surprising if personal preference played a main role in the choice of apps.

Average MARS quality scores for the 18 most cited apps varied greatly with 2 apps not meeting the minimum acceptability score of 3. None of the apps received the maximum score of 5. The lowest median score was for the aesthetics subscale, which asked questions about layout, graphics quality, and visual appeal. The Functionality subscale had a highest median score, with questions about performance, ease of use, navigation, and gestural design. This is in line with the results of previous studies using MARS for quality assessment of weight management apps [18], prevention of driving after drinking alcohol apps [19], and mindfulness-based apps [20], where functionality scores were also the highest.

In line with the current model of tinnitus management, where sound is the main component of the majority of tinnitus management strategies and programs, sound was also the main focus of the majority of apps. Current sound therapy options available on the NHS include various devices that play sounds, including sound generators and combination aids. However, those devices can only play a limited number of sounds; therefore, not all patients find an option that helps with their tinnitus. As the choice of sound options that can be delivered via many of the apps is large and very often sounds are customizable, it is much more likely that the individual patient's needs regarding sound therapy will be met via this option, allowing for personalization of a tinnitus management plan.

About 20% of people with tinnitus experience symptoms that affect their quality of life. They might experience disturbed sleep, hearing and concentration problems, social isolation, anxiety, depression, irritation, or stress. It is, therefore, not surprising that people listed apps addressing those problems through meditation and mindfulness, relaxation exercises, and elements of CBT. However, it was noted during the quality assessment using the MARS scale that some of the content might not be appropriate for people with tinnitus to access without guidance from a health care professional. Specifically, *Beltone Tinnitus Calmer* and *ReSound Relief* apps have a section giving examples of negative thoughts, which without a proper explanation might potentially have a negative impact on the user.

Strengths and Limitations

This study is the first one to review mobile phone apps for the management of tinnitus. It is the first study to assess the quality of apps used for tinnitus management using the MARS scale. Apps were tested on both iPhone iOS and Android platforms. Expert ratings on 30% of the reviewed apps had a high-level interrater reliability in this study.

Given the large number of apps for tinnitus management and the fact that people with tinnitus use both tinnitus-specific apps and apps developed for other problems, we have undertaken a bottom-up approach, rather than a systematic search in the apps stores. The strength of such approach is that we were able to identify apps not developed specifically for tinnitus that people use and that might potentially be useful for tinnitus management. Moreover, one of these apps that would not be identified by simple search was *Headspace*: Guided Meditation & Mindfulness. On the other hand, there might be some apps that were missed from our list. Given that this is the first study looking at apps for tinnitus, it seemed the best approach to, in the first instance, look at the apps currently used by people with tinnitus.

Future Research

Our study showed that people use apps for the management of their tinnitus; however, this is done mostly as a self-help option, without conjunction with management provided by hearing health care professionals. Future research should look at the possibility of incorporating apps into the management of tinnitus by health care professionals and creating guidelines for the use of apps as a part of a tinnitus management plan. Further research involving patients and clinicians on the desired content and usability features of apps for tinnitus management should be conducted. There is no evidence for the efficacy of apps for the management of tinnitus, and none of the listed apps were

assessed for efficacy for tinnitus management in a trial. Future research is needed to determine the efficacy of apps for management of tinnitus.

Conclusions

Further research should consider the place for apps in the tinnitus management (stand-alone self-management intervention vs part of the management by a hearing professional). As content of the apps varies in respect to sound options, information, and management strategies, it seems that the choice of the best management app should be guided by individual patient needs and preferences.

Acknowledgments

This study presents independent research supported by the NIHR. The views expressed in this study are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health and Social Care.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Checklist for Reporting Results of Internet E-Surveys (CHERRIES).

[PDF File (Adobe PDF File), 60KB - mhealth_v7i1e10353_app1.pdf]

Multimedia Appendix 2

Coding manual developed for content analysis of the apps.

[PDF File (Adobe PDF File), 42KB - mhealth_v7i1e10353_app2.pdf]

Multimedia Appendix 3

Characteristics of the 55 apps listed by respondents as those they have used for management of their tinnitus.

[PDF File (Adobe PDF File), 75KB - mhealth_v7i1e10353_app3.pdf]

Multimedia Appendix 4

Content analysis and features of the 18 apps that at least two respondents listed as those they have tried to manage their tinnitus: (1) White Noise Free, (2) Oticon Tinnitus Sound, (3) Relax Melodies: Sleep Sounds, (4) myNoise, (5) Tinnitus Therapy Lite, (6) Headspace: Guided Meditation & Mindfulness, (7) Sleep Bug: White Noise Soundscapes & Music Box, (8) Beltone Tinnitus Calmer, (9) Sleep Pillow, (10) Soothing Sounds Lite, (11) Tinnitus Aid: Nature sounds to mask ear ringing, (12) Tinnitus Balance, (13) Rain Rain Sleep Sounds, (14) Nature Sounds, (15) Relax Noise 3, (16) ReSound Relief, (17) Sleep Well Hypnosis, and (18) Zenways. V—content or feature present in an app.

[PDF File (Adobe PDF File), 103KB - mhealth_v7i1e10353_app4.pdf]

References

1. Hoare DJ, Searchfield GD, El RA, Henry JA. Sound therapy for tinnitus management: practicable options. J Am Acad Audiol 2014 Jan;25(1):62-75. [doi: 10.3766/jaaa.25.1.5] [Medline: 24622861]
2. Adams D, Lee P. Deloitte. State of the smart Consumer business usage patterns, Global Mobile consumer survey 2017 Uk Cut URL: https://www.deloitte.co.uk/mobileuk2017/assets/img/download/global-mobile-consumer-survey-2017_uk-cut.pdf [accessed 2018-11-22] [WebCite Cache ID 747WPgTT8]
3. The World Bank. 2018 Feb 23. Mobile Applications for the Health Sector URL: http://siteresources.worldbank.org/INFORMATIONANDCOMMUNICATIONANDTECHNOLOGIES/Resources/mHealth_report.pdf [accessed 2018-02-23] [WebCite Cache ID 6xRZITPjf]
4. Misra S, Lewis TL, Aungst TD. Medical application use and the need for further research and assessment for clinical practice: creation and integration of standards for best practice to alleviate poor application design. JAMA Dermatol 2013 Jun;149(6):661-662. [doi: 10.1001/jamadermatol.2013.606] [Medline: 23783150]
5. Lewis TL. A systematic self-certification model for mobile medical apps. J Med Internet Res 2013;15(4):e89 [FREE Full text] [doi: 10.2196/jmir.2446] [Medline: 23615332]
6. Murfin M. Know your apps: an evidence-based approach to evaluation of mobile clinical applications. J Physician Assist Educ 2013;24(3):38-40. [Medline: 24261171]

7. Charani E, Castro-Sánchez E, Moore LS, Holmes A. Do smartphone applications in healthcare require a governance and legal framework? It depends on the application!. BMC Med 2014;12:29 [FREE Full text] [doi: 10.1186/1741-7015-12-29] [Medline: 24524344]
8. mHealth Task Force. mHealth Task Force. 2012 Sep 24. Findings and Recommendations: Improving care delivery through enhanced communications among providers, patients, and payers URL: http://www2.itif.org/2012-mhealth-taskforce-recommendations.pdf [accessed 2018-02-23] [WebCite Cache ID 6xRYzdAnV]
9. Aitken M, Gauntlett C. IQVIA. 2013. Patient Apps for Improved Healthcare: From Novelty to Mainstream URL: https://www.iqvia.com/ [accessed 2018-02-23] [WebCite Cache ID 6xRbS2uXs]
10. NHS. NHS Apps Library URL: https://apps.beta.nhs.uk/ [accessed 2018-02-23] [WebCite Cache ID 6xRaRNtXY]
11. Paglialonga A, Tognola G, Pinciroli F. Apps for hearing science and care. Am J Audiol 2015 Sep;24(3):293-298. [doi: 10.1044/2015_AJA-14-0093] [Medline: 26649533]
12. Eysenbach G. Improving the quality of web surveys: the checklist for reporting results of Internet e-surveys (CHERRIES). J Med Internet Res 2004 Sep 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
13. 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;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
14. Martinez C, Wallenhorst C, McFerran D, Hall DA. Incidence rates of clinically significant tinnitus: 10-year trend from a cohort study in England. Ear Hear 2015;36(3):e69-e75 [FREE Full text] [doi: 10.1097/AUD.0000000000000121] [Medline: 25470370]
15. Sanchez TG, Medeiros IR, Levy CP, Ramalho JD, Bento RF. Tinnitus in normally hearing patients: clinical aspects and repercussions. Braz J Otorhinolaryngol 2005;71(4):427-431 [FREE Full text] [doi: 10.1016/S1808-8694(15)31194-0] [Medline: 16446955]
16. Schaette R, McAlpine D. Tinnitus with a normal audiogram: physiological evidence for hidden hearing loss and computational model. J Neurosci 2011 Sep 21;31(38):13452-13457 [FREE Full text] [doi: 10.1523/JNEUROSCI.2156-11.2011] [Medline: 21940438]
17. Office of National Statistics. United Kingdom; 2018. URL: https://www.ons.gov.uk/ [accessed 2018-02-23] [WebCite Cache ID 6xRZy8gRE]
18. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int J Behav Nutr Phys Act 2016;13:35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
19. Wilson H, Stoyanov SR, Gandabhai S, Baldwin A. The quality and accuracy of mobile apps to prevent driving after drinking alcohol. JMIR Mhealth Uhealth 2016 Aug 08;4(3):e98 [FREE Full text] [doi: 10.2196/mhealth.5961] [Medline: 27502956]
20. Mani M, Kavanagh DJ, Hides L, Stoyanov SR. Review and evaluation of mindfulness-based iPhone apps. JMIR Mhealth Uhealth 2015;3(3):e82 [FREE Full text] [doi: 10.2196/mhealth.4328] [Medline: 26290327]

Abbreviations

BRC: Biomedical Research Center
BTA: British Tinnitus Association
CBT: cognitive behavioral therapy
MARS: Mobile Apps Rating Scale
mHealth: mobile health
NHS: National Health Service
NIHR: National Institute for Health Research
TRT: tinnitus retraining therapy

Edited by G Eysenbach; submitted 21.03.18; peer-reviewed by B Arnoldussen, L Cheong; comments to author 25.08.18; revised version received 27.09.18; accepted 27.09.18; published 23.01.19.

Please cite as:
Sereda M, Smith S, Newton K, Stockdale D
Mobile Apps for Management of Tinnitus: Users' Survey, Quality Assessment, and Content Analysis
JMIR Mhealth Uhealth 2019;7(1):e10353
URL: http://mhealth.jmir.org/2019/1/e10353/
doi:10.2196/10353
PMID:30672739

Original Paper

Mobile Phone Apps Targeting Medication Adherence: Quality Assessment and Content Analysis of User Reviews

Jamie Yea Eun Park[1,2]; Jenny Li[1,2], BSc; Alyssa Howren[1,2,3], MSc; Nicole Wen Tsao[1,2,3], BSc, PhD; Mary De Vera[1,2,3], PhD

[1]Faculty of Pharmaceutical Sciences, University of British Columbia, Vancouver, BC, Canada

[2]Collaboration for Outcomes Research and Evaluation, Vancouver, BC, Canada

[3]Arthritis Research Canada, Vancouver, BC, Canada

Corresponding Author:
Mary De Vera, PhD
Faculty of Pharmaceutical Sciences
University of British Columbia
2405 Wesbrook Mall
Vancouver, BC, V6T 1Z3
Canada
Phone: 1 604 827 2138
Fax: 1 604 207 4059
Email: mdevera@mail.ubc.ca

Abstract

Background: With the accessibility and widespread use of mobile phones, mobile phone apps targeting medication adherence may be useful tools to help patients take medications as prescribed.

Objective: Our objectives were to (1) characterize and assess mobile phone medication adherence apps guided by a conceptual framework on the focus of adherence interventions and (2) conduct a content analysis of Web-based reviews to explore users' perspectives and experiences with mobile phone medication adherence apps.

Methods: We searched for mobile phone medication adherence apps using keyword searches in Apple and Android operating systems. We characterized all apps in terms of number of downloads, ratings, languages, cost, and disease target. We categorized apps according to 4 key features of (1) alerting to take medication, (2) tracking medication taking, (3) reminding to refill or indicating amount of medication left, and (4) storing medication information. We then selected representative apps from each operating system for detailed quality assessment and user testing. We also downloaded Web-based reviews for these selected apps and conducted a qualitative content analysis using an inductive approach involving steps of initial open coding, construction of categories, and abstraction into themes.

Results: We identified 704 apps (443 from Apple and 261 from Android). The majority of apps across both operating systems had 1 or 2 features—specifically, 37.2% (165/443) and 38.1% (169/443) of Apple apps, respectively, and 41.4% (108/261) and 31.4% (108/261) of Android apps, respectively. Quality assessment and user testing of 20 selected apps revealed apps varied in quality and commonly focused on behavioral strategies to enhance medication adherence through alerts, reminders, and logs. A total of 1323 eligible Web-based reviews from these 20 selected apps were analyzed, and the following themes emerged: (1) features and functions appreciated by users, which included the ability to set up customized medication regimen details and reminders, monitor other health information (eg, vitals, supplements, and manage multiple people or pets), support health care visits (eg, having a list of medications and necessary health information in 1 app); (2) negative user experiences that captured technical difficulties (glitches, confusing app navigation, and poor interoperability), dosage schedule, and reminder setup inflexibility; and (3) desired functions and features related to optimization of information input, improvement of reminders, and upgrading app performance (better synchronization or backup of data and interoperability).

Conclusions: A large number of mobile phone medication adherence apps are currently available. The majority of apps have features representing a behavioral approach to intervention. Findings of the content analysis offer mostly positive feedback as well as insights into current limitations and improvements that could be addressed in current and future medication adherence apps.

(JMIR Mhealth Uhealth 2019;7(1):e11919) doi:10.2196/11919

KEYWORDS

medication adherence; mobile apps; mHealth

Introduction

Background

For many patients living with lifelong diseases, taking medications as prescribed is challenging. The World Health Organization has declared medication nonadherence as an epidemic and has called for feasible, patient-tailored solutions [1]. Particularly relevant, there has been surging interest in the use of mobile phones in public health practice (mobile health, mHealth) to address medication nonadherence, given their widespread use [2,3]. Indeed, mobile phones may represent a patient-centered means of targeting medication adherence, with features such as sending alerts to take medications, tracking doses, and supporting medication instructions.

A number of prior reviews have identified and described mobile phone medication adherence apps. In 2013, Dayer et al identified 160 apps on the Apple, Android, and Blackberry operating systems and subsequently published an update in 2017 with 645 apps to include those on the Windows operating system [4,5]. In 2014, Bailey et al extracted information on functions of 424 apps from app descriptions [3]. In 2016, Heldenbrand et al and Santo et al identified 347 and 272 apps, respectively, and categorized them based on author-identified features [6,7]. In 2017, Haase et al found 30 apps and classified ideal app features used to improve medication adherence [8]. In 2018, Ahmed et al analyzed 681 identified apps using app repository overviews or websites and mentioned the lack of health care professional involvement in the development of medication adherence apps [9]. These prior studies have incorporated various evaluation methods to assess app features, including using author-created rating systems [4,5,7,9], existing rating scales (eg, Mobile App Rating Scale and checklist for developing health literate mHealth apps endorsed by Institute of Medicine) [3,5,6,7,10], and user testing [4,5,9].

Although prior assessments of mobile phone medication adherence apps have added insight into these tools, they remain limited for various reasons, including evaluations based on app descriptions, short periods of trial, and user testing based on free versions. Furthermore, evaluations of app reviews have been limited [11,12]. Indeed, app reviews posted by the target users are publicly accessible and add to a valuable, naturally generated pool of data that to date have not been fully utilized. Altogether, the constantly growing number of mobile phone users [11] along with greater recognition of the problem of medication adherence in recent years [12] necessitates an update to aforementioned prior studies. In addition, an expansion of the knowledge on user experiences is needed.

Objectives

As such, our objectives were to (1) characterize and assess mobile phone medication adherence apps guided by a conceptual framework on the focus of adherence interventions and (2) conduct a content analysis of Web-based reviews to explore users' perspectives and experiences with mobile phone medication adherence apps.

Methods

Identification of Mobile Phone Medication Adherence Apps

We searched for mobile phone medication adherence apps on Apple (iTunes store) and Android (Google Play) operating systems during the month of May 2017. We applied 8 keywords ("medication," "adherence," "compliance," "dose," "dosing," "drugs," "reminder," and "pills") and did not impose inclusion criteria to ensure the broadest capture possible. However, apps were excluded if they were associated with services of specific pharmacies or businesses or had a primary purpose of advertising or other similar commercial activities. We downloaded each included app and extracted information on the operating system, number of downloads, rating, language, cost to download, and disease target.

Characterizing and Assessing the Quality of Mobile Phone Medication Adherence Apps

To guide characterization of apps, we applied the conceptual framework on the theoretical targets of adherence interventions by assigning app features as *educational* (targets adherence by conveying information), *behavioral* (targets adherence by targeting, shaping or reinforcing specific behavior patterns), or *affective* (targets adherence through appeals to feelings and emotions or social relationships) [13,14] (see Multimedia Appendix 1). We then selected 10 representative apps from each operating system to conduct quality assessment and user testing. For Android, the primary selection criterion was number of downloads (≥100,000 downloads); in the event of a tie, the app with a higher average rating was selected. For Apple, as information on number of downloads is not provided, the selection criteria were on ratings followed by search retrieval order. Moreover, 2 authors who are clinically trained as pharmacists (JL and NWT) independently assessed the quality of selected apps using iPhone 5 (iOS9) and Samsung Galaxy Note 4 (Android Version 6.0.1) based on 12 features of [4] alerting to take medication; tracking medication taking (*behavioral*); reminding to refill or indicating amount of medication left (*behavioral*); storing medication information (*educational*); complex medication instructions or notes; database of medications; backup, cloud access, or means of access through another device; exportation or printing of data; free to download; alerts do not require internet connection; multiple profiles or patients; and multiple languages. In addition, ease of use was rated according to 3 levels: (1) easy—uses nontechnical language and involves functions that facilitate use (ie, drop-down menus) to minimize input, with usage of app usage learned in 5 min or less; (2) moderate—also involves functions that facilitate use but requiring greater degree of input, with usage of app learned in over 5 min but not less than 15 min; and (3) difficult—uses technical (medical or scientific) language, involves multiple functions, and requires substantial

input, with usage of app learned in over 15 min. Authors discussed independently conducted quality assessments to come to a consensus for final reporting.

Content Analysis of User Reviews

A qualitative content analysis was conducted on Web-based reviews for the 20 aforementioned apps. Specifically, we extracted user reviews submitted in English between January 1, 2017 and January 1, 2018 published on the official iOS app (Apple) and Google Play (Android) store and imported these into NVivo 11 (QSR International). We conducted a qualitative content analysis using an inductive approach and followed 3 main coding steps of (1) initial open coding, (2) construction of categories, and (3) abstraction into themes [15]. The constant comparative method was applied throughout the coding process [16]. We reached data saturation, a point of redundancy during the data analysis where no new concepts contributing to categories and themes arise [17], by the time the reviews for the twelfth app were coded.

Results

Identification and Quality Assessment of Mobile Phone Medication Adherence Apps

Our search strategy identified a total of 878 apps across both Apple and Android operating systems. After applying all exclusion criteria, 704 apps, with 443 from Apple and 261 from Android, were included (as shown in Figure 1).

Table 1 summarizes the characteristics of included mobile phone medication adherence apps, including number of downloads, rating, language, and cost of downloading. The majority of apps across both operating systems had 1 or 2 features—specifically, 37.2% (165/443) and 38.1% (169/443) of Apple apps, respectively, and 41.4% (108/261) and 31.4% (108/261) of Android apps, respectively. Four-set Venn diagrams showing the possible combination of the 4 key features for included Apple and Android apps are shown in Figure 2.

Figure 1. Flow of smartphone medication adherence apps included and most commonly used keywords (does not add up to number of apps since multiple keywords may be used to identify an app).

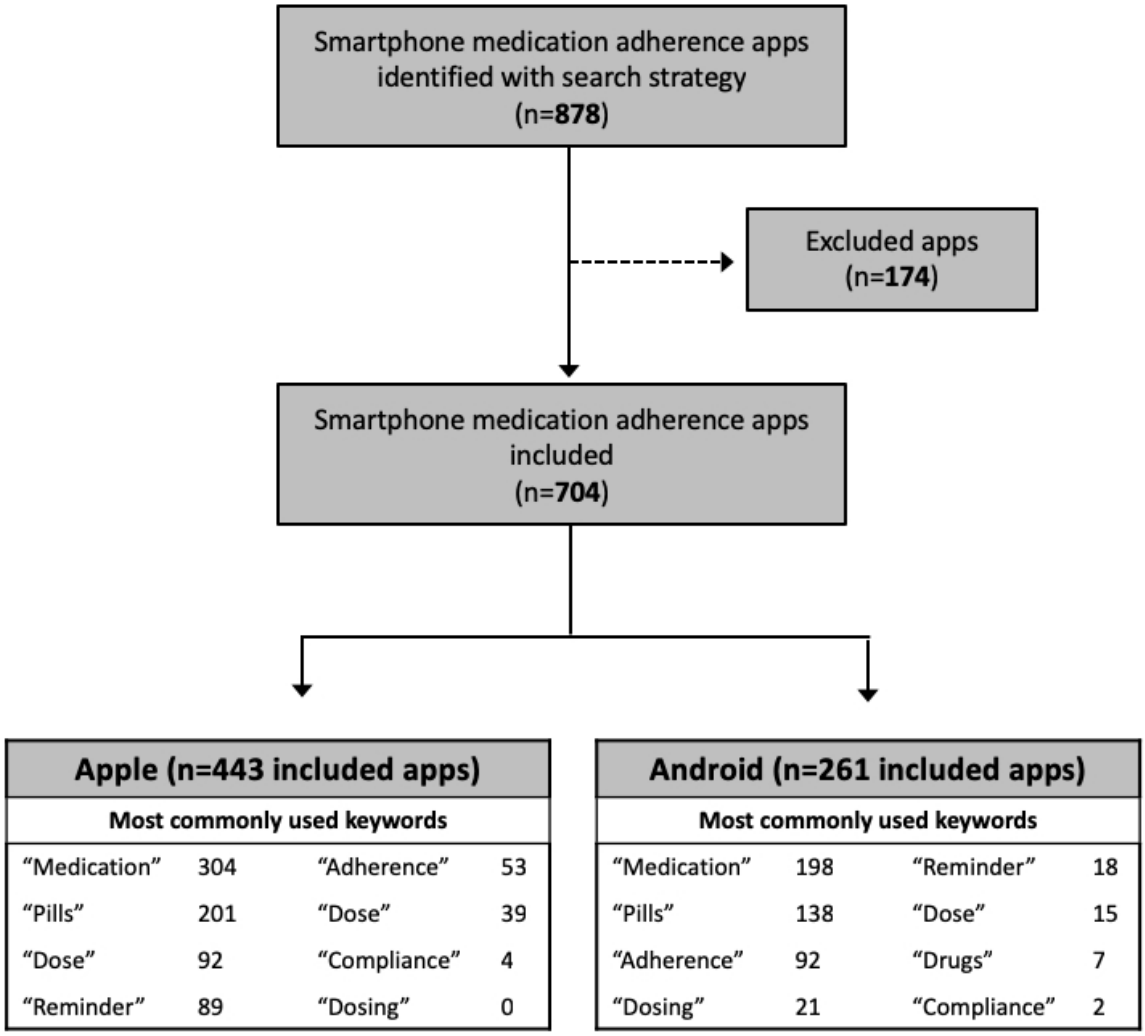


Table 1. Characteristics of included mobile phone medication adherence apps.

Characteristics	Apple (N=443), n (%)	Android (N=261), n (%)
Number of downloads[a]		
≤100 or unspecified	N/A[b]	83 (31.8)
Approximately 500 to 1000	N/A	79 (30.3)
Approximately 5000 to 10,000	N/A	66 (25.3)
Approximately 50,000 to 100,000	N/A	20 (7.7)
>100,000	N/A	13 (5.0)
Ratings[c]		
≤1/5	0 (0.0)	4 (1.5)
1/5<rating≤2/5	3 (0.7)	5 (1.9)
2/3<rating≤3/5	0 (0.0)	17 (6.5)
3/5<rating≤4/5	2 (0.5)	104 (4.0)
4/5<rating≤5/5	7 (1.6)	98 (37.5)
Unrated	431 (97.3)	39 (14.9)
Languages[d]		
English	443 (100.0)	261 (100.0)
German	83 (18.7)	0 (0.0)
Spanish	83 (18.7)	0 (0.0)
French	74 (16.7)	0 (0.0)
Japanese	41 (9.3)	0 (0.0)
Russian	40 (9.0)	0 (0.0)
Simplified Chinese	37 (8.4)	0 (0.0)
Traditional Chinese	24 (5.4)	0 (0.0)
Cost of download (US$)[e]		
0.00	347 (78.3)	225 (86.2)
0.00<cost≤1.00	2 (0.5)	4 (1.5)
1.00<cost≤5.0	76 (17.2)	28 (10.7)
cost>5.00	17 (3.8)	4 (1.5)
Target		
General	328 (74.0)	200 (76.6)
Contraceptives	32 (7.2)	35 (13.4)
Asthma or chronic obstructive pulmonary disease	12 (2.7)	4 (1.5)
Epilepsy	9 (2.0)	3 (1.1)
Psychiatry	7 (1.6)	3 (1.1)
Diabetes	3 (0.7)	3 (1.1)
Other[f]	52 (11.7)	13 (5.0)
Number of key features		
4	34 (7.7)	21 (8.0)
3	75 (16.9)	48 (18.4)
2	169 (38.1)	82 (31.4)
1	165 (37.2)	108 (41.4)
Other	0 (0.0)	2 (0.8)

[a]Numbers are approximated to "0", "1", or "5" in each digit.

[b]N/A: not applicable.

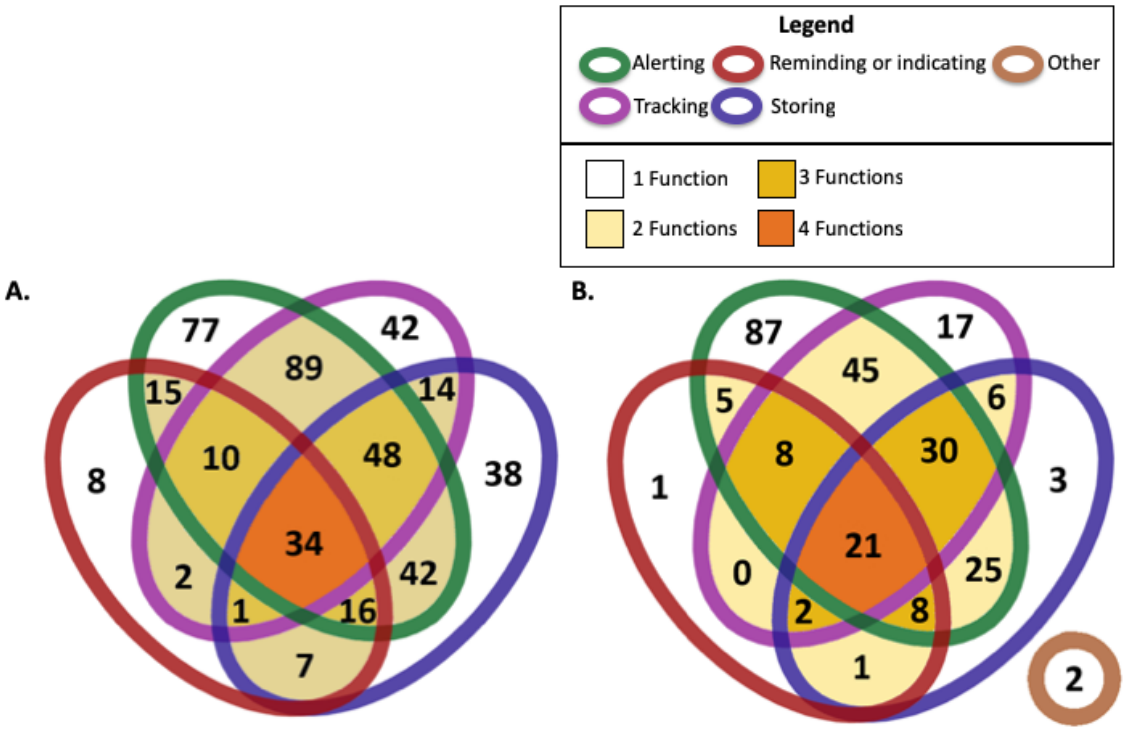
[c]Only apps that have ratings are included in this count.

[d]Each app can have multiple languages.

[e]Once the user begins using the app, the app may ask for additional costs not included in the cost of download.

[f]Other diseases included oncology, cardiology, and post-transplants.

Figure 2. Smartphone medication adherence apps according to 4 key features: alerting (to take medication), tracking (medication taking), reminding (to refill)/indicating (amount of medication left) and storing (medication information) for A. Apple and B. Android operating system.



With respect to apps with a single feature, the majority of which are involved in alerting, with 80 (17.9%, 80/443) in Apple and 87 (33.1%, 83/261) in Android. With respect to apps with dual features, the most predominant combination involved alerting and tracking with 89 (19.9%, 89/443) in Apple and 45 (17.1%, 45/261) in Android. The full list of mobile phone medication adherence apps identified is available in Multimedia Appendix 2.

Results of the detailed quality assessment with respect to additional features and user testing of 20 selected Apple and Android medication adherence apps are summarized in Table 2. For Apple apps, the 2 most common features (90%, 9/10) were alerting (to take medications) and alerts that do not require internet connection. With respect to user-friendliness, 40% (4/10) apps were rated as "easy," 50% (5/10) as "moderate," and only 10% (1/10) as "difficult." Similarly, for Android apps, the 2 most common features (100%, 10/10) were alerting (to take medications) and alerts that do not require internet connection. Furthermore, all 10 Android apps assessed were free to download. With respect to user-friendliness, 20% (2/10) apps were rated as "easy," 60% (6/10) as "moderate," and 20% (2/10) as "difficult."

Content Analysis of User Reviews

User reviews were available for 14 (6 Apple and 8 Android) of the 20 selected apps, and altogether, 1323 reviews (331 Apple and 992 Android) were analyzed. Content analyses resulted in 3 themes: (1) features and functions appreciated by users, (2) negative user experiences, and (3) desired features and functions. These themes and associated categories are summarized in Table 3 and described in detail below.

Theme 1: Features and Functions Appreciated by Users

Category 1.1: App Performance and Practical Aspects

About one-third of the reviews in this category (34.16%, 452/1323) provided positive comments such as "Love this app!" (Android App #4) and "Awesome app" (Apple App #3) without specific details regarding which aspect of the apps were valued. For reviews that provided further details, user-friendliness and ease of use were commonly mentioned:

> *Would be good for all ages, including elderly people once set up [...] So far it seems excellent, I really like the format, it's not made complicated by superfluous functions that never get used.* [Android App #5]

Some noted that they tried multiple apps before settling on their current app:

[...] I tried >8 till I settled on this one. [Apple App #3]

Use [the] app every single day...many times a day. [Apple App #2]

Reviews also suggested frequent use of apps by users:

Table 2. Quality assessment and user testing of selected mobile phone medication adherence apps in Apple and Android.

Features	Adherence intervention target	Apple (n=10 apps), n (%)	Android (n=10 apps), n (%)
Assessment of quality according to availability of features			
1. Alerting to take medication	Behavioral	9 (90)	10 (100)
2. Tracking medication taking	Behavioral	7 (70)	7 (70)
3. Reminding to refill or indication amount of medication left	Behavioral	5 (50)	7 (70)
4. Storing medication information	Educational	8 (80)	7 (70)
5. Complex medication instructions and/or notes	Educational	6 (60)	8 (80)
6. Database of medications	Educational	5 (50)	2 (20)
7. Backup, cloud access, or means of access through another device	N/A[a]	7 (70)	6 (60)
8. Exportation or printing of data	N/A	6 (60)	5 (50)
9. Free to download	N/A	8 (80)	10 (10)
10. Alerts do not require internet connection	N/A	9 (90)	10 (10)
11. Multiple profiles or patients	N/A	5 (50)	5 (50)
12. Multiple languages	N/A	3 (30)	2 (20)
Number of features, mean (range)	N/A	6.5 (3-9)	6.6 (2-10)
Assessment of user friendliness			
Easy	N/A	4 (40)	2 (20)
Moderate	N/A	5 (50)	6 (60)
Difficult	N/A	1 (10)	2 (20)

[a]N/A: not applicable.

Table 3. Themes and categories emerging from content analysis of Web-based reviews for selected mobile phone medication adherence apps (N=1323 user reviews).

Themes and categories	Reviews, n (%)[a]
1. Features and functions appreciated by users	
1.1 App performance and practical aspects	765 (57.82)
1.2 Helpful reminders and notifications	540 (40.81)
1.3 Monitoring other health information	80 (6.05)
1.4 Versatility of medication information input and display	79 (6.04)
1.5 Supports health care visits	63 (4.76)
2. Negative user experiences	
2.1 Technical difficulties	393 (29.70)
2.2 Challenges with medication information input and display	58 (4.38)
2.3 Problems with reminders and notifications	40 (3.02)
3. Desired functions and features	
3.1 Optimizing medication information display and input	73 (5.51)
3.2 Improving reminders and notifications	39 (2.94)
3.3 Upgrading app performance	25 (1.88)

[a]The percentages of the user reviews coded do not add up to 100% because user reviews can be coded to multiple categories or themes.

In addition, recurring positive comments referred to the willingness of users to pay for the pro-version of the app for desired features and functions:

> *Highly recommended, even if you take only one medication or vitamin a day. Worth paying for the upgrade to Premium version to get all the customize options - especially for those who need a visual of individual medications.* [Android App #5]

Users also appreciated the user interface being "elegant in a friendly minimalistic way" (Apple App #3). Commonly, reviews complimented the apps as having a "clean and simple format" (Apple App #3) and a "beautiful layout" (Apple App #2). Other practical aspects of the apps that users enjoyed include no advertisements, doing what the app says on its description, providing friendly and swift customer support, and constant upgrades to the app to fix existing problems and add new features.

Category 1.2: Helpful Reminders and Notifications

Around 29.47% (390/1323) of the reviews expressed that reminders supported users with remembering and keeping track of their medications. For users who recently got diagnosed with a condition and/or recently started to take a handful of medications, they found the reminder features valuable:

> *I recently had a medical issue that required me to be on several medications.I've never taken meds regularly and this app helped me make sure I took everything when I was supposed to. Highly recommend.* [Android App #6]

Frequently, reviews mentioned the benefits of setting reminders not only for the users' medications but also for refills:

> *Reminds me when to take it. Reminds me went to refill the prescription. Absolutely essential to those who take medication.* [Apple App #2]

Other aspects of helpful reminders and notifications include ability to customize alarm sounds and snooze feature to defer taking medications. For example:

> *I love that you can choose what alarm tone you can use. I have to take two medications and I am able to set two different tones.* [Android App #3]

Category 1.3: Monitoring Other Health Information

Alongside managing prescribed medications, apps helped with monitoring other health information such as nonprescription medications (eg, supplements), vital measurements (eg, blood pressure, blood sugar, and pulse), symptoms (eg, pain), and drug safety (eg, side effects). Other less commonly mentioned health information that users monitored with their apps included bowel movements, cigarette usage, diet, exercise, and water consumption. Many reviews praised the ability to manage multiple people or pets. Furthermore, 1 user mentioned the following:

> *Now I love the multiple patients feature [my dogs have no phones to track their meds :)], and it is quick to switch from one to another.* [Android App #5]

Category 1.4: Versatility of Medication Information Input and Display

Many reviews were complimentary of the customization of the medication input. These included the ability to embed medication details (eg, adding shape or color or type of medications, and inputting notes or pictures) and add personalized dosing options (eg, varying frequency settings and scheduling dosage change). The apps were able to accommodate unconventional dosing schedules such as medications that users take once a week, alternative weekend hours, and every 2 weeks. In terms of medication input display, many users commented on the useful features to track inventory, "help [users] keep track of all of what and when [users] have taken [their] pills" (Apple App #3) and generate summary reports to "produce a list of current medications and to graph my progress of selected vitals" (Android App #4).

Category 1.5: Supports Health Care Visits

Reviews frequently expressed the convenience of having their list of medications and necessary health information in 1 app. For instance:

> *This is an outstanding app. It's so helpful when going to the doctor. I just show them my phone, no more dragging pill bottles with me! So much you can do with this app. I love it!* [Android App #2]

Users mentioned a wide range of occasions when the app is handy including single or different health care providers visits, emergency rooms, and for self-medication management. In particular, users found keeping track of health care visit appointments and linking prescriptions to the corresponding physician or pharmacy helpful.

Theme 2: Negative User Experiences

Category 2.1: Technical Difficulties

Around 15.79% (209/1323) of the negative user experiences related to technical glitches and bugs from the apps that would disrupt the app features and functions, particularly the reminder feature (eg, not notifying user at the correct time and app crashes). Recurring complaints also referred to difficulty and confusion in app setup and use:

> *There is no tutorial or help section on the app, so you just blindly have to click on things to try to figure out how to set anything up for the first time.* [Apple App #4]

Reviews also commented on the "clunky" user interface (app visual design):

> *Popup menus have dark text on dark background, so they can't really be read.* [Android App #4]
>
> *Very slow menus and animations for no reasons, lots of extra buttons and sub menus.* [Android App #3]

Poor interoperability was another area that users noted. Users mentioned inadequate synchronization and backup with multiple devices and programs: inability to transfer data to Secure Digital card, to sync between multiple devices, and the lack of "website database or cloud to restore database files from" (Android App

#4). Particularly, for Apple watch, reviews mentioned the issue of reliability:

> *It links with Apple Watch but does not sink back with your iPhone. So if you click on 'take' it does not register with the iPhone app and thus keeps pinging you to take your meds.* [Apple App #2]

Users also referred to unsatisfactory customer service and updates. Many users mentioned they "submitted many feedback messages about [a] problem and no response or fixes [were made]" (Android App #5). Users commonly experienced long waits, with no response or outdated, unhelpful response from customer support. Regarding app updates, some reviews noted that there have not been any recent updates or that the new updates cause more problems (eg, glitches, more advertisements, and more confusing). For users that pay a subscription fee, reviews similarly criticized:

> *I really like this app, however there appears to not have been any updates since April. If there are to be no improvements should you be charging? And especially at the price you are asking?* [Android App #5]

Other technical difficulties identified in reviews were cost (eg, expensive and necessity of subscription of pro version to use the app), high volume of advertisements, burden on the device (eg, large device storage and high use of battery and data), discrepancy with app description and actual app features, and lack of security measures (eg, user confidentiality).

Category 2.2: Challenges With Medication Information Input and Display

Inflexible information input was a recurring complaint which consisted of difficulty inputting medications from different countries because of "inadequate medicines information base" (Apple App #2) and inability to customize dosing schedule (eg, scheduling different dosing schedule based on the day). Users often have personalized drug regimens that they need to adhere to, and they expressed that several apps do not reflect their or their loved one's correct drug regimen, for example:

> *The only issue that I didn't like was I have 1 medicine that I only take Sun-Thur and I don't have that option [...].* [Android App #6]

This problem was also applicable to special populations such as children because "[…] children's dosages are determined by weight so sometimes you get weird doses for customized medications." (Apple App #5).

Users mentioned the inability to "[…] update the time of the subsequent dosages throughout the day" (Android App #3) based on the time users take the medication, which may lead to taking medications at improper times. Inability to input supplementary notes or pictures (eg, adding "whether to take meds with food, or before or after meals" [Android App #3]) and fix mistakes in medications logs were challenges that users also expressed. In terms of medication information display, reviews referred to the inconsistency of units, unfavorable display in military time, inability "to view a list of dates and times meds were taken" (Android App #3), and inability to view and track balance of medications remaining. An example of a user review includes:

> *Was great until I tried added an oral suspension medicine: the app would only allow me to enter amounts in grams, instead of ml. I don't see a user-friendly way of choosing a unit.* [Apple App #5]

Category 2.3: Problems With Reminders and Notifications

Users most commonly expressed challenges with limited alert customization, especially in terms of the alarm loudness. Specifically, users commented that at night, "no matter what tone, the tone is too soft and doesn't ring long enough to pull me out of my sleep" (Android App #6). Users mentioned that the customization with ringtone is "a very important feature, since it is how you will be notified" (Android App #6). Some complaints evolved around the inability to prioritize reminders if multiple medications are taken at the same time:

> *[...] if you take 6 meds at 13:00, you'll get 6 notifications at the same time, sounds like your phone is having a seizure.* [Android App #3]

For 1 app in particular, users noted their frustration on the app's incapability to adjust the schedule based on the Daylight Savings Time: For users who "[...] take several medications every day, and is really a pain to have to go in and change the time on each one" (Android App #5). Some reviewers mentioned the hassle of unlocking their phones for their alarms to ring or to record as taken, for example, "USELESS. No alarm unless u open app!" [Android App #5]

Theme 3: Desired Functions and Features

Category 3.1: Optimizing Medication Information Display and Input

Flexibility of data input was a request that commonly appeared. Users wished for sections to add notes for details on their specific medication or dose, their symptoms, or mood. They also requested to add pictures of their medication bottles and pills. Frequently, users mentioned that "dose options need to reflect actual doses" (Apple App #2), particularly in terms of being able to schedule different dosing schedules and dosage change based on the day, for example:

> *Some prescriptions have you take one pill one day and two on another. Therefore, it would be good if you could set it to a different number of pills on specific days. I know I could just add the prescription more than once, but then tracking the number of pills remaining wouldn't be accurate.* [Apple App #2]

Many users desired to have the ability to readjust their drug schedule based on the actual time the drug was taken:

> *Would be nice if you could define a dose to be given X hours after the previous one, instead of strictly every X hours, in case a dose was given late.* [Android App #4]

Fewer user reviews requested for the ability to add 0.5 portions and to have a more adjustable dosing frequency, barcoding or scanning function to easily input their medications, and unit setting. Improvements to the medication history section include

being able to "summarize medication activity" (Android app #3) by having "an option to enter end date" (Android App #6) and by being "able to look back at the actual times when a medication dose is taken, not just that it was taken." (Android App #3).

Category 3.2: Improving Reminders and Notifications

Users mainly made requests on 2 particular aspects of reminders: customization of reminder setups and suggestions on new, beneficial features. Reviewers asked for alert customizations in terms of loudness and ringtones, reminder time frame, and involvement of their caregivers or family members in their care. Users also requested for more efficient methods to indicate medications as taken without opening the app. For example, a user specifically suggested "I only wish I could use voice command to 'take' medicine in the middle of the night when in pain without fumbling for my glasses." (Apple App #3).

Category 3.3: Upgrading App Performance

Better interoperability was the request that most frequently appeared in this category. Users wished for enhanced linkage to other devices (eg, Apple Watch) and programs (eg, Web version of the app, other Apple or Android devices, and Dropbox) to sync or manage their data and appointments. As 1 user summarized:

> *Synchronization between 2 devices, [in] other words, [require] the app [to] run on 2 devices and something entered on one device can also be seen on another one (i.e. phone and iPad).* [Apple App #2]

Discussion

Principal Findings

This study provides better understanding of medication adherence apps from dual perspectives. First, quality assessment and testing by pharmacist researchers provides a health care provider's lens, and second, content analysis of reviews provides a target user's lens. Major strengths of this study include an update to the current landscape of 704 medication adherence apps. The subsequent content analysis of user reviews conducted soon after the identification and quality assessment of apps adds uniqueness to our study. The apps analyzed were still available in their corresponding app stores and allowed the authors to compare results of our quality assessment. Indeed, although qualitative analyses of user reviews for disease-specific apps including for bipolar disorder and weight loss [18,19] have been previously published, target users' (patients') experiences with medication adherence apps have not been extensively studied. Previously, Stawarz et al in 2014 conducted a user review analysis of the top 50 reviews for 40 apps available only on the Android operating system [20]. Bailey et al in 2014 conducted a user review analysis of 26 eligible apps that appeared in their initial search results [3]. However, these were limited to an arbitrarily chosen top 75 "most helpful" reviews, and imposing preidentified themes to their analysis made it largely deductive instead of allowing themes to be inductively generated from the data. Therefore, our systematic strategy to identifying apps, replicable steps (eg, based on dates) for selecting reviews, and purposeful application of content analysis methodology provides a more in-depth, rigorous approach to understanding users' experiences and perspectives with mobile phone medication adherence apps.

Indeed, combining quality assessment and user testing of apps with qualitative analyses of corresponding Web-based reviews provided the opportunity to contextualize respective findings. For example, user reviews were mostly positive, and the main theme that emerged was features and functions appreciated by users. We noticed that the more commonly a feature was available (eg, alerting), the higher the number of user reviews were present in their corresponding category (eg, reminders and notifications). Most users found the reminders and notifications features helpful for multiple medications including nonprescription or as-needed medications. Users generally found apps to be user-friendly and appreciated the simple app design. These characteristics address existing barriers of difficult app navigation and the time consumption when inputting their medications [21,22]. Despite the previously expressed challenges, including the inability to create reminders without internet connection and for multiple people on numerous medications [21], our results reveal that currently, the majority of the apps (90% [9/10] for Apple and 100% [10/10] for Android) do not require internet connection and include the ability to manage medications for multiple people and pets.

A practical finding from our study is that users commonly expressed that they tried multiple apps before settling on one that they favored. Moreover, one of the reasons that could be associated with this frustration may involve search terms [21,23]. The quantity of adherence apps yielded by the search results varied significantly among keywords. The search terms "adherence" and "compliance" yielded relatively few relevant search results on both operating systems (53 and 4 apps on Apple and 39 and 2 apps on Android). On the contrary, the terms "medication" and "pills" yielded the most results on both operating systems (304 and 201 on Apple and 198 and 138 on Android). This reveals preferences the public, or at least the technology community, may have with regard to the language that is used to discuss the topic of medication adherence. Patient-friendly terms (eg, "pills") appear more frequently used than relatively jargon-like terms used by the medical community (eg, "compliance"). Health care providers who may be recommending these apps to patients may benefit from being aware of the types of language and terminology preferred by the specific patients they are caring for.

Limitations

During the app identification process, only single search terms were used. It is not known whether the usage of compound search terms would have yielded a larger number of results or perhaps more tailored results. We extracted user reviews for content analysis within the 1-year period (January 1, 2017, to January 1, 2018), and since then, there may be different versions of the apps that may have appeared or removed. Furthermore, given the sheer number of available apps, we limited content analyses to representative apps from each operating system and only those submitted in English, as such reviews may not accurately represent the entire population of app users. In addition, individuals providing reviews may be systematically

different from those who do not—in that they are likely those who strongly favor or dislike apps.

Conclusions

Our app quality assessment and content analysis of user review study provide a view of the available mobile apps for medication adherence and the target users' (patients') experiences with medication adherence mobile apps. Our findings can inform the future development of the next generation of medication adherence apps co-designed with patients, researchers, and technology companies.

Acknowledgments

MDV holds a Canada Research Chair in Medication Adherence, Utilization, and Outcomes and is a recipient of a Network Scholar Award from The Arthritis Society/Canadian Arthritis Network and a Scholar Award from the Michael Smith Foundation for Health Research.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Explanation of app features and application to conceptual framework.

[PDF File (Adobe PDF File), 50KB - mhealth_v7i1e11919_app1.pdf]

Multimedia Appendix 2

Complete list of medication adherence apps identified.

[PDF File (Adobe PDF File), 115KB - mhealth_v7i1e11919_app2.pdf]

References

1. World Health Organization. 2003. Adherence to long-term therapies- evidence for action URL: http://www.who.int/chp/knowledge/publications/adherence_full_report.pdf [accessed 2018-08-03] [WebCite Cache ID 71Vc0ZBE0]
2. Patrick K, Griswold WG, Raab F, Intille SS. Health and the mobile phone. Am J Prev Med 2008 Aug;35(2):177-181 [FREE Full text] [doi: 10.1016/j.amepre.2008.05.001] [Medline: 18550322]
3. Bailey SC, Belter LT, Pandit AU, Carpenter DM, Carlos E, Wolf MS. The availability, functionality, and quality of mobile applications supporting medication self-management. J Am Med Inform Assoc 2014;21(3):542-546 [FREE Full text] [doi: 10.1136/amiajnl-2013-002232] [Medline: 24163156]
4. Dayer L, Heldenbrand S, Anderson P, Gubbins PO, Martin BC. Smartphone medication adherence apps: potential benefits to patients and providers. J Am Pharm Assoc (2003) 2013;53(2):172-181 [FREE Full text] [doi: 10.1331/JAPhA.2013.12202] [Medline: 23571625]
5. Dayer LE, Shilling R, Van Valkenburg M, Martin BC, Gubbins PO, Hadden K, et al. Assessing the medication adherence app marketplace from the health professional and consumer vantage points. JMIR Mhealth Uhealth 2017 Apr 19;5(4):e45 [FREE Full text] [doi: 10.2196/mhealth.6582] [Medline: 28428169]
6. Heldenbrand S, Martin BC, Gubbins PO, Hadden K, Renna C, Shilling R, et al. Assessment of medication adherence app features, functionality, and health literacy level and the creation of a searchable Web-based adherence app resource for health care professionals and patients. J Am Pharm Assoc (2003) 2016;56(3):293-302. [doi: 10.1016/j.japh.2015.12.014] [Medline: 27067551]
7. 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]
8. Haase J, Farris KB, Dorsch MP. Mobile applications to improve medication adherence. Telemed J E Health 2017;23(2):75-79. [doi: 10.1089/tmj.2015.0227] [Medline: 27248315]
9. Ahmed I, Ahmad NS, Ali S, Ali S, George A, Saleem DH, et al. Medication adherence apps: review and content analysis. JMIR Mhealth Uhealth 2018 Mar 16;6(3):e62 [FREE Full text] [doi: 10.2196/mhealth.6432] [Medline: 29549075]
10. Broderick J, Devine T, Langhans E, Lemerise AJ, Lier S, Harris L. Designing health literate mobile apps. Washington, DC: Institute of Medicine; 2014 Jan 28. URL: https://nam.edu/wp-content/uploads/2015/06/HealthLiterateApps.pdf [accessed 2018-07-29] [WebCite Cache ID 71XO6uWs4]
11. Statista. 2016. Number of smartphone users worldwide from 2014 to 2020 (in billions) URL: https://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ [accessed 2018-07-25] [WebCite Cache ID 6zxLKoht0]
12. Zusman E, De Vera M. Contemporary Trends and Issues in Medication Non-Adherence: A Systematic Review. 2018 Jun 25 Presented at: Drug Information Association (DIA) 2018; June 24-28, 2018; Boston URL: https://www.diaglobal.org/

zh-cn/flagship/dia-2018/program/posters/Poster-Presentations/Poster-Presentations-Details?ParentProductID=6589208&ProductID=7326319&AbstractID=79284
13. Galo JS, Mehat P, Rai SK, Avina-Zubieta A, De Vera MA. What are the effects of medication adherence interventions in rheumatic diseases: a systematic review. Ann Rheum Dis 2016 Apr;75(4):667-673. [doi: 10.1136/annrheumdis-2014-206593] [Medline: 25667208]
14. Roter DL, Hall JA, Merisca R, Nordstrom B, Cretin D, Svarstad B. Effectiveness of interventions to improve patient compliance: a meta-analysis. Med Care 1998 Aug;36(8):1138-1161. [Medline: 9708588]
15. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
16. Glaser B, Strauss A. The Discovery of Grounded Theory: Strategies for Qualitative Research. Chicago, IL: Aldine Publishers; 1967.
17. Bowen GA. Naturalistic inquiry and the saturation concept: a research note. Qual Res 2008 Feb;8(1):137-152. [doi: 10.1177/1468794107085301]
18. Frie K, Hartmann-Boyce J, Jebb S, Albury C, Nourse R, Aveyard P. Insights from Google Play Store user reviews for the development of weight loss apps: mixed-method analysis. JMIR Mhealth Uhealth 2017 Dec 22;5(12):e203 [FREE Full text] [doi: 10.2196/mhealth.8791] [Medline: 29273575]
19. Nicholas J, Fogarty AS, Boydell K, Christensen H. The reviews are in: a qualitative content analysis of consumer perspectives on apps for bipolar disorder. J Med Internet Res 2017 Apr 07;19(4):e105 [FREE Full text] [doi: 10.2196/jmir.7273] [Medline: 28389420]
20. Stawarz K, Cox A, Blandford A. Don't forget your pill! Designing effective medication reminder apps that support users' daily routines. 2014 Presented at: SIGCHI Conference on Human Factors in Computing Systems; April 26-May 1, 2018; Toronto, Ontario, Canada p. 2269-2278 URL: https://dl.acm.org/citation.cfm?doid=2556288.2557079
21. Choi A, Lovett AW, Kang J, Lee K, Choi L. Mobile applications to improve medication adherence: existing apps, quality of life and future directions. Adv Pharmacol Pharma 2015 Sep;3(3):64-74. [doi: 10.13189/app.2015.030302]
22. Grindrod KA, Li M, Gates A. Evaluating user perceptions of mobile medication management applications with older adults: a usability study. JMIR Mhealth Uhealth 2014;2(1):e11 [FREE Full text] [doi: 10.2196/mhealth.3048] [Medline: 25099993]
23. Boudreaux ED, Waring ME, Hayes RB, Sadasivam RS, Mullen S, Pagoto S. Evaluating and selecting mobile health apps: strategies for healthcare providers and healthcare organizations. Transl Behav Med 2014 Dec;4(4):363-371 [FREE Full text] [doi: 10.1007/s13142-014-0293-9] [Medline: 25584085]

Abbreviations

mHealth: mobile health

Edited by G Eysenbach; submitted 11.08.18; peer-reviewed by E Jeong, T Patel; comments to author 14.10.18; revised version received 17.10.18; accepted 17.10.18; published 31.01.19.

Please cite as:
Park JYE, Li J, Howren A, Tsao NW, De Vera M
Mobile Phone Apps Targeting Medication Adherence: Quality Assessment and Content Analysis of User Reviews
JMIR Mhealth Uhealth 2019;7(1):e11919
URL: http://mhealth.jmir.org/2019/1/e11919/
doi:10.2196/11919
PMID:30702435

Original Paper

Mobile Phone Ownership, Health Apps, and Tablet Use in US Adults With a Self-Reported History of Hypertension: Cross-Sectional Study

Aisha T Langford[1], PhD; Craig A Solid[2], PhD; Ebony Scott[1], MSEd; Meeki Lad[3], MPH; Eli Maayan[3], BS; Stephen K Williams[4], MD; Azizi A Seixas[4], PhD

[1]Division of Comparative Effectiveness and Decision Science, Department of Population Health, NYU School of Medicine, New York, NY, United States

[2]Solid Research Group, LLC, St. Paul, MN, United States

[3]The Langford Lab, Department of Population Health, NYU School of Medicine, New York, NY, United States

[4]Center for Healthful Behavior Change, Department of Population Health, NYU School of Medicine, New York, NY, United States

Corresponding Author:
Aisha T Langford, PhD
Division of Comparative Effectiveness and Decision Science
Department of Population Health
NYU School of Medicine
227 E 30th Street #645
New York, NY, 10016
United States
Phone: 1 646 501 2914
Email: aisha.langford@nyulangone.org

Abstract

Background: Mobile phone and tablet ownership have increased in the United States over the last decade, contributing to the growing use of mobile health (mHealth) interventions to help patients manage chronic health conditions like diabetes. However, few studies have characterized mobile device ownership and the presence of health-related apps on mobile devices in people with a self-reported history of hypertension.

Objective: This study aimed to describe the prevalence of smartphone, tablet, and basic mobile phone ownership and the presence of health apps by sociodemographic factors and self-reported hypertension status (ie, history) in a nationally representative sample of US adults, and to describe whether mobile devices are associated with health goal achievement, medical decision making, and patient-provider communication.

Methods: Data from 3285 respondents from the 2017 Health Information National Trends Survey were analyzed. Participants were asked if they owned a smartphone, tablet, or basic mobile phone and if they had health apps on a smartphone or tablet. Participants were also asked if their smartphones or tablets helped them achieve a health-related goal like losing weight, make a decision about how to treat an illness, or talk with their health care providers. Chi-square analyses were conducted to test for differences in mobile device ownership, health app presence, and app helpfulness by patient characteristics.

Results: Approximately 1460 (37.6% weighted prevalence) participants reported a history of hypertension. Tablet and smartphone ownership were lower in participants with a history of hypertension than in those without a history of hypertension (55% vs 66%, P=.001, and 86% vs 68%, P<.001, respectively). Participants with a history of hypertension were more likely to own a basic mobile phone only as compared to those without a history of hypertension (16% vs 9%, P<.001). Among those with a history of hypertension exclusively, basic mobile phone, smartphone, and tablet ownership were associated with age and education, but not race or sex. Older adults were more likely to report having a basic mobile phone only, whereas those with higher education were more likely to report owning a tablet or smartphone. Compared to those without a history of hypertension, participants with a history of hypertension were less likely to have health-related apps on their smartphones or tablets (45% vs 30%, P<.001) and report that mobile devices helped them achieve a health-related goal (72% vs 63%, P=.01).

Conclusions: Despite the increasing use of smartphones, tablets, and health-related apps, these tools are used less among people with a self-reported history of hypertension. To reach the widest cross-section of patients, a mix of novel mHealth interventions

and traditional health communication strategies (eg, print, web based, and in person) are needed to support the diverse needs of people with a history of hypertension.

(JMIR Mhealth Uhealth 2019;7(1):e12228) doi:10.2196/12228

KEYWORDS

smartphone; text messaging; health communication; ownership; goals; cell phone; telemedicine; hypertension; tablets; chronic disease

Introduction

In recent years, mobile health (mHealth) interventions have been proposed as promising strategies for delivering health interventions to people with chronic health conditions [1,2]. For example, mHealth interventions for type 2 diabetes account for a sizable proportion of published articles [1-5]. However, the literature on mHealth interventions for cardiovascular disease generally and hypertension specifically is less robust [6-8]. This gap in the literature is notable, given that heart disease is the primary cause of death among adults in the United States [9,10] and hypertension is a major risk factor for heart disease [10-13]. Recent reports estimate that the proportion of US adults with hypertension is approximately 46% [13,14]. If people in the prehypertension range are considered, hypertension becomes a concern for more than half of US adults, thereby highlighting the need to initiate more mHealth approaches to help prevent or manage hypertension (eg, medication management and lifestyle change).

Broadly, mHealth is conceptualized as an area of electronic health (eHealth) that uses mobile technologies such as mobile phones and wireless devices for health research and healthcare delivery [15,16]. Examples of mHealth interventions include short message service text messaging, telephone-delivered interventions (eg with a nurse or health coach), Bluetooth-enabled pill boxes and fitness monitors, and health-related smartphone apps [17-20]. A report published by the Pew Research Center in 2015 noted that text messaging was the most widely used smartphone basic feature or app, with approximately 97% of smartphone owners reporting that they used text messaging at least once over the course of 1 week [21]. This finding was part of an "experience sampling" survey conducted by Pew Research Center in which smartphone owners were contacted twice a day for a week and queried about how they used their smartphone in the hour immediately before answering survey questions [21].

The use of mHealth interventions over time has largely mirrored the rapid growth in ownership of smartphones and other devices over the last two decades [2,22,23]. In 2018, approximately 95% of Americans owned a mobile phone of some kind [23]. Between 2010 and 2016, the Pew Research Center reported that smartphone ownership among Americans increased from 35% to 77% and tablet ownership increased from 3% to 51% [24]. Among adults who own basic mobile phones only, those aged ≥65 years comprise the largest proportion (40%) compared to all other age groups. With regard to race, the proportion of basic mobile phone ownership for white, black, and Hispanic people was 17%, 23%, and 20%, respectively [23]. With regard to mobile health apps, more than 300,000 apps are available to the general public to download from Google Play and the Apple Store [25]. Some early evidence suggests that apps are associated with behavior change intentions as well as actual behavior change related to diet, physical activity, weight loss, and smoking cessation [26,27]. However, few health-related apps are evidence based [28] and few are consistently used over time [27,29].

Despite the ubiquitous use of mobile phones in the United States and growing interest among researchers to develop mHealth interventions for people with chronic diseases, the data are still limited with regard to mobile phone ownership and the use of mHealth interventions among people with a self-reported history of hypertension. Thus, the objectives of this study were to answer four key questions:

1. Do people with a self-reported history of hypertension differ from those without a self-reported history of hypertension in terms of basic mobile phone, smartphone, and tablet ownership?
2. Among people with a self-reported history of hypertension exclusively, does mobile device ownership differ by age, gender, race/ethnicity, or education?
3. Are there differences between people with and those without a self-reported history of hypertension with regard to having health-related apps on their smartphones and tablets?
4. How do people with a self-reported history of hypertension differ from those without a self-reported history of hypertension with regard to the role that smartphones and tablets play in helping them achieve health-related goals, make medical decisions, or establish patient-provider communication?

Given the large number of US adults affected by hypertension [13] and growing interest in the promise of mHealth interventions, our findings may help inform future efforts to develop hypertension-focused mHealth interventions for patients in clinical settings and the general public.

Methods

Overview of the Health Information National Trends Survey

The Health Information National Trends Survey (HINTS) is a probability-based, nationally representative survey of US noninstitutionalized adults aged ≥18 years [30], sponsored by the US Department of Health and Human Services. HINTS has been administered approximately every 1-3 years since 2003, with the goal of collecting nationally representative data that track changes in health communication and information technology. The sample design for the 2017 HINTS 5, Cycle

1, consisted of a single-mode mail survey using the next-birthday method for respondent selection and comprised two stages. In the first stage, a stratified sample of addresses was selected from a file of residential addresses. In the second stage, one adult was selected within each sampled household. The sampling frame consisted of a database of addresses used by Marketing Systems Group to provide random samples of addresses. HINTS is approved by the Office of Management and Budget (approval number, 0925-0538), the office that reviews all federally sponsored surveys. Since the present study involves a secondary analysis of a publicly available dataset, it was exempt from institutional review board approval at the authors' home institution. Full details about the HINTS methodology are available on the HINTS website [31].

Study Design and Participants

The present study used a cross-sectional design to evaluate participant data from HINTS 5, Cycle 1 (N=3285). Survey responses were collected between January 25, 2017, and May 5, 2017, with complete data from 3191 respondents. The self-reported hypertension status was assessed using the question, "Has a doctor or other health professional ever told you that you had high blood pressure or hypertension (yes/no)?" Of the complete sample, approximately 1460 participants self-reported a history of hypertension (37.6% weighted prevalence).

Measures

Mobile Device Ownership

Participants were asked if they owned a tablet computer like iPad, Samsung Galaxy, or Motorola Xoom; smartphone such as iPhone, Android, Blackberry, or Windows phone; or a basic mobile phone only. Response options were yes and no.

Apps Related to Health and Wellness

If participants responded "yes" to owning a tablet or smartphone, they were asked if they had any "apps" related to health and wellness, with response options of yes, no, and don't know.

Tablet or Smartphone Helpfulness

Participants were asked if their tablet or smartphone ever helped them track progress on a health-related goal such as quitting smoking, losing weight, or increasing physical activity; make a decision about how to treat an illness or condition; and have discussions with their health care provider. The response options were yes and no.

Sociodemographic Factors

Covariates were selected based on known associations with hypertension, and mHealth broadly and included age, race/ethnicity, gender, and education [19,26,27,32,33]. Age was assessed as a categorical and continuous variable with age ranges 18-34 (reference), 35-49, 50-64, 65-74, and ≥75 years. Race/ethnicity was categorized into five categories: non-Hispanic white (reference), non-Hispanic black/African American, Hispanic, Non-Hispanic Asian, and non-Hispanic other. Gender was assessed as male (reference) and female. Education was classified into four categories: less than high school (reference), 12 years or completed high school, some college, and college degree. Income was not evaluated due to its collinearity with education.

Statistical Analyses

Survey weights provided in the HINTS data were utilized to calculate weighted percentages of subject characteristics. To test for differences in outcomes across relevant patient characteristics, unadjusted chi-square tests were conducted on the weighted percent in agreement with various measures. Due to the survey weights, a Rao-Scott F-adjusted chi-square statistic was used, which yields a more conservative interpretation than the traditional Wald chi-square test used for unweighted analyses [34]. We used SAS software, version 9.4 of the SAS System, to conduct all analyses (SAS Institute Inc, Cary, NC), employing procedures that can account for the sampling design of HINTS, such as PROC SURVEYFREQ and PROC SURVEYREG. These procedures utilize the survey weights available in the HINTS data to obtain population-level point estimates and bootstrap accurate estimates of standard errors. To account for multiple comparisons, we applied methods to limit the false-discovery rate to 0.05 [35,36]. *P* values<.05 were considered statistically significant.

Results

Demographic Characteristics by Hypertension Status

Compared to participants without a self-reported history of hypertension, those with a history of hypertension were significantly more likely to be older, male, and black and have received less formal education (Table 1).

Smartphone, Tablet, and Basic Mobile Phone Ownership by Hypertension Status

Of the full sample, approximately 62%, 79%, and 22% of HINTS participants reported owning a tablet, smartphone, and basic mobile phone only, respectively, and 84% reported that they had either a tablet or smartphone. When evaluating device ownership by hypertension status, fewer participants with a history of hypertension reported owning a tablet (55% vs 66%) or smartphone (68% vs 86%) than those without a history hypertension. Additionally, those with a self-reported history of hypertension were almost twice as likely to own a basic mobile phone only as compared to those without a history of hypertension (16.3% vs 8.5%).

Table 1. Participant characteristics according to the hypertension status.

Characteristics	All (N=3285), weighted %	Hypertension (N=1460), weighted % (95% CI)	Non-hypertension (N=1749), weighted % (95% CI)	*P* value
Age group (N=3095), years				<.001
18-34	21.96	5.11 (2.3-7.9)	31.97 (27.7-36.3)	
35-49	28.77	22.05 (17.4-26.7)	32.76 (28.0-37.5)	
50-64	30.25	39.89 (35.5-44.3)	24.52 (21.8-27.2)	
65-74	10.92	17.13 (15.5-18.8)	7.23 (6.3-8.2)	
≥75	8.1	15.82 (14.0-17.7)	3.52 (2.7-4.3)	
Race/ethnicity (N=2906)				<.001
Non-Hispanic white	65.61	66.77 (63.2-70.3)	64.95 (62.9-67.0)	
Non-Hispanic black	10.39	14.25 (12.0-16.5)	8.2 (6.7-9.7)	
Hispanic	15.64	11.77 (9.3-14.3)	17.83 (16.2-19.4)	
Non-Hispanic Asian	5.61	3.25 (1.7-4.8)	6.95 (5.9-8.0)	
Non-Hispanic other	2.76	3.96 (2.8-5.1)	2.07 (1.5-2.7)	
Gender (N=3161)				0.03
Male	49	52.6 (49.4-55.8)	46.85 (45.0-48.7)	
Female	51	47.4 (44.2-50.6)	53.15 (51.3-55.0)	
Education (N=3125)				<.001
Less than high school	8.6	11.28 (8.1-14.4)	6.99 (4.7-9.3)	
High school	22.91	30.15 (25.7-34.6)	18.54 (15.9-21.2)	
Some college	32.66	33.02 (29.0-37.0)	32.44 (29.9-35.0)	
College degree or higher	35.83	25.55 (22.6-28.5)	42.02 (40.0-44.1)	
Have a tablet (N=3196)[a]	61.6	54.7 (50.7-58.8)	65.7 (61.0-70.3)	0.001
Have a smartphone (N=3178)[a]	79	67.7 (64.1-71.4)	85.8 (82.9-88.6)	<.001
Have a basic mobile phone (N=3124)[a]	21.5	28.9 (25.6-32.2)	17 (13.5-20.6)	<.001
Have either a tablet or smartphone (N=3209)[a]	84.1	75.3 (71.6-79.0)	89.4 (87.1-91.7)	<.001
Have a basic mobile phone, but neither a tablet nor a smartphone (N=3108)[a]	11.4	16.3 (13.4-19.2)	8.5 (6.4-10.5)	0.001

[a]Values for these variables represent weighted percent of those who answered "Yes."

Smartphone, Tablet, and Basic Mobile Phone Ownership Among Participants With a History of Hypertension Only

There were significant differences in device ownership by age and education, but not race or gender (Tables 2 and 3). For example, participants aged 18-24 years were more likely to own a tablet or smartphone than those aged 65-74 years (90% vs 51% and 90% vs 61%, respectively; *P*<.001). Of the people aged ≥75 years, the majority owned a basic mobile phone only (47%) compared to a tablet (29%) or smartphone (30%). As the level of education of participants increased, so did the likelihood of owning a smartphone or tablet. Compared to participants with less than a high school education, those with a college degree or higher education were more likely to own a tablet and smartphone (41% vs 71% and 47% vs. 87%, respectively, *P*<.001).

Table 2. Differences in device ownership and use among patients with hypertension according to demographics. Sample sizes reflect the number of participants with nonmissing values for device question and characteristics.

Demographic	"Yes" to having a tablet (%)	"Yes" to having a smartphone (%)	"Yes" to having a mobile phone (%)	"Yes" to having a tablet or smartphone[a] (%)	"Yes" to having a tablet and smartphone[b] (%)
Age group, years	N=1389	N=1379	N=1350	N=1396	N=1377
18-34	89.6	90.2	1.0	99.5	80.4
35-49	61.1	86.7	15.5	93.9	46.1
50-64	60.3	74.2	26.7	81.6	52.7
65-74	51.4	60.9	40.4	68.2	43.3
≥75	29.0	29.9	47.3	41.7	16.5
Race/ethnicity	N=1274	N=1267	N=1246	N=1277	N=1266
Non-Hispanic white	58.1	71.7	26.9	79.5	50.0
Non-Hispanic black	59.7	75.4	27.2	82.1	52.9
Hispanic	48.9	74.8	24.5	80.1	43.3
Non-Hispanic Asian	55.5	75.0	21.1	75.0	55.5
Non-Hispanic other	78.8	60.2	44.5	83.4	55.7
Gender	N=1432	N=1419	N=1389	N=1440	N=1417
Male	52.9	67.1	27.4	74.9	44.9
Female	56.9	68.8	29.8	75.8	49.2
Education	N=1410	N=1398	N=1370	N=1417	N=1396
Less than high school	40.7	46.5	29.5	52.8	34.3
High school	44.9	53.4	43.6	66.4	31.1
Some college	56.7	74.4	23.1	79.6	51.1
College degree or higher	70.8	86.6	17.6	91.9	65.4

[a]Includes participants with nonmissing values for either device question (tablet or smartphone).

[b]Includes participants with nonmissing values for both device questions (tablet and smartphone).

Table 3. Statistical results of device ownership and use among patients with hypertension. Sample sizes reflect the number of participants with nonmissing values for device questions and characteristics.

Parameter	Wald chi-square value	F-adjusted chi-square value	*P* value
Age group			
"Yes" to having a tablet (N=1389)	62.566	14.68	<.001
"Yes" to having a smartphone (N=1379)	140.397	32.95	<.001
"Yes" to having a mobile phone (N=1350)	52.49	12.32	<.001
"Yes" to having a tablet or smartphone[a] (N=1396)	121.111	28.42	<.001
"Yes" to having a tablet and smartphone[b] (N=1377)	75.7	17.77	<.001
Race/ethnicity			
"Yes" to having a tablet (N=1274)	6.638	1.56	.20
"Yes" to having a smartphone (N=1267)	1.69	0.40	.81
"Yes" to having a mobile phone (N=1246)	3.29	0.77	.55
"Yes" to having a tablet or smartphone[a] (N=1277)	0.95	0.22	.92
"Yes" to having a tablet and smartphone[b] (N=1266)	1.57	0.369	.83
Gender			
"Yes" to having a tablet (N=1432)	0.9871	—	.33
"Yes" to having a smartphone (N=1419)	0.2128	—	.65
"Yes" to having a mobile phone (N=1389)	0.5645	—	.46
"Yes" to having a tablet or smartphone[a] (N=1440)	0.0587	—	.81
"Yes" to having a tablet and smartphone[b] (N=1417)	1.134	—	.29
Education			
"Yes" to having a tablet (N=1410)	55.912	17.88	<.001
"Yes" to having a smartphone (N=1398)	64.93	20.76	<.001
"Yes" to having a mobile phone (N=1370)	28.07	8.975	<.001
"Yes" to having a tablet or smartphone[a] (N=1417)	63.70	20.365	<.001
"Yes" to having a tablet and smartphone[b] (N=1396)	43.498	13.91	<.001

[a]Includes participants with nonmissing values for either device question (tablet or smartphone).

[b]Includes participants with nonmissing values for both device questions (tablet and smartphone).

Presence of Health-Related Apps According to Hypertension Status

Significant differences were observed between participants with and those without a self-reported history of hypertension with regard to having health-related apps on their tablet or smartphone. For example, only 36.5% of participants with a history of hypertension reported having health-related apps compared to 49.2% of those without a history of hypertension (Figure 1; *P*<.001). Participants with a self-reported history of hypertension were also more likely to report that they did not know if they had a health-related app compared to those without a history of hypertension (ie, 6.2% vs 2.9%).

Among participants who owned a tablet or smartphone (Figure 2), those with a self-reported history of hypertension were less likely to report that their tablet or smartphone helped them reached a health-related goal like quitting smoking, losing weight, or increasing physical activity (62.6% vs 71.7%, *P*=.02). With regard to whether tablets or smartphones helped people make a decision about an illness or talk to health care providers, no differences were observed between participants with and those without a history of hypertension.

Presence of Health-Related Apps According to Sociodemographics Among Participants With a History of Hypertension Only

General trends were observed with regard to age when examining HINTS participants with a self-reported history of hypertension exclusively (Figure 3). Participants with health-related apps on their tablets or smartphones were mostly in age categories of 18-34 years (53%) and 35-49 years (51%), whereas percentages for age groups of 50-64, 65-74, and ≥75 years were lower (32%, 35%, and 21%, respectively). Participants who were ≥50 years of age were more likely to report that they did not know if they had any health-related apps on their tablets and smartphones.

Figure 1. Use of health and wellness apps according to self-reported history of hypertension. Excludes those who answered “do not own a tablet or smartphone”. HTN: hypertension.

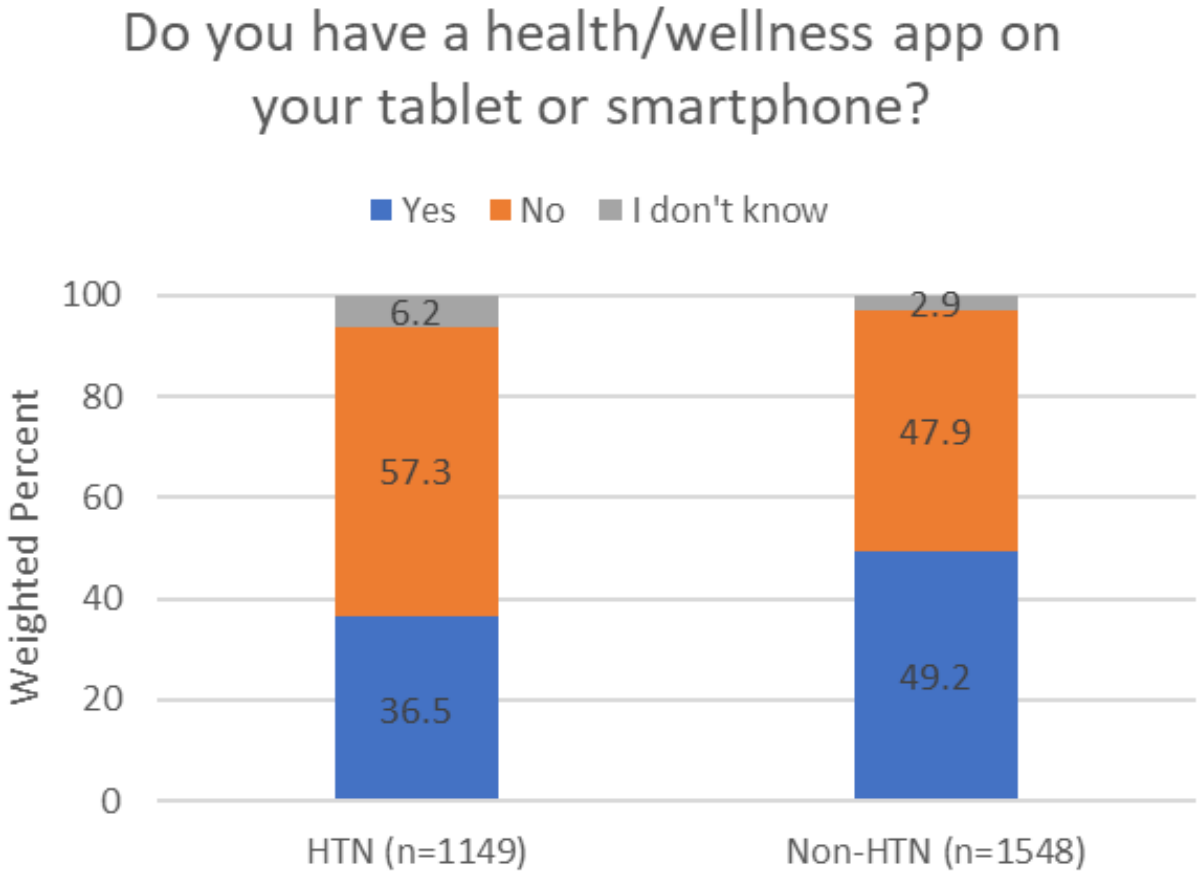


Figure 2. How tablets and smartphones help people reach goals, make decisions, and talk to health care providers according to self-reported history of hypertension. Excludes those who answered “do not own a tablet or smartphone." HTN: hypertension.

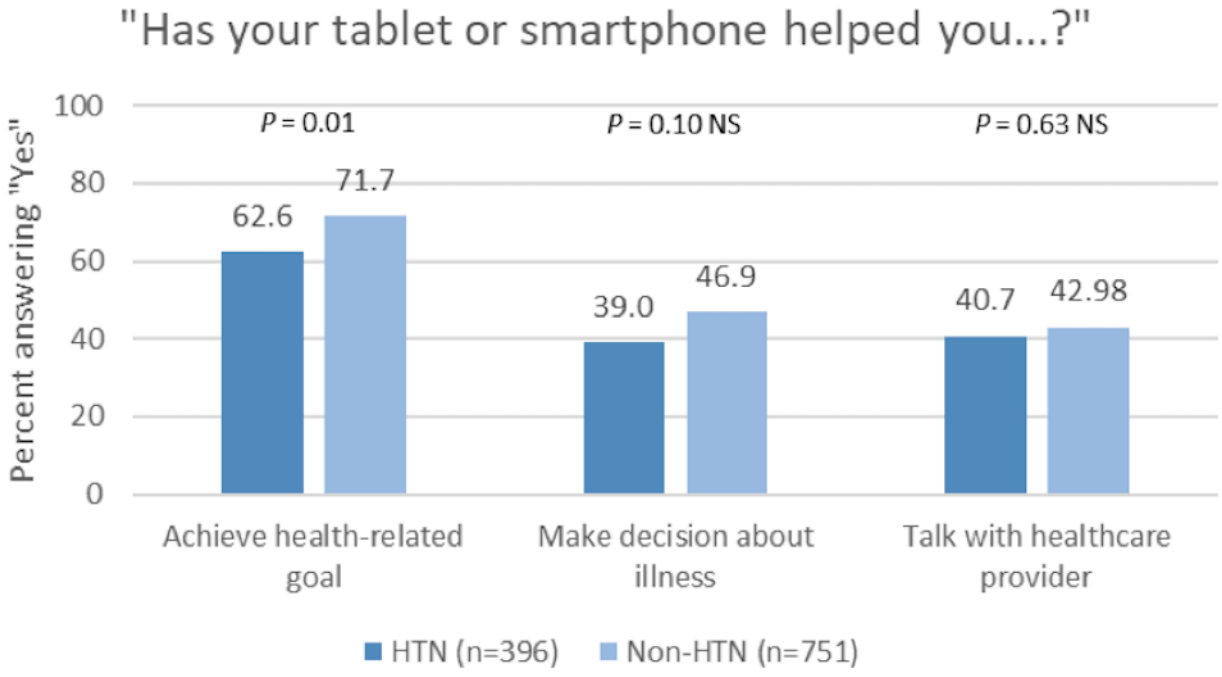
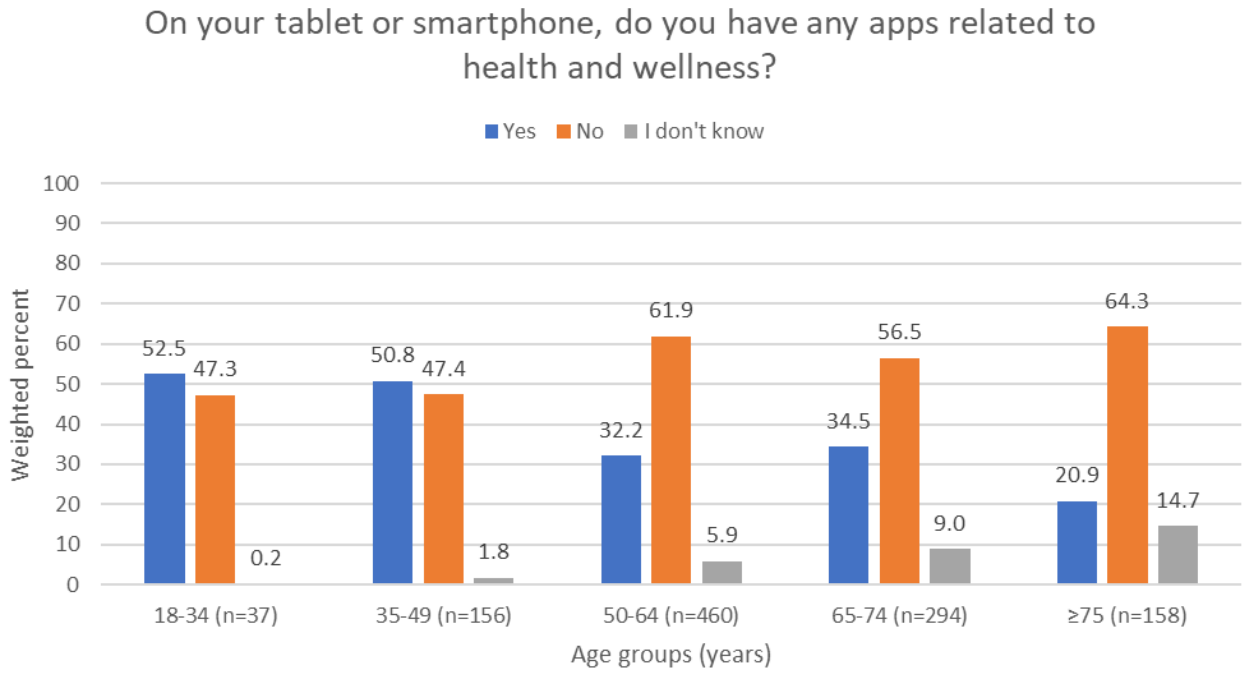


Figure 3. Among those with a self-reported history of hypertension exclusively, differences in health/wellness app use according to age. P=.001 for the test of whether the distribution of responses ("yes", "no", or "I don't know") differs by age group. Excludes those who answered "do not own a tablet or smartphone.".



Discussion

Principal Findings

The purpose of this study was to describe the prevalence of mobile device ownership and presence of health-related apps according to the self-reported hypertension status and sociodemographic factors in a nationally representative sample of US adults. We also examined whether smartphones and tablets helped people improve health goal achievement, medical decision making, and patient-provider communication. In summary, we found that people with a self-reported history of hypertension were less likely to own a smartphone or tablet, have health apps on their mobile devices, and report that smartphone and tablets helped them achieve a health-related goal compared to those without a self-reported history of hypertension. We also found that among people with a history of hypertension exclusively, smartphone ownership was associated with age and education, but not race or sex. Specifically, younger people and people with a college education are more likely to have smartphones, whereas older adults and people with a less formal education are more likely to have basic mobile phones.

A key finding from our study was that participants with a self-reported history of hypertension were significantly less likely to have health-related apps on their smartphones or tablets than those without a self-reported history of hypertension. This finding differs from other research that has shown no difference in health app downloads between people with and those without a chronic illness [18]. It is possible that people without a self-reported history of hypertension had other medical conditions for which they were using health-related apps (eg, type 2 diabetes) or they used general health-promotion apps more often (eg, apps to help track diet and physical activity). Researchers should be cautious about solely relying on health apps to support hypertension self-management by patients, as such reliance may cause them to overlook patients who do not use health apps. It is also possible that some patients were not using health-related apps due to limited eHealth literacy [37], attitudes about technology [38], or challenges with self-regulation and self-control [39,40]. Nevertheless, health-related apps may still be promising tools for a subset of the population that likes to engage with mHealth tools and prefers technology-based strategies for managing health.

Regarding the various ways that smartphones or tablets may help people, participants with a self-reported history of hypertension were less likely to report that their smartphones and tablets helped them reach a health-related goal like quitting smoking, losing weight, or increasing physical activity. This finding is notable, given that lifestyle behaviors can help people manage hypertension [41] and that tracking behaviors is a common feature of many health apps [42].

Strengths and Limitations

A strength of this study was our use of the HINTS data set. HINTS collects nationally representative data every few years from a large sample of US adults. This routine data collection allows researchers to monitor health communication trends over time. Given that many survey items are unique to HINTS and not available in other publicly available data sets, HINTS is a rich resource for evaluating various aspects of health behavior, information seeking, and trends in health communication. Despite its strengths, some limitations of this study must be noted. The hypertension status was self-reported in HINTS and not confirmed through a clinical diagnosis. Over- or underreporting of hypertension may have affected the results

in either direction; however, we are unable to quantify the degree of over- or underreporting in the HINTS data set. We have no information about whether people who took the survey were currently receiving antihypertensive medication; the level of blood pressure control among medication users, which may have implications for health app use generally; and how people may use mobile devices for health and wellness. We are unable to distinguish between the various types of hypertension that a person may have experienced (eg, pregnancy related, primary hypertension, or secondary hypertension caused by another medical problem). We do not have information about when participants were ever informed that they had hypertension. It is possible that people who were recently informed (eg, within the last 12 months) would be more motivated to use health-related apps than people who have been living with hypertension for many years, or vice versa.

Future Directions for Research

Given the small but growing amount of data on mHealth interventions for hypertension, our findings raise several questions for future investigation. For example, because there are many ways to support hypertension prevention and management (eg, pharmacological, nonpharmacological, mHealth, eHealth, print, and in person), more research is needed to determine patients' preferences for various interventions and whether these preferences are associated with long-term engagement. For example, it is possible that some people may prefer telephone-based counseling or text messaging interventions over health-related apps when given the choice. Further evaluation of which behavioral theory (or combination of theories) best predicts hypertension-related mHealth intervention uptake is also needed. A recent systematic review noted that few mHealth interventions are based on behavioral theories [43]. Of the limited studies that highlight such theories, the Health Belief Model is commonly selected [43]. Moreover, as more health apps are developed, further examination of the features that matter most to patients and the quality of these components warrant more attention [44,45]. For example, Khalid et al found that privacy and ethics concerns, hidden costs, interface design, and app crashes were commonly reported complaints among mobile app users [46]. Finally, further research on the correlations between an individual's eHealth literacy and mHealth use may inform how patients will respond to mHealth interventions for hypertension in the future [47].

Conclusions

Smartphone and tablet ownership and the presence of health apps on mobile devices are less common in people with a self-reported history of hypertension compared to those without a history of hypertension. Future studies should examine how to disseminate and implement mHealth interventions in the populations most affected by hypertension. Moving forward, a combination of novel mHealth interventions and traditional health communication strategies (eg, print, web based, in person, and telephone based) may be needed to reach a wide cross-section of patients with a self-reported history of hypertension.

Conflicts of Interest

None declared.

References

1. Lunde P, Nilsson BB, Bergland A, Kværner KJ, Bye A. The Effectiveness of Smartphone Apps for Lifestyle Improvement in Noncommunicable Diseases: Systematic Review and Meta-Analyses. J Med Internet Res 2018 May 04;20(5):e162 [FREE Full text] [doi: 10.2196/jmir.9751] [Medline: 29728346]
2. 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;12:CD007459. [doi: 10.1002/14651858.CD007459.pub2] [Medline: 23235644]
3. Veazie S, Winchell K, Gilbert J, Paynter R, Ivlev I, Eden KB, et al. Rapid Evidence Review of Mobile Applications for Self-management of Diabetes. J Gen Intern Med 2018 Jul;33(7):1167-1176. [doi: 10.1007/s11606-018-4410-1] [Medline: 29740786]
4. Lee J, Choi M, Lee SA, Jiang N. Effective behavioral intervention strategies using mobile health applications for chronic disease management: a systematic review. BMC Med Inform Decis Mak 2018 Dec 20;18(1):12 [FREE Full text] [doi: 10.1186/s12911-018-0591-0] [Medline: 29458358]
5. Holmen H, Wahl AK, Cvancarova SM, Ribu L. Tailored Communication Within Mobile Apps for Diabetes Self-Management: A Systematic Review. J Med Internet Res 2017 Jun 23;19(6):e227 [FREE Full text] [doi: 10.2196/jmir.7045] [Medline: 28645890]
6. Palmer MJ, Barnard S, Perel P, Free C. Mobile phone-based interventions for improving adherence to medication prescribed for the primary prevention of cardiovascular disease in adults. Cochrane Database Syst Rev 2018 Dec 22;6:CD012675. [doi: 10.1002/14651858.CD012675.pub2] [Medline: 29932455]
7. Mohammadi R, Ayatolahi Tafti M, Hoveidamanesh S, Ghanavati R, Pournik O. Reflection on Mobile Applications for Blood Pressure Management: A Systematic Review on Potential Effects and Initiatives. Stud Health Technol Inform 2018;247:306-310. [Medline: 29677972]

8. Rehman H, Kamal AK, Morris PB, Sayani S, Merchant AT, Virani SS. Mobile Health (mHealth) Technology for the Management of Hypertension and Hyperlipidemia: Slow Start but Loads of Potential. Curr Atheroscler Rep 2017 Mar;19(3):12. [doi: 10.1007/s11883-017-0649-y] [Medline: 28210974]
9. Kochanek K, Murphy S, Xu J, Arias E. Centers for Disease Control and Prevention. 2016. Mortality in the United States, 2016 URL: https://www.cdc.gov/nchs/products/databriefs/db293.htm [accessed 2018-12-24] [WebCite Cache ID 74tmK1Rtj]
10. Benjamin EJ, Virani SS, Callaway CW, Chamberlain AM, Chang AR, Cheng S, American Heart Association Council on EpidemiologyPrevention Statistics CommitteeStroke Statistics Subcommittee. Heart Disease and Stroke Statistics-2018 Update: A Report From the American Heart Association. Circulation 2018 Mar 20;137(12):e67-e492. [doi: 10.1161/CIR.0000000000000558] [Medline: 29386200]
11. Lewington S, Clarke R, Qizilbash N, Peto R, Collins R, Prospective SC. Age-specific relevance of usual blood pressure to vascular mortality: a meta-analysis of individual data for one million adults in 61 prospective studies. Lancet 2002 Dec 14;360(9349):1903-1913. [Medline: 12493255]
12. Rapsomaniki E, Timmis A, George J, Pujades-Rodriguez M, Shah AD, Denaxas S, et al. Blood pressure and incidence of twelve cardiovascular diseases: lifetime risks, healthy life-years lost, and age-specific associations in 1·25 million people. Lancet 2014 May 31;383(9932):1899-1911 [FREE Full text] [doi: 10.1016/S0140-6736(14)60685-1] [Medline: 24881994]
13. Whelton PK, Carey RM, Aronow WS, Casey DE, Collins KJ, Dennison Himmelfarb C, et al. 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. J Am Coll Cardiol 2018 May 15;71(19):e127-e248. [doi: 10.1016/j.jacc.2017.11.006] [Medline: 29146535]
14. Merai R, Siegel C, Rakotz M, Basch P, Wright J, Wong B, DHSc, et al. CDC Grand Rounds: A Public Health Approach to Detect and Control Hypertension. MMWR Morb Mortal Wkly Rep 2016 Nov 18;65(45):1261-1264 [FREE Full text] [doi: 10.15585/mmwr.mm6545a3] [Medline: 27855138]
15. van Heerden A, Tomlinson M, Swartz L. Point of care in your pocket: a research agenda for the field of m-health. Bull World Health Organ 2012 May 1;90(5):393-394 [FREE Full text] [doi: 10.2471/BLT.11.099788] [Medline: 22589575]
16. Healthcare Information and Management Systems Society.: Healthcare Information and Management Systems Society (HIMSS); 2012 Jan 05. Definitions of mHealth Internet URL: https://www.himss.org/definitions-mhealth [accessed 2018-12-24] [WebCite Cache ID 74tmhKa5F]
17. Anderson-Lewis C, Darville G, Mercado RE, Howell S, Di Maggio S. mHealth Technology Use and Implications in Historically Underserved and Minority Populations in the United States: Systematic Literature Review. JMIR Mhealth Uhealth 2018 Jun 18;6(6):e128 [FREE Full text] [doi: 10.2196/mhealth.8383] [Medline: 29914860]
18. 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]
19. Friedberg JP, Robinaugh DJ, Wang B, Allegrante JP, Lipsitz SR, Natarajan S. Who is being reached for a telephone-delivered intervention for patients with uncontrolled hypertension? Telemed J E Health 2014 Mar;20(3):229-234. [doi: 10.1089/tmj.2013.0071] [Medline: 24386927]
20. Urrea B, Misra S, Plante TB, Kelli HM, Misra S, Blaha MJ, et al. Mobile Health Initiatives to Improve Outcomes in Primary Prevention of Cardiovascular Disease. Curr Treat Options Cardiovasc Med 2015 Dec;17(12):59. [doi: 10.1007/s11936-015-0417-7] [Medline: 26474892]
21. Smith A, Page D. Pew Research Center.: Pew Research Center Internet; 2015. US smartphone use in 2015 URL: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/ [accessed 2018-12-24] [WebCite Cache ID 74tmluP3L]
22. Klasnja P, Pratt W. Healthcare in the pocket: mapping the space of mobile-phone health interventions. J Biomed Inform 2012 Feb;45(1):184-198 [FREE Full text] [doi: 10.1016/j.jbi.2011.08.017] [Medline: 21925288]
23. Pew Research Center. 2018 Feb 05. Mobile Fact Sheet Internet URL: http://www.pewinternet.org/fact-sheet/mobile/ [accessed 2018-12-24] [WebCite Cache ID 74tmvB22G]
24. Smith A. Pew Research Center. 2017. Record shares of Americans now own smartphones, have home broadband URL: http://www.pewresearch.org/fact-tank/2017/01/12/evolution-of-technology/ [accessed 2018-12-24] [WebCite Cache ID 74tnJ6BtU]
25. IQVIA. 2017 Nov 07. The Growing Value of Digital Health: Evidence and Impact on Human Health and the Healthcare System [URL: https://www.iqvia.com/institute/reports/the-growing-value-of-digital-health [accessed 2019-01-04] [WebCite Cache ID 75B6r4hXt]
26. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who Uses Mobile Phone Health Apps and Does Use Matter? A Secondary Data Analytics Approach. J Med Internet Res 2017 Apr 19;19(4):e125 [FREE Full text] [doi: 10.2196/jmir.5604] [Medline: 28428170]
27. Ernsting C, Dombrowski SU, Oedekoven M, Kanzler M, Kuhlmey A, Gellert P. Using Smartphones and Health Apps to Change and Manage Health Behaviors: A Population-Based Survey. J Med Internet Res 2017 Apr 05;19(4):e101 [FREE Full text] [doi: 10.2196/jmir.6838] [Medline: 28381394]

28. 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]
29. 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]
30. National Cancer Institute. Frequently Asked Questions | HINTS URL: https://hints.cancer.gov/about-hints/frequently-asked-questions.aspx [accessed 2018-12-24] [WebCite Cache ID 74tneaxee]
31. Finney RLJ, Davis T, Beckjord EB, Blake K, Moser RP, Hesse BW. Picking up the pace: changes in method and frame for the health information national trends survey (2011-2014). J Health Commun 2012;17(8):979-989 [FREE Full text] [doi: 10.1080/10810730.2012.700998] [Medline: 23020763]
32. Bersamin A, Stafford RS, Winkleby MA. Predictors of hypertension awareness, treatment, and control among Mexican American women and men. J Gen Intern Med 2009 Nov;24 Suppl 3:521-527 [FREE Full text] [doi: 10.1007/s11606-009-1094-6] [Medline: 19842001]
33. Majernick TG, Zacker C, Madden NA, Belletti DA, Arcona S. Correlates of hypertension control in a primary care setting. Am J Hypertens 2004 Oct;17(10):915-920. [doi: 10.1016/j.amjhyper.2004.05.016] [Medline: 15485754]
34. Centers for Disease Control and Prevention. Continuous NHANES Web Tutorial: Hypothesis Testing: Task 3b URL: https://www.cdc.gov/nchs/tutorials/NHANES/NHANESAnalyses/HypothesisTesting/Task3b.htm [accessed 2018-12-24] [WebCite Cache ID 74tnlYQFL]
35. Benjamini Y, Hochberg Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society. Series B (Methodological) 1995;57(1):289-300.
36. Columbia University Mailman School of Public Health. False Discovery Rate URL: https://www.mailman.columbia.edu/research/population-health-methods/false-discovery-rate [accessed 2018-12-24] [WebCite Cache ID 74tnriHbz]
37. Karnoe A, Furstrand D, Christensen KB, Norgaard O, Kayser L. Assessing Competencies Needed to Engage With Digital Health Services: Development of the eHealth Literacy Assessment Toolkit. J Med Internet Res 2018 May 10;20(5):e178 [FREE Full text] [doi: 10.2196/jmir.8347] [Medline: 29748163]
38. O'Leary K, Vizer L, Eschler J, Ralston J, Pratt W. Understanding patients' health and technology attitudes for tailoring self-management interventions. AMIA Annu Symp Proc 2015;2015:991-1000 [FREE Full text] [Medline: 26958236]
39. Bandura A. Social cognitive theory of self-regulation. Organizational Behavior and Human Decision Processes 1991 Dec;50(2):248-287. [doi: 10.1016/0749-5978(91)90022-L]
40. Muraven M, Baumeister RF. Self-regulation and depletion of limited resources: does self-control resemble a muscle? Psychol Bull 2000 Mar;126(2):247-259. [Medline: 10748642]
41. Booth JN, Abdalla M, Tanner RM, Diaz KM, Bromfield SG, Tajeu GS, et al. Cardiovascular Health and Incident Hypertension in Blacks: JHS (The Jackson Heart Study). Hypertension 2017 Dec;70(2):285-292 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.117.09278] [Medline: 28652461]
42. Dennison L, Morrison L, Conway G, Yardley L. Opportunities and challenges for smartphone applications in supporting health behavior change: qualitative study. J Med Internet Res 2013;15(4):e86 [FREE Full text] [doi: 10.2196/jmir.2583] [Medline: 23598614]
43. Cho Y, Lee S, Islam SMS, Kim S. Theories Applied to m-Health Interventions for Behavior Change in Low- and Middle-Income Countries: A Systematic Review. Telemed J E Health 2018 Feb 13. [doi: 10.1089/tmj.2017.0249] [Medline: 29437546]
44. Kumar N, Khunger M, Gupta A, Garg N. A content analysis of smartphone-based applications for hypertension management. J Am Soc Hypertens 2015 Feb;9(2):130-136. [doi: 10.1016/j.jash.2014.12.001] [Medline: 25660364]
45. 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;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
46. Khalid H, Shihab E, Nagappan M, Hassan AE. What Do Mobile App Users Complain About? IEEE Softw 2015 May;32(3):70-77. [doi: 10.1109/MS.2014.50]
47. Norman CD, Skinner HA. eHEALS: The eHealth Literacy Scale. J Med Internet Res 2006 Nov;8(4):e27 [FREE Full text] [doi: 10.2196/jmir.8.4.e27] [Medline: 17213046]

Abbreviations

eHealth: electronic health
HINTS: Health Information National Trends Survey
mHealth: mobile health

Edited by G Eysenbach; submitted 16.09.18; peer-reviewed by X Shen, J Edwards; comments to author 07.10.18; revised version received 23.11.18; accepted 09.12.18; published 14.01.19.

Please cite as:
Langford AT, Solid CA, Scott E, Lad M, Maayan E, Williams SK, Seixas AA
Mobile Phone Ownership, Health Apps, and Tablet Use in US Adults With a Self-Reported History of Hypertension: Cross-Sectional Study
JMIR Mhealth Uhealth 2019;7(1):e12228
URL: http://mhealth.jmir.org/2019/1/e12228/
doi:10.2196/12228
PMID:31344667

Original Paper

Use of Health Apps and Wearable Devices: Survey Among Italian Associations for Patient Advocacy

Paola Mosconi[1*], Biol Sci D; Silvia Radrezza[1*], Pharm D; Emanuele Lettieri[2], PhD Management Engineering; Eugenio Santoro[3*], Comp Sci D

[1]Laboratory for Medical Research and Consumer Involvement, Department of Public Health, Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milan, Italy

[2]Department of Management, Economics and Industrial Engineering, Politecnico of Milan, Milan, Italy

[3]Laboratory of Medical Informatics, Department of Public Health, Istituto di Ricerche Farmacologiche Mario Negri IRCCS, Milan, Italy

[*]these authors contributed equally

Corresponding Author:
Eugenio Santoro, Comp Sci D
Laboratory of Medical Informatics
Department of Public Health
Istituto di Ricerche Farmacologiche Mario Negri IRCCS
19, Via Giuseppe la Masa
Milan, 20156
Italy
Phone: 39 02 39014 ext 562
Fax: 39 02 33200231
Email: eugenio.santoro@marionegri.it

Abstract

Background: Technological tools such as Web-based social networks, telemedicine, apps, or wearable devices are becoming more widespread in health care like elsewhere. Although patients are the main users, for example, to monitor symptoms and clinical parameters or to communicate with the doctor, their perspective is seldom analyzed, and to the best of our knowledge, no one has focused on the patients' health care advocacy associations' point of view.

Objective: The objective of this study was to assess patients' health care advocacy associations' opinions about the use, usefulness, obstacles, negative aspects, and impact of health apps and wearable devices through a Web-based survey.

Methods: We conducted a Web-based survey through SurveyMonkey over nearly 3 months. Participants were contacted via an email explaining the aims of the survey and providing a link to complete the Web-based questionnaire. All the 20 items were mandatory, and the anonymized data were collected automatically into a database. Only fully completed questionnaires were considered for analysis.

Results: We contacted 1998 patients' health care advocacy associations; a total of 258 questionnaires were received back (response rate 12.91%), and 227 of the received questionnaires were fully completed (completion rate 88.0%). Informative apps, hospital apps for viewing medical reports or booking visits, and those for monitoring physical activity are the most used. They are considered especially useful to improve patients' engagement and compliance with treatment. Wearable devices to check physical activity and glycemia are the most widespread considering, again, their benefits in increasing patients' involvement and treatment compliance. For health apps and wearable devices, the main obstacles to their use are personal and technical reasons; the risk of overmedicalization is considered the most negative aspect of their constant use, while privacy and confidentiality of data are not rated a limitation. No statistical difference was found on stratifying the answers by responders' technological level (P=.30), age (P=.10), and the composition of the association's advisory board (P=.15).

Conclusions: According to responders, health apps and wearable devices are sufficiently known and used and are considered potential supports for greater involvement in health management. However, there are still obstacles to their adoption, and the developers need to work to make them more accessible and more useful. The involvement of patients and their associations in planning services and products based on these technologies (as well as others) would be desirable to overcome these barriers and boost awareness about privacy and the confidentiality of data.

(JMIR Mhealth Uhealth 2019;7(1):e10242) doi:10.2196/10242

KEYWORDS

health apps; mobile phone; Web-based survey; patients' health care advocacy associations; wearable devices

Introduction

New technologies, connectivity, and availability of the internet have boosted the development of mobile apps and wearable devices in the last decade. Those for health and care focus mainly on the management of chronic conditions such as diabetes [1], cardiovascular diseases [2], and specific populations or conditions [3,4]. Moreover, these devices have gained ground in health self-monitoring and preventive medicine [5]. Italy is the second most frequent worldwide user of new technology, particularly wearables, after the United States and before Germany and France [6].

Electronic health (eHealth) and medical devices such as apps and wearable devices are becoming more and more important in health debate. In the past year alone, health apps have grown by nearly 78,000 units, reaching a total of 325,000 registered [7]. Despite their spread, however, the use of these new technologies arouses discussions about the collection and sharing of data, focused on their protection, privacy, accuracy, and reliability [8,9].

A recent study about the most downloaded "mobile health (mHealth) apps" found that only 30.5% had a privacy policy. However, many of these documents used technical language not accessible to most lay people or they did not focus enough on the app itself [10,11]. Moreover, health data are often stored in the "cloud" so we cannot know where it is exactly, making it not completely safe [12].

There is still debate in the literature on the use and efficacy of apps and wearable devices to help patients collect and access their clinical data [13]. To date, patients' engagement is considered an important clinical outcome, and these tools could make them more responsible for treatment management, monitoring symptoms, identifying risk factors, and how best to prevent other diseases, for example, by changing their lifestyle [14]. Health apps or wearable devices could be suitable above all for chronic diseases. Evidence of this comes from a systematic review based on randomized controlled trials, where Yuan et al reported that the use of mobile apps by diabetic adults lowered glycemia more than standard care alone [15].

Several surveys on consumer perspective [16,17], use of mHealth app [18], and wearable devices [19] have been conducted regarding the use of technological tools in health care; however, to the best of our knowledge, none has been related to patients' health care advocacy associations. These associations, in Italy, like in other industrialized countries, are a growing reference point in the public health debate and innovation, influencing research and the political agenda [20]. The associations are spokespersons for a multiplicity of opinions and collecting their points of view is the best way to actually engage consumers and patients better. We conducted an observational study through a Web-based survey to analyze the representatives of health care advocacy associations' point of view.

Methods

Methodology

The Web-based survey was close, voluntary, and took only a few minutes to complete. No incentive was given. The SurveyMonkey Web-based survey service [21] was used to design the questionnaire, manage the survey, and collect data. The Web questionnaire was 5 pages long, and all data were anonymized and protected by Norton and TRUSTe.

An email was sent to 1998 contacts, describing the survey and including a link inviting them to complete the questionnaire. Reminder emails were sent after 3, 5, and 7 weeks. The questionnaire was organized in a series of linked pages (multiple-item screens) with electronic instructions to facilitate the flow. A progress indicator was permanently visible, and compilers could enter a personal comment for each question. The questions were in bold type, and the answers had optional buttons.

According to the "Checklist for Reporting Results of Internet E-Surveys" (CHERRIES), all the questions, except the last one, were mandatory to obtain more solid data on the endpoints of the survey. To guarantee the possibility of answering, most questions provided a nonresponse option, that is, "I don't know." Again, based on the CHERRIES checklist, we analyzed only completed questionnaires, excluding questionnaires that had missing data due to the responder stopping early and leaving the website.

If all answers were not completed, it was not possible to continue and confirm the questionnaire. A "back button" was provided to change answers before submitting them, but thereafter, no further changes were allowed. The answers were collected automatically in the SurveyMonkey database. Overall, 6 questionnaires were filled manually in the database by the coordinator center to clean up any technical problems of the responders. Of note, under Italian law, ethical approval was not required for this kind of survey.

Development of the Questionnaire

We conducted a literature review in PubMed using the following keywords: "digital technologies," "survey," "questionnaire," "eHealth," "mHealth," "digital health," "digital innovation," and "development." We retrieved a number of surveys about "digital innovation" involving citizens but none involved patients' health care advocacy associations. Based on the material collected, we drafted 20 items about the use, usefulness, obstacles, negative aspects, and impact related to eHealth technologies, focusing on health apps and wearable devices.

In addition, we recorded the characteristics of the patients' health care advocacy associations contacted (year of foundation,

setting, geographic distribution, and current composition of the advisory board) and of responders (sex, age, education, personal use of health apps, or wearable devices) to have a complete framework of the sample (see Multimedia Appendix 1). In this regard, some previous studies showed that age and individual technological level influenced a person's interest in the adoption of health apps also in clinical situations [22].

We pretested the questionnaire with the collaboration of 16 representatives of patients' health care advocacy associations to assess its readability, clarity, and completeness and to collect suggestions. Seven representatives answered, and a general positive consensus was gathered; hence, no major changes were required.

Recruitment

We started from a database of patients' health care advocacy associations available at the "Laboratory for Medical Research and Consumer Involvement" (n=2087); duplicate and wrong email addresses were removed (n=98). This database, developed in 2004, includes contacts of members of the board of associations—no single member—who had participated in previous research projects, training courses, and initiatives.

We sent one email invitation to each association, usually to the president of the association or a member of the board. The questionnaire asked for his or her personal opinion (see questions 7-10 and 14 in the Multimedia Appendix 1) and to report the common belief of the members represented (see questions 11-13 in the Multimedia Appendix 1).

The survey was announced through Web on the Mario Negri Institute websites and 2 patients' health care association websites; 9 health care associations requested to participate by contacting the coordinator center directly. Overall, 1998 patients' health care advocacy associations were reached via emails. A unique site visitor was permitted by checking the internet protocol address of responder. Results were presented according to CHERRIES" [23] and a similar guideline formulated by Bennett et al [24].

Statistical Analysis

Descriptive analyses were conducted using SAS version 9.4 (SAS Institute Inc) to cross variables according to several criteria. On the basis of the previous studies suggestions, we analyzed the pattern of answers by responders' age and their technological level. In addition, we considered the composition of the advisory board of the patients' health care advocacy associations as a possible confounding factor. A statistical test (*P* value) was expected only for results that gave significant patterns using a predetermined alpha level of .05.

Results

The survey was open from March 23 to June 8, 2017. A total of 258 answers were collected, giving a response rate of 12.91% (258/1998). We analyzed 227 completed questionnaires, as 31 responders dropped out before completing the questionnaire, giving a completion rate of 88.0% (227/258).

Tables 1 and 2 summarize the main characteristics of the patients' health care advocacy associations and responders, respectively. Out of 227 responders, 121 (53.3%) worked locally, while 43 (18.9%) and 63 (27.8%) worked on the regional and national levels, respectively. Most associations were based in the north of Italy. The advisory board comprised all or most of the patients or relatives for 58.6% (133/227) answers; furthermore, 18.1% (41/227) responders stated that there were no patients or relatives on their board. Respondents were aged 25-88 (mean age 56, SD 12) years; females and males were fairly represented, and 46.3% (105/227) respondents reported an appreciable level of technology, using at least one health app or wearable device, or both.

According to the 227 participants, Web-based social networks were used for health communication, promotion, or grouping patients with a specific disease, and telemedicine had a great impact on health care out of all technological innovations, with 183 (80.6%) responders and 178 (78.4%) of preferences, respectively; these were followed by wearable devices, such as smart-watches or wristbands, and health apps, with 148 (65.2%) and 146 (64.3%) responders, respectively. However, they foresaw that in the next 3 years, the impact would increase more for health apps, wearable devices, and telemedicine services than for Web-based social networks (Table 3).

Among health apps, those providing health and disease information, hospital apps for services such as viewing medical reports or booking visits, and those for tracking fitness or physical activity were the most used by members of health care advocacy associations, with 149 (65.6%), 118 (52.0%), and 116 (51.1%) of 227 participants, respectively. Apps to improve treatment and medication compliance, health or condition trackers, and diet or nutrition apps followed with 106 (46.7%), 96 (42.2%), and 80 (35.2%) of 227 responders, respectively. Symptom-checker apps for self-diagnosis were the least used, with 22.0% (50/227) positive answers. Wearable devices monitoring physical activity and glycemia were the most widespread among 227 respondents—110 (48.5%) and 108 (47.6%), respectively—followed by those monitoring heart rate and blood pressure, weight, and sleep tracking, with 106 (46.7%), 88 (38.8%), and 64 (28.2%) respondents, respectively.

From the point of view of patients' health care advocacy associations' representatives, health apps and wearable devices are not only useful to improve patients' engagement in their own health and treatment compliance but also to help them understand their health status and conditions, enhancing the communication between patients and physicians, and reducing health care costs (Table 4).

Table 1. Characteristics of patients' health care advocacy associations (N=227).

Characteristics	Associations, n (%)
Year of foundation	
1940-1999	119 (52.4)
2000-2016	108 (47.6)
Types	
Oncology	38 (16.7)
Diabetes	34 (15.0)
Rare diseases	28 (12.3)
Neurology	19 (8.4)
Cardiovascular	15 (6.6)
Disability	12 (5.3)
Breast cancer	13 (5.7)
Pediatric	7 (3.1)
AIDS	5 (2.2)
Brain-injured	4 (1.8)
Autism	2 (0.9)
Asthma	1 (0.4)
Other	49 (21.6)
Geographic distribution	
North	126 (55.5)
Center	50 (22.0)
South-Islands	51 (22.5)
Weblink	
Website	203 (89.4)
Facebook account	186 (81.9)
Twitter account	80 (35.2)
YouTube channel	74 (32.6)
Blog	43 (18.9)
Advisory board	
All members are patients or their relatives	78 (34.4)
Most members are patients	55 (24.2)
Patients' representatives are nearly half	27 (11.9)
Patients' representatives are a minority	26 (11.5)
There are no patients' representatives	41 (18.1)

Table 2. Characteristics of responders (N=227).

Characteristics	Responders, n (%)
Sex	
Males	100 (44.1)
Females	127 (56.0)
Age in years	
≤50	67 (29.5)
51-65	106 (46.7)
>65	54 (23.8)
Education level	
Elementary school	1 (0.4)
Secondary school	10 (4.4)
High school	93 (41.0)
Degree or superior	121 (53.3)
Other	2 (0.9)
Personal use of health app or wearable	
Both	38 (16.7)
Only health app	50 (22.0)
Only wearable	17 (7.5)
None	122 (53.7)

Table 3. Impact of technological tools on medical care and health (N=227).

Technological tool	Responders, n (%)			
	Current impact		Future impact	
	Yes	No	Yes	No
Health app	146 (64.3)	81 (35.7)	192 (84.6)	35 (15.4)
Wearable	148 (65.2)	79 (34.8)	183 (80.6)	44 (19.4)
Telemedicine	178 (78.4)	49 (21.6)	211 (93.0)	16 (7.1)
Social network	183 (80.6)	44 (19.4)	197 (86.8)	30 (13.2)

Table 4. Utility of health apps and wearables in health care (N=227).

Utility	Responders, n (%)	
	Positive effect	Negative or no effect
Empowerment in own health	204 (89.9)	23 (10.1)
Improve doctor-patient communication	167 (73.6)	60 (26.4)
Understand own health condition	179 (78.9)	48 (21.1)
Reduce public health costs	127 (55.9)	100 (44.1)
Improve compliance	186 (81.9)	41 (18.1)

Responders reported that the main obstacles to the adoption of health apps and wearable devices among members of their health care advocacy associations were personal motivations, such as concern about not being able to use them, or technical reasons, including owning an unsuitable smartphone, followed by the lack of evidence of their usefulness, accuracy, and reliability. Conversely, the lack of confidence in data protection and confidentiality seemed to restrict their use for only 31.3% (71/227) responders. It was interesting to highlight that about a quarter of responders had no clear opinion about the privacy and accuracy of data collected through apps and wearable devices (Table 5).

Table 5. Obstacles related to the use of health apps and wearables (N=227).

Obstacles	Yes, n (%)	No, n (%)	I don't know, n (%)
Technical barrier	144 (63.4)	52 (22.9)	31 (13.7)
Personal opinion	144 (63.4)	48 (21.2)	35 (15.4)
Low trust in recorded data utility	73 (32.2)	89 (39.2)	65 (28.6)
Low trust in data reservation and privacy	70 (30.8)	100 (44.1)	57 (25.1)
Low trust in recorded data reliability and quality	71 (31.3)	80 (35.2)	76 (33.5)
Low example of their utility in public health	90 (39.6)	59 (26.0)	78 (34.4)

Table 6. Negative aspects of constant adoption of health apps and wearable devices (N=227).

Negative aspects	Yes, n (%)	No, n (%)
Become dependent	148 (65.2)	79 (34.8)
No privacy	70 (30.8)	157 (69.2)
Excessive control of own health	149 (65.6)	78 (34.4)
Overmedicalization	160 (70.5)	67 (29.5)
Weaken doctor-patient communication	101 (44.5)	126 (55.5)

Among negative aspects or risks about the use of health apps and wearable devices, respondents considered first overmedicalization (ie, the overuse of drugs, supplements, and medical devices or scheduling medical examinations even if they are not really needed), followed by the risk of becoming dependent on technology, or weakening patient-doctor communication (Table 6). The lack of confidence in the protection and confidentiality of data was considered a negative aspect only by 30.8% (70/227) responders in agreement with the answers related to the obstacles reported in Table 5.

When we asked about the kind of health apps on which developers should focus in future to improve medical care and health, respondents put first those improving and disseminating health services and tools, with out of 227, 186 (81.9%) preferences, followed by those boosting compliance (175, 77.1%) and those monitoring vital signs (156, 68.7%), diet and nutrition (147, 64.8%), and physical activity (137, 60.4%). Of the 227 respondents, about two-third (154, 67.8%) asked to focus the efforts on informative health apps, while only 101 (44.5%) wanted more attention to symptom-checker apps. For wearable devices, the respondents stated that more attention could be paid to heart rate and blood pressure monitoring (167, 73.4%), glycemia (165, 72.7%), weight monitoring (139, 61.2%), and fitness trackers (136/227, 60.0%); less interest was shown in sleep trackers (112, 49.3%).

Data about the use, usefulness for health care, obstacles, and negative aspects of health apps and wearable devices were stratified by the respondents' age (≥58 years), technological level (users of one health app or one wearable device at least compared to not users), and the composition of the association's advisory board (all or most are patients compared to few or no patients). There were no statistically significant associations, although there was a slight influence of the technological level of responders. High technological level responders reported a greater positive impact of all technological tools (health apps, wearable devices, telemedicine, and Web-based social networks) on health and health care, now and in the future than the lower technological level responders. In addition, attitudes were different toward the usefulness of health apps and wearable devices. According to high technological responders, patients' engagement and compliance with treatment related to these tools were most important. In addition, higher technological level responders considered the risk of becoming dependent and excessive control of one's own health less important than the less technological ones.

Discussion

This study offers an exclusive view of patients' health care advocacy associations' opinions about eHealth technological tools that have not yet been well explored. The results suggest that the most commonly used health apps appear to be informative apps, apps providing access to hospital services, and fitness or physical activity tracking apps, while the favorite wearable devices are those for fitness, blood glucose, and heart rate monitoring. Our findings are different from those reported in a recent study based on citizens where fitness, diet or nutrition, and symptom navigator apps were at the top of the rankings with 59%, 52%, and 36% of use, respectively [17]. In addition, our results differ from those reported in a recent study focused on consumer's perceived attitudes about wearable devices in health monitoring, in which exercise coaching (61%) and location tracking (59%) came first [19].

Despite this, our results are similar to those from a study conducted on patients and confirm how both health apps and wearable devices can be used for patient engagement. In fact, a recent survey found that the main reasons for their adoption were chronic disease management (81%), support for medical adherence (66%), and fitness tracking (46%) [13].

Respondents are optimistic about the future of health apps and wearable devices. They suggest that in the near future, developers should focus on apps providing more services,

increasing drug and therapy compliance, and monitoring vital signs.

The gaps between current opinion and needs may be explained considering that today's responders are concerned about the reliability of the data collected by health apps and wearable devices but are confident that in the future, more evidence will be provided to support their use and efficacy in collecting data for diseases monitoring and management. On the other hand, our survey shows that the lack of evidence of reliability and accuracy of health apps and wearable devices, along with difficulty in adopting them, are the main obstacles to their use.

About the accuracy—but also the effectiveness—of these tools, researchers need to raise the overall quality of interventional trials conducted on patients, focusing on mobile apps and wearable devices. These trials are still limited due to the short follow-up, low recruitment rate, and high proportion of withdrawals before the scheduled time. It is, therefore, hard to see whether they might become valuable for helping patients with their own health [25].

Other studies raised the question of discontinuing the use of health apps and wearable devices when the real setting is not taken into consideration. For example, a recent study showed that the majority of users increased their physical activity after purchasing a wearable device, but nearly one-third stopped tracking after just 6 months [26]. Another study found that 55% of the 325,000 health apps available in the app stores are downloaded <5000 times and only 2% of all health apps count >500,000 monthly active users [7]. Fuller involvement of patients and their associations to identify specific needs could probably narrow this gap and make wearable devices and health apps more appealing and useful for engagement [27].

The lack of confidence in data protection and confidentiality seemed not to limit the use of health apps and wearable devices and did not appear to be one of the main negative aspects. This is very important because the adoption of apps and wearable devices in health care might give rise to challenges about security, data protection, and data reuse [8,9]. For example, a study on 79 apps certified as clinically safe and trustworthy by the "UK-National Health Service Health Apps Library" revealed that 89% of them transmitted information to Web-based services, and none encrypted personal information stored locally. Two-thirds of the apps sent personal information over the internet without encryption and 20% did not have a specific privacy policy [28]. However, it is interesting that about a quarter of our responders knew little about the question of data privacy and confidentiality; a possible explanation is that patients' representatives themselves do not know enough, necessitating more awareness and information. An alternative reason is that participants do not care much about these topics and are willing to exchange part of their data for a potential health care gain.

This survey has several limitations. First, respondents are representatives of health care advocacy associations and answered reporting their point of view and the patients' perspective. Second, as Italy has no central database of consumers' and patients' associations, we contacted a limited number of associations; hence, this sample may not be representative of all Italian situations and attitudes. The response rate, 12.91% (258/1998), may have increased the selection bias, but it agrees with other similar studies. The response rate of Web-based surveys is often low, and many strategies have been investigated to increase it [29]. Finally, the heterogeneity of the health care advocacy associations that participated could have added further selection bias.

In conclusion, the survey shows that health apps and wearable devices are sufficiently used and appreciated by patients as potential supports for greater engagement in their health. There are still obstacles to their use, however, on which developers should work to make them more accessible and more useful. The involvement of patients and their health care advocacy associations in designing services and products based on these technologies—and others—is desirable to overcome barriers and make their development and acceptance easier and more competitive.

Acknowledgments

Associations participating in the pretested questionnaire are as follows:

AMA Associazione malattia di Alzheimer, Associazione Centro Donna, Associazione Progetto Luna, ATRACTO Onlus, AVULSS di Palermo, Tarlov Italia Onlus, and Voglia di Vivere onlus.

Associations that have provided consent to be cited are as follows:

Familiari Cardiopatici Onlus, Un cuore per amico, ADA Associazione Diabetici Alcamese, ADOCM Crisalide, AFADOC Onlus, AGAD associazione Giovani e Adulti con Diabete Onlus, Mantova, AGD Catanzaro Associazione Giovani Diabetici, AGOP Onlus Associazione Genitori Oncologia Pediatrica, AIDO Associazione Italiana Donatori Organi tessuti e cellule, AITA Abruzzo Onlus, ALICe Sardegna, ANIAD Onlus Associazione Nazionale Italiana Atleti Diabetici, ANIO Onlus Associazione Nazionale per le Infezioni Osteoarticolari, ACLTI Associazione contro le Leucemie e i Tumori nell'Infanzia, AEL Onlus Associazione Emofilici Lazio, AGD Piemonte e Valle d'Aosta, AILS, AIMA Associazione Italiana Malattia di Alzheimer, AIRP Associazione Italiana Rene Policistico, AISA Sezione Lombarda Onlus, AISM, AISTOM Associazione Italiana Stomizzati, AITSaM Onlus Associazione Italiana Tutela Salute Mentale, ALA Milano Onlus, Albinit, ALOMAR Associazione Lombarda Malati Reumatici, Alteg, Alzheimer Uniti Roma Onlus, Amici del Cuore "A Borin" Padova, Amici del Cuore di Camposampiero e della Strada del Santo, Amici del Cuore Verbania, Amici di Gianni Brundu, AMOC Onlus Associazione Malati Oncologici Colonretto, ANED Associazione Nazionale Emodializzati Dialisi e Trapianto Onlus, ANGEA Associazione Nazionale Genitori Eczema Atopico E

Allergie Alimentari Onlus, ANGSA La Spezia, ANLAIDS Onlus, ANVOLT Associazione Nazionale Volontari per la Lotta contro i Tumori Udine, APMAR Onlus Associazione Persone Con Malattie Reumatologiche e Rare, ASLIDIA Associazione Ligure per la lotta contro il Diabete Sanremo, ASNET Associazione Sarda Emodializzati e Trapiantati della Sardegna, ASRI Associazione Solidarietà Rifugiati e Immigrati, Associazione Sedna dei Diabetici Onlus, Associazione Traumi Cranici provincia di Reggio Emilia e Modena, Associazione Acromati Italiani, Associazione Aldo Perini Onlus, Associazione Amici Del Ceppo, Associazione Anastasis Onlus, Associazione Comunità Progetto Sud Onlus, Associazione culturale La Città di Pulcinella, Associazione di promozione sociale I Girasoli, Associazione Diabete Iglesias Carbonia ADIC Onlus, Associazione Diabete Versilia, Associazione Diabetici Area Pratese Onlus, Associazione Diabetici Basso Molise, Associazione Diabetici Copparo ADICO, Associazione Diabetici della provincia di Grosseto, Associazione Diabetici di Cinisello e comuni Limitrofi ADCL, Associazione Diabetici Marsicana, Associazione Genitori La Nostra Famiglia, Associazione Gianmarco De Maria, Associazione Giuliana Cerretti per l'Oncologia Onlus, Associazione Interzona, Associazione Iosempredonna Onlus, Associazione Italiana Celiachia Lombardia Onlus, Associazione Italiana Cistite Interstiziale, Associazione Italiana Dislipidemie Ereditarie, Associazione Italiana Estrofia Vescicale Epispadia Onlus, Associazione Italiana Laringectomizzati AILAR Onlus, Associazione Italiana Pazienti Anticoagulati, Firenze, Associazione Italiana Pazienti BPCO, Associazione Italiana Persone Down, Belluno, Associazione Italiana Stomizzati Sicilia, Associazione Ligure Sindrome x-fragile onlus, Associazione Nazionale Alfa1 AT, Associazione Nazionale Porpora Trombotica Trombocitopenica Onlus, Associazione Onlus Carmine Speranza, Associazione Parkinson Rovigo & Amici Onlus, Associazione Progetto Endometriosi, Associazione Reggiana per la Lotta e Cura dell'AIDS Onlus, Associazione Sarda Paratetraplegici, Associazione Scientifica Culturale Alter Ego, Associazione Serena a Palermo Onlus, Associazione Traumi Cranici Toscani ATRACTO Onlus, Associazione Vitadidonna Onlus, Associazione Vittorio Lodini progetto Seno, AVULSS Associazione di Volontariato nelle Unità Locali dei Servizi Sociosanitari, Palermo, Cerignola per l'oncologia Onlus, CFS Associazione Italiana Onlus, Cibo Amico allergia alimentare e anafilassi, CIDP Italia Onlus, CLEO Club Epatologi Ospedalieri, Diabete Brescia Onlus Brescia e Provincia, Diabete Zero Onlus, Difendiamoci dal Diabete Cittanova, Donna Per Donna Onlus, ESA Educazione alla Salute Attiva, Europa Donna Italia, Fand Ogliastra, Fand Milano, FAVO, Federasma e allergie, Federazione Alzheimer Italia, Federazione Diabete Sicilia, Federazione Italiana Incontinenti e Disfunzioni del Pavimento Pelvico Fincopp, Federazione Regionale Associazioni Toscane Diabetici Onlus, Fondazione Alessandra Bisceglia W Ale Onlus, Fondazione ANT Italia Onlus, Fondazione Attilia Pofferi Onlus, GILS Gruppo Italiano per la Lotta alla Sclerodermia Onlus, Gruppo di discussione e azione "Italia Glioblastoma Multiforme cancro al cervello", GSD Non Vedenti Milano Onlus, INSU' Associazione Giovani Diabetici Onlus, Inversa Onlus, InVIta la vita Onlus, La Lampada di Aladino Onlus, Le Donne Scelgono, Lega Italiana per la Lotta contro i Tumori Imperia Sanremo, Legaconsumatori Lucca, LIFC Piemonte Onlus, LILT Lega Italiana Per La Lotta Contro I Tumori, LILT Lega Italiana Per La Lotta Contro I Tumori Forli' Cesena, LILT Lega Italiana Per La Lotta Contro I Tumori Ragusa, LILT Lega Italiana Per La Lotta Contro I Tumori Rimini, MEDeA Onlus, Pavia nel Cuore Onlus, Per Andare Oltre Onlus, Plus Onlus, PRODES Progetto Diabete e Salute Fand Roma, Progetto Luna, Progetto Luna Onlus, Progetto Vita Onlus, Salute Donna Onlus, SAMOT Onlus, Sardegna Medicina, Speranza Onlus Associazione Familiari Diversabili Psichici, Tarlov Italia Onlus, UFHa unione famiglie handicappati, Unione Italiana Lotta alla Distrofia Muscolare Pisa,Verso Il Sereno Onlus, Vivere senza stomaco si può, Voglia di Vivere Onlus, and WALCE Onlus.

We thank Dr Alessandra Lugo for assistance in data analysis and Gianna Costa for secretarial support. We are grateful to JD Baggott for editing the manuscript.

This study was funded by an unconditional grant received by the Mario Negri Institute IRCCS from the Permanent Observatory on Digital Innovation in Healthcare at the Politecnico di Milano.

Conflicts of Interest

None declared.

Authors' Contributions

PM and ES conceived the study. PM, SR, and ES participated in the design of the survey and the questionnaire and collected, managed, and analyzed data. All the authors contributed with data interpretation. PM, SR, and ES drafted the manuscript; PM, SR, EL, and ES critically reviewed the manuscript and approved the final version.

Multimedia Appendix 1

Questionnaire provided to participants (English version).

[PDF File (Adobe PDF File), 400KB - mhealth_v7i1e10242_app1.pdf]

References

1. Jacques Rose K, Petrut C, L'Heveder R, de Sabata S. IDF Europe position on mobile applications in diabetes. Diabetes Res Clin Pract 2017 Sep 07. [doi: 10.1016/j.diabres.2017.08.020] [Medline: 28951337]

2. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M, Herreros-González J. Mobile apps in cardiology: review. JMIR Mhealth Uhealth 2013 Jul 24;1(2):e15 [FREE Full text] [doi: 10.2196/mhealth.2737] [Medline: 25098320]
3. Pinto S, Caldeira S, Martins JC. e-Health in palliative care: review of literature, Google Play and App Store. Int J Palliat Nurs 2017 Aug 02;23(8):394-401. [doi: 10.12968/ijpn.2017.23.8.394] [Medline: 28854054]
4. Santoro E, Castelnuovo G, Zoppis I, Mauri G, Sicurello F. Social media and mobile applications in chronic disease prevention and management. Front Psychol 2015;6:567 [FREE Full text] [doi: 10.3389/fpsyg.2015.00567] [Medline: 25999884]
5. Piwek L, Ellis DA, Andrews S, Joinson A. The Rise of Consumer Health Wearables: Promises and Barriers. PLoS Med 2016 Feb;13(2):e1001953 [FREE Full text] [doi: 10.1371/journal.pmed.1001953] [Medline: 26836780]
6. Kantar Worldpanel ComTech. Marketwired. 2016. Kantar Launches Quarterly Report for Wearable Technology URL: http://www.marketwired.com/press-release/kantar-launches-quarterly-report-for-wearable-technology-2121234.htm [accessed 2018-11-14] [WebCite Cache ID 73wHYcjno]
7. Research2Guidance. 2017 Nov. mHealth App Economics 2017: Current Status and Future Trends in Mobile Health URL: https://research2guidance.com/product/mhealth-economics-2017-current-status-and-future-trends-in-mobile-health/ [accessed 2018-11-18]
8. Privacy Rights Clearinghouse. Jul 01. Mobile health and fitness apps: what are the privacy risks? URL: https://www.privacyrights.org/consumer-guides/mobile-health-and-fitness-apps-what-are-privacy-risks [accessed 2018-11-15] [WebCite Cache ID 73wOmzuYl]
9. BinDhim NF, Trevena L. Health-related smartphone apps: regulations, safety, privacy and quality. BMJ Innov 2015 Mar 05;1(2):43-45. [doi: 10.1136/bmjinnov-2014-000019]
10. Torous J, Roberts LW. The Ethical Use of Mobile Health Technology in Clinical Psychiatry. J Nerv Ment Dis 2017 Jan;205(1):4-8. [doi: 10.1097/NMD.0000000000000596] [Medline: 28005647]
11. 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. [doi: 10.1136/amiajnl-2013-002605] [Medline: 25147247]
12. Leenes R, van Barkel R, Gutwirth S, De Hert P. Data Protection and Privacy: (In)visibilities and Infrastructures. Cham, Switzerland: Springer International Publishing; 2017:796.
13. Volp KG, Mohta NS. NEJM Catalyst. 2017 Jul 10. Patient Engagement Survey: Technology Tools Gain Support — But Cost Is a Hurdle URL: https://catalyst.nejm.org/patient-engagement-technology-tools-gain-support/
14. European Commission. 2014 Apr 10. Green Paper on mobile health (“mHealth”) URL: https://ec.europa.eu/digital-single-market/en/news/green-paper-mobile-health-mhealth [accessed 2018-11-15] [WebCite Cache ID 73wbTUeQt]
15. Wu Y, Yao X, Vespasiani G, Nicolucci A, Dong Y, Kwong J, et al. Mobile App-Based Interventions to Support Diabetes Self-Management: A Systematic Review of Randomized Controlled Trials to Identify Functions Associated with Glycemic Efficacy. JMIR Mhealth Uhealth 2017 Mar 14;5(3):e35 [FREE Full text] [doi: 10.2196/mhealth.6522] [Medline: 28292740]
16. European Commission. 2014 Oct. Public Perceptions of Science, Research and Innovation URL: http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_419_en.pdf
17. Accenture Consulting. Patients Want a Heavy Dose of Digital URL: https://www.accenture.com/_acnmedia/PDF-8/Accenture-Patients-Want-A-Heavy-Dose-of-Digital-Infographic-v2.pdf [accessed 2018-11-15] [WebCite Cache ID 73we9Azyx]
18. Carroll JK, Moorhead A, Bond R, LeBlanc WG, Petrella RJ, Fiscella K. Who Uses Mobile Phone Health Apps and Does Use Matter? A Secondary Data Analytics Approach. J Med Internet Res 2017 Dec 19;19(4):e125 [FREE Full text] [doi: 10.2196/jmir.5604] [Medline: 28428170]
19. Wen D, Zhang X, Lei J. Consumers' perceived attitudes to wearable devices in health monitoring in China: A survey study. Comput Methods Programs Biomed 2017 Mar;140:131-137. [doi: 10.1016/j.cmpb.2016.12.009] [Medline: 28254069]
20. Richards T, Montori VM, Godlee F, Lapsley P, Paul D. Let the patient revolution begin. BMJ 2013 May 14;346(may14 1):f2614-f2614. [doi: 10.1136/bmj.f2614] [Medline: 23674136]
21. SurveyMonkey. URL: https://www.surveymonkey.com/ [accessed 2018-11-18]
22. de Vries ST, Denig P, Lasheras Ruiz C, Houÿez F, Wong L, Sutcliffe A, IMI Web-RADR Work Package 3b Consortium. Interest in a Mobile App for Two-Way Risk Communication: A Survey Study Among European Healthcare Professionals and Patients. Drug Saf 2018 Dec;41(7):697-712 [FREE Full text] [doi: 10.1007/s40264-018-0648-0] [Medline: 29500800]
23. Eysenbach G. Improving the quality of Web surveys: the Checklist for Reporting Results of Internet E-Surveys (CHERRIES). J Med Internet Res 2004 Dec 29;6(3):e34 [FREE Full text] [doi: 10.2196/jmir.6.3.e34] [Medline: 15471760]
24. Bennett C, Khangura S, Brehaut JC, Graham ID, Moher D, Potter BK, et al. Reporting guidelines for survey research: an analysis of published guidance and reporting practices. PLoS Med 2010 Aug;8(8):e1001069 [FREE Full text] [doi: 10.1371/journal.pmed.1001069] [Medline: 21829330]
25. Covolo L, Ceretti E, Moneda M, Castaldi S, Gelatti U. Does evidence support the use of mobile phone apps as a driver for promoting healthy lifestyles from a public health perspective? A systematic review of Randomized Control Trials. Patient Educ Couns 2017 Dec;100(12):2231-2243. [doi: 10.1016/j.pec.2017.07.032] [Medline: 28855063]
26. Fox G, Garland S, Keibel A, Saxon LA. NEJM Catalyst. 2017 Sep 14. Why People Stick With or Abandon Wearable Devices URL: https://catalyst.nejm.org/stay-abandon-wearable-devices/ [accessed 2018-11-15] [WebCite Cache ID 73wcgyoYb]

27. Machado G, Pinheiro MB, Lee H, Ahmed OH, Hendrick P, Williams C, et al. Smartphone apps for the self-management of low back pain: A systematic review. Best Pract Res Clin Rheumatol 2016 Dec;30(6):1098-1109 [FREE Full text] [doi: 10.1016/j.berh.2017.04.002] [Medline: 29103552]
28. Huckvale K, Prieto JT, Tilney M, Benghozi P, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Med 2015 Sep 07;13:214 [FREE Full text] [doi: 10.1186/s12916-015-0444-y] [Medline: 26404673]
29. Fan W, Yan Z. Factors affecting response rates of the web survey: A systematic review. Computers in Human Behavior 2010 Mar;26(2):132-139. [doi: 10.1016/j.chb.2009.10.015]

Abbreviations

CHERRIES: Checklist for Reporting Results of Internet E-Surveys
eHealth: electronic health
mHealth: mobile health

Edited by G Eysenbach; submitted 27.02.18; peer-reviewed by E Bergner, P Petaipimol; comments to author 23.07.18; revised version received 27.09.18; accepted 27.09.18; published 15.01.19.

Please cite as:
Mosconi P, Radrezza S, Lettieri E, Santoro E
Use of Health Apps and Wearable Devices: Survey Among Italian Associations for Patient Advocacy
JMIR Mhealth Uhealth 2019;7(1):e10242
URL: http://mhealth.jmir.org/2019/1/e10242/
doi:10.2196/10242
PMID:30664455

Original Paper

Mobile Phone–Based Telemedicine Practice in Older Chinese Patients with Type 2 Diabetes Mellitus: Randomized Controlled Trial

Chenglin Sun[1*], MD, PhD; Lin Sun[1*], MD; Shugang Xi[1*], MD, PhD; Hong Zhang[1], MD; Huan Wang[1], MD; Yakun Feng[2], MD; Yufeng Deng[1], MD; Haimin Wang[1], MD; Xianchao Xiao[1], MD, PhD; Gang Wang[1], MD; Yuan Gao[1], MD; Guixia Wang[1], MD, PhD

[1]Department of Endocrinology, First Hospital of Jilin University, Changchun, China

[2]Health Management Center, First Affiliated Hospital of Zhengzhou University, Zhengzhou, China

[*]these authors contributed equally

Corresponding Author:
Chenglin Sun, MD, PhD
Department of Endocrinology
First Hospital of Jilin University
71 Xinmin Street
Changchun,
China
Phone: 86 0431 88783212
Fax: 86 0431 88786066
Email: clsun213@163.com

Abstract

Background: Previous studies on telemedicine interventions have shown that older diabetic patients experience difficulty in using computers, which is a barrier to remote communication between medical teams and older diabetic patients. However, older people in China tend to find it easy to use mobile phones and personal messaging apps that have a user-friendly interface. Therefore, we designed a mobile health (mHealth) system for older people with diabetes that is based on mobile phones, has a streamlined operation interface, and incorporates maximum automation.

Objective: The goal of the research was to investigate the use of mobile phone–based telemedicine apps for management of older Chinese patients with type 2 diabetes mellitus (T2DM). Variables of interest included efficacy and safety.

Methods: A total of 91 older (aged over 65 years) patients with T2DM who presented to our department were randomly assigned to one of two groups. Patients in the intervention group (n=44) were provided glucometers capable of data transmission and received advice pertaining to medication, diet, and exercise via the mHealth telemedicine system. Patients assigned to the control group (n=47) received routine outpatient care with no additional intervention. Patients in both groups were followed up at regular 3-month intervals.

Results: After 3 months, patients in the intervention group showed significant (P<.05) improvement in postprandial plasma glucose level. After 6 months, patients in the intervention group exhibited a decreasing trend in postprandial plasma glucose and glycated hemoglobin levels compared with the baseline and those in the control group (P<.05).

Conclusions: Mobile phone–based telemedicine apps help improve glycemic control in older Chinese patients with T2DM.

Trial Registration: China Clinical Trial Registration Center ChiCTR 1800015214; http://www.chictr.org.cn/showprojen.aspx?proj=25949 (Archived by WebCite at http://www.webcitation.org/73wKj1GMq).

(JMIR Mhealth Uhealth 2019;7(1):e10664) doi:10.2196/10664

KEYWORDS

telemedicine; type 2 diabetes; health management

Introduction

Diabetes mellitus is one among the top three chronic, noninfectious diseases in the world [1]. Older people with diabetes constitute a high-risk group, and approximately one in five older patients with type 2 diabetes mellitus (T2DM) develops severe complications [2] such as diabetic neuropathy, nephropathy, retinopathy, or vasculopathy. Thus, these patients are susceptible to renal failure, loss of sight, loss of lower limbs [3], and the risk of severe hyperglycemia and hypoglycemia [4], which impairs their quality of life, imposes financial burden on the patients and community health care systems [2], and decreases life expectancy. Self-management of diabetes includes dietary monitoring, exercise, self-monitoring of blood glucose levels, and adjustments in mental status [5]. Telemedicine management systems allow for remote medical consultations and provision of personalized medical advice, including dietary and lifestyle-related advice from qualified care providers [6,7]. These systems offer an advantage for older patients as they help overcome the distance barrier and the loss of medical treatment opportunities [8,9]. Therefore, research on the applicability of telemedicine systems to older patients is of much clinical relevance.

A previously conducted computer-based telemedicine study [10] involving subjects who have had diabetes for more than 1 year showed that 6 months of telemedicine intervention obviously improved fasting and postprandial blood glucose (PBG) and triglyceride levels in the intervention group. However, as it was a computer-based study, we found in the course of the study that there were many older patients with diabetes who were unfamiliar with computer operations. In that study, a computer was designed to automatically transmit blood glucose meter readings. However, as the computer is not a portable device, transfer of information from the patient to the computer and then to the medical team was not found to be highly realistic [11]. The increasing popularity of mobile phones and user-friendly personal messaging apps has promoted their use in a subset of older patients [12]. We also found that older people in China are better at using mobile phones than computers. This type of special phenomenon is related to the economic development in China. The popularity of computers has occurred relatively late in China, and there are only a few older people who are familiar with computer operations. Nevertheless, with the rapid improvement in living standards in recent years, the popularity of mobile phones has maintained pace with that across the globe, and some older people have skipped the era of computer use. In fact, people are more open to learn to use a mobile phone than a computer, which requires a higher level of proficiency [13]. Considering the known obstacles in mobile phone use, we designed a mobile phone–based mobile health (mHealth) management platform to encourage mobile phone use by older patients; the user interface was designed to provide maximum possible simplicity and automation [14,15].

In this study, we conducted a mobile medical intervention experiment lasting for half a year to determine whether a diabetes mHealth management system based on mobile phones is suitable for older patients. We also evaluated the impact of using this system on glycemic control, treatment adherence, and the rate of occurrence of adverse events (for example, hypoglycemia), as well as overall satisfaction.

Methods

Ethical Considerations

The study protocol was designed in accordance with the Declaration of Helsinki and was approved by the ethics committee of the First Affiliated Hospital of Jilin University. This trial was registered at the China Clinical Trial Registration Center (ChiCTR 1800015214). Written informed consent was obtained from all patients prior to their enrollment.

Participants and Recruitment

Patients who attended the outpatient endocrinology department of the First Affiliated Hospital of Jilin University between March and September 2016 were eligible for inclusion in this randomized controlled trial.

Inclusion criteria in this study were age older than 65 years, glycated hemoglobin (HbA_{1c}) level 7.0% to 10.0%, and the ability to use a mobile phone. Exclusion criteria were illiteracy, abnormal liver and kidney function, severe diabetic complications, use of insulin pumps, and participation in other clinical trials.

A total of 91 patients were enrolled: 44 (19 males) in the intervention group and 47 (18 males) in the control group.

Study Design and Randomization

Patients were randomly assigned to the intervention and control groups using the random number sequence generated by SPSS Statistics version 17.0 (IBM Corp) in batches of 6 patients at a time. Patients in the intervention group were provided training to independently use the mHealth management app and upload the glucometer data, which was then automatically transmitted to the medical server (glucometer was connected to the mobile phone via Bluetooth). The medical teams logged on to the system and sent medical advice and reminders to patients to monitor their glucose levels via the personal messaging app or telephonically every 2 weeks. Patients in the control group received a free glucometer and were followed up through conventional outpatient clinic appointments. For the control group patients, no limitations were imposed to the number of visits; however, they were instructed to monitor and record their blood glucose data regularly.

The study dietitian offered guidance for blood glucose monitoring and provided dietary advice based on the individual blood glucose levels. Patients in the intervention group used the app-based diet management software to input daily dietary intake. The dietitian received the daily dietary record of each patient via the mHealth app. On the basis of the analysis of this information, once-monthly dietary recommendations were sent from the dietitian to patients in the intervention group. The control group received dietary guidance from dietitians during face-to-face meetings at baseline and at the conclusion of all study-related procedures.

Information pertaining to physical activity (daily calorie expenditure) was obtained from patients in the intervention group via text message. The patients were instructed on how to text pedometer data to the study personnel. This information was analyzed, and each patient in the intervention group was provided with guidance related to aerobic and resistance-based exercises. In the control group, guidance related to exercise was provided during face-to-face dietary counseling session during clinic visits.

All patients were followed up in the outpatient clinic at 3-month intervals. Patients in both groups underwent physical examination, blood biochemical tests, follow-up clinic visits, and ambulatory therapy by the same medical team.

Data Collection and Measurements

In order to assess the condition of patients, the following data were reviewed at baseline and every 3 months until the end of the experiment (a total of 3 times): medical history, treatment details, physical examination, and laboratory investigations. Patient compliance was assessed by the frequency of uploading blood glucose data in the intervention group. At the end of the experiment, all patients completed a satisfaction questionnaire, which contained 7 questions, with each question awarded a score of 1 and the highest possible total score being 7 points. Higher total score indicated better satisfaction.

Statistical Analysis

Data were processed using SPSS Statistics version 17.0 (IBM Corp). Normally distributed variables are presented as mean and standard deviation; nonnormally distributed variables are presented as median and interquartile range. Between-group differences to normally distributed variables were assessed using an independent sample *t* test, whereas those to nonnormally distributed variables were assessed using a Mann-Whitney *U* test. For intragroup comparison, normally distributed variables were tested by paired *t* test and nonnormally distributed variables were tested by Wilcoxon rank-sum test. $P<.05$ was considered indicative of a statistically significant difference. Study figures were created using SigmaPlot (Systat Software Inc).

Results

Baseline characteristics of the study population are summarized in Table 1. No significant between-group differences were observed with age, physical findings, or biochemical indices.

HbA_{1c} level reflects the level of glycemic control over the past 3 months. After the first 3 months, we noted a significant improvement in HbA_{1c} levels over the baseline level in both the control group (7.18% [SD 0.85%] vs 7.88% [SD 0.64%], $P<.001$; Table 2 and Figure 1) and the intervention group (6.97% [SD 0.65%] vs 7.84% [SD 0.73%], $P<.001$; Table 2 and Figure 1). Since patients in both groups were given medication, diet, and exercise guidance at the beginning of the trial, a reduction in HbA_{1c} level was observed in both groups at the completion of 3 months, and there was no significant difference between the two groups ($P=.25$; Table 2 and Figure 2). Patients in the intervention group exhibited a decrease in PBG levels relative to baseline; a significant between-group difference was observed in this respect ($P=.04$; Table 2 and Figure 3).

At 6 months, the HbA_{1c} level in the intervention group was significantly lower than that at baseline (6.84% [SD 0.765%] vs 7.84% [SD 0.73%], $P<.001$; Figure 1) and that in the control group at 6 months (6.84% [SD 0.765%] vs 7.22% [SD 0.87%], $P=.02$; Table 2 and Figure 1). The extent of decrease in HbA_{1c} level from baseline level in the intervention group was more than that in the control group (1.07% [SD 0.89%] vs 0.62% [SD 1.00%], $P=.045$; Figure 2). After the 6 months, PBG levels in the intervention group demonstrated continuous improvement as compared with baseline level (10.62 [SD 2.07] mmol/L vs 13.10 [SD 4.13] mmol/L, $P=.002$; Figure 3) and that at 3 months (10.62 [SD 2.07] mmol/L vs 12.09 [SD 3.35] mmol/L, $P=.03$; Table 2 and Figure 3) and were also significantly lower than that in the control group at 6 months (10.62 [SD2.07] mmol/L vs 12.19 [SD 2.54] mmol/L, $P=.004$; Table 2 and Figure 3).

After 6 months, we obtained satisfactory results from the survey of patients in the intervention group. Higher total scores indicated better satisfaction. The average satisfaction score was 6.3 (SD 0.78). Individual questions measured details regarding whether the intervention improved the self-monitoring of patients' blood glucose levels (0.93 [SD 0.14]), diet, exercise and other self-management skills (0.85 [SD 0.20]), and knowledge of diabetes (0.98 [SD 0.08]), as well as the effect on their psychological status (0.96 [SD 0.12]; Multimedia Appendix 1).

Table 1. Baseline characteristics of the two groups.

Characteristic	Control (n=47)	Intervention (n=44)	*P* value
Age in years, median (IQR)[a]	68.04 (66-72)	67.9 (66-71)	.85
Gender, male, n (%)	18 (38)	19 (43)	__[b]
Diabetes mellitus, duration in years, mean (SD)	11.52 (7.73)	11.19 (6.39)	.80
FBG[c] (mmol/L), mean (SD)	7.78 (1.85)	8.0 (2.54)	.41
PBG[d] (mmol/L), mean (SD)	12.44 (3.37)	13.10 (4.13)	.46
HbA_{1c}[e] (%), mean (SD)	7.88 (0.64)	7.84 (0.73)	.53
TC[f] (mmol/L), mean (SD)	4.92 (1.24)	5.00 (0.97)	.76
TG[g] (mmol/L), mean (SD)	2.31 (1.85)	2.41 (1.82)	.80
HDL-C[h] (mmol/L), median (IQR)	1.21 (1.05-1.40)	1.09 (0.85-1.25)	.28
LDL-C[i] (mmol/L), median (IQR)	2.86 (2.28-3.67)	2.92 (2.37-3.29)	.84
BUN[j] (mmol/L), median (IQR)	5.79 (4.76-6.69)	5.62 (5.13-7.05)	.39
Cr[k] (mmol/L), median (IQR)	59.1 (52.58-69.98)	65.05 (54.28-76.58)	.26
AST[l] (U/L), median (IQR)	21.00 (17.50-24.00)	21.30 (17.75-24.25)	.53
ALT[m] (U/L), median (IQR)	20.00 (13.00-32.25)	20.50 (14.70-30.00)	.83
r-GT[n] (U/L), median (IQR)	20.00 (16.00-26.75)	24.5 (19.00-36.00)	.80
Body mass index, median (IQR)	23.30 (21.93-25.88)	23.60 (22.48-26.38)	.63
Blood pressure (mm Hg), systolic, mean (SD)	136.04 (19.37)	132.55 (11.82)	.55
Blood pressure (mm Hg), diastolic, median (IQR)	80.00 (73.50-90.00)	83.00 (74.00-87.75)	.99

[a]IQR: interquartile range.
[b]Indicates a range of values.
[c]FBG: fasting blood glucose.
[d]PBG: postprandial blood glucose.
[e]HbA_{1c}: glycated hemoglobin.
[f]TC: total cholesterol.
[g]TG: triglyceride.
[h]HDL-C: high-density lipoprotein–cholesterol.
[i]LDL-C: low-density lipoprotein–cholesterol.
[j]BUN: blood urea nitrogen.
[k]Cr: creatinine.
[l]AST: aspertate aminotransferase.
[m]ALT: alanine aminotransferase.
[n]r-GT: r-glutamyltransferase.

Table 2. The follow-up data of the two groups.

Characteristics	3 months			6 months		
	Control	Intervention	*P* value	Control	Intervention	*P* value
FBG[a] (mmol/L), mean (SD)	7.57 (2.15)	7.20 (1.70)	.41	7.24 (2.49)	7.26 (2.17)	.96
PBG[b] (mmol/L), mean (SD)	13.15 (3.64)	12.09 (3.35)	.04	12.19 (2.54)	10.62 (2.07)[c]	.004
HbA_{1c}[d] (%), mean (SD)	7.18 (0.85)[e]	6.97 (0.65)[e]	.25	7.22 (0.87)	6.84 (0.76)[e]	.02
TC[f] (mmol/L), mean (SD)	4.84 (1.08)	4.94 (0.80)	.57	4.66 (1.19)	4.63 (0.70)	.88
TG[g] (mmol/L), mean (SD)	1.69 (0.97)	1.66 (0.84)[e]	.86	1.75 (0.86)	1.79 (0.87)	.80
HDL-C[h] (mmol/L), median (IQR[i])	1.34 (1.12-1.51)	1.30 (1.07-1.45)	.39	1.30 (1.15-1.49)	1.2 (1.02-1.35)	.46
LDL-C[j] (mmol/L), median (IQR)	2.99 (2.08-3.52)	2.87 (2.64-3.27)	.56	2.85 (2.03-3.61)	2.88 (2.43-3.14)	.68
BMI[k], median (IQR)	23.25 (22.13-26.23)	23 (22.68-27.43)	.07	22.62 (21.55-24.45)	23.8 (22.5-27.3)	.30
Blood pressure (mm Hg), systolic, mean (SD)	140.61 (14.433)	137.05 (15.07)	.40	130.69 (11.22)	134.48 (9.08)	.22
Blood pressure (mm Hg), diastolic, median (IQR)	80 (69-86.75)	79 (73.75-84.25)	.86	79 (75-84)	80 (78-84)	.78

[a]FBG: fasting blood glucose.

[b]PBG: postprandial blood glucose.

[c]*P*<.01 versus baseline.

[d]HbA_{1c}: glycated hemoglobin.

[e]*P*<.05.

[f]TC: total cholesterol.

[g]TG: triglyceride.

[h]HDL-C: high-density lipoprotein–cholesterol.

[i]IQR: interquartile range.

[j]LDL-C: low-density-lipoprotein–cholesterol.

[k]BMI: body mass index.

Figure 1. The changes in HbA_{1c} levels after follow-up in both groups. After 3 months, HbA_{1c} levels in both groups were significantly improved compared with baseline data (*P*<.01). Six months later, intervention group HbA_{1c} was lower than baseline (*P*<.01), as were the control group HbA_{1c} levels (*P*<.05). "a" indicates *P*<.05 versus baseline and asterisk indicates *P*<.05 versus control group.

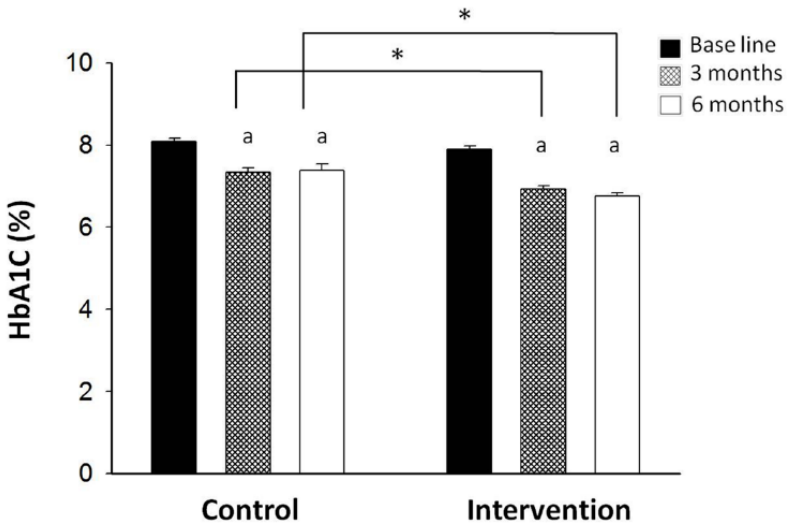


Figure 2. The comparison of the amplitude of change of HbA_{1c} levels in both groups. The mean change in HbA_{1c} levels from baseline to 6 months in the intervention group was significantly higher than that in the control group (P<.05). Asterisk indicates P<.05 versus control group.

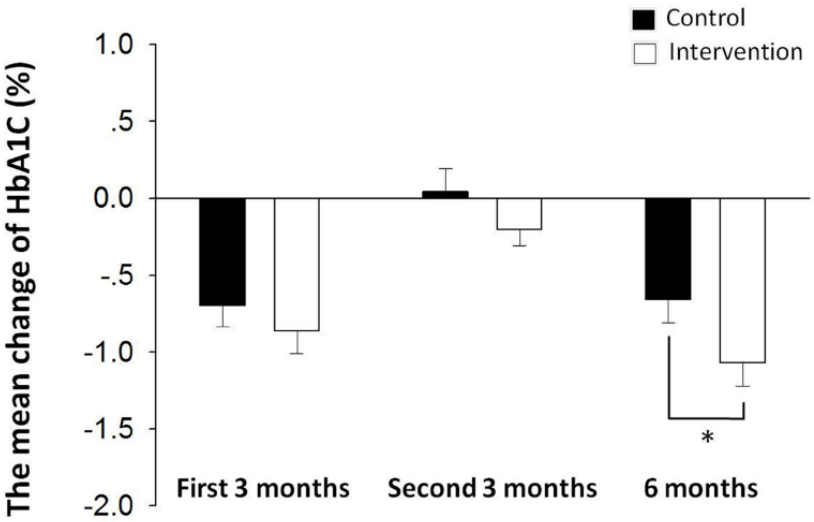
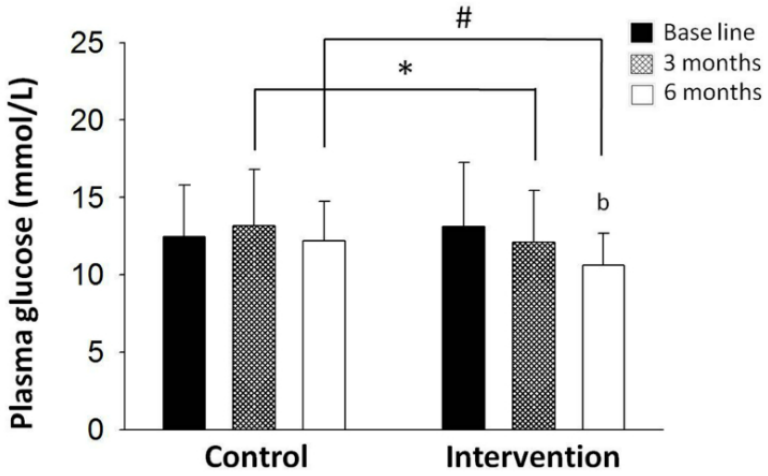


Figure 3. The changes in postprandial blood glucose levels after follow-up in both groups. At the end of 3 and 6 months, the intervention group postprandial blood glucose was significantly lower than the control group postprandial blood glucose (P<.05 and P<.01). "b" indicates P<.01 versus baseline; asterisk indicates P<.05; and # indicates P<.01 versus control group.



Discussion

Principal Findings

We completed a 6-month, prospective, randomized controlled trial of the mHealth telemedicine system in patients with T2DM aged over 65 years. The results showed that the PBG and HbA_{1c} levels in the intervention group were significantly lower than those in the control group; our results are very similar to those reported by Lim et al [16] and Egede et al [17]. At the completion of 3 months, the PBG level in the intervention group was 1.02 mmol/L lower than that at baseline; by the completion of 6 months, the PBG levels showed a progressive decrease of approximately 1.21 mmol/L relative to that at the completion of 3 months. After 6 months, we also observed a significant difference between the intervention and control groups in this respect. At the completion of 3 months, the HbA_{1c} level in the intervention group had decreased by 1% as against 0.66% in the control group. Although the between-group difference was not statistically significant, the HbA_{1c} levels in the intervention group at 6 months showed a further decline of 0.13% as against an increase of 0.04% in the control group. After the sixth month, the intervention group showed continuous improvement in HbA_{1c} levels and the between-group difference was statistically significant; this finding is consistent with the results of Cho et al [18], who also demonstrated the efficacy of telemedicine interventions after a certain period of time. These findings suggest that older patients require time to familiarize themselves with the mHealth system. However, after self-training and remote support from the medical team, the patients started independent use of the portable smart device, which reflected in the positive effects [19]. In this research, both groups showed improved blood glucose and HbA_{1c} levels. This may be attributable to personalized medicine and dietary and exercise plans provided to all subjects [20]. However, without remote supervision and ongoing support, it is difficult to achieve sustained efficacy in the long term; thus, long-term follow-up is essential for older patients with diabetes [21]. During the study, 13 subjects in the control group (7 with hyperglycemia and 6 with hypoglycemia) and 5 subjects in the intervention

group (4 with hyperglycemia and 1 with hypoglycemia) required adjustment of drug dosage for titration of glycemic control; however, none of the patients in either group experienced any serious adverse events or aggravation of complications. The intervention group had significantly fewer hypoglycemic events compared with the control group. This was likely attributable to prompt identification of the risk of hypoglycemia in the intervention group by the medical team via the mHealth platform and the consequent implementation of timely corrective actions [10,22,23]. After the trial, over 89% of patients in the intervention group continued to measure their blood glucose level 2 to 3 days each week. Intervention group satisfaction survey responses indicated that frequent communication with the medical team via the mHealth platform enhanced patient understanding of diabetes, increased their awareness, and helped alleviate depressive symptoms [8,24].

Comparison With Prior Work

Previous studies have shown that telemedicine interventions can improve blood glucose control in patients with diabetes [25,26]; however, the applicability of telemedicine for older patients has rarely been discussed.

The study by Quinn et al [24] and Kim et al [27] showed that middle-aged and older patients with diabetes have good interaction in a mobile phone–based diabetes education environment and that it significantly improves the self-management of blood glucose levels. However, the study was conducted over a period of 1 month, which is too short to determine the compliance of older patients over the long term with remote intervention. However, Egede et al [17] conducted a 12-month-long study involving remote psychotherapy intervention for older diabetic patients. Older diabetic patients not only maintained good compliance but also achieved long-term glycemic control. However, the intervention involved only psychotherapy, and there was no routine medication-, diet- and exercise-related intervention. Therefore, the study was not designed to determine the advantages of remote intervention with regular outpatient treatment. Cho et al [18] performed a 6-month comparative study of telemedicine and traditional outpatient treatment; although the benefits of telemedicine were not found at 3 months, HbA_{1c} levels were significantly improved at 6 months and the benefit was mainly found among women aged over 40 years. Williams et al [28] conducted a 12-month long-distance interventional clinical trial among African Americans aged over 21 years; the results suggested that long-term remote interventions can improve long-term glycemic control. However, all the above studies involved remote interventions in a wide range of age groups. Since the cognitive ability of older patients is relatively low, the operation interface used by middle-aged patients cannot be expected to be equally effective in older patients. Therefore, we greatly simplified the user interface of our telemedicine system to make it suitable for use by older patients. This enhanced the confidence of patients and their ability to follow the advice and provided us with valuable data that can be analyzed.

Limitations

In general, telemedicine facilitates good glycemic control in older diabetic patients. In this study, the personal and family medical history, smoking history, history of alcohol intake, birth history, history of drug allergy, and personal living environment were not included in the analysis [6]. However, these factors can potentially affect the nutritional status and function of major organs; in addition, this information is important for the assessment of the quality of life of patients [29,30]. Moreover, data collected from dietary caloric intake and expenditure are not as accurate as blood glucose data; therefore, the effect of dietary and exercise-related guidance on glycemic control was not reliably measured; it is necessary to develop an accurate data collection method for calorie intake and consumption [31]. When this was achieved, telemedicine assisted the medical team and allowed the team to provide timely warnings of the risk of hypoglycemia or hyperglycemia as well as encouraged patients to continue their diet and exercise plan.

Conclusions

In this study, PBG level in the intervention group was significantly lower than that in the control group after the first 3 months. The improvement in glycemic control was sustained after 6 months and showed a significant difference from that in the control group. Our results suggest that the improved glycemic control in the intervention group was attributable to improved communication between doctors and patients with real-time tracking of older diabetic patients by the mHealth system and improved patient compliance after implementation of mHealth monitoring. On the basis of our findings, we can conclude that telemedicine is effective and safe for older diabetic patients.

Acknowledgments

This study was supported by the Science Technology Department of Jilin Province (20180623006TC) and the Interdisciplinary Project of First Hospital of Jilin University (JDYYJC010).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Satisfaction survey.

[PDF File (Adobe PDF File), 48KB - mhealth_v7i1e10664_app1.pdf]

Multimedia Appendix 2

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - mhealth_v7i1e10664_app2.pdf]

References

1. Xu Y, Wang L, He J, Bi Y, Li M, Wang T, 2010 China Noncommunicable Disease Surveillance Group. Prevalence and control of diabetes in Chinese adults. JAMA 2013 Sep 04;310(9):948-959. [doi: 10.1001/jama.2013.168118] [Medline: 24002281]
2. Remler DK, Teresi JA, Weinstock RS, Ramirez M, Eimicke JP, Silver S, et al. Health care utilization and self-care behaviors of Medicare beneficiaries with diabetes: comparison of national and ethnically diverse underserved populations. Popul Health Manag 2011 Feb;14(1):11-20 [FREE Full text] [doi: 10.1089/pop.2010.0003] [Medline: 21241171]
3. Pratley RE, Heller SR, Miller MA. Treatment of type 2 diabetes mellitus in the older adult: a review. Endocr Pract 2014 Jul;20(7):722-736. [doi: 10.4158/EP13192.RA] [Medline: 24518176]
4. Kirkman MS, Briscoe VJ, Clark N, Florez H, Haas LB, Halter JB, Consensus Development Conference on Diabetes and Older Adults. Diabetes in older adults: a consensus report. J Am Geriatr Soc 2012 Dec;60(12):2342-2356 [FREE Full text] [doi: 10.1111/jgs.12035] [Medline: 23106132]
5. Inzucchi SE, Bergenstal RM, Buse JB, Diamant M, Ferrannini E, Nauck M, American Diabetes Association (ADA), European Association for the Study of Diabetes (EASD). Management of hyperglycemia in type 2 diabetes: a patient-centered approach: position statement of the American Diabetes Association (ADA) and the European Association for the Study of Diabetes (EASD). Diabetes Care 2012 Jun;35(6):1364-1379 [FREE Full text] [doi: 10.2337/dc12-0413] [Medline: 22517736]
6. Picton P, Wiljer D, Urowitz S, Cafazzo JA. Engaging patients in online self-care technologies for chronic disease management. Healthc Q 2016;18(4):55-61. [Medline: 27009709]
7. Fottrell E, Jennings H, Kuddus A, Ahmed N, Morrison J, Akter K, et al. The effect of community groups and mobile phone messages on the prevention and control of diabetes in rural Bangladesh: study protocol for a three-arm cluster randomised controlled trial. Trials 2016 Dec 19;17(1):600 [FREE Full text] [doi: 10.1186/s13063-016-1738-x] [Medline: 27993166]
8. Walker CL, Kopp M, Binford RM, Bowers CJ. Home telehealth interventions for older adults with diabetes. Home Healthc Now 2017 Apr;35(4):202-210. [doi: 10.1097/NHH.0000000000000522] [Medline: 28353510]
9. Isaković M, Sedlar U, Volk M, Bešter J. Usability pitfalls of diabetes mHealth apps for the elderly. J Diabetes Res 2016;2016:1604609 [FREE Full text] [doi: 10.1155/2016/1604609] [Medline: 27034957]
10. Kim H, Sun C, Yang S, Sun L, Li F, Choi I, et al. Randomized, open-label, parallel group study to evaluate the effect of internet-based glucose management system on subjects with diabetes in China. Telemed J E Health 2016 Dec;22(8):666-674. [doi: 10.1089/tmj.2015.0170] [Medline: 26938489]
11. Matthew-Maich N, Harris L, Ploeg J, Markle-Reid M, Valaitis R, Ibrahim S, et al. Designing, implementing, and evaluating mobile health technologies for managing chronic conditions in older adults: a scoping review. JMIR mHealth uHealth 2016 Jun 09;4(2):e29. [doi: 10.2196/mhealth.5127] [Medline: 27282195]
12. Kim BY, Lee J. Smart devices for older adults managing chronic disease: a scoping review. JMIR Mhealth Uhealth 2017 May 23;5(5):e69. [doi: 10.2196/mhealth.7141] [Medline: 28536089]
13. Holmen H, Torbjørnsen A, Wahl AK, Jenum AK, Småstuen MC, Arsand E, et al. A mobile health intervention for self-management and lifestyle change for persons with type 2 diabetes, part 2: one-year results From the Norwegian randomized controlled Trial RENEWING HEALTH. JMIR Mhealth Uhealth 2014 Dec;2(4):e57 [FREE Full text] [doi: 10.2196/mhealth.3882] [Medline: 25499872]
14. Arnhold M, Quade M, Kirch W. Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. J Med Internet Res 2014 Apr;16(4):e104 [FREE Full text] [doi: 10.2196/jmir.2968] [Medline: 24718852]
15. van den Berg N, Schumann M, Kraft K, Hoffmann W. Telemedicine and telecare for older patients—a systematic review. Maturitas 2012 Oct;73(2):94-114. [doi: 10.1016/j.maturitas.2012.06.010] [Medline: 22809497]
16. Lim S, Kang SM, Kim KM, Moon JH, Choi SH, Hwang H, et al. Multifactorial intervention in diabetes care using real-time monitoring and tailored feedback in type 2 diabetes. Acta Diabetol 2016 Apr;53(2):189-198. [doi: 10.1007/s00592-015-0754-8] [Medline: 25936739]
17. Egede LE, Walker RJ, Payne EH, Knapp RG, Acierno R, Frueh BC. Effect of psychotherapy for depression via home telehealth on glycemic control in adults with type 2 diabetes: subgroup analysis of a randomized clinical trial. J Telemed Telecare 2017 Jan 01:1357633X17730419. [doi: 10.1177/1357633X17730419] [Medline: 28945160]
18. Cho JH, Kim H, Yoo SH, Jung CH, Lee WJ, Park CY, et al. An Internet-based health gateway device for interactive communication and automatic data uploading: clinical efficacy for type 2 diabetes in a multi-centre trial. J Telemed Telecare 2017 Jul;23(6):595-604. [doi: 10.1177/1357633X16657500] [Medline: 27381040]
19. Archer N, Keshavjee K, Demers C, Lee R. Online self-management interventions for chronically ill patients: cognitive impairment and technology issues. Int J Med Inform 2014 Apr;83(4):264-272. [doi: 10.1016/j.ijmedinf.2014.01.005] [Medline: 24507762]

20. Zhou P, Xu L, Liu X, Huang J, Xu W, Chen W. Web-based telemedicine for management of type 2 diabetes through glucose uploads: a randomized controlled trial. Int J Clin Exp Pathol 2014;7(12):8848-8854 [FREE Full text] [Medline: 25674254]
21. Lee T, Yeh Y, Liu C, Chen P. Development and evaluation of a patient-oriented education system for diabetes management. Int J Med Inform 2007 Sep;76(9):655-663. [doi: 10.1016/j.ijmedinf.2006.05.030] [Medline: 16815741]
22. Hu Y, Wen X, Wang F, Yang D, Liu S, Li P, et al. Effect of telemedicine intervention on hypoglycaemia in diabetes patients: a systematic review and meta-analysis of randomised controlled trials. J Telemed Telecare 2018 Jan 01:1357633X18776823. [doi: 10.1177/1357633X18776823] [Medline: 29909748]
23. Jeong JY, Jeon J, Bae K, Choi Y, Park K, Kim J, et al. Smart care based on telemonitoring and telemedicine for type 2 diabetes care: multi-center randomized controlled trial. Telemed J E Health 2018 Aug;24(8):604-613. [doi: 10.1089/tmj.2017.0203] [Medline: 29341843]
24. Quinn CC, Khokhar B, Weed K, Barr E, Gruber-Baldini AL. Older adult self-efficacy study of mobile phone diabetes management. Diabetes Technol Ther 2015 Jul;17(7):455-461 [FREE Full text] [doi: 10.1089/dia.2014.0341] [Medline: 25692373]
25. Weinstock RS, Teresi JA, Goland R, Izquierdo R, Palmas W, Eimicke JP, IDEATel Consortium. Glycemic control and health disparities in older ethnically diverse underserved adults with diabetes: five-year results from the Informatics for Diabetes Education and Telemedicine (IDEATel) study. Diabetes Care 2011 Feb;34(2):274-279 [FREE Full text] [doi: 10.2337/dc10-1346] [Medline: 21270184]
26. Quinn CC, Shardell MD, Terrin ML, Barr EA, Ballew SH, Gruber-Baldini AL. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. Diabetes Care 2011 Sep;34(9):1934-1942 [FREE Full text] [doi: 10.2337/dc11-0366] [Medline: 21788632]
27. Kim HS, Yang SJ, Jeong YJ, Kim YE, Hong SW, Cho JH. Satisfaction survey on information technology-based glucose monitoring system targeting diabetes mellitus in private local clinics in Korea. Diabetes Metab J 2017 Jun;41(3):213-222 [FREE Full text] [doi: 10.4093/dmj.2017.41.3.213] [Medline: 28657235]
28. Williams JS, Lynch CP, Knapp RG, Egede LE. Technology-Intensified Diabetes Education Study (TIDES) in African Americans with type 2 diabetes: study protocol for a randomized controlled trial. Trials 2014 Nov 25;15:460 [FREE Full text] [doi: 10.1186/1745-6215-15-460] [Medline: 25425504]
29. Wakabayashi I. Smoking and lipid-related indices in patients with diabetes mellitus. Diabet Med 2014 Jul;31(7):868-878. [doi: 10.1111/dme.12430] [Medline: 24606550]
30. Wu Y, Ding Y, Tanaka Y, Zhang W. Risk factors contributing to type 2 diabetes and recent advances in the treatment and prevention. Int J Med Sci 2014;11(11):1185-1200 [FREE Full text] [doi: 10.7150/ijms.10001] [Medline: 25249787]
31. Kearns JW, Bowerman D, Kemmis K, Izquierdo RE, Wade M, Weinstock RS. Group diabetes education administered through telemedicine: tools used and lessons learned. Telemed J E Health 2012 Jun;18(5):347-353. [doi: 10.1089/tmj.2011.0165] [Medline: 22468984]

Abbreviations

HbA_{1c}: glycated hemoglobin
PBG: postprandial blood glucose
T2DM: type 2 diabetes mellitus

Edited by G Eysenbach; submitted 01.04.18; peer-reviewed by K Schäfer, P Rasche; comments to author 06.08.18; revised version received 19.09.18; accepted 07.10.18; published 04.01.19.

Please cite as:
Sun C, Sun L, Xi S, Zhang H, Wang H, Feng Y, Deng Y, Wang H, Xiao X, Wang G, Gao Y, Wang G
Mobile Phone–Based Telemedicine Practice in Older Chinese Patients with Type 2 Diabetes Mellitus: Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(1):e10664
URL: https://mhealth.jmir.org/2019/1/e10664/
doi:10.2196/10664
PMID:30609983

Original Paper

Women's Attitudes Toward Self-Monitoring of Their Pregnancy Using Noninvasive Electronic Devices: Cross-Sectional Multicenter Study

Katharina Schramm[1*]; Niklas Grassl[1*], MSc; Juliane Nees[1], MD; Janine Hoffmann[2], MD; Holger Stepan[2], MD, PhD, Prof Dr; Thomas Bruckner[3], PhD; Markus W Haun[4], MD; Imad Maatouk[4], MD; Markus Haist[5], MD; Timm C Schott[6], DMD, PhD, Prof Dr; Christof Sohn[1], MD, PhD, Prof Dr; Sarah Schott[1], MD, PhD, Prof Dr

[1]Department of Gynecology and Obstetrics, University Women's Clinic Heidelberg, Heidelberg, Germany

[2]Department of Gynecology and Obstetrics, University Women's Clinic Leipzig, Leipzig, Germany

[3]Institute of Medical Biometry and Informatics, University of Heidelberg, Heidelberg, Germany

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

[5]Frauenarztpraxis Markus Haist & Anja Ritthaler, Pforzheim, Germany

[6]Centre of Dentistry, Department of Orthodontics and Orofacial Orthopedics, University of Tuebingen, Tuebingen, Germany

[*]these authors contributed equally

Corresponding Author:
Sarah Schott, MD, PhD, Prof Dr
Department of Gynecology and Obstetrics
University Women's Clinic Heidelberg
Im Neuenheimer Feld 440
Heidelberg, 69120
Germany
Phone: 49 6221567906
Email: sarah.schott@med.uni-heidelberg.de

Abstract

Background: Pregnancy can be distressing, particularly if expectant mothers are worried about the well-being of their fetus. Consequently, the desire for reassurance and frequent fetal monitoring is often pronounced. Smart wearable devices and telemedicine are promising tools that could assist women in self-monitoring their pregnancy at home, hence disburdening emergency departments (EDs). They present the possibility to clarify the need for urgent care remotely and offer tighter pregnancy monitoring. However, patients' acceptance of such new technologies for fetal monitoring has not yet been explored extensively.

Objective: This survey aimed to elucidate the attitudes of women toward self-monitoring of their pregnancy using noninvasive electronic devices. The technical details of the proposed devices were not specified.

Methods: A cross-sectional multicenter study was conducted at the departments of obstetrics of the University Hospitals of Heidelberg and Leipzig, Germany. All patients seen in the obstetrics clinic who were above 18 years were offered participation. We designed a survey questionnaire including validated instruments covering population characteristics, issues in current and past pregnancies, as well as attitudes toward self-monitoring of pregnancy with smart devices.

Results: A total of 509 pregnant women with no previous experience in telemedicine participated. Only a small minority of 5.9% (29/493) regarded self-monitoring with wearable devices as an alternative to consulting their physicians. Along these lines, only 7.7% (38/496) strongly believed they would visit the ED less often if such devices were readily available. However, if the procedure were combined with a Web-based telemetric physician consult, 13.5% (66/487) would be highly motivated to use the devices. Furthermore, significantly more women regarded it as an alternative prior to seeing a doctor when they perceived a decline in fetal movements (P<.001). Interestingly, women with university degrees had a higher propensity to engage in pregnancy self-monitoring compared with women without one (37% vs 23%; P=.001). Of the participants, 77.9% (381/489) would like smart wearable devices to measure fetal heart sounds, and 62.6% (306/489) wished to use the devices on their own. Feedback from a doctor or midwife was also very important in their choice of such devices (61.8%, 301/487 wished feedback). The intended frequency of use differed vastly among women, ranging from 13.8% (65/471) who would like to use such a device several times per day to 31.6% (149/471) who favored once per week at most.

Conclusions: Our results point to a skeptical attitude toward pregnancy self-monitoring among pregnant women. Nevertheless, many women are open to using devices for pregnancy monitoring in parallel to consulting their physician. The intention to use such devices several times daily or weekly, expressed by more than half of the participants, highlights the potential of such technologies.

(JMIR Mhealth Uhealth 2019;7(1):e11458) doi:10.2196/11458

KEYWORDS

eHealth; fetal monitoring; pregnancy; telemedicine

Introduction

Health surveillance apps and devices are becoming more and more popular [1-5]. Many companies offer personalized health trackers and market them as lifestyle products suitable for everyday application. Approximately 80% of women in their fertile period in the United States own a smartphone [6], and around 23% of female smartphone owners use mobile health apps [7]. Consequently, nutrition and menstrual cycle monitoring, as well as pregnancy diaries, have become quite common [8]. However, no medically reliable, cost-effective wearable device is on the market for pregnancy self-monitoring. The use of such devices by pregnant women at home could enable closer observation, provide reassurance to concerned expectant mothers, and better identify high-risk patients. Such devices would constitute a milestone for obstetric care [9]. Since electronic health (eHealth) devices bring new challenges in terms of reliability and medicolegal responsibility, the attitude and comfort of both patients and medical staff are highly relevant [10]. More generally, there is still a surprising lack of data surrounding eHealth and mobile health devices despite its widely recognized potential [2,11,12].

Giving pregnant women the possibility to monitor their fetus at home does not only offer advantages for themselves but also for their physicians and midwives. In obstetrics as well as other medical disciplines, urgent and nonurgent emergency visits are increasing [13]. Concerns and worries about the unborn child are common reasons for emergency consultations in obstetrics. In these cases, it is often challenging for patients to judge whether a given condition is pathological or physiological, resulting in avoidable consultations. Several studies indicate that pregnant women are frequently dismissed upon entering the emergency room since no pathological finding was detected despite a high sense of urgency felt by the patient [14,15]. Additional data from prolonged home monitoring could potentially add diagnostic value in such situations. It is also conceivable that in the future, the additional data might enable medical staff to reassure patients and their partners remotely. For this scenario to become a reality, evidence of the advantages of home monitoring would be required, and legal responsibilities would need to be clarified first.

Prolonged or even continuous monitoring of pregnancy with smart wearable devices that assess several fetal parameters could provide a very comprehensive picture of fetal well-being based on an extensive collection of data. Such recordings might, therefore, be of great value for observation and diagnostics. However, women's willingness to use remote devices at home to monitor pregnancy has hardly been explored. We therefore designed a survey among expectant mothers to elucidate women's attitudes toward pregnancy monitoring with smart devices.

Methods

Survey Design and Questionnaire

Patients from 4 doctors were included in this prospective cross-sectional multicenter survey at the obstetric emergency departments (EDs) of the University Hospital Women's Clinics in Heidelberg and Leipzig, Germany or cooperating obstetric outpatient clinics. German-speaking women, aged between 18 and 55 years, were eligible to contribute if they showed the capacity for consent and provided informed consent. Approximately 800 eligible patients were offered participation. This accounts for approximately 10% of the obstetrics patients in the participating centers in the inclusion period, and this subset can be considered a random sample. The questionnaire was completed by 509 women, resulting in a participation rate of 63.9% (509/796).

Patients completed the questionnaires pseudonymously on paper while in the clinic. Participants needed approximately 15 minutes for completing the survey and did not receive any compensation. Exclusion criteria consisted of an inadequate understanding of the German language or refusing to participate. The survey was approved by the ethics committees of both medical faculties (Heidelberg S-525/2016, Leipzig 092/17-lk).

The questionnaire consisted of 21 closed-ended questions to allow for quantitative statistical analysis. In addition, we also asked patients to provide their year of birth and their estimated due date. The questions were conceived by 2 experienced obstetricians and a clinical psychologist following a literature review and an interdisciplinary discussion with diverse members of the labor and delivery staff of the University Hospital in Heidelberg. The questionnaire was subsequently tested on 10 volunteers of our target population that were not part of the final study.

The introductory questions captured characteristics of our study population such as their highest education degree, marital status, employment type, and health care plan. These questions were followed by an assessment of their current and previous pregnancies including the number of previous pregnancies, deliveries, previous modes of delivery, planned mode of delivery in this pregnancy, complications in this pregnancy, whether conception was natural, and whether the current pregnancy was a multiple pregnancy. The remaining questions focused on eHealth devices for pregnancy monitoring. The description of the devices in question did not cover technical details but

specified that they were noninvasive and that women would be able to put them on autonomously. The first block of these questions inquired under which circumstances and how often the respondent was willing to engage in pregnancy monitoring at home, with possible responses like "if I felt my baby less..." The second block focused on the expectations toward devices to monitor a pregnancy at home such as display of recordings, design, and functionality. The original questions are provided in Multimedia Appendix 1. Further results regarding emergency visits are published elsewhere (Schramm et al, under review).

Statistical Analysis

The statistical analyses were performed in Microsoft Excel version 15.31 and SAS 9.1 Documentation. The relative frequencies of the replies of close-ended questions were calculated and stated as the percentage of the total numbers of replies. For questions that asked participants to rank their agreement with certain statements on a Likert-scale ranging from 1 to 5, the percentages of replies and the weighted mean were computed (Tables 1 and 2). Inferential statistics comprised chi-square tests for categorical data. For all analyses, statistical significance was at a type 1 error of 5% (2-tailed).

Results

Study Population

Sample characteristics are shown in Table 1 and Multimedia Appendix 2. The questionnaire was completed by 509 women, resulting in a participation rate of 63.9% (509/796) for all women that were offered participation. The completeness rate of these data was 96.3% (490/509). The sample is by and large representative of pregnant women in Germany [16], with an overrepresentation of high-risk pregnancies due to the care structure of the participating centers.

Perception of Telemedicine as Alternative to Physician Consult

In the first step, the questionnaire explored if devices for self-monitoring could reduce physician consultations. When asked whether they regarded self-monitoring of their pregnancy as an alternative to consulting a physician, only 5.9% (29/493) of our study participants strongly agreed, 36.7% (181/493) of participants strongly disagreed, and the remaining 57.4% (283/493) of women favored an intermediate to skeptical standpoint toward this statement (Table 2). However, significantly more women regarded it as an alternative prior to seeing a doctor when they felt fewer baby movements (P<.001). Still, only 7.7% (38/496) strongly believed they would visit the ED less often during pregnancy if such devices were in place.

The picture was slightly different if the assumed use of mobile devices to monitor pregnancy were combined with a telemedical consultation with a physician. In this scenario, 13.5% (66/487) of participants strongly believed they would use such devices, and only 12.5% (61/487) categorically rejected that notion (Multimedia Appendix 3). A small minority of patients felt insecure using such technologies if Web-based contact with a physician was established. However, we registered a strong agreement of 41.8% (192/471) to the statement that cardiotocography (CTG) provides more certainty than self-monitoring.

Expected Properties of Mobile Devices for Pregnancy Monitoring

Preferences for the readout of mobile devices for pregnancy monitoring varied. While 29.8% (134/450) of the participants preferred a simple binary reading stating either that everything is normal or that consulting a physician is recommended, a majority of 39.6% (178/450) were in favor of more detailed information allowing graduating fetal well-being and providing information on fetal status. The remaining 30.7% (138/450) even wanted such devices to display as much information as possible.

The expectations of different features for pregnancy monitoring are summarized in Figure 1. Patients were asked to indicate which features of mobile devices they regarded as particularly important. Results are displayed as the percentage of total replies. Multiple answers were possible.

The most important feature recommended by our study participants was to enable mothers to listen to their baby's heartbeat (78.2%, 381/487). An independent application, an endorsement by physicians or midwives, and feedback about proper utilization was also important to potential users (62.8%, 306/487; 50.5%, 246/487; and 61.8%, 301/487, respectively). Properties of comparatively less importance to the users included wearing comfort and secure positioning (34.5%, 168/487 and 35.7%, 174/487, respectively). Around 1-third of participants would have liked the devices to allow measurements in different body positions and during movement. The least important features were the possibility to mute the fetal heart sound and the optical design.

Finally, we also addressed how frequently a mobile device for pregnancy self-monitoring would be used. Interestingly, 13.8% (65/471) of the participants indicated that they would perform pregnancy monitoring at home several times per day and 4.8% (23/471) even would use it "all the time if possible." A frequency of once per day was preferred by 22.1% (104/471). The number of participants who were inclined to use it 1-4 times a week was 27.6% (130/471), and 31.6% (149/471) would use it less than once per week.

The attitudes toward pregnancy monitoring with eHealth devices differed substantially depending on socioeconomic status. Among patients with a university degree, 37.1% (77/207) would have consulted their obstetrician less often if they had the chance to monitor their fetus at home compared with 23.1% (66/286) without a university degree (P=.001). At the same time, 75.8% (157/207) of academics preferred a more detailed readout over a binary readout of monitoring devices compared with 66.3% (161/243) of nonacademics (P=.03). Both the attitude toward pregnancy monitoring with wearable devices and preferences for features of mobile devices did not depend on age, the number of previous pregnancies, marital status, or health care plan.

Table 1. Sample characteristics of the study population.

Characteristic	Responses, n (%)
Age (years)	
18-20	12 (2.4)
21-25	47 (9.3)
26-30	168 (33.2)
31-35	178 (35.2)
36-40	84 (16.6)
>41	17 (3.4)
Highest education degree	
Dropped out of school	11 (2.2)
Secondary education ending with ninth grade	41 (8.1)
Secondary education ending with tenth grade	147 (29.0)
University entrance diploma	101 (19.9)
University degree	207 (40.8)
Marital status	
Married, living with spouse	470 (92.7)
Married, living separated from spouse	9 (1.8)
Single, without children	16 (3.2)
Single, with children	11 (2.2)
Widowed	1 (0.2)
Employment	
Full-time (>35 hours per week)	143 (28.3)
Part-time (15-34 hours per week)	68 (13.4)
By the hour (1-14 hours per week)	1 (0.2)
Educational training (student)	20 (4.0)
Housewife	41 (8.1)
Unemployed	14 (2.8)
Leave of absence (ie, maternity leave)	219 (43.3)
Health care plan	
Public health care	391 (89.9)
Private health care	44 (10.1)
Public family health care	43 (9.9)
Supplementary insurance	29 (6.7)

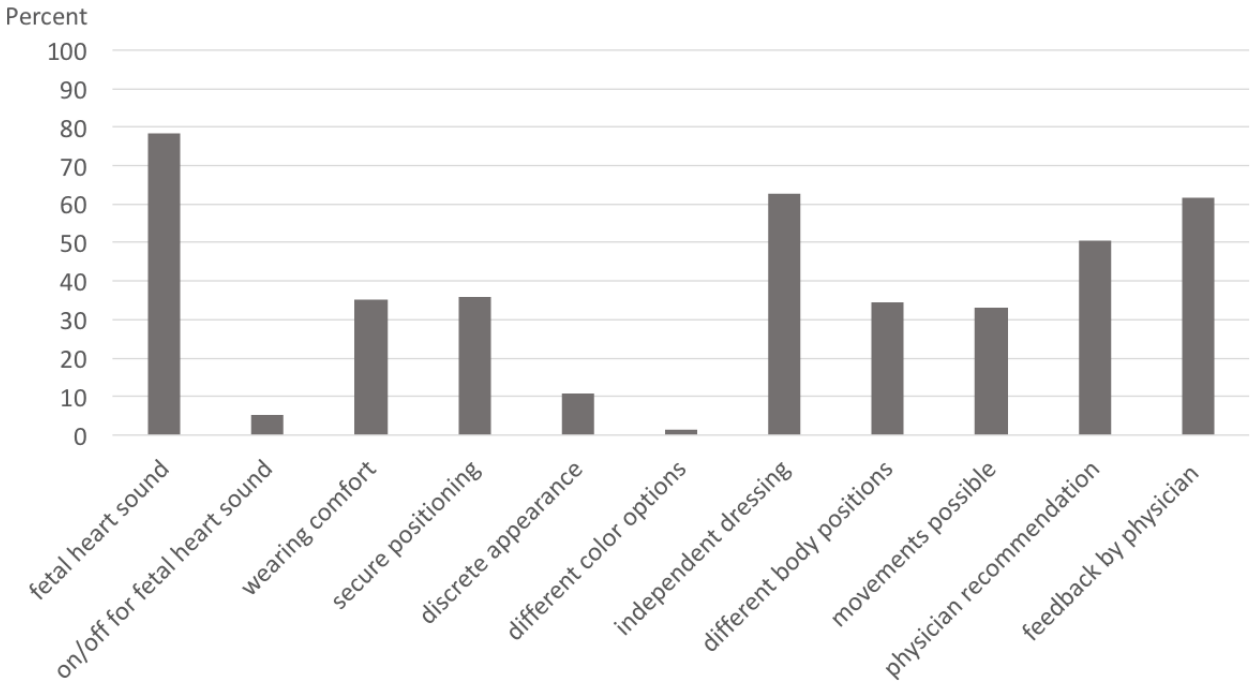
Table 2. Pregnancy monitoring at home as an alternative to direct physician consultation.

If I had the possibility to monitor your baby at home…	Score=1[a]	Score=2	Score=3	Score=4	Score=5	Weighted mean[b]
...I would regard this as an alternative to consulting a physician.	181 (36.7)	102 (20.7)	118 (23.9)	63 (12.8)	29 (5.9)	2.3
...I would regard this as an alternative prior to visiting a physician if I felt my baby less.	132 (26.8)	82 (16.7)	102 (20.7)	112 (22.8)	64 (13.0)	2.8
...I cannot imagine this and would always prefer a direct consult with a physician or midwife.	48 (9.8)	60 (12.2)	115 (23.4)	126 (25.7)	141 (28.8)	3.5
...I would visit the emergency department less often.	120 (24.1)	110 (22.2)	117 (23.6)	111 (22.4)	38 (7.7)	2.7

[a]Participants were asked to indicate their agreement to the following statements on a scale from 1 to 5 (1 signifies strong disagreement and 5 strong agreement); absolute numbers are shown, and percentages or replies are indicated in brackets.

[b]Weighted means are shown for each statement.

Figure 1. Patient preferences for features of mobile devices for pregnancy monitoring.



The number of pregnancies did not have any impact on the willingness to use eHealth devices prior to seeing a physician (44% in first pregnancy vs 42% in women with previous pregnancies, *P*=.77). Also, we did not detect any significant differences between the attitude toward the use of eHealth devices between patients that had several emergency visits during their pregnancy compared with patients that had no or just 1 emergency visit (47% vs 52%, *P*=.71).

Discussion

Principal Findings

Our comprehensive study provides a detailed account of patients' attitude toward pregnancy monitoring with eHealth devices when faced with several scenarios. A large proportion of women are open to the idea of monitoring their pregnancy using eHealth devices but prefer to use them in addition to rather than instead of consultations with their physician. The data indicate that patients that have never used eHealth devices can imagine consulting a physician using these new technologies. However, the skepticism toward these eHealth devices is greatest in scenarios where these tools are employed to replace a doctor's consult. This is in line with results from a study on telemedicine in postoperative care that found patients to be afraid of losing their personal relationship with their doctor when engaging in telemedicine [17]. If a scenario is offered to monitor with a device combined with a Web-based consultation with a physician, significantly more patients would feel comfortable using it. Many patients might also underestimate the extent to which telemedicine has already become part of professional medical practice [18].

Also, 36.7% (181/493) of our cohort could imagine using eHealth devices prior to visiting their doctor if they felt fewer baby movements. It seems that a physician's or midwife's judgment is regarded as a lot more trustworthy than the reading of a monitoring device. However, it has been well established that face-to-face interactions are not superior to telemedical

interactions on professional practice and health care outcomes in several medical specialties [19-21].

High trust in the opinion of a health care professional is also reflected in the features of pregnancy monitoring devices that are most important to patients. Recommendation by a physician or midwife and their feedback about the proper use of the device are found to be very important. Only the ability to record the fetal heart rate and to apply the device independently was discovered to be more essential to patients. All in all, the functionality of such devices is the key to patients compared with design aspects like different color options or discrete appearance. More women with university degrees prefer a detailed over a binary readout of such devices compared with nonacademics. In general, this seems to be the population that feels most inclined to engage in pregnancy monitoring at home. A possible explanation might be that women with a university degree feel more comfortable about autonomous fetal monitoring or expect greater benefits from the use of new technologies. Of note, the proportion of participants with university degrees in our study was 40.8% compared with 14.8% in the general population [22].

Limitations

The study population was by and large representative of pregnant women in Germany, with a slight overrepresentation of academics and high-risk pregnancies due to the fact that participants were mainly seen at university hospitals with a maximal level perinatal care. Given that high-risk pregnancies imply more extensive monitoring and more frequent antenatal consultations, this patient population is likely to benefit the most from telemedical pregnancy monitoring [23].

Furthermore, this study was carried out by only a few doctors in service. Therefore, not all possible women meeting the inclusion criteria were reached but rather a random cross-section thereof. Hence, we expect selection bias to be limited, especially as this was a bicentric study.

The study is an account of attitudes toward telemedicine in obstetrics at a time when telemedicine is barely used. In interpreting the findings of this study, one has to bear in mind that all study participants had no prior experience with remote pregnancy monitoring. Thus, the application of eHealth devices and telemedicine was left to women´s imagination. Consequently, our study focuses on intent rather than actual behavior. Previous studies suggest that knowledge of telemedicine in the general population is limited, and people who are not familiar with it tend to reject it [24,25]. We assume that the experience and more widespread use of such devices will have a profound impact on those attitudes.

Outlook for Pregnancy Self-monitoring

The fact that almost 20% of women would be willing to wear devices for pregnancy monitoring all the time or several times daily highlights the desire of some women to closely monitor their pregnancy. In fact, monitoring frequencies of once per week or more—preferred by 2-thirds of our participants—seem only practical using self-monitoring devices. Hence, these responses highlight pregnant women's desire to get frequent updates on fetal well-being and their self-evolvement. The desired monitoring frequencies require the new devices to be handy and noninvasive. So, far, the CTG does not open options for that kind of monitoring, as it requires a health care professional to set it up correctly. In addition, it is considered an invasive procedure that exposes the fetus to ultrasound. Whether permanent fetal monitoring via CTG causes fetal harm has not been investigated yet, but experts in ultrasound medicine recommend following the as low as reasonably achievable principle for the use of ultrasound in obstetrics [26,27]. Finally, the reading of CTGs has a high intra- and interrater variability [28]. These results fueled the development of computer-based CTG analysis [29] to increase the reliability of CTG interpretation in the future.

Consequently, pregnancy monitoring devices could gain great popularity, though as a supplement rather than as a replacement for pregnancy monitoring by physicians. Nonetheless, it should also be remembered that many women are opposed to extensive monitoring of their pregnancy and the medicalization of the female body [30]. Other disciplines have proven that the acceptance increases with knowledge and a more widespread use [31-35]. Whether such devices will be able to provide reliable diagnoses in the future that are equally as trustworthy and reassuring as the judgment of a physician remains to be elucidated. Prospective studies are needed to address feasibility, safety issues, and effectiveness. For the time being, our study highlights that self-monitoring devices have the potential to become a valuable supplement in antepartum care. However, from a current standpoint, it seems unlikely that devices for pregnancy self-monitoring will relieve EDs from consultations by pregnant women for nonurgent indications any time soon.

Conclusions

Our study provides a first comprehensive picture of the attitudes of women toward pregnancy self-monitoring at a time when the use of such technology is not established in Germany. The majority of study participants seem reserved toward any form of pregnancy monitoring that does not involve close interactions with health care professionals. However, at the same time, a vast majority expressed interest in frequent fetal monitoring if reliable and easy-to-use devices were available. This suggests that devices for fetal self-monitoring could become a valuable supplement to physicians' and midwives' obstetrics care and ought to be investigated in clinical studies soon.

Acknowledgments

The authors thank all study participants and colleagues who supported this study. This study was supported by German Society of Psychosomatic Obstetrics and Gynecology with a research fellowship to KS.

Conflicts of Interest

None declared

Multimedia Appendix 1

English translation of survey questions.

[PDF File (Adobe PDF File), 68KB - mhealth_v7i1e11458_app1.pdf]

Multimedia Appendix 2

Patient characterisitcs with respect to previous pregnancies and deliveries.

[PDF File (Adobe PDF File), 12KB - mhealth_v7i1e11458_app2.pdf]

Multimedia Appendix 3

Pregnancy monitoring with online consultation of physician.

[PDF File (Adobe PDF File), 23KB - mhealth_v7i1e11458_app3.pdf]

References

1. Krebs P, Duncan DT. Health App Use Among US Mobile Phone Owners: A National Survey. JMIR Mhealth Uhealth 2015;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
2. Tuckson RV, Edmunds M, Hodgkins ML. Telehealth. N Engl J Med 2017 Oct 19;377(16):1585-1592. [doi: 10.1056/NEJMsr1503323] [Medline: 29045204]
3. Jamison RN, Ross EL. Utilization, Reliability, and Validity of a Smartphone App for Chronic Pain Management: A Randomized Controlled Trial. iproc 2016 Dec 23;2(1):e20 [FREE Full text] [doi: 10.2196/iproc.6060]
4. Torous J, Friedman R, Keshavan M. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. JMIR Mhealth Uhealth 2014;2(1):e2 [FREE Full text] [doi: 10.2196/mhealth.2994] [Medline: 25098314]
5. Shen C, Wang MP, Chu JT, Wan A, Viswanath K, Chan SSC, et al. Health App Possession Among Smartphone or Tablet Owners in Hong Kong: Population-Based Survey. JMIR Mhealth Uhealth 2017 Jun 05;5(6):e77 [FREE Full text] [doi: 10.2196/mhealth.7628] [Medline: 28583905]
6. Wartella E, Rideout V, Montague H, Beaudoin-Ryan L, Lauricella A. Teens, Health and Technology: A National Survey. MaC 2016 Jun 16;4(3):13. [doi: 10.17645/mac.v4i3.515]
7. Pew Research Center. 2012. Mobile health 2012 Internet URL: http:// www.pewinternet.org/files/old-media/ /Files/Reports/2012/PIP_ MobileHealth2012_FINAL.pdf [accessed 2018-11-08] [WebCite Cache ID 73mYd0eCk]
8. Goetz M, Müller M, Matthies LM, Hansen J, Doster A, Szabo A, et al. Perceptions of Patient Engagement Applications During Pregnancy: A Qualitative Assessment of the Patient's Perspective. JMIR Mhealth Uhealth 2017 May 26;5(5):e73 [FREE Full text] [doi: 10.2196/mhealth.7040] [Medline: 28550005]
9. Lanssens D, Vandenberk T, Thijs IM, Grieten L, Gyselaers W. Effectiveness of Telemonitoring in Obstetrics: Scoping Review. J Med Internet Res 2017 Sep 27;19(9):e327 [FREE Full text] [doi: 10.2196/jmir.7266] [Medline: 28954715]
10. Grassl N, Nees J, Schramm K, Spratte J, Sohn C, Schott TC, et al. A Web-Based Survey Assessing the Attitudes of Health Care Professionals in Germany Toward the Use of Telemedicine in Pregnancy Monitoring: Cross-Sectional Study. JMIR Mhealth Uhealth 2018 Aug 08;6(8):e10063 [FREE Full text] [doi: 10.2196/10063] [Medline: 30089606]
11. Roess A. The Promise, Growth, and Reality of Mobile Health - Another Data-free Zone. N Engl J Med 2017 Nov 23;377(21):2010-2011. [doi: 10.1056/NEJMp1713180] [Medline: 29116869]
12. Vashist SK, Luppa PB, Yeo LY, Ozcan A, Luong JHT. Emerging Technologies for Next-Generation Point-of-Care Testing. Trends Biotechnol 2015 Nov;33(11):692-705. [doi: 10.1016/j.tibtech.2015.09.001] [Medline: 26463722]
13. Uscher-Pines L, Pines J, Kellermann A, Gillen E, Mehrotra A. Emergency department visits for nonurgent conditions: systematic literature review. Am J Manag Care 2013 Jan;19(1):47-59 [FREE Full text] [Medline: 23379744]
14. Baur C, Thangarajah F, Mallmann P, Kirn V. Notfallmäßige Vorstellung im Kreißsaal – klare Diagnose oder maternale Sorge? Geburtshilfe Frauenheilkd 2016 Oct 13;76(10). [doi: 10.1055/s-0036-1593200]
15. Kilfoyle KA, Vrees R, Raker CA, Matteson KA. Nonurgent and urgent emergency department use during pregnancy: an observational study. Am J Obstet Gynecol 2017 Feb;216(2):181.e1-181.e7 [FREE Full text] [doi: 10.1016/j.ajog.2016.10.013] [Medline: 27773714]
16. Pötzsch O. Geburten in Deutschland. Wiesbaden: Statistisches Bundesamt; 2012.
17. Abelson JS, Kaufman E, Symer M, Peters A, Charlson M, Yeo H. Barriers and benefits to using mobile health technology after operation: A qualitative study. Surgery 2017 Dec;162(3):605-611. [doi: 10.1016/j.surg.2017.05.007] [Medline: 28651777]

18. Duffy S, Lee TH. In-Person Health Care as Option B. N Engl J Med 2018 Jan 11;378(2):104-106. [doi: 10.1056/NEJMp1710735] [Medline: 29320653]
19. Currell R, Urquhart C, Wainwright P, Lewis R. Telemedicine versus face to face patient care: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2000(2):CD002098. [doi: 10.1002/14651858.CD002098] [Medline: 10796678]
20. Isetta V, León C, Torres M, Embid C, Roca J, Navajas D, et al. Telemedicine-based approach for obstructive sleep apnea management: building evidence. Interact J Med Res 2014 Feb 19;3(1):e6 [FREE Full text] [doi: 10.2196/ijmr.3060] [Medline: 24554392]
21. Kummervold PE, Johnsen JK, Skrøvseth SO, Wynn R. Using noninferiority tests to evaluate telemedicine and e-health services: systematic review. J Med Internet Res 2012 Sep 28;14(5):e132 [FREE Full text] [doi: 10.2196/jmir.2169] [Medline: 23022989]
22. Statistisches Bundesamt. 2015. Bildungsstand: Bevölkerung nach Bildungsabschluss in Deutschland URL: https://www.destatis.de/DE/ZahlenFakten/GesellschaftStaat/BildungForschungKultur/Bildungsstand/Tabellen/Bildungsabschluss.html [accessed 2018-11-08] [WebCite Cache ID 73ma7W89y]
23. Buysse H, De Moor G, Van Maele G, Baert E, Thienpont G, Temmerman M. Cost-effectiveness of telemonitoring for high-risk pregnant women. Int J Med Inform 2008 Jul;77(7):470-476. [doi: 10.1016/j.ijmedinf.2007.08.009] [Medline: 17923433]
24. Sun N, Rau PLP. The acceptance of personal health devices among patients with chronic conditions. Int J Med Inform 2015 Apr;84(4):288-297. [doi: 10.1016/j.ijmedinf.2015.01.002] [Medline: 25655783]
25. Meade PT, Rabelo L. The technology adoption life cycle attractor: Understanding the dynamics of high-tech markets. Technological Forecasting and Social Change 2004 Sep;71(7):667-684. [doi: 10.1016/j.techfore.2004.01.008] [Medline: 26610052]
26. Abramowicz JS, Kremkau FW, Merz E. [Obstetrical ultrasound: can the fetus hear the wave and feel the heat?]. Ultraschall Med 2012 Jun;33(3):215-217. [doi: 10.1055/s-0032-1312759] [Medline: 22700164]
27. Sheiner E, Freeman J, Abramowicz JS. Acoustic output as measured by mechanical and thermal indices during routine obstetric ultrasound examinations. J Ultrasound Med 2005 Dec;24(12):1665-1670. [Medline: 16301723]
28. Reif P, Schott S, Boyon C, Richter J, Kavšek G, Timoh KN, et al. Does knowledge of fetal outcome influence the interpretation of intrapartum cardiotocography and subsequent clinical management? A multicentre European study. BJOG 2016 Dec;123(13):2208-2217 [FREE Full text] [doi: 10.1111/1471-0528.13882] [Medline: 26888657]
29. Kouskouti C, Regner K, Knabl J, Kainer F. Cardiotocography and the evolution into computerised cardiotocography in the management of intrauterine growth restriction. Arch Gynecol Obstet 2017 Apr;295(4):811-816. [doi: 10.1007/s00404-016-4282-8] [Medline: 28180962]
30. Shaw JCA. The medicalization of birth and midwifery as resistance. Health Care Women Int 2013;34(6):522-536. [doi: 10.1080/07399332.2012.736569] [Medline: 23514572]
31. McIntosh S, Cirillo D, Wood N, Dozier AM, Alarie C, McConnochie KM. Patient evaluation of an acute care pediatric telemedicine service in urban neighborhoods. Telemed J E Health 2014 Dec;20(12):1121-1126 [FREE Full text] [doi: 10.1089/tmj.2014.0032] [Medline: 25290233]
32. Cranen K, Groothuis-Oudshoorn CG, Vollenbroek-Hutten MM, IJzerman MJ. Toward Patient-Centered Telerehabilitation Design: Understanding Chronic Pain Patients' Preferences for Web-Based Exercise Telerehabilitation Using a Discrete Choice Experiment. J Med Internet Res 2017 Jan 20;19(1):e26 [FREE Full text] [doi: 10.2196/jmir.5951] [Medline: 28108429]
33. Nissen L, Lindhardt T. A qualitative study of COPD-patients' experience of a telemedicine intervention. Int J Med Inform 2017 Dec;107:11-17. [doi: 10.1016/j.ijmedinf.2017.08.004] [Medline: 29029687]
34. McGrail KM, Ahuja MA, Leaver CA. Virtual Visits and Patient-Centered Care: Results of a Patient Survey and Observational Study. J Med Internet Res 2017 May 26;19(5):e177 [FREE Full text] [doi: 10.2196/jmir.7374] [Medline: 28550006]
35. Bonoto BC, de Araújo VE, Godói IP, de Lemos LLP, Godman B, Bennie M, et al. Efficacy of Mobile Apps to Support the Care of Patients With Diabetes Mellitus: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. JMIR Mhealth Uhealth 2017 Mar 01;5(3):e4 [FREE Full text] [doi: 10.2196/mhealth.6309] [Medline: 28249834]

Abbreviations

CTG: cardiotocography
ED: emergency department
eHealth: electronic health

Edited by G Eysenbach; submitted 01.07.18; peer-reviewed by D Lanssens, D Gootjes; comments to author 16.08.18; revised version received 08.09.18; accepted 27.09.18; published 07.01.19.

Please cite as:
Schramm K, Grassl N, Nees J, Hoffmann J, Stepan H, Bruckner T, Haun MW, Maatouk I, Haist M, Schott TC, Sohn C, Schott S Women's Attitudes Toward Self-Monitoring of Their Pregnancy Using Noninvasive Electronic Devices: Cross-Sectional Multicenter Study
JMIR Mhealth Uhealth 2019;7(1):e11458
URL: https://mhealth.jmir.org/2019/1/e11458/
doi:10.2196/11458
PMID:30617040

Original Paper

In-Home Cardiovascular Monitoring System for Heart Failure: Comparative Study

Nicholas J Conn[1], PhD; Karl Q Schwarz[2], MD; David A Borkholder[1], PhD

[1]Microsystems Engineering, Rochester Institute of Technology, Rochester, NY, United States

[2]University of Rochester Medical Center, School of Medicine and Dentistry, University of Rochester, Rochester, NY, United States

Corresponding Author:
David A Borkholder, PhD
Microsystems Engineering
Rochester Institute of Technology
168 Lomb Memorial Drive
Rochester, NY,
United States
Phone: 1 585 475 6067
Email: david.borkholder@rit.edu

Abstract

Background: There is a pressing need to reduce the hospitalization rate of heart failure patients to limit rising health care costs and improve outcomes. Tracking physiologic changes to detect early deterioration in the home has the potential to reduce hospitalization rates through early intervention. However, classical approaches to in-home monitoring have had limited success, with patient adherence cited as a major barrier. This work presents a toilet seat–based cardiovascular monitoring system that has the potential to address low patient adherence as it does not require any change in habit or behavior.

Objective: The objective of this work was to demonstrate that a toilet seat–based cardiovascular monitoring system with an integrated electrocardiogram, ballistocardiogram, and photoplethysmogram is capable of clinical-grade measurements of systolic and diastolic blood pressure, stroke volume, and peripheral blood oxygenation.

Methods: The toilet seat–based estimates of blood pressure and peripheral blood oxygenation were compared to a hospital-grade vital signs monitor for 18 subjects over an 8-week period. The estimated stroke volume was validated on 38 normative subjects and 111 subjects undergoing a standard echocardiogram at a hospital clinic for any underlying condition, including heart failure.

Results: Clinical grade accuracy was achieved for all of the seat measurements when compared to their respective gold standards. The accuracy of diastolic blood pressure and systolic blood pressure is 1.2 (SD 6.0) mm Hg (N=112) and –2.7 (SD 6.6) mm Hg (N=89), respectively. Stroke volume has an accuracy of –2.5 (SD 15.5) mL (N=149) compared to an echocardiogram gold standard. Peripheral blood oxygenation had an RMS error of 2.3% (N=91).

Conclusions: A toilet seat–based cardiovascular monitoring system has been successfully demonstrated with blood pressure, stroke volume, and blood oxygenation accuracy consistent with gold standard measures. This system will be uniquely positioned to capture trend data in the home that has been previously unattainable. Demonstration of the clinical benefit of the technology requires additional algorithm development and future clinical trials, including those targeting a reduction in heart failure hospitalizations.

(JMIR Mhealth Uhealth 2019;7(1):e12419) doi:10.2196/12419

KEYWORDS

ballistocardiogram; BCG; blood pressure; ECG; electrocardiogram; heart failure; Internet of Things; IoT; photoplethysmogram; PPG; remote monitoring; SpO2; stroke volume

Introduction

The Burden of Heart Failure

In-home monitoring technologies have the potential to transform the health care system by enabling the transition from reactive care to proactive and preventive care. This is especially important for cardiovascular disease (CVD), the leading cause of death worldwide. Heart failure (HF), a type of CVD characterized by a weakened heart muscle, impacts approximately 6.5 million Americans with over 960,000 new

cases each year [1]. HF costs the United States an estimated $30.7 billion annually and is expected to increase 127% to $69.7 billion by 2030 [2]. With approximately 80% of the total cost associated with HF due to hospitalization [1], there is an opportunity to reduce the cost of HF by lowering hospitalization rates.

To contain costs, the Centers for Medicare and Medicaid Services is both penalizing hospitals with excess readmissions and moving to Bundled Payments for Care Improvement [3,4]. Despite increasing penalties [5], readmission rates remain high for HF, with over 20% of patients readmitted within 30 days and up to 50% by 6 months [6]. To successfully reduce readmissions, early detection of deterioration and subsequent intervention is required. Since patient awareness of symptomatology often lags behind deterioration, successfully tracking physiologic changes in the home is a critical component of an early intervention strategy.

In-Home Monitoring to Reduce Heart Failure Hospitalization Rates

Current models for reducing HF hospitalizations through in-home monitoring have had mixed success due to delays in the analysis of potentially important clinical data [7], studies with insufficient power for drawing conclusions [8], and low adherence [9,10]. As an example, the Telemonitoring to Improve Heart Failure Outcomes trial for automated telemonitoring did not show a reduction in hospitalizations, due in part to low adherence [9]. In this large-scale study involving 826 subjects with telemonitoring, only 55% of the patients were using the system at the end of the trial. Similar results were found in the Baroreflex Activation Therapy for Heart Failure (BEAT-HF) study, where adherence was cited as a critical factor for not showing any change in HF hospitalizations with in-home monitoring of the electrocardiogram (ECG), weight, and blood pressure (BP) [10]. In this study, the subsection of the patients who had better adherence to monitoring had a significantly lower rate of hospital readmissions [11]. The BEAT-HF investigators stated “there remain difficulties in getting heart failure patients even to perform basic aspects of self-care, such as daily weight and BP monitoring” [11].

In contrast to these studies, the CardioMEMS Heart Sensor Allows Monitoring of Pressure to Improve Outcomes in NYHA Class III Heart Failure Patients (CHAMPION) trial demonstrated a 37% reduction in class III HF hospitalizations, where pulmonary artery pressures were measured daily through an implantable device (CardioMEMS) incorporating a patient-initiated data transfer using a bed-based system [12]. In comparison with prior studies that did not show a reduction in hospitalization, only 1.5% of the treatment group was noncompliant. In the CHAMPION Trial, the high level of adherence was due to the patient selection criteria, a preprocedure monitoring agreement, and the use of a nurse telephone intervention system [13]. In addition to high levels of adherence, the success of the CHAMPION trial can be attributed to the robust and clinically relevant measurements captured by the CardioMEMS device [12].

To have successful in-home monitoring with a nonimplantable system that is applicable to the broader patient population, a novel approach must be taken to bypass adherence issues. Furthermore, such a system must provide a sufficiently diverse and relevant set of measurements to practitioners that enable early detection of deterioration and intervention, as the epidemiology of HF is extremely complex.

Measurements for Monitoring Heart Failure

HF occurs when the heart muscle is weakened and unable to maintain sufficient blood flow to meet the body’s needs. Consistent monitoring of BP is critical throughout the entire management and treatment of HF [4], as optimal BP control is a primary goal for HF [14]. In part, this is because lower systolic BP is associated with increase readmission and mortality rates [15,16], and uncorrelated high systolic BP typically precipitates acute decompensation [14]. Furthermore, according to the Framingham Study, early and continuous control of elevated BP appears to be the primary method for preventing chronic HF in the general population, with pulse and systolic pressure associated with significant risk for HF [17,18].

As HF is characterized by poor cardiac performance, cardiac output (CO) is an important component in diagnosis and management of HF [19]. CO is defined as the product of stroke volume (SV) and heart rate (HR) and is typically measured using an echocardiogram, a cardiac magnetic resonance imaging (MRI), or catheterization. In acute HF, the myocardium is unable to maintain sufficient CO and if untreated leads to chronic HF and death [20]. During diuretic treatment, CO and SV must not be significantly reduced and therefore must be monitored whenever possible to ensure optimal treatment [21]. Currently there is no in-home solution for remotely and accurately monitoring CO and SV. As such, the benefits and predictive value of monitoring CO and SV on a daily basis remain unproven.

The epidemiology of HF is complex, and a limited set of measurements is often insufficient for making clinical decisions. Peripheral oxygenation saturation (SpO_2) has value as a supporting measure that can be used to determine the best course for treatment during the initial diagnosis, acute HF, and acute decompensation [20,22]. During acute decompensation, patients presenting with hypoxemia (oxygen saturation <90%) warrant hospitalization and require daily monitoring of oxygen saturation and other vital signs until stabilization [23].

The objective of this study was to demonstrate that a toilet seat–based cardiovascular monitoring system is capable of measuring systolic and diastolic blood pressures, stroke volume, and blood oxygenation.

Methods

A Toilet Seat–Based Cardiovascular Monitoring System for In-Home Monitoring

A toilet seat–based cardiovascular monitoring system can be integrated into a subject’s natural daily routine with no change in habit, enabling measurements to be taken at one or more times each day. Issues with subject preparation and subject error are greatly reduced, since skin contact is automatic and highly repeatable each use. While a toilet seat–based monitoring system

is intermittent in nature, ensured adherence will enable long-term daily trend monitoring of parameters that do not need to be captured continuously, such as BP.

This work demonstrates that a toilet seat–based monitoring system (Figure 1) is capable of accurately capturing the following clinically relevant parameters: BP, SV, and blood oxygenation. The single-lead ECG measured from the seat has been previously correlated to the 12-lead ECG and validated against standard lead II for HR, heart rate variability, QRS duration, and the corrected QT interval [24] and will not be discussed herein. This set of measurements, gathered from a single device, provides a broad view of a patient's cardiovascular health.

An Integrated System for Cardiovascular Monitoring

The proposed cardiovascular monitoring system installs directly on a standard toilet, is battery powered, wireless, waterproof, and requires no additional connections or user interaction (Figure 2). This monitoring system unobtrusively captures cardiovascular data automatically whenever the user sits on the toilet. Requiring no direct user actions for measurement, patient adherence is enhanced. The seat incorporates a single-lead ECG for measuring the electrical activity of the heart and as a reference for ensemble averaging [24], a ballistocardiogram (BCG) for measuring the mechanical forces associated with the cardiac cycle, and a photoplethysmogram (PPG) for measuring SpO_2 and pulse transit time (PTT) (Figure 2).

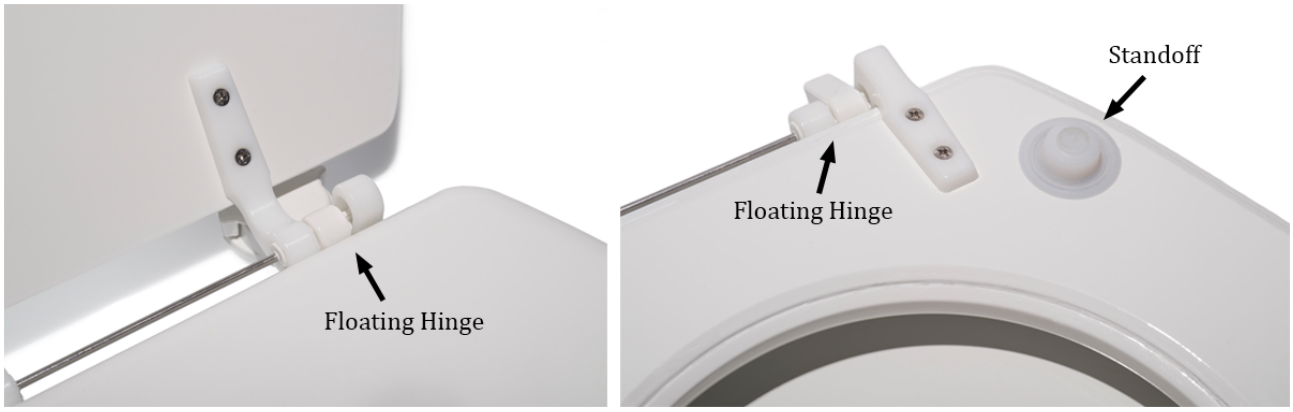
While many methods (eg, echocardiography) are available to clinicians for measuring ventricular performance (eg, CO), all require costly dedicated equipment that is generally deployed in a formal medical setting. These methods require expert technicians and interpretation that can be subject to reviewer bias or geometric errors. Because of these limitations, none of these techniques can be used for day-to-day monitoring of cardiac function or be incorporated into nonimplantable, in-home, medical diagnostic devices. The BCG fills this gap and enables devices to be created that are capable of monitoring cardiac function in the home in an inexpensive way [25-27].

The force present on the seat is measured through 4 independent load cells underneath 4 standoffs placed on the bottom of the seat (Figures 2 and 3). To accurately measure the small forces that correspond to the BCG on a toilet seat, a floating hinge (Figure 3) is required to ensure that all of the load is captured by the load cells [28].

Figure 1. A toilet seat–based cardiovascular monitoring system (left) is integrated into an individual's daily routine without requiring any change in habit, thereby addressing patient adherence. The system captures a comprehensive set of clinically relevant measurements automatically (right).

Figure 2. The toilet seat–based cardiovascular monitoring system is completely self-contained, battery-powered, wireless, and cleanable with all sensors and electronics instrumentation integrated inside of the seat. It can measure the electrocardiogram (ECG), photoplethysmogram (PPG), and the ballistocardiogram (BCG).

Figure 3. A floating hinge ensures that the weight on the seat is completely captured by the load cells under each standoff rather than having a portion of it carried by the hinge. This is a requirement for accurate ballistocardiogram monitoring on a toilet seat.



The BCG waveform has been shown to correlate with the pre-ejection period and CO [25,27,29]. To demonstrate the correlation between BCG amplitude and cardiac function, differences in the ensemble averaged BCG waveforms measured from the seat for a typical HF subject and for a typical normative subject at both rest and post-stress are shown in Figure 4. In combination with the ECG, the BCG is used to estimate SV and calculate the starting point for the pulse transit time for BP estimation.

The PPG is an optical measure of local blood volume [30]. When captured at 2 wavelengths (eg, red and infrared), SpO_2 can be estimated [30]. The ratio between the waveform amplitudes from each wavelength is used to calculate an R-value, which can then be converted to SpO_2 through a device-specific calibration curve (R-curve) [30,31]. Typically, SpO_2 is measured on the finger or the earlobe, where intersubject variability in the local tissue is low, allowing for a universal R-curve to be used across the entire population. When measuring SpO_2 from the back of the thigh, as is the case with the seat, a single point calibration is required for each subject due to the variability in local tissue. Additionally, the single point calibration for each subject will mitigate the sensitivity of the SpO_2 estimate to skin pigment, which becomes a potential source of error during hypoxia (<70%) [32].

A controlled desaturation test was performed on 2 subjects of varying weight and body type to demonstrate the difference in R-curves between subjects as measured on the seat. Each subject slowly reduced their SpO_2 to approximately 80% while data were captured on the seat. An R-value was calculated from the seat data for various saturation levels as measured by a gold standard pulse oximeter, which was taken from the finger with a hospital grade vital signs monitor (ProCare 400 Vital Signs Monitor, General Electric Company). This was used to generate an R-curve for each subject (Figure 5), where the shaded region is the potential error of the hospital grade vital signs monitor.

The slope of each subject's R-curve (–36.9 and –33.1) varies by 11% and matches that expected in literature (–33.3) [31]. While the variation in slopes is small, the difference in offset between the R-curves was significant. This indicates that while the seat is capable of measuring relative changes in SpO_2 without a per-subject calibration, a single point calibration is required for providing an absolute measure of SpO_2.

One of the key benefits of the proposed system is the ability to use a combination of the ECG, BCG, and PPG to extract meaningful parameters such as BP. Literature shows that it is possible to estimate BP from pulse wave velocity (PWV) [33-36], which is the speed at which the pressure wave propagates through the arterial system. An aortic PWV can be measured from the seat using an estimate of the subject's aortic length and the PTT. The aortic PTT is defined as the time it takes for the pressure wave to transit the aorta. The seat calculates this by determining the time interval between the BCG feature that relates to ejection and the appropriate peripheral PPG wave feature.

There are two key differences between the seat and the majority of other PWV-based estimates of BP that allow for a more accurate and robust determination of BP. Many examples in literature use the pulse arrival time instead of the PTT, which uses the ECG as the proximal timing point for calculating the time of propagation [37-39] and includes a portion of the pre-ejection period in the overall time estimate. This results in an inaccurate estimate of transit time as the ECG timing is a poor surrogate for ejection timing [36,37]. Additionally, the peripheral timing point for the seat-based PTT is measured on the back of the thigh, which is in close proximity to the end of the aorta. This is in contrast to other peripheral measurement sites, such as the finger or foot. In these cases, the PTT, and the subsequent PWV measure, is not dominated by the aorta resulting in a less robust correlation to BP.

Human Subject Testing

Diastolic BP, systolic BP, SV, and SpO_2 are validated with human subject data obtained from studies at the Rochester Institute of Technology and the University of Rochester Medical Center. Studies were performed under informed consent and used protocols approved by each institution's Institutional Review Board for Protection of Human Subjects. General exclusion criteria for all of the studies include subjects who are less than 18 years of age, pregnant, weigh more than 180 kg, cannot follow instructions in English, or have mechanical circulatory support or impaired cognitive or functional status. Each controlled study compares the capabilities of the seat to

a clinical grade gold standard, quantitatively comparing the accuracy of each measure.

For each of the following studies, recordings were captured in a lab or clinical setting. Subjects were instructed not to urinate or defecate, not to talk, and to sit as they normally would in their home when recordings were captured. No other instructions were given. As urination and defecation can shift BP, SV, and HR, it is not the intention of this work to analyze the physiologic changes during urination or defecation but rather to track daily trends at steady state. In future in-home studies, algorithms will be developed to identify and reject periods of urination and defecation through classification of motion artifacts and the physiologic shifts associated with this change in state.

Prior to parameter estimation and feature extraction, each of the signals undergoes a continuous signal quality check where entire recordings may be rejected, and beats with abnormal intervals are removed, since changes in the diastolic duration and ventricular filling can significantly shift beat-by-beat BP and SV, as described in Conn et al [24,28]. The demographic information for the remaining subjects in each cohort, grouped by measurement, are shown in Table 1.

Figure 4. The ballistocardiogram (BCG) amplitude and timing vary greatly based on the cardiovascular state. Heart failure (HF) BCG waveforms have a much smaller amplitude when compared with the normal BCG waveform at rest and poststress.

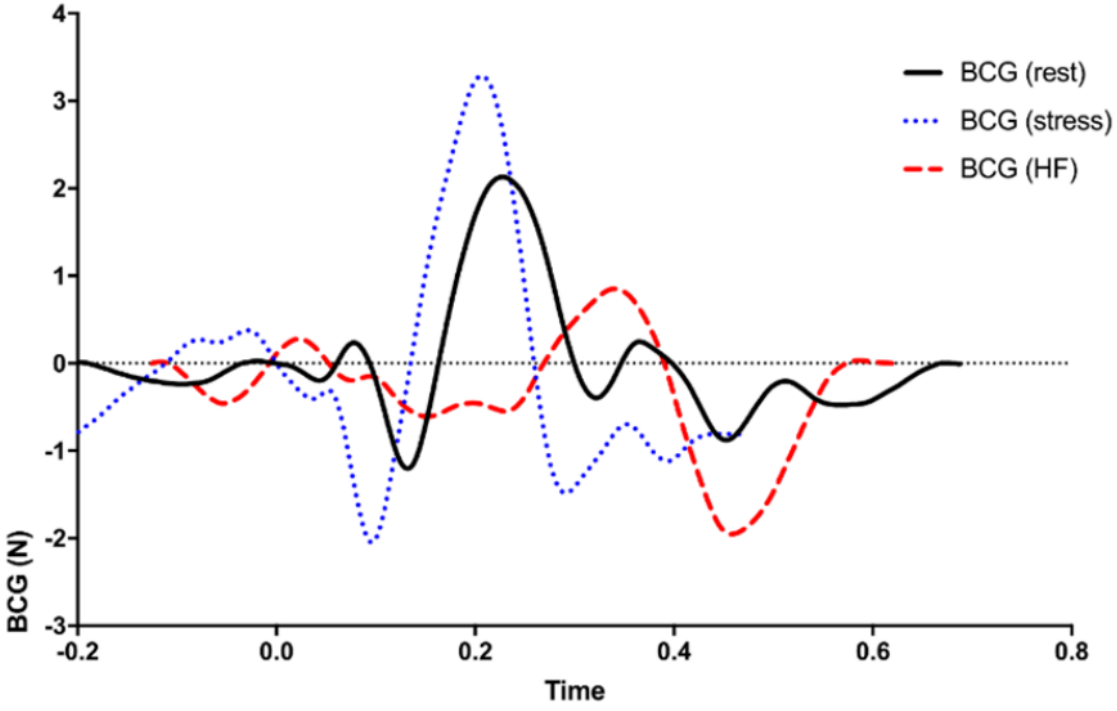


Figure 5. When characterizing the R-value for SpO_2 (peripheral oxygen saturation) estimation on the seat with controlled desaturation testing, the R-curve slope matches literature and is the same across subjects. A different offset necessitates the use of a per-subject calibration for absolute SpO_2 estimation. The shaded regions represent the acceptable level of error around the best fit line according to the ISO standard for pulse oximetry.

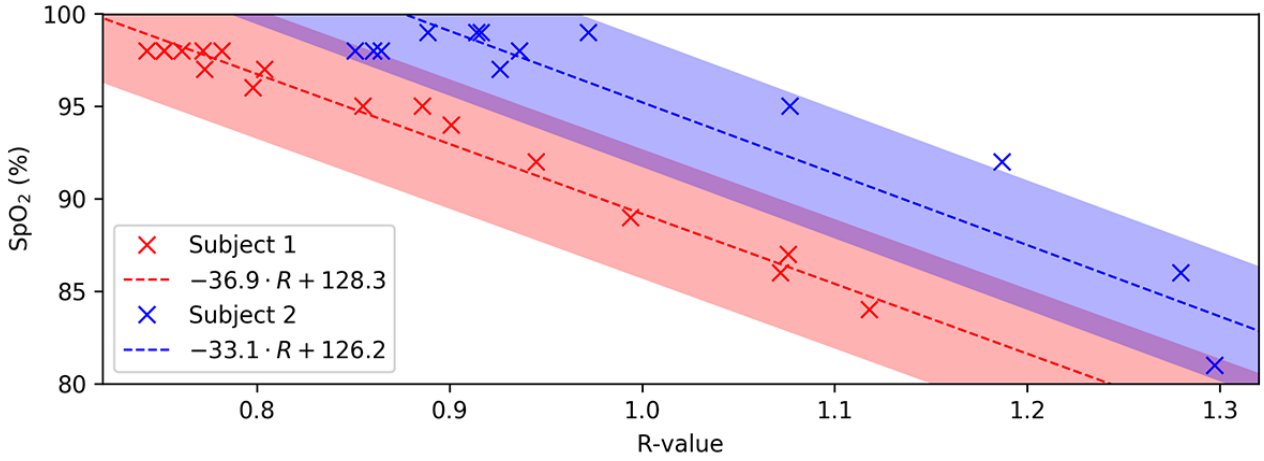


Table 1. Demographic information for the cohorts used to validate each of the seat measures. These statistics only include data that have passed the automated signal quality check.

Cohort	Male (n)	Age (years)	Weight (kg)	Height (cm)	Body mass index (kg/m^2)
Diastolic BP[a] (n=12)	9	22.5 (2.8)	76.1 (13.5)	173.6 (9.3)	25.2 (3.6)
Systolic BP (n=11)	9	22.6 (2.9)	76.7 (13.9)	175.2 (7.9)	24.9 (3.5)
SV[b] Normative (n=38)	22	24.5 (5.6)	73.3 (21.1)	171.8 (7.6)	24.7 (6.2)
SV In-Clinic (n=111)	59	55.7 (16.4)	81.3 (19.5)	170.6 (10.3)	27.8 (5.6)
SpO_2[c] (n=11)	10	22.5 (2.9)	77.6 (12.9)	173.9 (9.5)	25.6 (3.4)

[a]BP: blood pressure.

[b]SV: stroke volume.

[c]SpO_2: peripheral oxygen saturation.

Systolic and diastolic BP was validated on 12 and 11 normative (healthy) subjects, respectively, with no history of heart disease, chronic obstructive pulmonary disease, diabetes, or peripheral vascular disease. A single measurement session was used for calibrating the blood pressure estimator. For each subject, a maximum of 5 recordings was taken per week for a total of 8 weeks. A maximum of 2 recordings was allowed per day with a minimum of 4 hours separating subsequent recordings.

During in-home use of the seat, it cannot be guaranteed that the subject will follow the recommendations of the American Heart Association, where BP should be measured after the subject has sat at rest for at least 5 minutes with their back supported by a chair [40]. To ensure that the seat accurately measures the subject's BP during typical use, the gold standard measure of brachial BP was captured from a hospital-grade vital signs monitor (ProCare 400 Vital Signs Monitor, General Electric Company) before, during, and after every seat recording, with the average value compared to the seat estimate. As the PTT measured from the seat is dominated by the aorta, the resulting BP estimate is most related to central BP. While the differences between central pressures and brachial pressures are variable between individuals [41], the shifts in pressure will track together within a single subject, allowing for a comparison to brachial BP over time. This potential source of error is mitigated by performing a per-subject calibration of the PWV model to brachial pressures.

SV was validated on 38 normative subjects and 111 patients undergoing a standard echocardiogram at the University of Rochester Medical Center for any underlying condition, including HF. For each of the subjects in this cohort, a reference measure of SV was gathered at rest while the subject was supine using an echocardiogram (Vivid i, General Electric Company). SV was calculated from the velocity time interval at the left ventricular outflow tract using Doppler mode with all estimations performed by a single cardiologist to eliminate interobserver variability. The seat data were gathered during the same session while at rest and compared to the echocardiogram measure of SV.

SpO_2 was validated on 11 normative subjects with no history of heart disease, chronic obstructive pulmonary disease, diabetes, or peripheral vascular disease. A single calibration session was used to calculate the subject-specific R-curve offset. For each subject, a maximum of 5 recordings was taken per week for a total of 8 weeks. A maximum of 2 recordings was allowed per day, with a minimum of 4 hours separating subsequent recordings. A gold standard measure of SpO_2 was captured from a hospital-grade vital signs monitor (ProCare 400 Vital Signs Monitor) before and after every recording. The average SpO_2 value was used to determine the accuracy of the seat estimate.

Results

Blood Pressure Estimates Robustly Correlate to Gold Standard

The Bland-Altman plots in Figure 6 demonstrate that diastolic (left) and systolic (right) BP estimates compare favorably to a clinical gold standard. Additional recordings were automatically rejected for systolic BP (N=89) compared to diastolic BP (N=112) because signal quality requirements are more stringent for systolic BP. The resulting error for the diastolic BP is 1.2 (SD 6.0) mm Hg and the resulting error for the systolic BP is –2.7 (SD 6.6) mm Hg. These results exceed the Association for the Advancement of Medical Instrumentation (AAMI) standards, which require all measurements across every subject to have an accuracy lower than ±5 (SD 8) mm Hg [42]. For visualization on the Bland-Altman plots in Figure 6, the required SD has been converted to a limits of agreement (SD multiplied by 1.96) and is shown as a shaded region around the mean. These data demonstrate that the seat is capable of tracking shifts in blood pressure over time without recalibration.

When the estimation error is stratified by body mass index (BMI), the seat's estimate of BP slightly overestimates diastolic BP and underestimates systolic compared to the gold standard for larger BMIs. This trend is more significant for the systolic BP estimate. The error in estimating both systolic and diastolic BP stratified by BMI for this cohort is shown in Table 2.

Accurate Stroke Volume Estimation Across a Large and Diverse Population

The seat estimate of SV is compared to the echocardiogram in a Bland-Altman plot (Figure 7). The mean and SD of the heart rate across all subjects in this cohort is 74.3 (SD 12.4) bpm with a range of 47.8 bpm to 113.2 bpm. Literature indicates that the limits of agreement (1.96 SD) for an echocardiogram

Doppler-based measure of SV is 35.2 mL when calculated from the velocity time integral [43,44]. This is shown as a shaded region in Figure 7. Similarly, the limits of agreement for SV calculated using a cardiac MRI compared to thermodilution is 22 mL [45]. These compare favorably with the seat estimate of SV across a diverse population with limits of agreement of 30.4 mL when compared to an echocardiogram (Figure 7). When stratified by BMI, the limits of agreement are as follows: 15.3 mL (N=4) for a BMI less than 18.5, 26.6 mL (N=56) for a BMI between 18.5 and 25, 34.6 mL (N=56) for a BMI between 25 and 30, and 29.5 mL (N=33) for a BMI over 30. This indicates that the ability of the seat to estimate SV is not significantly impacted by BMI.

Consistent Estimation of Peripheral Blood Oxygenation

The Bland-Altman plot in Figure 8 shows the accuracy of the seat's estimate of SpO_2 compared to the gold standard for 91 data points collected from 11 subjects over a period of 8 weeks. The resulting root mean square error (A_{RMS}) of the seat estimate of SpO_2 compared to the gold standard is 2.3%. This exceeds the accuracy required by the ISO standard for SpO_2 ($A_{RMS,MAX}$=3.5%) [46]. Assuming a zero mean for visualization purposes, the required $A_{RMS,MAX}$ can be converted to a limits of agreement (shaded region in Figure 8) by multiplying by 1.96.

Figure 6. The toilet seated–based cardiovascular monitoring system has been shown to accurately measure blood pressure over an 8-week period. Both the diastolic (left) and systolic (right) blood pressure (BP) estimates from the seat exceed the accuracy required by the Association for the Advancement of Medical Instrumentation (AAMI) standard converted to a limits of agreement (shaded regions).

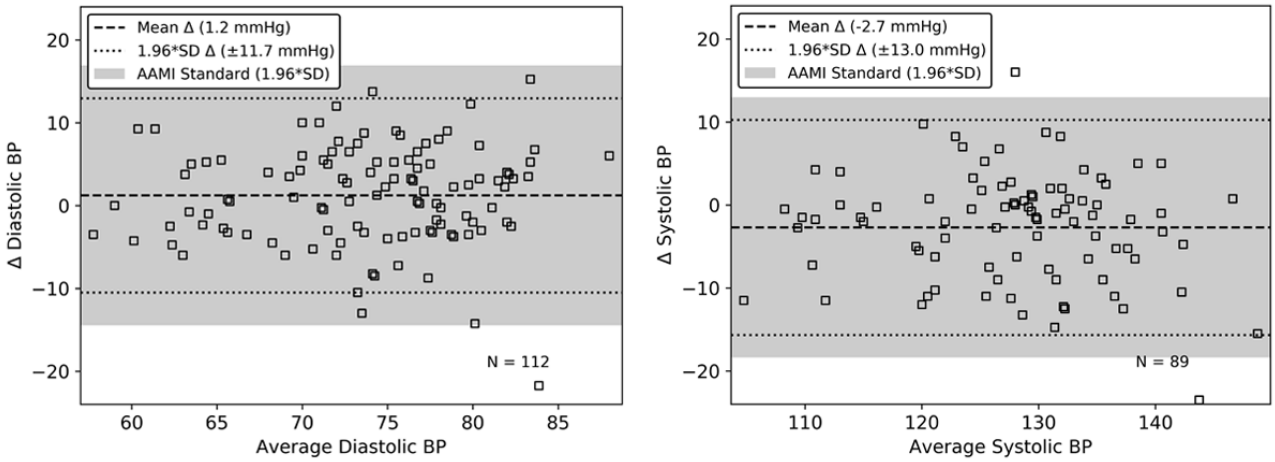


Table 2. Error in blood pressure estimation for this cohort stratified by body mass index at the time of calibration (no subjects in this cohort had a body mass index lower than 18.5 kg/m^2).

BMI[a], range	Diastolic (mm Hg)		Systolic (mm Hg)	
	n	mean (SD)	n	mean (SD)
18.5≤BMI<25.0	52	–0.1 (6.5)	49	–1.8 (7.1)
25.0≤BMI<30.0	50	2.0 (5.4)	34	–3.5 (5.8)
30.0≤BMI	10	3.6 (4.4)	6	–5.6 (5.3)

[a]BMI: body mass index; expressed in kg/m^2.

Figure 7. The seat estimate of stroke volume (SV) strongly correlates to the echocardiography measure of SV measured from the left ventricular outflow tract (LVOT). In comparison, the literature shows that the echocardiogram SV measure has a limits of agreement of 35.2 mL (shaded region) compared to an arterial gold standard.

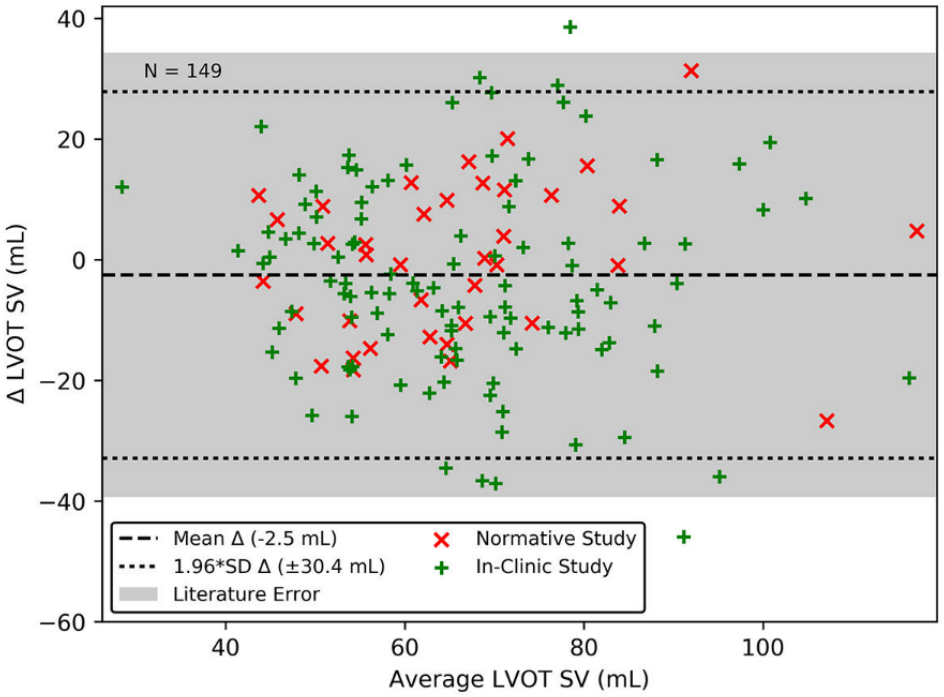
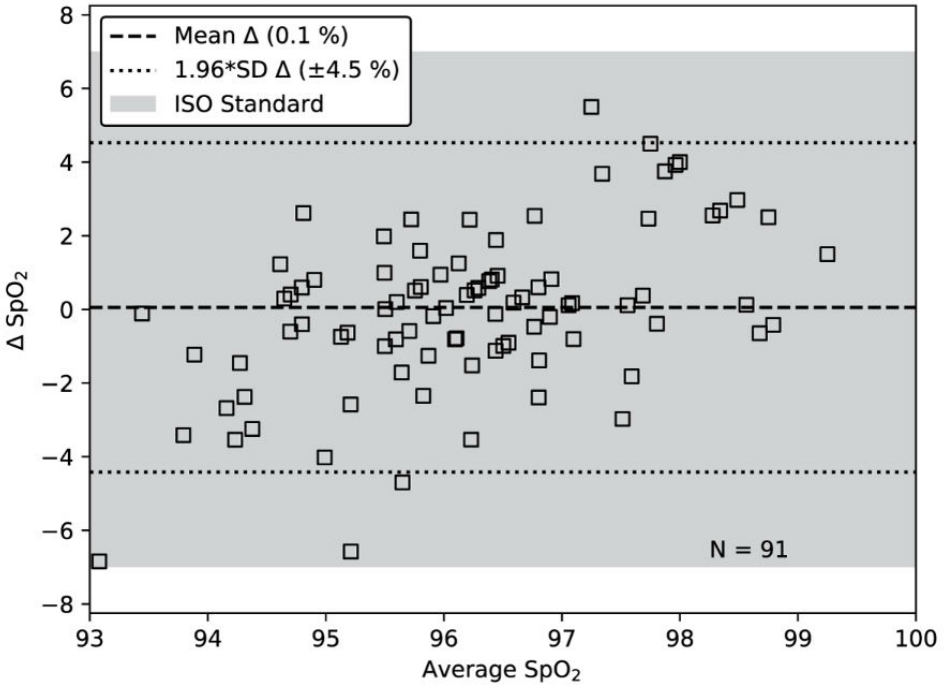


Figure 8. The limits of agreement for SpO_2 (peripheral oxygen saturation) is 4.5% with an A_{RMS} (root mean square error) of 2.3%. This exceeds the accuracy required by the ISO standard for SpO_2 where $A_{RMS, MAX}$ is 3.5%, which equates to a limits of agreement of 6.9% (shaded region).



Discussion

Principal Results

This work demonstrates that a toilet seat–based cardiovascular monitoring system can robustly measure systolic and diastolic BP, SV, and SpO_2 compared to their respective clinical gold standards (Table 3). Results show that SV can be estimated absolutely with an accuracy comparable to the echocardiogram, which is the most commonly used method for measuring SV.

Table 3. Principal results for each of the seat measures compared to their respective clinical gold standards.

Metric	Diastolic BP[a] (mm Hg; n=112)	Systolic BP (mm Hg; n=89)	SV[b] (mL; n=149)	A_{RMS}[c] SpO_2[d] (%; n=91)
Error	1.2 (6.0)	–2.7 (6.6)	–2.5 (30.4)[e]	2.3
Target	±5.0 (8.0)	±5.0 (8.0)	±2.0 (35.2)[e]	3.5

[a]BP: blood pressure.

[b]SV: stroke volume.

[c]A_{RMS}: root mean square error.

[d]SpO_2: peripheral oxygen saturation.

[e]Error and target presented as limits of agreement (1.96*SD).

Both systolic and diastolic BP can be measured in a normative population over a period of 8 weeks using the ECG, BCG, and PPG with a per-subject calibration within the AAMI standards [42]. The seat is capable of measuring a subject's SpO_2 using a reflectance mode pulse oximeter positioned to make contact with the upper thigh with a single-point calibration for each subject. The single-lead ECG, HR, HRV, QRS duration, and corrected QT interval have been previously validated [24].

Limitations and Future Work

One limitation of this work is that the seat-based system has not been used extensively in the home setting. In each of these studies, the seat was set up for automatic detection and data acquisition upon subject use with unattended data transmission to a cloud database. There was no use of lab or clinic infrastructure for gathering seat data, and the seat was deployed exactly as it would be in the home. While BP and SpO_2 were measured over multiple weeks, the majority of subjects did not show significant shifts within this time frame. Another limitation of the work is that the limits of agreement for SV is similar to that of the echocardiogram. It is unknown whether the errors in the estimation of SV are due to the seat or the echocardiographic measure of the left ventricular outflow tract area. As such, future studies will compare the seat's estimate of SV to a gold standard cardiac MRI.

Future work will initially focus on larger scale studies to validate BP and SpO_2 across a broader population. As urination and defecation can shift BP, SV, and HR, detection algorithms will be developed and validated to remove these states from the analysis, ensuring that accurate daily trends at steady state are captured. Subject identification using biometrics to enable truly passive monitoring in a multiple person household will also be developed. Additionally, the possibility of removing the per-subject calibration required for BP and SpO_2 estimation will be investigated through the incorporation of easily measured subject-specific information (eg, thigh thickness). Body weight measurements on the seat have not yet been verified. Future studies will investigate the ability of machine learning to estimate body weight based on posture and body type. This work and the aforementioned future studies will lead to a 2-phased clinical trial where an alert-based system for early detection of deterioration will be developed with in-home seat data and subsequently validated with the goal of demonstrating a reduction in HF hospitalization rates.

Broad Impact

The toilet seat–based cardiovascular monitoring system has the potential fill a gap in patient monitoring by capturing trend data that has been previously unattainable. This system has the potential to address many of the challenges with in-home monitoring in a form factor that integrates into the daily routine of patients, bypassing barriers to adherence and providing a comprehensive and accurate set of clinically relevant measurements. In addition to in-home monitoring, a secondary use for this device includes monitoring of patients in the hospital. While BP and SpO_2 are routinely monitored, daily measurements of SV and CO can be used to provide additional insights into the effectiveness of ongoing treatments.

Such a device may enable new approaches and capabilities in the diagnosis and treatment of cardiovascular disease, including but not limited to those with HF. After further demonstration of the measurement capabilities of the seat, this device will be uniquely positioned to advance an automated alert-based system that could be part of a broader interventional strategy in future clinical trials, with the goal of reducing HF hospitalizations. If successful, this strategy has the potential to reduce the burden of HF and cardiovascular disease on the health care industry as well as improve the quality of life for patients. Through the successful development, deployment, and integration with clinical practice, this device could facilitate the transition from a reactive to proactive-based approach to health care.

Acknowledgments

The authors thank Jennifer Forbes for assisting with all subject testing.

Conflicts of Interest

None declared.

References

1. Benjamin EJ, Blaha MJ, Chiuve SE, Cushman M, Floyd J, Fornage M, et al. Heart disease and stroke statistics—2017 update: a report from the American Heart Association. Circulation 2017 Mar 07;135(10):e146-e603. [doi: 10.1161/CIR.0000000000000485] [Medline: 28122885]
2. Heidenreich PA, Trogdon JG, Khavjou OA, Butler J, Dracup K, Ezekowitz MD, et al. Forecasting the future of cardiovascular disease in the United States: a policy statement from the American Heart Association. Circulation 2011 Mar 1;123(8):933-944 [FREE Full text] [doi: 10.1161/CIR.0b013e31820a55f5] [Medline: 21262990]
3. Centers for Medicare & Medicaid Services. 2018. Bundled Payments for Care Improvements—Advanced URL: https://innovation.cms.gov/initiatives/bpci-advanced/ [accessed 2018-06-02] [WebCite Cache ID 74knrQs1U]
4. Centers for Medicare & Medicaid Services. 2016. Bundled Payments for Care Improvement (BPCI) Initiative: general information URL: https://innovation.cms.gov/initiatives/bundled-payments/ [accessed 2018-12-18] [WebCite Cache ID 74ko1b4Ap]
5. McIlvennan CK, Eapen ZJ, Allen LA. Hospital readmissions reduction program. Circulation 2015 May 19;131(20):1796-1803 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.114.010270] [Medline: 25986448]
6. O'Connor CM. High heart failure readmission rates: is it the health system's fault? JACC Heart Fail 2017 Dec;5(5):393 [FREE Full text] [doi: 10.1016/j.jchf.2017.03.011] [Medline: 28449800]
7. Anker SD, Koehler F, Abraham WT. Telemedicine and remote management of patients with heart failure. Lancet 2011 Aug 20;378(9792):731-739. [doi: 10.1016/S0140-6736(11)61229-4] [Medline: 21856487]
8. Koehler F, Winkler S, Schieber M, Sechtem U, Stangl K, Böhm M, et al. Impact of remote telemedical management on mortality and hospitalizations in ambulatory patients with chronic heart failure: the telemedical interventional monitoring in heart failure study. Circulation 2011 May 3;123(17):1873-1880 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.111.018473] [Medline: 21444883]
9. Chaudhry SI, Mattera JA, Curtis JP, Spertus JA, Herrin J, Lin Z, et al. Telemonitoring in patients with heart failure. N Engl J Med 2010 Dec 9;363(24):2301-2309 [FREE Full text] [doi: 10.1056/NEJMoa1010029] [Medline: 21080835]
10. Ong MK, Romano PS, Edgington S, Aronow HU, Auerbach AD, Black JT, Better Effectiveness After Transition–Heart Failure (BEAT-HF) Research Group. Effectiveness of remote patient monitoring after discharge of hospitalized patients with heart failure: the better effectiveness after transition—Heart Failure (BEAT-HF) randomized clinical trial. JAMA Intern Med 2016 Mar;176(3):310-318 [FREE Full text] [doi: 10.1001/jamainternmed.2015.7712] [Medline: 26857383]
11. O'Riordan M. Heartwire from Medscape. 2015. BEAT-HF: remote telemonitoring fails to prevent HF rehospitalizations (transcript) URL: https://login.medscape.com/login/sso/getlogin?urlCache=aHR0cHM6Ly93d3cubWVkc2NhcGUuY29tL3ZpZXdhcnRpY2xlLzg1NDExMg==&ac=401 [accessed 2018-12-18] [WebCite Cache ID 74koYCG7Q]
12. Abraham WT, Adamson PB, Bourge RC, Aaron MF, Costanzo MR, Stevenson LW, et al. Wireless pulmonary artery haemodynamic monitoring in chronic heart failure: a randomised controlled trial. Lancet 2011 Feb 19;377(9766):658-666. [doi: 10.1016/S0140-6736(11)60101-3] [Medline: 21315441]
13. Rathman LD, Fiorini DM, Nissley KM, Kurtz KM, Small RS, Roberts JD. Holding up both ends of the bargain: ambulatory hemodynamic monitoring using CardioMEMS. J Cardiac Fail 2016 Aug;22(8):S56. [doi: 10.1016/j.cardfail.2016.06.167]
14. Yancy CW, Jessup M, Bozkurt B, Butler J, Casey DE, Drazner MH, et al. 2013 ACCF/AHA guideline for the management of heart failure: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines. J Am Coll Cardiol 2013 Oct 15;62(16):e147-e239 [FREE Full text] [doi: 10.1016/j.jacc.2013.05.019] [Medline: 23747642]
15. Gheorghiade M, Abraham WT, Albert NM, Greenberg BH, O'Connor CM, She L, OPTIMIZE-HF Investigators and Coordinators. Systolic blood pressure at admission, clinical characteristics, and outcomes in patients hospitalized with acute heart failure. JAMA 2006 Nov 08;296(18):2217-2226. [doi: 10.1001/jama.296.18.2217] [Medline: 17090768]
16. Chin MH, Goldman L. Correlates of early hospital readmission or death in patients with congestive heart failure. Am J Cardiol 1997 Jun 15;79(12):1640-1644. [Medline: 9202355]
17. Kannel WB, Castelli WP, McNamara PM, McKee PA, Feinleib M. Role of blood pressure in the development of congestive heart failure. The Framingham study. N Engl J Med 1972 Oct 19;287(16):781-787. [doi: 10.1056/NEJM197210192871601] [Medline: 4262573]
18. Haider AW, Larson MG, Franklin SS, Levy D, Framingham Heart Study. Systolic blood pressure, diastolic blood pressure, and pulse pressure as predictors of risk for congestive heart failure in the Framingham Heart Study. Ann Intern Med 2003 Jan 07;138(1):10-16. [Medline: 12513039]
19. Mann D, Zipes D, Libby P, Bonow R, editors. Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine, Third Edition. Philadelphia: Saunders; 1988.
20. Nieminen MS, Böhm M, Cowie MR, Drexler H, Filippatos GS, Jondeau G, ESC Committe for Practice Guideline (CPG). Executive summary of the guidelines on the diagnosis and treatment of acute heart failure: the Task Force on Acute Heart Failure of the European Society of Cardiology. Eur Heart J 2005 Feb;26(4):384-416. [doi: 10.1093/eurheartj/ehi044] [Medline: 15681577]

21. Remme WJ, Swedberg K, Task Force for the Diagnosis and Treatment of Chronic Heart Failure, European Society of Cardiology. Guidelines for the diagnosis and treatment of chronic heart failure. Eur Heart J 2001 Sep;22(17):1527-1560. [doi: 10.1053/euhj.2001.2783] [Medline: 11492984]
22. Mebazaa A, Yilmaz MB, Levy P, Ponikowski P, Peacock WF, Laribi S, et al. Recommendations on pre-hospital & early hospital management of acute heart failure: a consensus paper from the Heart Failure Association of the European Society of Cardiology, the European Society of Emergency Medicine and the Society of Academic Emergency Medicine. Eur J Heart Fail 2015 Jun;17(6):544-558 [FREE Full text] [doi: 10.1002/ejhf.289] [Medline: 25999021]
23. Joseph SM, Cedars AM, Ewald GA, Geltman EM, Mann DL. Acute decompensated heart failure: contemporary medical management. Tex Heart Inst J 2009;36(6):510-520 [FREE Full text] [Medline: 20069075]
24. Conn NJ, Schwarz KQ, Borkholder DA. Nontraditional electrocardiogram and algorithms for inconspicuous in-home monitoring: comparative study. JMIR Mhealth Uhealth 2018 May 28;6(5):e120 [FREE Full text] [doi: 10.2196/mhealth.9604] [Medline: 29807881]
25. Inan OT, Etemadi M, Paloma A, Giovangrandi L, Kovacs GTA. Non-invasive cardiac output trending during exercise recovery on a bathroom-scale-based ballistocardiograph. Physiol Meas 2009 Mar;30(3):261-274. [doi: 10.1088/0967-3334/30/3/003] [Medline: 19202234]
26. Inan OT, Etemadi M, Wiard RM, Giovangrandi L, Kovacs GTA. Robust ballistocardiogram acquisition for home monitoring. Physiol Meas 2009 Feb;30(2):169-185. [doi: 10.1088/0967-3334/30/2/005] [Medline: 19147897]
27. Giovangrandi L, Inan OT, Wiard RM, Etemadi M, Kovacs GTA. Ballistocardiography—a method worth revisiting. Conf Proc IEEE Eng Med Biol Soc 2011;2011:4279-4282 [FREE Full text] [doi: 10.1109/IEMBS.2011.6091062] [Medline: 22255285]
28. Conn N. Robust Algorithms for Unattended Monitoring of Cardiovascular Health [Dissertation]. Rochester: Rochester Institute of Technology; 2016.
29. Inan OT, Etemadi M, Wiard RM, Kovacs GTA, Giovangrandi L. Non-invasive measurement of Valsalva-induced hemodynamic changes on a bathroom scale ballistocardiograph. Conf Proc IEEE Eng Med Biol Soc 2008;2008:674-677. [doi: 10.1109/IEMBS.2008.4649242] [Medline: 19162745]
30. Webster J, editor. Design of Pulse Oximeters. Boca Raton: CRC Publishers; 1997.
31. Azmal G. Continuous measurement of oxygen saturation level using photoplethysmography signal. 2006 Presented at: 2006 International Conference on Biomedical and Pharmaceutical Engineering; 2006; Singapore p. 504-507.
32. Bickler PE, Feiner JR, Severinghaus JW. Effects of skin pigmentation on pulse oximeter accuracy at low saturation. Anesthesiology 2005 Apr;102(4):715-719. [Medline: 15791098]
33. Lillie JS, Liberson AS, Mix D, Schwarz KQ, Chandra A, Phillips DB, et al. Pulse wave velocity prediction and compliance assessment in elastic arterial segments. Cardiovasc Eng Technol 2015 Mar;6(1):49-58. [doi: 10.1007/s13239-014-0202-x] [Medline: 26577102]
34. Lillie JS, Liberson AS, Borkholder DA. Quantification of hemodynamic pulse wave velocity based on a thick wall multi-layer model for blood vessels. JFFHMT 2016. [doi: 10.11159/jffhmt.2016.007]
35. Gribbin B, Steptoe A, Sleight P. Pulse wave velocity as a measure of blood pressure change. Psychophysiology 1976 Jan;13(1):86-90. [Medline: 1244636]
36. Martin SL, Carek AM, Kim C, Ashouri H, Inan OT, Hahn J, et al. Weighing scale-based pulse transit time is a superior marker of blood pressure than conventional pulse arrival time. Sci Rep 2016 Dec 15;6(1):39273 [FREE Full text] [doi: 10.1038/srep39273] [Medline: 27976741]
37. Zhang G, Gao M, Xu D, Olivier NB, Mukkamala R. Pulse arrival time is not an adequate surrogate for pulse transit time as a marker of blood pressure. J Appl Physiol (1985) 2011 Dec;111(6):1681-1686 [FREE Full text] [doi: 10.1152/japplphysiol.00980.2011] [Medline: 21960657]
38. Pinheiro E. Pulse arrival time and ballistocardiogram application to blood pressure variability estimation. 2009 Presented at: 2009 IEEE International Workshop on Medical Measurements and Applications; 2009; Cetraro. [doi: 10.1109/MEMEA.2009.5167970]
39. Kim C, Carek AM, Mukkamala R, Inan OT, Hahn J. Ballistocardiogram as proximal timing reference for pulse transit time measurement: potential for cuffless blood pressure monitoring. IEEE Trans Biomed Eng 2015 Nov;62(11):2657-2664 [FREE Full text] [doi: 10.1109/TBME.2015.2440291] [Medline: 26054058]
40. American Heart Association. Monitoring your blood pressure at home URL: https://www.heart.org/en/health-topics/high-blood-pressure/understanding-blood-pressure-readings/monitoring-your-blood-pressure-at-home [accessed 2018-12-18] [WebCite Cache ID 74ks29fXX]
41. McEniery CM, Cockcroft JR, Roman MJ, Franklin SS, Wilkinson IB. Central blood pressure: current evidence and clinical importance. Eur Heart J 2014 Jul;35(26):1719-1725 [FREE Full text] [doi: 10.1093/eurheartj/eht565] [Medline: 24459197]
42. White WB, Berson AS, Robbins C, Jamieson MJ, Prisant LM, Roccella E, et al. National standard for measurement of resting and ambulatory blood pressures with automated sphygmomanometers. Hypertension 1993 Apr;21(4):504-509. [Medline: 8458649]

43. Espersen K, Jensen EW, Rosenborg D, Thomsen JK, Eliasen K, Olsen NV, et al. Comparison of cardiac output measurement techniques: thermodilution, Doppler, CO2-rebreathing and the direct Fick method. Acta Anaesthesiol Scand 1995 Feb;39(2):245-251. [Medline: 7793193]
44. Shahgaldi K, Manouras A, Brodin L, Winter R. Direct measurement of left ventricular outflow tract area using three-dimensional echocardiography in biplane mode improves accuracy of stroke volume assessment. Echocardiography 2010 Oct;27(9):1078-1085. [doi: 10.1111/j.1540-8175.2010.01197.x] [Medline: 20546012]
45. Hundley WG, Li HF, Hillis LD, Meshack BM, Lange RA, Willard JE, et al. Quantitation of cardiac output with velocity-encoded, phase-difference magnetic resonance imaging. Am J Cardiol 1995 Jun 15;75(17):1250-1255. [Medline: 7778549]
46. Center for Devices and Radiological Health, Food and Drug Administration. 2013 Mar 04. Pulse Oximeters—Premarket Notification Submissions [510(k)s]: Guidance for Industry and Food and Drug Administration Staff URL: https://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM081352.pdf [accessed 2018-12-18] [WebCite Cache ID 74krkq8W4]

Abbreviations

AAMI: Association for the Advancement of Medical Instrumentation
A_{RMS}: root mean square error
BCG: ballistocardiogram
BEAT-HF: Baroreflex Activation Therapy for Heart Failure
BMI: body mass index
BP: blood pressure
CHAMPION: CardioMEMS Heart Sensor Allows Monitoring of Pressure to Improve Outcomes in NYHA Class III Heart Failure Patients
CO: cardiac output
CVD: cardiovascular disease
ECG: electrocardiogram
HF: heart failure
HR: heart rate
MRI: magnetic resonance imaging
PPG: photoplethysmogram
PWV: pulse wave velocity
PTT: pulse transit time
SpO_2: peripheral oxygen saturation
SV: stroke volume

Edited by G Eysenbach; submitted 04.10.18; peer-reviewed by O Steichen, I Kedan, JB Park, M Giner-Soriano; comments to author 23.10.18; revised version received 28.10.18; accepted 16.11.18; published 18.01.19.

Please cite as:
Conn NJ, Schwarz KQ, Borkholder DA
In-Home Cardiovascular Monitoring System for Heart Failure: Comparative Study
JMIR Mhealth Uhealth 2019;7(1):e12419
URL: http://mhealth.jmir.org/2019/1/e12419/
doi:10.2196/12419
PMID:30664492

Original Paper

Perspectives of Nonphysician Clinical Students and Medical Lecturers on Tablet-Based Health Care Practice Support for Medical Education in Zambia, Africa: Qualitative Study

Sandra Barteit[1], MA; Florian Neuhann[1], MD; Till Bärnighausen[1,2,3], MSc, MD, ScD, PhD; Annel Bowa[4], MSc; Sigrid Lüders[5], MD; Gregory Malunga[4], MSc; Geoffrey Chileshe[4], BSc; Clemence Marimo[6], MD; Albrecht Jahn[1], MSc, MD, PhD

[1]Heidelberg Institute of Global Health, Heidelberg, Germany

[2]Harvard TH Chan School of Public Health, Department of Global Health and Population, Boston, MA, United States

[3]Africa Health Research Institute (AHRI), KwaZulu-Natal, South Africa

[4]Chainama College of Health Sciences, Lusaka, Zambia

[5]SolidarMed, Lusaka, Zambia

[6]School of Medicine, University of Zambia, Lusaka, Zambia

Corresponding Author:
Sandra Barteit, MA
Heidelberg Institute of Global Health
Im Neuenheimer Feld 130.3
Heidelberg, 69120
Germany
Phone: 49 06221 ext 5634030
Email: barteit@uni-heidelberg.de

Related Article:

This is a corrected version. See correction statement: http://mhealth.jmir.org/2019/4/e13431/

Abstract

Background: Zambia is faced with a severe shortage of health workers and challenges in national health financing. This burdens the medical licentiate practitioner (MLP) program for training nonphysician clinical students in Zambia because of the shortage of qualified medical lecturers and learning resources at training sites. To address this shortage and strengthen the MLP program, a self-directed electronic health (eHealth) platform was introduced, comprising technology-supported learning (e-learning) for medical education and support for health care practice. MLP students were provided with tablets that were preloaded with content for offline access.

Objective: This study aimed to explore MLP students' and medical lecturers' perceptions of the self-directed eHealth platform with an offline-based tablet as a training and health care practice support tool during the first year of full implementation.

Methods: We conducted in-depth qualitative interviews with 8 MLP students and 5 lecturers and 2 focus group discussions with 16 students to gain insights on perceptions of the usefulness, ease of use, and adequacy of self-directed e-learning and health care practice support accessible through the offline-based tablet. Participants were purposively sampled. Verbatim transcripts were analyzed following hypothesis coding.

Results: The eHealth platform (e-platform), comprising e-learning for medical education and health care practice support, was positively received by students and medical lecturers and was seen as a step toward modernizing the MLP program. Tablets enabled equal access to offline learning contents, thus bridging the gap of slow or no internet connections. The study results indicated that the e-platform appears adequate to strengthen medical education within this low-resource setting. However, student self-reported usage was low, and medical lecturer usage was even lower. One stated reason was the lack of training in tablet usage and another was the quality of the tablets. The mediocre quality and quantity of most e-learning contents were perceived as a primary concern as materials were reported to be outdated, missing multimedia features, and addressing only part of the curriculum.

Medical lecturers were noted to have little commitment to updating or creating new learning materials. Suggestions for improving the e-platform were given.

Conclusions: To address identified major challenges, we plan to (1) introduce half-day training sessions at the beginning of each study year to better prepare users for tablet usage, (2) further update and expand e-learning content by fostering collaborations with MLP program stakeholders and nominating an e-platform coordinator, (3) set up an e-platform steering committee including medical lecturers, (4) incorporate e-learning and e-based health care practice support across the curriculum, as well as (5) implement processes to promote user-generated content. With these measures, we aim to sustainably strengthen the MLP program by implementing the tablet-based e-platform as a serious learning technology for medical education and health care practice support.

(JMIR Mhealth Uhealth 2019;7(1):e12637) doi:10.2196/12637

KEYWORDS

educational technology; computers, handheld; computer-assisted decision making; mobile apps; information dissemination; education; allied health occupations; Africa, Southern; Zambia

Introduction

Background

The severe shortage of health workers in sub-Saharan Africa is projected to worsen until 2030 and is aggravated by significant population growth [1]. This shortage is profoundly burdening the Zambian health care system. Although Zambia launched efforts to reduce this shortage [2], remote and rural areas are still severely underserved with only 7 clinicians per 10,000 people (urban areas: 16/10,000 people) [2-4] and insufficient coverage of health facilities [5]. Health workers are essential for closing the gap in delivering essential health services [3] and for moving forward toward universal health coverage and meeting health-related objectives of the sustainable development goals [6]. The Zambian National Health Strategic Plan for 2017-2021 [7] aims "to provide equitable access to cost-effective, quality health services as close to the family as possible" and specifically recognizes the need to "increase the proportion of rural households living within 5 km of the nearest health facility from the current 50% to 80% in 2030" [7]. Furthermore, efforts have been made to improve the allocation of current health workers throughout Zambia's districts [2]. However, Zambia's medical training institutions do not currently have the capacity to train the necessary number of health workers to sufficiently staff underserved areas [8].

In 2002, to alleviate the severe health worker shortage, the Ministry of Health initiated a medical licentiate practitioner (MLP) program at the Chainama College of Health Sciences (CCHS), Lusaka [9]. MLPs are nonphysician clinicians who receive shorter, skills-oriented medical training compared with physicians but can perform many traditionally physician-designated diagnostic and therapeutic tasks [10]. At first, the MLP program was implemented as a 2-year upgrade training for the cadre of clinical officers [4,11] and later transformed from a diploma-level degree to a Bachelor of Science degree in Clinical Sciences in 2013 [9]. A 1-year bridging course allowed diploma-holding MLPs to also acquire a Bachelor of Science degree at CCHS.

The importance of MLPs for Zambia's health system has been shown, particularly in rural areas at the district level where the most critical shortages exist [9,10,12]. MLPs are trained in 4 main specialties: pediatrics, surgery, internal medicine, and obstetrics and gynecology. The last 2 years of the MLP training are focused on practical skills, whereby MLP students rotate every 8 weeks across health facilities, each with a medical focus in 1 of the 4 main specialties.

Objectives

To strengthen the MLP program and specifically address the lack of medical lecturers and learning resources in practicum sites, a self-directed electronic health (eHealth) [13] platform (e-platform) that comprised technology-supported learning (e-learning) for medical education and health care practice support was introduced in 2016 as part of a more extensive blended learning approach [14]. The e-learning provides Web-based and offline access to static and interactive medical e-learning materials assembled and developed according to the existing curriculum, such as lecture notes, medical books, virtual patient cases, medical pictures, and videos on medical procedures. The health care practice support component provides access to standard treatment guidelines and medical algorithms to diagnose and treat patients. E-platform materials were generally pre-existing and not explicitly developed as electronically based materials. As training in the third and fourth study years includes clinical rotations in remote areas throughout Zambia, tablets were distributed (7 inches, Android-based) to all third- and fourth-year and bridging students to enable ubiquitous access to tablet-based content. The e-platform was implemented for Web access with the open-source Moodle software (Moodle HQ, Moodle Community), and MLP students were able to access contents with provided tablets or their own mobile devices. The content was downloaded to the tablets for offline usage via the Moodle mobile app; no internet connection was necessary to access the tablet's contents, even in rural areas.

A pilot phase from January 2016 to August 2016 provided insight into shortcomings of the e-learning and eHealth platform (e-platform) that led to adaptations. The most significant change was an increase in available e-learning materials (including interactive learning materials), followed by a change in the tablet type that allowed mobile data access, which was highly requested by users. The first year of full implementation of the e-platform (September 2016 to August 2017) was evaluated qualitatively and quantitatively.

We present the results of the qualitative evaluation based on success factors for e-learning implementation as a framework

to categorize themes. The objective was to explore students' and medical lecturers' perceptions of the self-directed e-platform as a tool for medical learning and teaching within the MLP program in the low-resource setting of Zambia to adapt and further develop the e-platform.

Methods

Research Context

A qualitative study design was used to explore the perceptions of MLP students and medical lecturers and evaluate the success of the implementation of the offline-based, self-directed e-platform in strengthening medical education in Zambia. The development of the evaluation framework was based on a prior pilot evaluation study of the MLP e-platform. The research methodology underpinning this study was content analysis [15]. Pilot evaluation results had identified shortcomings of the blended learning approach that needed improvement and are reported elsewhere [14]. In-depth interviews (IDIs) and focus group discussions (FGDs) allowed participants to state their opinions and experiences without the restrictions of predefined categories or terms. Bhuasiri et al's research framework [16] provided the structure to categorize the identified themes within dimensions of a successful implementation of e-learning and particularly in low-resource settings (see Multimedia Appendix 1). We believed this framework was appropriate for the health care practice support component as health care practice support is also part of MLP education. With Delphi and analytic hierarchy process methods, Bhuasiri et al [16] identified the most relevant factors for e-learning success in low-resource settings and incorporated social cognitive and motivational theory, critical success factors, DeLone and McLean information system success, and the technology acceptance model. This framework was further adapted to reflect CCHS's local setting and include factors that emerged during project implementation and this study's data analysis. Bhuasiri et al's research framework [16] includes 3 main dimensions of a successful e-learning implementation: personal, environmental, and system-related. The research framework unfolds into 7 subdimensions (learners and instructors' characteristics, extrinsic motivation, e-learning environment, infrastructure, system quality, course and information quality, and institution and service quality). We further adapted and expanded these subdivisions (see Table 1) to incorporate the tablet-based e-platform in the research framework. Mobile learning in this context describes access to the MLP e-learning as well as health care practice support materials offline on mobile devices.

Data Collection

We conducted 2 FGDs with MLP students, approximately 80 min each (see Multimedia Appendix 2 for the 17-item interview guide), and 13 IDIs with students and medical lecturers, approximately 20 to 40 min each (see Multimedia Appendix 3 for the 19-item discussion guide). In the IDIs, the e-platform is referred to as the e-learning platform. Interview guides for both the IDIs and FGDs complied with the questionnaire's development guide for education research by Artino et al [17] and were based on a concise literature review, peer-based feedback rounds, and pilot testing. Moreover, 5 of the IDIs were conducted with medical lecturers and 8 with students. To reflect the study population, respondents were also purposively selected to proportionally represent gender and age groups. For each FGD, more students (n=10) than estimated necessary (n=6) were invited to ensure sufficient turnout to produce relevant information [18]. Age was categorized in 3 age groups: ≤35, >35 and ≤45, and >45 years (see Table 2 for details).

The group of medical lecturers was purposively drawn from the group of active lecturers in the third, fourth, and bridging study years within the MLP program. The IDIs with students were held in a secluded office within the MLP administration block on the CCHS campus and were conducted by the principal researcher, the first author of this manuscript, (German, female, doctoral student) in English, who was trained in information technology and global health and with over 10 years of experience in the field. The principal researcher engaged with study participants over short periods during project visits before the study's commencement as part of the intervention implementation within the overall blended learning project. The 2 FGDs were conducted with students in lecture halls on the CCHS campus. The IDIs with medical lecturers were primarily conducted via phone (n=4) and 1 was held at the University Teaching Hospital in Lusaka. IDIs with medical lecturers and FGDs with students were guided by a facilitator (Zambian, female, master degree-level in Social Sciences) in English. The facilitator had no engagement with study participants before the study's commencement.

At the time of the study, all students and medical lecturers were actively engaged in the MLP program. Furthermore, 2 student FGDs were considered sufficient to elicit the majority of prevalent themes [19]. Data saturation for IDIs was assumed according to the 10+3 criterion [20]. Field notes were taken during the interviews, and IDIs and FGDs were audio-recorded. During the IDIs and FGDs, questions were asked according to the interview and discussion protocol, thus prompting interviewees to provide further details until each line of inquiry was sufficiently covered.

The study began with the distribution of the tablets and the quantitative data collection in September 2016, after which IDIs and FGDs were held at the CCHS campus in July to August 2017 (end of the study year); therefore, the maximum exposure to the use of tablets was 12 months (see Figure 1, timeline of study).

Table 1. Bhuasiri et al's research framework [16] for successful technology-supported learning (e-learning) implementation, with our additions for the tablet-based e-platform (marked with [add] in the table). The research framework has 3 major themes (individual dimension, environmental dimension, and system dimension) that unfold into subdimensions.

Dimensions	Term definitions
1. Individual dimension	
1.1 Learner's characteristics	
Attitude toward tablet-based e-platform[a]	"Learners' impression of participating in [m-learning[b] or mHealth[c]] activities through [tablet] usage" [16]
Focus on interaction	"The degree of contact and educational exchange among learners and between learners and instructors" [16] from the student's perspective
1.2 Instructor's characteristics (medical lecturers)	
Attitude toward tablet-based e-platform	Instructor's "impression of participating in [m-learning/mHealth] activities through [tablet] usage" [16]
Interaction fairness	"The extent to which the learner feels having been treated fairly regarding his or her interaction with the instructor throughout the [m-learning/mHealth] process" [16]
Focus on interaction	"The degree of contact and educational exchange [...] between learners and instructors" [16] from the instructor's perspective
1.3 Extrinsic motivation	
Perceived usefulness	"The degree to which a person believes that using [an m-learning/mHealth] system would enhance his or her learning performance" [16]
Technological flexibility	The degree of flexibility that the technology is providing to users in a given setting [add]
Expandability	The degree to which the provided m-learning and mHealth system and technology can be expanded according to user needs [add]
Saving resources	The degree to which the provided m-learning and mHealth system and technology are saving users' resources as measured by monetary spending, time, and additional characteristics [add]
Punishment/restriction	The degree to which the provided m-learning and mHealth system and technology is restricting or punishing the user
2. Environmental dimension	
2.1 Interaction opportunities	"Learner's perceived interactions with others" [16] through m-learning and mHealth
3. System dimension	
3.1 Infrastructure and system quality	
Ease of use	"Refers to the degree to which the prospective user expects the use of [m-learning/mHealth] to be free of effort" [16]
System functionality	"The perceived ability of [m-learning/mHealth] to provide flexible access to instructional and assessment media" [16]
Technological adequacy	Refers to the degree to which the user expects the provided device to fit the setting and area of use [add]
Technological quality	The quality of the provided device as measured by battery runtime, hardware reliability, operating system quality, and other characteristics [add]
Internet quality	"The quality of the internet that can be measured by transmission rate, error rates, and other characteristics" [16]
3.2 Course and information quality	
Reliability	"Concerned with the degree of accuracy, dependability, and consistency of the information" [16]
Relevant content	"The degree of congruence between what the learner wants or requires and what is provided by the information, course content, and services" [16]
3.3 Institution and service quality	

Dimensions	Term definitions
Sustainability of the e-platform	The degree to which m-learning and mHealth is implemented sustainably within the educational infrastructure [add]
Tablet and e-platform training	"The amount of specialized instruction and practice that is afforded to the learner to increase the learner's proficiency in utilizing [m-learning/mHealth] [...]." [16]
Service quality	The quality of the service provided for m-learning and mHealth and the provided device

[a]e-platform: e-learning platform with an electronic health component.

[b]m-learning: mobile learning (with tablets and other mobile devices).

[c]mHealth: mobile health.

Table 2. Study participants of in-depth interviews and focus group discussions according to age groups and gender.

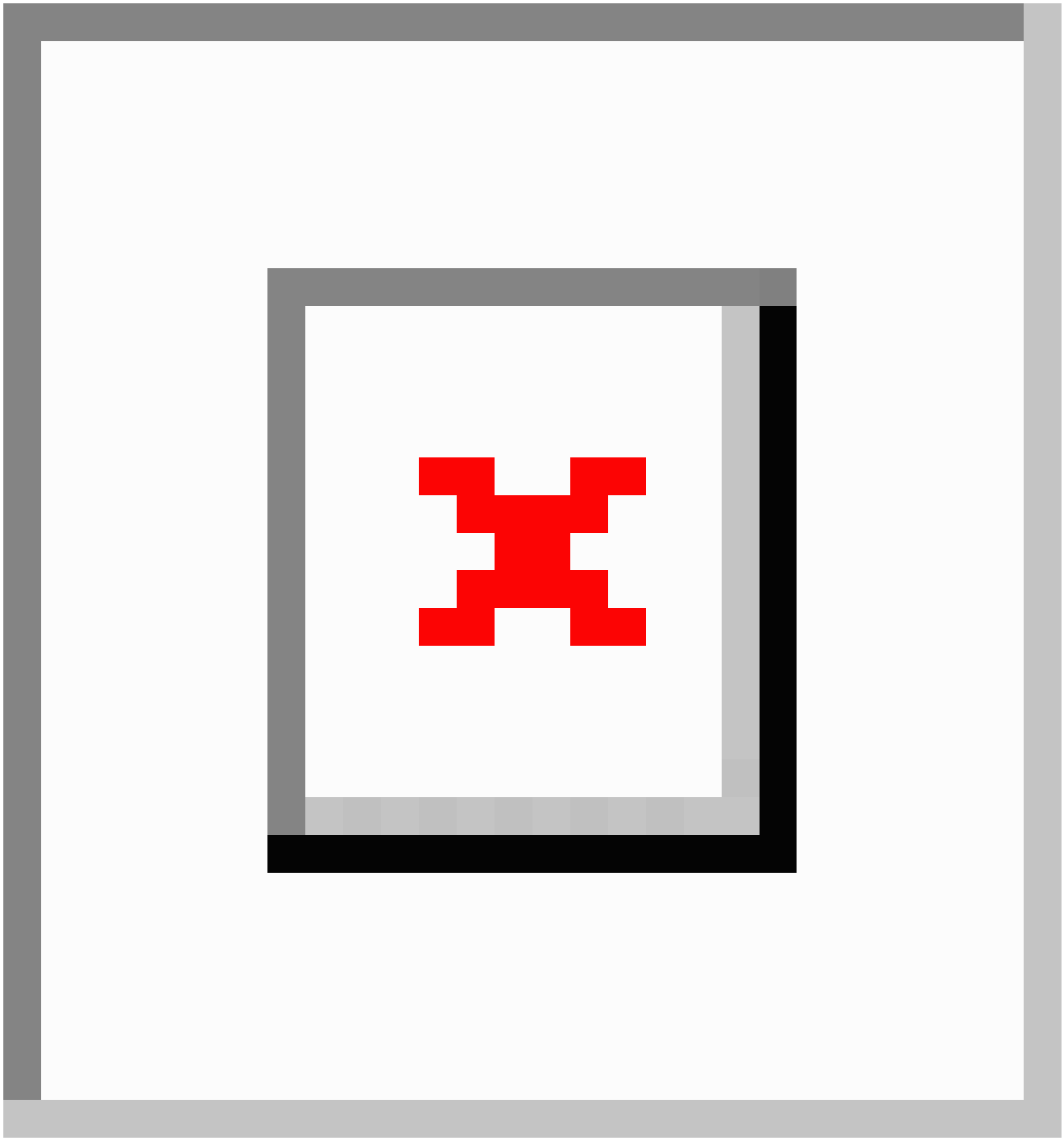
Age group (in years)	MLP[a] students				Medical lecturers' IDIs[b]	
	Female		Male		Female	Male
	IDIs	FGDs[c]	IDIs	FGDs		
≤35	2	3	0	6	0	0
>35 and ≤45	1	1	3	2	0	4
>45	1	1	1	1	0	1

[a]MLP: medical licentiate practitioner.

[b]IDIs: in-depth interviews.

[c]FGDs: focus group discussions.

Figure 1. Timeline of study events in chronological order. MLP: medical licentiate practitioner; e-platform: e-learning platform with an eHealth component; IDIs: in-depth interviews; FGDs: focus group discussions.



Data Analysis

The principal researcher recorded and then transcribed the students' IDIs. Responses from all IDIs were coded by the principal researcher in 2 coding cycles to determine initial concepts with a commercial, qualitative data analysis program (NVivo, QSR International). Hypothesis coding was applied "to assess the researcher-generated hypothesis" [21]. The adapted research framework of Bhuasiri et al [16] directed the "researcher-generated, predetermined list of codes to qualitative data specifically to assess a researcher-generated hypothesis" [21], and the themes were, therefore, defined in advance (see the coding tree in Table 1, thematic areas from 1.1-3.3). The students' and medical lecturers' responses in the transcripts were coded according to the concepts and subsequent codes from Bhuasiri's adapted research framework [21]. Coded elements were exported to a word processor where they were assigned to 1 theme. The principal researcher decided upon the most representative quotations to reflect the respective themes.

The FGD facilitator recorded and transcribed the student FGDs and medical lecturer IDIs. The principal researcher coded data for each participant in the student FGDs in sequence. The data of a randomly selected participant were coded first and only then were the data of another randomly selected participant coded. The hypothesis coding included 3 major dimensions (or themes): individual, environmental, and system. A total of 7 categories emerged in the students' and medical lecturers'

responses: (1) learner's characteristics, (2) instructor's characteristics, (3) extrinsic motivation, (4) interaction opportunities, (5) infrastructure and system quality, (6) course and information quality, and (7) institution and service quality.

Transcripts were not returned to participants for comments or corrections as it was not feasible in the given setting. However, the interview answers were validated directly after the IDIs as interview answers were run through question-by-question and answer-by-answer with each interviewee. Furthermore, IDI responses were validated with responses from the FGDs. All data were pseudonymized before data analysis. This study adhered to the consolidated criteria for reporting qualitative research checklist [22].

Ethical Considerations

The study protocol was approved by the Biomedical Research Ethics Committee of the University of Zambia and the ethical committee of the University Hospital Heidelberg, Germany. Before taking part in this study, all participating students and medical lecturers agreed to an informed consent that explained the scope and purpose of this study and the right to withdraw at any point. All participants gave written consent and were treated with respect.

Results

In-Depth Interview Characteristics

Out of a total of 81 MLP students and 23 medical lecturers, 8 MLP students (female n=4 and male n=4) and 5 medical lecturers (male n=5) had IDIs. IDIs lasted from 26 to 47 min, with an average of 33 min. Out of 10 invited participants per student FGD, the first FGD included 3 female and 6 male students from the third study year (duration: 78 min) and the second had 3 female and 5 male students from the fourth and bridging study years (duration: 88 min). Results are presented according to the respective coding themes of the hypothesis coding, and dimension themes are indicated in italics in the text that follows.

The interviewees are referred to in the text as follows: for IDIs, *interviewees* 1 to 8 are student participants and *lecturers* 1 to 5 are medical lecturers and for FGDs, *respondents* 1 to 9 are third study year students, for example, respondent 1 (third) and respondents 1 to 7 are fourth study year and bridging students, for example, respondent 1 (fourth/bridging).

Individual Dimension

Learners' Characteristics (Students)

The *attitude* of students *toward the tablet-based e-platform* was generally quite positive as interviewee 1 (third) expressed:

> *I think we were the lucky ones, it [tablet-based e-platform] came at the right time.*

Some students requested more *focus on interaction* by lecturers regarding the integration of tablets in their lectures, as respondent 6 (third) stated:

> *We would have lectures, then she [lecturer] would like to post whatever she would teach or any new guidelines. She would make sure that it is made available to us on the e-learning platform. So, if the other lecturers follow suit and try and take her example.*

Instructors' Characteristics (Medical Lecturers)

The *attitude* of medical lecturers *toward the tablet-based e-platform* was also quite positive, as lecturer 1 said:

> *I think it's a platform [that] ought to move forward, a platform that we ought to support.*

In general, the e-platform was well perceived; however, some lecturers recognized themselves as nonusers or low-users, such as lecturer 4:

> *Yes [my] general view is it is a very good program, but I think it is still meeting challenges whereby we [medical lecturers] have very few accesses, may be due to our time*

Lecturer 1, identified as a low-user who was only involved at one point in time, said:

> *Yes I have done it [uploading] two times. Well, because when I uploaded that material, I thought it was up-to-date, but I have not been visiting the site.*

Reasons for instructor low usage and participation were explained by lecturer 1:

> *First of all, the lecturers need to have tablets; number two, they need to have connectivity, internet connectivity; number three, there needs to be a platform where lecturers and students and other lecturers interact and discuss this.*

Lecturer 3 identified the need for:

> *...office space. If you are given an opportunity to go there and sit in the staff room or lecturers' room, then that willhelp you to have more preparations and access to the platform.*

Lecturer 2 attributed the low usage to the newness of the e-learning program and the need for face-to-face training:

> *I don't think they [medical lecturers] are using it [tablet-based e-platform] actively; some are, some are not. I think it is a developing program. So, I think it is still in the early phases...Lecture notes, there will always be need to explain in person. The videos I think will never be enough, so there will always be a need actually to explain and explain and explain in person. You must actually be there in the ward and do things on the patient...*

To increase the level of *interaction*, lecturer 3 suggested regular content review meetings and the introduction of e-based assignments:

> *...the only way we can improve on that is to have a regular review of the material that we have contributed...If assignments are scheduled at specific intervals, and both, the students and the lecturers are expected to either review or revise those assignments,*

> *probably that may stimulate the lecturers to use the platform much more regularly, including the students.*

Lecturer 1 recognized that a question-and-answer section would be beneficial:

> *If I am given an opportunity to interact with students, I will be answering their questions or their concerns. Then, I will log in specifically for those questions and try to answer them...it should be a two-way thing.*

Extrinsic Motivation

Many students *perceived* the tablet as *useful* for medical practice as a quick reference tool. They welcomed the *technological flexibility* for patient care, as respondent 7 (third) explained:

> *Basically, a tablet is, you will have everything. You have books, it is like everything is compressed into that, and it has been very useful, very useful with the patient you need to consult. There are treatment guidelines. If you are trying to calculate something, it is a very useful gadget.*

Students reported that the tablet provided them with greater study flexibility, as interviewee 5 described:

> *Wherever I am, I'm walking on the road, I'm on the bus, I can easily open - then do studies. Now with that[tablet], it has encouraged me to study.*

Interviewee 6 specified:

> *I get access where there's internet or not. It's [tablet] really contributing immensely.*

Students appreciated tablet *expandability* also, as respondent 8 (third) highlighted:

> *Actually these tablets they also allow students to put [upload] their own books, so it is very useful.*

A few students stated that the tablet *saved them money*, as interviewee 1 said:

> *I don't have to pay so much for books anymore. I don't have to always have internet bundles to access information when I need it.*

A few students did not find the tablet useful in *saving time*, as interviewee 7 explained:

> *There are a lot of things. The period that I came as a student, I wasn't treated as a student, I was treated as manpower. So, I didn't have much time, besides for an exam, which I was focusing on...My fear is that I might waste time. So, I just concentrate on what is important at present for the sake of exams because my period is short.*

The tablet was also perceived as a *restriction*, as interviewee 7 stated:

> *Sometimes I used to leave it [tablet] in the room, for security purposes. Especially if you lose the tablet, you'll pay. So, I was extra careful. So, it was a bit risky, but sometimes I would go with it.*

Environmental Dimension

Interaction Opportunities

Some students requested an addition to the e-platform to foster *interaction*, for example, respondent 9 (third) said:

> *What I would like to see is student participation. Could there be like a forum where students can add questions and lecturers will respond with teaching materials based on what the student is requesting, or at least good feedback?Learning,it is student-based.*

System Dimension

Infrastructure and System Quality

In general, opinions on the *ease of use* varied among individuals. However, the majority of students seemed to use the tablets with ease, as respondent 7 (fourth/bridging) exemplified:

> *I only had challenges just a few weeks after we were given the tablets. I don't have much challenge. I think certain challenges are being expressed maybe at the individual level.*

Some students perceived the tablets as an additional burden, as respondent 3 (fourth/bridging) stated:

> *I don't know why the tablets have to be made so complicated, such that we need training. Already we have got a lot of work, a lot of things to study. So, [it] is like we are taking another profession in IT, those things.*

To enhance the ease of use for students, lecturer 2 suggested keeping the e-platform materials clear and concise:

> *I think to make it [e-platform] more user-friendly is [to] remember that the students do not have a lot of time. So, like anybody else, you want to go to a material,which is easily accessible, easily understood without a lot of homework. Welive in an age where there is too much information elsewhere. If I just read one lecture on the tablet and I don't need to resort to a textbook, then I will go to that particular lecture. So, it's a question of how user-friendly are thematerials on the tablets.*

System functionality was perceived as challenging in regard to comprehensive learning materials, as interviewee 2 mentioned:

> *Because you'll find that, you scroll, scroll, scroll, scroll.*

To this end, interviewee 1 suggested:

> *most of the presentations will have to be converted to a tablet-acceptable version.*

In general, the tablet as a learning and health care practice support technology was found to be *adequate* within the given setting, especially for the hospital setting, as respondent 5 (third) said:

> *with the coming of the e-learning tablets it is something that is easy to carry whenever you are faced with a challenge when you are maybe with a patient.*

Many students identified the quality of the tablets as inadequate. The primary concern was the tablet battery, as respondent 2 (fourth/bridging) stated:

> *within a few hours my tablet drains. And sometimes even if you are not using it, it heats up and just after a few hours, the battery finishes.*

Some students perceived the tablet as too unreliable for medical practice, as interviewee 4 explained:

> *When you need it [tablet], if the patient is in front of you, you need to clarify something, it [e-platform tablet application] fails to open, you get stuck.*

Other concerns regarded repairs and spare parts, as a respondent (third) summarized:

> *When the battery finishes, we don't know where to get these batteries. So, next time when they are buying these things they should at least consider things [tablets] that we can easily find here locally.*

Respondent 4 (fourth/bridging) was concerned about:

> *what happens if three, four months from now the tablet may cease, is there anything that will happen or that's the end of everything?*

Internet quality was not perceived as crucial, as lecturer 3 stated:

> *It [e-platform] takes off the burden of the need for the students to look for internet access, especially in their remote areas where internet access is a challenge.*

However, access to the internet was still perceived as necessary beyond what was available as materials on the tablets, as respondent 8 (fourth/bridging) highlighted:

> *The e-learning platform worked with the loaded materials. But now out of the academic situation, we need to go and explore so that with every situation that we may encounter, we may be able to go online.*

Course and Information Quality

With regard to the subtopic of reliability, a majority of the students perceived the *reliability* of the available materials on the e-platform as inferior, as interviewee 7 expressed:

> *I was very disappointed with the tablets the time they gave us. Why I got so disappointed is some of the notes, which they had put in the tablet, they were old notes. Some of them as early or as late as 2005. Medicine is dynamic. It keeps on changing. So, this is 2017. If you put notes for 2005, 2006 it's not fair enough for the student.*

Interviewee 6 also highlighted the need for updates but was more content with the currently available learning materials:

> *Currently, I think it's enough. So, sometimes I feel maybe we are reading what is not latest and out of the bulletin medical journals.*

As a way to increase the quality and timeliness of the learning materials, lecturer 5 suggested the following:

> *have at least two [review meetings] in a quarter so that our contacts with the coordinators of the program and the Chainama [CCHS] staff also is increased, having more interaction and all viewpoints shared.*

Despite the identified shortcomings in reliability, many students found the information that was available on the e-platform to be generally *relevant* to their studies and medical practice. Interviewee 5 said that:

> *...especially when we are on the ward, we are doing our clinical work, it's easy to access information from the e-learning platform. And, most of the things that we were doing and what we are doing is actually there. It has made my life easier, in that I'm able to access the information I need without moving with the laptop, everywhere I go.*

Respondent 2 (fourth) revealed that he mainly engaged with the e-platform for exam preparations:

> *But then, of late just before the exams, that's when I could hear my friends say, they could tell me there is this there, that a lot of material that's when I got interested.*

Although the majority of students identified a lack of practical materials and requested additional e-learning materials, for example, respondent 9 (third) stated:

> *It lacks the practical information. So, I feel we need to add more lecture notes, which should come from the lecturers who are actually teaching us from here. They should put something that will help the student to know how to answer the examination questions. Some people are doing these OSCE exams for the first time.*

Respondent 1 (third) felt that she did not find relevant content:

> *...I just use it as a skeleton. It doesn't have enough meat that I need. So, we really need to put a lot of data in the Moodle [e-platform] for us to be able to use them effectively.*

Respondent 5 (fourth/bridging) remarked that it would be beneficial:

> *if I were [in a rural health facility], you find that certain equipment is not there, but I need to save a life, so I should also be taught alternatives when I am faced with such situations.*

Institution and Service Quality

To reach a *sustainablee-platform* at CCHS, lecturer 3 suggested involving the Ministry of Health and a fee-based system:

> *...to sell the idea [e-platform] with strong advocacy to the training directorate at the Ministry of Health. By doing so probably it gives Chainama [CCHS] as a college much more privilege when it comes to asking for extra funding. I am not sure in terms of how much students contribute towards the tablets,...just a small percentage to contribute towards the acquisition of replacement tablets.*

Speaking in regard to operationalization, lecturer 2 emphasized:

> *The ownership should sit in the college, especially in the department of medical licentiate training. That's where the ownership should be, that's the coordinating center, that's the connecting center for both those that teach and the students that pass through the program. Somebody should be identified in the department who can be the main person coordinating the e-learning platform because a lot of reviews, a lot of materials need to be developed over time.*

Training for the tablet-based e-platform was clearly stated as a necessity by the majority of students and medical lecturers as respondent 5 (third) mentioned:

> *I have had this tablet for some time. I don't know anything on how to access it because I would log in and try to click somewhere and it will ask for a password, and I don't even know how to press that. There is a need for training.*

More in-depth training from information technology (IT) support was also requested by respondent 3 (third):

> *I think at some point it would be right if we had people from the IT to come and just explain to us to say that okay when you see it behaving like this, the best way you are supposed to do is A-B-C. In that way, we are able to understand how to manage and operate the gadget nicely.*

Training formats were suggested as a 1-day class-based session with a concluding test or as individual training sessions with trained student experts for the e-platform, as interviewee 5 proposed:

> *Certain individuals might not have the computer literacy orexperience.But for those, I think, maybe it can be better and cost-effective if you can arrange with individuals. They can easily contact their colleagues who know how to do it. They can be taught, yes. Better to go to friends they're close to.*

A tutorial was also suggested by respondent 6 (fourth/bridging):

> *...maybe we come up with some form of a booklet where you put those steps that someone needs to follow, to open this and that, I think it would help.*

Medical lecturers suggested a training session incorporating a content revision, for example, lecturer 1 proposed:

> *I think a refresher. So that people are reminded of uploading materials.*

Lecturer 5 proposed the training:

> *as a review process also, to get feedback from them [medical lecturers] on how it [e-platform] is working and how best it can be made use of.*

The *service quality* provided by the CCHS IT staff was perceived as helpful and accessible as respondent 3 (third) stated:

> *I think the IT department has been very helpful...For me personally, every time that it [tablet] had malfunctioned, they try all their best to fix it.*

Students welcomed the available support, as respondent 7 (third) mentioned:

> *When you are using electronic gadgets, there are always few challenges. After upgrading it [tablet] stopped functioning, and we didn't know how to repair them. Until the people from the IT came.*

One concern regarded IT support after graduation, as respondent 9 (third) remarked:

> *Right now, when the tablet is not working, there is an internet problem the IT guys are available. Or at least we've got access channels. Probably in the next coming years, there will still be problems. So how do we access them [IT] once we are no longer here?*

Discussion

Principal Findings

Participants in FGDs and IDIs offered a range of insights into their perceptions of the tablet-based e-platform that comprised e-learning for medical education with an eHealth component. Overall, the perceptions were positive and varied toward the tablet-based e-platform as a tool for medical education and health care practice support at CCHS. The results of the evaluation proved useful for the further development of the e-platform as students' and instructors' perceptions of the needs and shortcomings of the tablets and the e-platform were identified.

A significant advantage was seen in the tablet-based offline component as it only required an internet connection for content updates. The tablet was adequate for the setting (especially for medical practice) as it fitted into the doctors' coat pockets and served as a handy reference. Quality weaknesses of the tablet were identified as fragility, fast-draining battery, and *buggy* operating system that frequently crashed or froze. The specific tablet model had been chosen as the best quality at an affordable price. Although it would be beneficial to use tablets available on the Zambian market, long-term tablet support will remain a challenge. Changes in technology and tablet manufacturer priorities require a continued investment to keep pace with changing technologies.

Some students believed they were missing curriculum content and tablet functionality. A few were overwhelmed by the tablet and perceived it as an extra burden to an already demanding study schedule. Comprehensive half-day training sessions at the beginning of the study year may enable productive usage of the tablet and the e-platform.

Both students and medical lecturers agreed that updated learning materials was one of the most significant needs, although the overall view on available learning materials of the e-platform was generally positive. The learning materials available at the time of the evaluation were gathered from various sources but were not optimized or designed as e-learning materials. Providing materials that follow multimedia principles could be beneficial as they were shown [23] to provide better learning outcomes and potentially improve the quality of the e-learning platform and its eHealth component. Incorporating teaching

methods that foster learning from surface to long-term memory, overlearning strategies such as practice tests, giving and receiving feedback, and spacing practice over time may all be beneficial [24]. Addressing student evaluations may increase self-evaluation and empower students to become more independent learners [24]. However, the limited number of lecturers poses a bottleneck. To this end, student engagement in creating content, so-called user-generated content, may be a potential solution [25]. Students could employ available mobile devices to generate content such as short videos on medical procedures. If, when preparing a topic for a short video, students read through the current medical guidelines and prepare the topic as an e-learning multimedia content, then they will also practice their pedagogical skills. Hence, user-generated content fosters medical learning and creates much-needed content for the e-platform.

Although lecturers were positive about the e-platform, they stated they were low-users or nonusers who rarely made use of the e-platform. Thus, their expressed attitude does not seem to be a reliable indicator of their actual involvement and usage. Moreover, their answers may have been biased by social desirability [26] and an expected openness to technologies. However, their overall positive answers may indicate a certain willingness and preparedness for more active, future involvement in the e-platform, and medical lecturer engagement is pivotal for updating and creating learning materials. As the numbers of medical lecturers are unlikely to substantially increase in the next few years [27,28], an e-platform coordinator could be key for the MLP program [29]. Regular lecturer review meetings could potentially foster content updating and the creation of new learning materials and also strengthen lecturer ownership of the content and pedagogy of learning materials [30]. Changes introduced with new learning technology should be constructively discussed, and barriers for implementation should be identified. Such operational structures may promote and support a sustainable e-platform.

Furthermore, we want to employ design thinking methods in the ongoing evaluation framework [25] to explore user habits and better understand the user and their challenges and increase adoption of digital technologies, especially by medical lecturers. As a next step for the MLP e-platform, e-learning modules are to be developed for all respective mandatory curriculum items. The construction of the initial e-learning modules is a substantial one-time effort mainly involving medical lecturers and IT. The quality and quantity of course modules on the e-platform and their respective content benefit from regular reviews and corresponding adjustments. Medical lecturers potentially may be more mindful of these e-platform tasks if these were included in the medical teaching schedule and teaching responsibilities, as well as if these topics were part of their continuous training and assessment. To empower medical lecturers to actively teach with the e-platform, we consider powerful training contents to comprise learning about engaging teaching strategies on how to best employ e-learning in a blended learning setting for medical education, as well as more technical skills like to work with the learning management system. Continuous costs for the MLP e-platform include training of involved staff, licensing costs of learning materials, authoring software for content and course development, as well as running and maintenance costs of IT infrastructure.

E-learning promotes a shift during which lecturers "become facilitators of learning and assessors of competency" instead of "distributors of content" [31]. This shift should be reflected in the curriculum [32]. To strengthen the e-learning infrastructure at CCHS, training for all involved need to be increased, which could be realized as weekly meetings with IT technicians to support a community of practice [33]. In the future, training for medical lecturers could be implemented as small learning units, so-called micro-learning sessions [34], which include short instructional videos on the e-platform and which can be more easily integrated into the clinical workday. Other methods to increase adoption may constitute so-called e-platform champions [35,36] who become points of reference as they are trained more intensely in e-learning and eHealth methodologies. Currently, the IT offices are separated from the administrative and medical lecturers' offices. Moving the IT staff physically closer to the MLP program administration's offices could increase their involvement in the educational processes of the e-platform and strengthen their role beyond simple IT support [33]. E-learning and eHealth can provide fertile learning environments that strengthen medical education and quality of care, but they need to be embedded in the curriculum with recognition of their strengths and shortcomings.

Limitations

The study has several limitations. First, as there are infrastructural restrictions on the total number of students per year, the study population based on the general MLP student population comprises only a small number of students. Second, we cannot rule out that the study participants' behavior would have differed if the researcher and discussion assistants had not been present. We tried to minimize this limitation by involving an outside person to conduct the FGDs and IDIs with the medical lecturers, in addition to the principal researcher who conducted the student IDIs. Third, it was not possible in the framework of this study to have study participants verify interview transcripts. We attempted to minimize this limitation by having the researchers review each question and respective respondent's answer at the end of each interview to verify the accuracy of the respondents' statements. Fourth, FGDs in this evaluation period could only be conducted with students, potentially diminishing the insight from the medical lecturers. Fifth, the study employs Bhuasiri's [16] framework, which potentially does not account for all involved dimensions and characteristics of a technology-enhanced intervention for medical education, and thus, further research to refine, adapt, and develop this framework accordingly may be required. Sixth, this study took place in the Zambian MLP program, so there are limitations on how far findings can be generalized to other clinical settings.

Conclusions

A self-directed, tablet-based e-platform that comprised e-learning with an eHealth component was introduced and proved feasible in the low-resource setting of Zambia. The qualitative evaluation included FGDs and IDIs that identified shortcomings, limiting the broad adoption of the e-platform

among students and medical lecturers. To improve on the shortcomings of the e-platform in this setting, comprehensive half-day training sessions at the beginning of the study year may increase the ease of use with provided tablets for students and medical lecturers. User-generated content, the nomination of an e-platform coordinator and an e-platform steering committee, as well as an e-platform embedded in the curriculum could improve the quality and quantity of learning content that has been perceived by students as outdated and insufficient. Furthermore, supporting regular meetings for medical lecturers to review and discuss the contents of the e-platform may increase engagement for technology-enhanced learning and teaching as well as use as a health care practice support tool. Institutional changes such as moving the IT department physically closer to the MLP program's medical lecturers and administration may foster vital exchanges that increase understanding and adoption of e-learning for medical education.

The implementation of digital learning environments such as e-learning and eHealth is a multidimensional process that ideally is cyclically iterated [25] to identify and understand the users' needs and actuators. A clear objective and a serious commitment are required within the implementing institution, and the effort of all involved is necessary to sustainably implement an e-platform as a beneficial learning and teaching method and a health care practice support tool for students and medical lecturers.

Acknowledgments

The authors wish to thank Else Kröner-Fresenius-Stiftung [2016_HA67] for supporting this study. The authors would also like to thank Agness Busiku for her support to this study. The authors would like to thank Julia M Challinor for her editorial support.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Table of Bhuasiri et al's research framework with 3 main dimensions and subdimensions adapted to the setting of the medical licentiate program.

[PDF File (Adobe PDF File), 41KB - mhealth_v7i1e12637_app1.pdf]

Multimedia Appendix 2

Interview guide: in-depth interviews.

[PDF File (Adobe PDF File), 57KB - mhealth_v7i1e12637_app2.pdf]

Multimedia Appendix 3

Interview guide: focus group discussions.

[PDF File (Adobe PDF File), 56KB - mhealth_v7i1e12637_app3.pdf]

References

1. World Health Organization (WHO). Geneva, Switzerland: World Health Organization; 2016. Global Strategy on Human Resources for Health: Workforce 2030 URL: https://www.who.int/hrh/resources/globstrathrh-2030/en/ [accessed 2019-01-09] [WebCite Cache ID 75IOnCC9b]
2. Walsh FJ, Musonda M, Mwila J, Prust ML, Vosburg KB, Fink G, et al. Improving allocation and management of the health workforce in Zambia. Health Aff (Millwood) 2017 Dec 1;36(5):931-937. [doi: 10.1377/hlthaff.2016.0679] [Medline: 28461362]
3. Tjoa A, Kapihya M, Libetwa M, Lee J, Pattinson C, McCarthy E, et al. Doubling the number of health graduates in Zambia: estimating feasibility and costs. Human Resources for Health 2010 Sep 22;8(1):22. [doi: 10.1186/1478-4491-8-22] [Medline: 20860827]
4. Ferrinho P, Siziya S, Goma F, Dussault G. The human resource for health situation in Zambia: deficit and maldistribution. Hum Resour Health 2011 Dec 19;9(1):30 [FREE Full text] [doi: 10.1186/1478-4491-9-30] [Medline: 22182366]
5. Gabrysch S, Cousens S, Cox J, Campbell OM. The influence of distance and level of care on delivery place in rural Zambia: a study of linked national data in a geographic information system. PLoS Med 2011 Jan 25;8(1):e1000394 [FREE Full text] [doi: 10.1371/journal.pmed.1000394] [Medline: 21283606]
6. World Health Organization. Increasing Access to Health Workers in Remote and Rural Areas Through Improved Retention. Global Policy Recommendations. Geneva, Switzerland: World Health Organization; 2010:79.

7. Ministry of Health, Republic of Zambia. Lusaka, Zambia: Ministry of Health, Republic of Zambia; 2016. Zambia National Health Strategic Plan 2017-2021 URL: http://dspace.unza.zm:8080/xmlui/handle/123456789/5113 [accessed 2019-01-08] [WebCite Cache ID 75HLKgRBk]
8. Ministry of Health, Republic of Zambia. Lusaka, Zambia: Ministry of Health, Republic of Zambia; 2018. National Human Resources for Health Strategic Plan 2018 - 2024 URL: https://tropmed25-hyg.urz.uni-heidelberg.de/seafile/f/b8b487cfb42b45779dc5/ [accessed 2019-01-08] [WebCite Cache ID 75HMFVR95]
9. Haggblom A, Bowa C, Thieme K, Graf U. Strengthening Zambia's health workforce: the Medical Licentiate training programme. Geneva, Switzerland: Geneva Health Forum; 2014 Presented at: Geneva Health Forum 2014; April 15-17, 2014; Geneva, Switzerland URL: http://ghf.g2hp.net/2013/09/30/solidarmed-support-to-zambias-medical-licentiate-training-programme/
10. Gajewski J, Mweemba C, Cheelo M, McCauley T, Kachimba J, Borgstein E, et al. Non-physician clinicians in rural Africa: lessons from the Medical Licentiate programme in Zambia. Hum Resour Health 2017 Aug 22;15(1):53 [FREE Full text] [doi: 10.1186/s12960-017-0233-0] [Medline: 28830528]
11. The Africa Working Group of the Joint Learning Initiative on Human Resources for Health and Development. World health organization (WHO). Geneva, Switzerland: The Africa Working Group of the Joint Learning Initiative; 2006 Sep. The health workforce in Africa: challenges and prospects The Africa Working Group of the Joint Learning initiative on Human Resources for Health and Development, report URL: https://www.who.int/workforcealliance/knowledge/resources/africawglearning/en/ [accessed 2019-01-09] [WebCite Cache ID 75IDy6GQV]
12. Lusale D, Cobb NM. Transformative Education for Health Professionals.: Transformative Education for Health Professionals; 2014 Nov 21. Zambian Associate Clinicians URL: https://whoeducationguidelines.org/content/zambian-associate-clinicians [accessed 2019-01-08] [WebCite Cache ID 75HNNzo1i]
13. Eysenbach G. What is e-health? J Med Internet Res 2001;3(2):E20 [FREE Full text] [doi: 10.2196/jmir.3.2.e20] [Medline: 11720962]
14. Barteit S, Jahn A, Bowa A, Lüders S, Malunga G, Marimo C, et al. How self-directed e-learning contributes to training for medical licentiate practitioners in Zambia: evaluation of the pilot phase of a mixed-methods study. JMIR Med Educ 2018 Nov 27;4(2):e10222 [FREE Full text] [doi: 10.2196/10222] [Medline: 30482744]
15. Neuendorf KA. The Content Analysis Guidebook. California, USA: Sage; 2017:456.
16. Bhuasiri W, Xaymoungkhoun O, Zo H, Rho JJ, Ciganek AP. Critical success factors for e-learning in developing countries: a comparative analysis between ICT experts and faculty. Comput Educ 2012 Feb;58(2):843-855. [doi: 10.1016/j.compedu.2011.10.010]
17. Artino Jr AR, La Rochelle JS, Dezee KJ, Gehlbach H. Developing questionnaires for educational research: AMEE Guide No. 87. Med Teach 2014 Jun;36(6):463-474 [FREE Full text] [doi: 10.3109/0142159X.2014.889814] [Medline: 24661014]
18. Cleary M, Horsfall J, Hayter M. Data collection and sampling in qualitative research: does size matter? J Adv Nurs 2014 Mar;70(3):473-475. [doi: 10.1111/jan.12163] [Medline: 24450874]
19. Guest G, Namey E, McKenna K. How many focus groups are enough? Building an evidence base for nonprobability sample sizes. Field Methods 2016 Apr 28;29(1):3-22. [doi: 10.1177/1525822X16639015]
20. Francis JJ, Johnston M, Robertson C, Glidewell L, Entwistle V, Eccles MP, et al. What is an adequate sample size? Operationalising data saturation for theory-based interview studies. Psychol Health 2010 Dec;25(10):1229-1245. [doi: 10.1080/08870440903194015] [Medline: 20204937]
21. Saldana J. The Coding Manual For Qualitative Researchers. London, UK: Sage; 2016:368.
22. 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]
23. Issa N, Schuller M, Santacaterina S, Shapiro M, Wang E, Mayer RE, et al. Applying multimedia design principles enhances learning in medical education. Med Educ 2011 Aug;45(8):818-826. [doi: 10.1111/j.1365-2923.2011.03988.x] [Medline: 21752078]
24. Hattie JA, Donoghue GM. Learning strategies: a synthesis and conceptual model. npj Sci Learn 2016 Aug 10;1(1):16013. [doi: 10.1038/npjscilearn.2016.13]
25. Lewrick M. Design Thinking: Radikale Innovationen in einer digitalisierten Welt [Design Thinking: Radical Innovations in a Digitized World]. Munich, Germany: C.H.Beck; Mar 9, 2018:143.
26. Krumpal I. Determinants of social desirability bias in sensitive surveys: a literature review. Qual Quant 2011 Nov 19;47(4):2025-2047 [FREE Full text] [doi: 10.1007/s11135-011-9640-9]
27. Ministry of Health, Republic of Zambia.: Ministry of Health, Republic of Zambia; 2017 Jul 7. Human Resources for Health Planning & Development Strategy Framework URL: http://www.moh.gov.zm/docs/NationalHRHPlanningAndDevelopmentStrategyFramework.pdf [accessed 2019-01-08] [WebCite Cache ID 75HOVIUi2]
28. Ferrinho P, Siziya S, Goma F, Dussault G. The human resource for health situation in Zambia: deficit and maldistribution. Hum Resour Health 2011 Dec 19;9:30 [FREE Full text] [doi: 10.1186/1478-4491-9-30] [Medline: 22182366]
29. Lin CC, Ma Z, Lin CP. Re-examining the critical success factors of e-learning from the EU perspective. Int J Manag Educ 2011 Dec;5(1):44-62. [doi: 10.1504/IJMIE.2011.037754]

30. World Health Organization (WHO). Geneva, Switzerland: World Health Organization; 2015. eLearning for undergraduate health professional education: a systematic review informing a radical transformation of health workforce development URL: https://www.who.int/hrh/documents/14126-eLearningReport.pdf [accessed 2019-01-09] [WebCite Cache ID 75IOMaUEO]
31. Ruiz JG, Mintzer MJ, Leipzig RM. The impact of E-learning in medical education. Acad Med 2006 Mar;81(3):207-212. [Medline: 16501260]
32. Frenk J, Chen L, Bhutta ZA, Cohen J, Crisp N, Evans T, et al. Health professionals for a new century: transforming education to strengthen health systems in an interdependent world. Lancet 2010 Dec 4;376(9756):1923-1958. [doi: 10.1016/S0140-6736(10)61854-5] [Medline: 21112623]
33. Langer AM. Information Technology and Organizational Learning: Managing Behavioral Change through Technology and Education. Florida, USA: CRC Press, Taylor & Francis Group; Nov 12, 2010:384.
34. Bruck PA, Motiwalla L, Foerster F. Mobile learning: A framework and evaluation. In: BLED 2012 Proceedings. 2012 Presented at: 25th Bled eConference; June 17-20, 2012; Bled, Slovenia p. 581-596 URL: https://aisel.aisnet.org/cgi/viewcontent.cgi?article=1041&context=bled2012 [doi: 10.1016/j.compedu.2005.10.011]
35. King E, Boyatt R. Exploring factors that influence adoption of e-learning within higher education. Br J Educ Technol 2015 Nov;46(6):1272-1280. [doi: 10.1111/bjet.12195]
36. Czerniewicz L, Brown C. A study of the relationship between institutional policy, organisational culture and e-learning use in four South African universities. Comput Educ 2009 Aug;53(1):121-131. [doi: 10.1016/j.compedu.2009.01.006]

Abbreviations

CCHS: Chainama College of Health Sciences
eHealth: electronic health
e-learning: technology-supported learning
e-platform: e-learning platform with an eHealth component
FGD: focus group discussion
IDIs: in-depth interviews
IT: information technology
MLP: medical licentiate practitioner

Edited by G Eysenbach; submitted 31.10.18; peer-reviewed by A Goffard, G Duftschmid; comments to author 21.11.18; revised version received 27.11.18; accepted 10.12.18; published 15.01.19.

Please cite as:
Barteit S, Neuhann F, Bärnighausen T, Bowa A, Lüders S, Malunga G, Chileshe G, Marimo C, Jahn A
Perspectives of Nonphysician Clinical Students and Medical Lecturers on Tablet-Based Health Care Practice Support for Medical Education in Zambia, Africa: Qualitative Study
JMIR Mhealth Uhealth 2019;7(1):e12637
URL: http://mhealth.jmir.org/2019/1/e12637/
doi:10.2196/12637
PMID:

Original Paper

Assessing the Validity of the MyJump2 App for Measuring Different Jumps in Professional Cerebral Palsy Football Players: An Experimental Study

Victor Coswig[1], PhD; Anselmo De Athayde Costa E Silva[1], PhD; Matheus Barbalho[2], PT; Fernando Rosch De Faria[3], MSc; Claudio D Nogueira[3], MSc; Mariane Borges[3], MSc; Jéssica R Buratti[3], BPhEd; Ivaldo B Vieira[4], MSc; Francisco Javier López Román[5], PhD, MD; José I Gorla[3], PhD

[1]Faculty of Physical Education, Federal University of Pará, Campus Castanhal, Castanhal, Brazil

[2]Faculty of Physical Education and Dance, Federal University of Goiás, Goiânia, Brazil

[3]Faculty of Physical Education, Department of Adapted Physical Education Studies, University of Campinas, Campinas, Brazil

[4]Doctoral Program in Health Sciences, Health Sciences, Universidad Católica San Antonio de Murcia, Múrcia, Spain

[5]Director de la Cátedra de Fisiología del Ejercicio, Health Sciences, Universidad Católica San Antonio de Murcia, Múrcia, Spain

Corresponding Author:
Victor Coswig, PhD
Faculty of Physical Education
Federal University of Pará
Campus Castanhal
Rua Major Wilson 3094
Casa 180 qd 8
Castanhal, 68742-89
Brazil
Phone: 55 91984273036
Email: vcoswig@gmail.com

Abstract

Background: Vertical jumps can be used to assess neuromuscular status in sports performance. This is particularly important in Cerebral Palsy Football (CP Football) because players are exposed to high injury risk, but it may be complicated because the gold standard for assessing jump performance is scarce in field evaluation. Thus, field techniques, such as mobile apps, have been proposed as an alternative method for solving this problem.

Objective: This study aims to evaluate the reliability of the measures of the MyJump2 app to assess vertical jump performance in professional CP Football.

Methods: We assessed 40 male CP Football athletes (age 28.1 [SD 1.4] years, weight 72.5 [SD 6.2] kg, and height 176 [SD 4.2] cm) through the countermovement jump (CMJ) and squat jump (SJ) using a contact mat. At the same time, we assessed the athletes using the MyJump2 app.

Results: There were no significant differences between the instruments in SJ height (P=.12) and flight time (P=.15). Additionally, there were no significant differences between the instruments for CMJ in jump height (P=.16) and flight time (P=.13). In addition, it was observed that there were significant and strong intraclass correlations in all SJ variables varying from 0.86 to 0.89 (both P<.001), which was classified as “almost perfect.” Similar results were observed in all variables from the CMJ, varying from 0.92 to 0.96 (both $P \leq .001$).

Conclusions: We conclude that the MyJump2 app presents high validity and reliability for measuring jump height and flight time of the SJ and CMJ in CP Football athletes.

(JMIR Mhealth Uhealth 2019;7(1):e11099) doi:10.2196/11099

KEYWORDS

cerebral palsy football; jump performance; mobile apps; mobile phone; paralympic sports

Introduction

Cerebral Palsy Football (CP Football) is a paralympic sport played exclusively by athletes with central neurological injuries, such as cerebral palsy (CP), traumatic brain injuries, or stroke [1]. CP Football players may have ataxia, hypertonia, or dystonia and are divided into classes based on their functional profile levels [1]. The concepts of the sport are quite similar to mainstream football, except for the use of 7 players, smaller field measurements, shorter duration of matches (30 minutes per time), and the lack of the offside rule [1]. Due the intermittent characteristic of the sport, and the fact that moments of high intensity are decisive in the game [2], CP Football has similar strength and power demands to mainstream soccer [3,4].

In addition, CP Football shows a relatively high injury rate. For example, injury rates for CP Football in the Rio 2016 Paralympic Games were 12.7/1000 athletes/day [5], in which it was noted that injuries by overuse had the second highest prevalence (4.5%) and higher injury rate during competition (34.5%). This is quite similar to the findings by Derman et al [6] who presented larger incidences of injury rates in CP Football (18.8/1000 athlete days) in comparison to all sports (12.1/1000 athlete days) in the London 2012 Paralympic Games.

There are many factors related to injury rates and sports performance in CP Football. From these, strength and power have been highlighted as relevant parameters for injury prevention [7,8] and training monitoring [9]. In this sense, it is important to establish reliable options for measuring and analyzing these variables, since vertical jumps are considered an ecological alternative provided in field assessments of lower-limb strength and power. Furthermore, increasing vertical jump assessment options could encourage and help coaches and trainers monitor and prescribe training based on this objective external load parameter, which has already been demonstrated to be a relevant factor in CP Football [5].

To accurately evaluate vertical jumps, expensive instruments with low portability are often needed, such as force platforms [10] and contact mats [11-12]. In this way, the demand for evaluation methods using mobile devices has increased. In addition to the technological advancement of smartphones, mobile apps have rapidly evolved from a trend [13] to a well-stablished part of sports and exercise medicine in constant and fast evolution [14]. Therefore, the MyJump2 app has become a viable option for evaluating jump heights. In addition, this app promises to provide quick and immediate data acquisition, allowing for easy monitoring in virtually any environment [15], and has already shown high reliability and reproducibility in vertical jumps when compared to force platforms and high-speed video cameras [16-18]. However, due to the specificities of CP athletes, such as asymmetries or involuntary spasms [6], it became necessary to test the validity and reliability of this mobile app in CP populations and sports. In this context, this study aimed to evaluate the reliability of the measures of the MyJump2 app to evaluate vertical jump performance in professional CP soccer athletes.

Methods

Experimental Design

To evaluate the reliability of the app, we recruited professional CP Football athletes during the 2017 Brazilian CP Football Championship. For this, the board of the National Association of Sport for the Disabled (NASD, Brazil) authorized the study. The directors of each team were clarified about the research proposal and methods and allowed the researchers to perform data collection with their athletes. First, we informed all subjects about the risks, benefits, and discomforts of participation by signing the consent form, and data collection occurred in one session. The study followed the ethical principles stated in the Declaration of Helsinki and was approved by the local ethics committee (#2.475.044). The athletes performed the countermovement jump (CMJ) and squat jump (SJ) and were assessed by 2 instruments simultaneously (a contact mat and MyJump2). Each participant performed 3 repetitions of each jump, and the highest height values were used in the analyses. The CMJ and SJ tests were performed in a circuit arrangement to guarantee similar rest time between the efforts.

Before data collection, subjects performed a specific warm-up for 5 minutes, which involved the execution of vertical jumps similar to those applied in the evaluations, in order to learn the how the jump would be executed and with stimulation of slow and fast cycles of stretching and shortening. After the specific warm-up period, participants were instructed to perform 3 CMJs with their hands fixed at the waist, performing the jump at the highest possible height [9]. All participants received verbal guidance to make the highest jump possible. There was no additional verbal stimulus to avoid differences between subjects.

Sample

A sample size calculation from an earlier investigation indicated that 7 participants would be needed, considering P=.001 and a power of 90% [16]. Thus, for this study, 40 male athletes (28.1 [SD 1.4] years, 72.5 [SD 6.2] kg, and 176 [SD 4.2] cm) without presenting acute or chronic conditions that prevented them from performing the jump protocol were included. From these, 30 were hemiplegic, 9 were diplegic, and 1 was monoplegic. Inclusion criteria required that participants have neurologic injuries at the central nervous system, be engaged in official professional CP Football competitions, and have prior experience in the vertical jumping exercise, which means that they performed plyometric exercises during training and were engaged in frequent jump tests during periodic evaluations. We excluded participants who did not complete the jumps for any reason, presented pain or injury, did vigorous physical activity, or had ingested central nervous system stimulants (ie, caffeine beverages) during the data collection phase (n=0).

Procedures

Countermovement Jump

In the CMJ, the individual starts in an orthostatic position with the hands fixed at the waist and, at the evaluators' command, performs a squat until the knees reach an angle of 90°. Then, the participant extends the hips and the knees to project the body vertically with the greatest speed and strength possible to

reach the maximum possible height [19]. Participants were instructed to not flex the knee or dorsiflex the ankle during the flight phase. All participants received verbal stimuli for a better performance. A 30-second rest interval between each jump was set.

Squat Jump

In the SJ, the individual starts in an orthostatic position and, at the evaluator´s command, performs a squat until the knees reach a 90° angle. This position is maintained in isometric contraction for 3 seconds, after which the individual extends the hips and knees to project the body and the load vertically at the highest possible speed and strength, that is, to achieve maximum power during the execution [19].

Contact Mat

The Jump System Pro Contact Mat (Cefise) evaluates the power output through flight time. Output data were collected by the Jump System Pro Software, version 1.0 (Cefise). The contact map showed high reliability for jump height, with an intraclass correlation coefficient (ICC) of 0.91 and a coefficient of variation of 10% [20].

MyJump2 Application

The app for the iOS operating system (Apple Inc) [17] was developed using the XCode0.5 software for Mac (OSX 10.9.2, Apple Inc) and installed on the iPhone 6s (Apple Inc). The evaluation required a high-speed camera (120 Hz) with a minimum resolution of 720p. The app analyzed the height of the vertical jumps by calculating the time between 2 frames (in ms) selected by the evaluator, corresponding to the loss and return of contact to the ground. For this instrument, the same evaluator performed all the collections and was always in the same position (at the front) and at the same distance (1.5 m) from the material being evaluated. For interevaluator and intraevaluator reliability, recorded videos were analyzed by 2 evaluators (inter) and one of them repeated the procedure after 1 week (intra).

Statistical Analysis

Data are presented in mean (SD). The Shapiro-Wilk test was used to assess normality, and the 2-tailed paired *t* test was used to compare instruments. For the reproducibility of the test measurements, the ICC, SE of measurement, that is, SEM=SD × √(1−ICC), and minimal detectable change, that is, MDC=SEM × 1.96 × (√2), were applied [21]. The Pearson correlation coefficient was applied for the correlations, and Bland-Altman plots were applied to test the level of agreement between instruments. All analyses were performed using SPSS software (IBM SPSS Statistics, version 20.0). For all variables, statistical significance was set at $P \leq .05$.

Results

The mean and SD of jump height and flight time of squat jump and countermovement jump were assessed using MyJump2 and a contact mat in 40 male CP Football athletes. Table 1 shows the values of the absolute comparison and ICC between the instruments for the jump height and flight time of the SJ and CMJ. There were no significant differences between the instruments in the jump height (P=.12) and flight time (P=.15) variables. The effect size of the 2 variables was trivial for jump height and flight time, according to the Cohen classification [22]. There were no significant differences between the instruments for the CMJ in the jump height (P=.16) and flight time (P=.13) variables. The effect size of the 2 variables was also trivial for jump height. In addition, it was observed that there were significant intraclass correlations in all SJ variables (P<.001). Strong correlations were found between jump height and flight time (ICC=0.89 and 0.86, respectively), being classified as “almost perfect” [22]. Similar significant intraclass correlations were observed in all variables from the CMJ ($P \leq .05$), where jump height and flight time presented excellent levels varying from 0.92 to 0.96, classified as “near perfect” [22].

Table 2 presents reliability data between evaluators, while Table 3 presents values between test and retest. For both, high values of intraclass correlation and very low values of standard error of measurements and minimal detectable changes were found. Bland-Altman and correlation analysis are presented in Figures 1 and 2 for SJ and in Figures 3 and 4 for CMJ, respectively. For both jumps, high levels of agreement were found, and the differences were similar for all ranges of heights.

Table 1. Values of absolute comparison and intraclass correlation coefficient between the instruments for the jump height and flight time of squat jump and countermovement jump.

Jump type	MyJump2, mean (SD)	Contact mat, mean (SD)	*t (df)*	*P* value	Cohen *d* effect size	ICC[a]	P_{ICC}[b]
Squat jump							
Jump height (cm)	25.1 (7.4)	26.2 (6.2)	−1.59 (35)	.12	0.17	0.89	<.001
Flight time (ms)	448.5 (70.0)	458.67 (55.6)	−1.47 (35)	.15	0.17	0.86	<.001
Countermovement jump							
Jump height (cm)	28.4 (6.5)	27.8 (6.1)	1.42 (33)	.16	0.09	0.92	<.001
Flight time (ms)	477.7 (56.1)	473.2 (52.5)	1.27 (33)	.21	0.02	0.96	<.001

[a]ICC: intraclass correlation coefficient.

[b]P_{ICC}: Level of significance of intraclass correlation coefficient, which was set at $P \leq .05$.

Table 2. Interevaluator intraclass correlation coefficient of squat jump and countermovement jump measurements in the MyJump2TM App.

Jump type	Evaluator 1, mean (SD)	Evaluator 2, mean (SD)	ICC[a]	*P* value[b]	Standard error of measurement	Minimal detectable change (%)
Squat jump						
Jump height (cm)	25.5 (7.0)	23.3 (7.1)	0.93	<.001	0.56	0.79 (3.09)
Flight time (ms)	452.5 (66.0)	442.52 (83.5)	0.90	<.001	7.45	10.53 (2.33)
Countermovement jump						
Jump height (cm)	28.4 (6.7)	27.1 (7.8)	0.95	<.001	0.36	0.51 (1.81)
Flight time (ms)	477.7 (57.5)	468.7 (71.3)	0.92	<.001	5.12	7.24 (1.51)

[a]ICC: intraclass correlation coefficient.

[b]Level of significance set at $P \leq .05$.

Table 3. Intraevaluator intraclass correlation coefficient of the MyJump2 in squat jump and countermovement jump.

Jump type	1st analysis	2nd analysis	ICC[a]a	*P* value[b]	Standard error of measurement	Minimal detectable change (%)
Squat jump						
Jump Height (cm)	25.5 (7.0)	24.8 (7.4)	0.99	<.001	0.07	0.10 (0.39)
Flight Time (ms)	452.48 (66.02)	447.8 (70.2)	0.95	<.001	4.08	5.77 (1.27)
Countermovement jump						
Jump Height (cm)	28.4 (6.7)	28.1 (6.6)	0.99	<.001	0.04	0.06 (0.20)
Flight Time (ms)	477.7 (57.5)	477.4 (56.1)	0.99	<.001	0.06	0.08 (0.01)

[a]ICC: intraclass correlation coefficient.

[b]Level of significance set at $P \leq .05$.

Figure 1. Bland-Altman for agreement analysis of squat jumps.

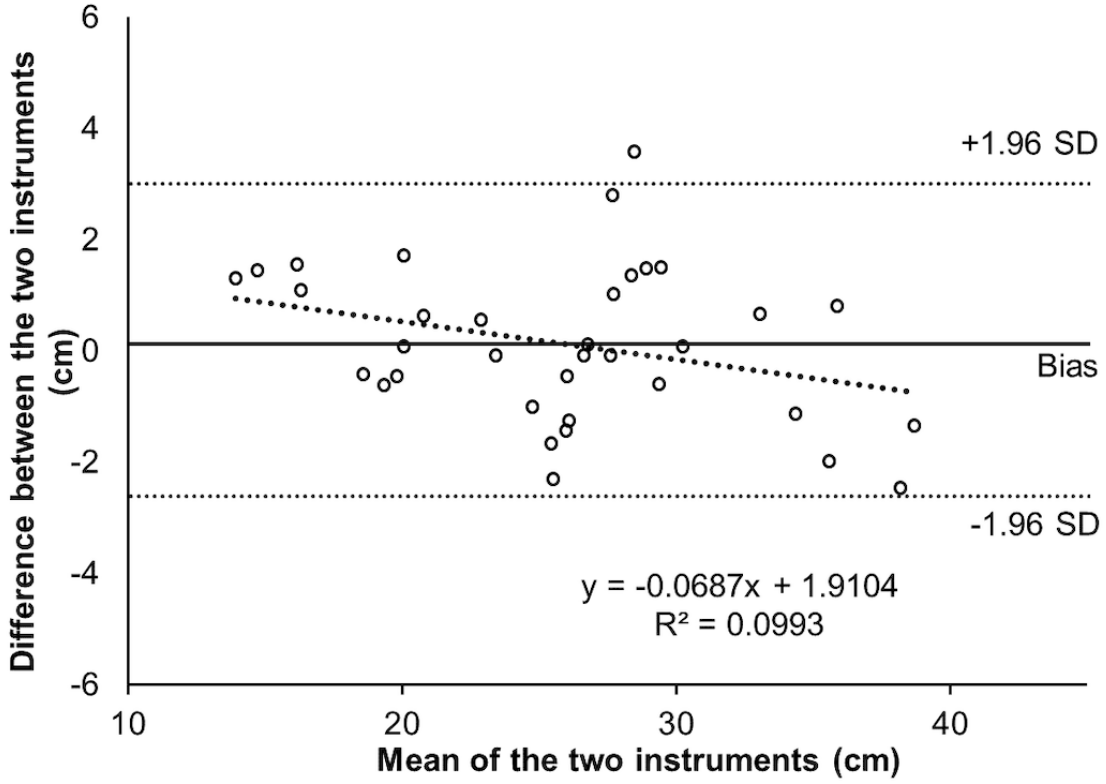


Figure 2. Correlation for agreement analysis of squat jumps.

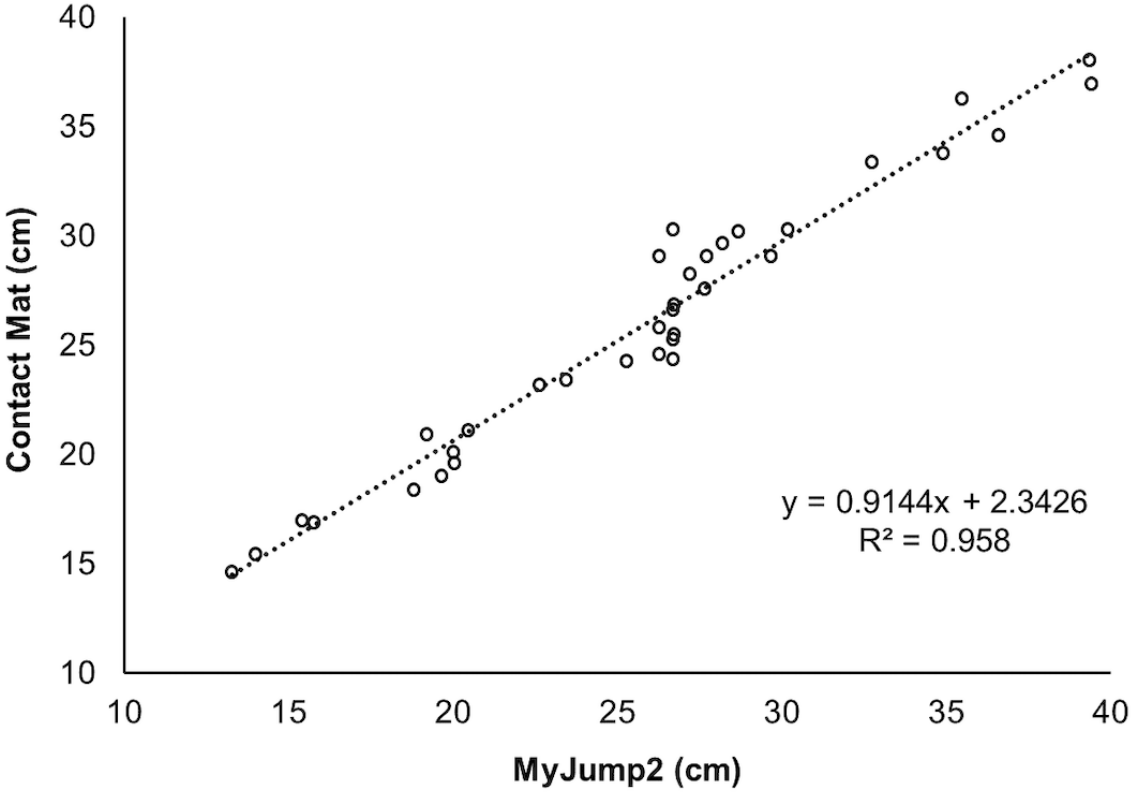


Figure 3. Bland-Altman for agreement analysis of countermovement jumps.

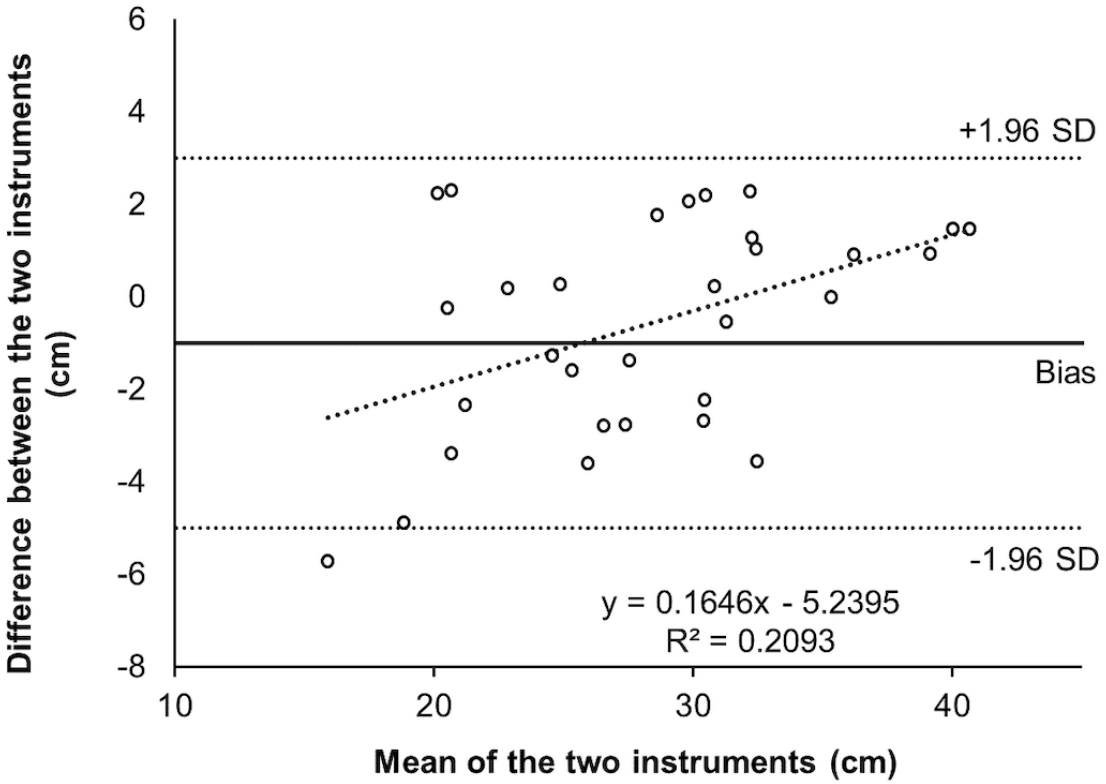
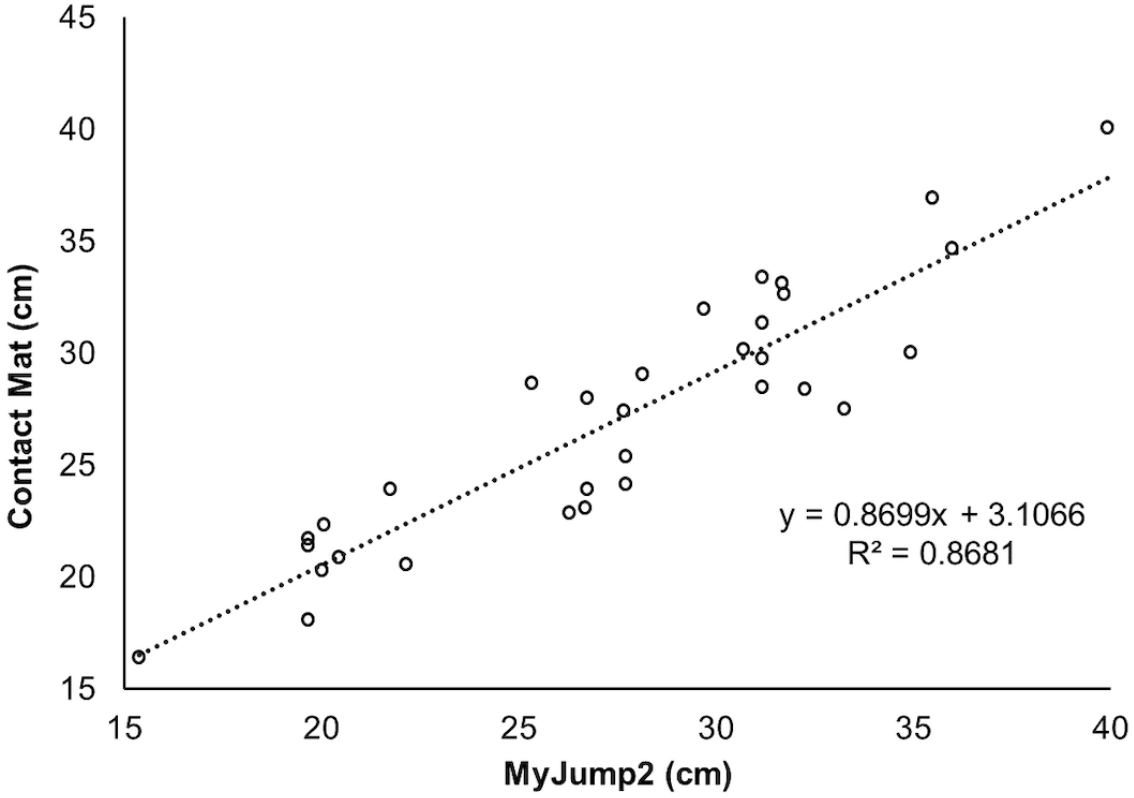


Figure 4. Correlation for agreement analysis of countermovement jumps.



Discussion

Principal Findings

The main objective of this study was to assess the reliability of a mobile app that measures jumping performance in CP Football athletes to establish a reliable field evaluation possibility for jump analysis. In this sense, we assessed 40 male, physically disabled players using the MyJump2 app and a contact mat simultaneously. Additionally, 2 evaluators made the measurements with the app, and one of them repeated the measurement 1 week later to analyze the interevaluator and intraevaluator variability. Our main results showed that MyJump2 is a reliable and valid method compared with the contact mat when assessing jump height and flight time.

Comparison With Prior Work

MyJump2 seems to be reliable for assessing jump height and flying time in CP Football. Our results are in agreement with 2 other studies that investigated MyJump reproducibility in vertical jumps [16,17]. Balsalobre-Fernandez et al [17] evaluated the app's validity, compared with a force platform, to assess CMJ in 20 recreationally active healthy men and observed a near perfect correlation in jump height. Thus, the authors indicated that CMJ could be easily measured and was reliable and reproducible through the app.

The other study on the theme [16] had analyzed different jumps (drop jump, SJ, and CMJ) in a sample of 21 male and female athletes, and the authors compared the app with a contact platform and a high-speed video camera method. In all jumps, there was a strong and significant correlation between the instruments. Similarly, our results showed that the CMJ has a strong correlation with the reference method in jump height (r=0.95).

Other relevant results are about the interevaluator and intraevaluator reliability. Our results show that the MyJump2 is reproducible when used by different subjects and at different occasions. These results are in agreement with literature about MyJump2 reproducibility [17].

Additionally, in competitive periods when it is not possible to use the gold standard equipment, the app seems to be a good alternative for evaluating athletes with neurologic damage who would benefit from frequent monitoring for training loads and adaptations [23-24] and also for soft-tissue lesion risk [25,26]. Given the portability and practicality of MyJump2, smartphones can quickly become a standard method for assessing physical performance in the field with great precision in CP Football.

The comparison between MyJump2 and other methods, such as force platforms [14-16] and the field method (Vertec) [27], is important to consolidate the app. This is justified by the use of a few force platforms in field evaluations, which raises the importance of the comparison between MyJump2 and other field methods [27]. In this regard, our study compares the app with one of the most used field techniques, the contact mat. Compared with force platforms, contact mats can be used in a wide variety of scenarios, and this can be considered a more ecologic option in agreement with the study that analyzed a comparison between MyJump2 and Vertec [27]. Another factor that deserves further comment is the type of jumps assessed here. The CMJ and SJ are consistent with the current literature [9,16,18,28,29].

Another important aspect of this study is that it appears to be the first work to use the app in a paralympic sport. In general, paralympic sports has a particular aspect beyond those related to training: the functional classification. To reduce subjectivity, it is necessary that the functional classification be evidence based [30]. In this sense, physical assessments are lacking and should be increased, including that for strength and power [31]. Thus, the implementation of mobile apps seems to be interesting in the evidence-based classification of paralympic sports. Another interesting characteristic of this paper concerns the sample. In the paralympic context, it is particularly challenging to do investigations with a large number of subjects, especially at higher competitive levels. Therefore, this study used a significant paralympic sample from a top national ranking team, which reinforces the originality and relevance of our findings.

Limitations

The study has some potential limitations. Due to the fact that data collection was conducted during a Brazilian CP Football Championship, it was not possible to use a better reference method, such as the force platform. Despite that, it could be considered a methodological choice in order to raise the ecological validity of our findings. With our study design, it was not possible to assess the use of the app in other conditions, such as before and after the games, to analyze the applicability of the app, which we suggest for further studies.

Future Directions

Since the MyJump2 app is a reliable method for assessing jump height in CP Football, future investigations may be designed in some topics differing from validation studies. With the limitations of laboratory methods, there is a trend toward using small samples in investigations, and less field assessments are observed. Therefore, one possibility will be to assess a wide sample to establish reference values of jump height and flight time in CP Football, using as large a sample as possible. Other possibilities may be to use prospective assessments of jump performance and its association with injury prevention parameters in this population, as well as assessments of performance levels and pregame, postgame, and competition recoveries. Regarding validation studies, it seems to be an important possibility for the validation of MyJump2 to assess other types of jumps besides CMJ and SJ. For example, the asymmetry jump test, drop jump test, and horizontal jump are three kinds of skills that can be used to assess jump performance and, to this date, have not been validated in CP Football evaluation.

Conclusions

Thus, we conclude that the MyJump2 app presents high validity and reliability to measure the jump height and flight time of the SJ and CMJ in elite CP Football athletes. Our findings suggest that this tool can be very useful in jump performance analysis of the Paralympics. In addition, we believe that our findings could encourage trainers, coaches, and athletes to monitor jump performance, which is relevant information to improve decision making in training control and prescription.

Acknowledgments

VC, ADACES, and JIG established the research design, conducted data analysis, discussed the results, and wrote the final version of the manuscript. M Barbalho, M Borges, FRDF, CDN, JRB, IBV, and FJLR took part in the data collection and writing of the manuscript. This research was sponsored by National Association of Sport for Disable (NASD, Brazil). Additionally, we thank the Pro-Rectory for Research and Post-Graduation of Federal University of Pará for financial support.

Conflicts of Interest

None declared.

References

1. International Federation of CP Football. IFCPF Sport Rules and Regulations. South Africa: IFCPF; 2018 Jan 01. CP Football: Laws of the Game, Classification and rules URL: https://www.ifcpf.com/rules [accessed 2018-11-27] [WebCite Cache ID 74FXR0hHJ]
2. Stølen T, Chamari K, Castagna C, Wisløff U. Physiology of soccer: an update. Sports Med 2005;35(6):501-536. [Medline: 15974635]
3. Reilly T, Williams AM, Nevill A, Franks A. A multidisciplinary approach to talent identification in soccer. J Sports Sci 2000 Sep;18(9):695-702. [doi: 10.1080/02640410050120078] [Medline: 11043895]
4. Paul DJ, Nassis GP. Testing strength and power in soccer players: the application of conventional and traditional methods of assessment. J Strength Cond Res 2015 Jun;29(6):1748-1758. [doi: 10.1519/JSC.0000000000000807] [Medline: 25546446]
5. Willick SE, Webborn N, Emery C, Blauwet CA, Pit-Grosheide P, Stomphorst J, et al. The epidemiology of injuries at the London 2012 Paralympic Games. Br J Sports Med 2013 May;47(7):426-432. [doi: 10.1136/bjsports-2013-092374] [Medline: 23515713]
6. Derman W, Runciman P, Schwellnus M, Jordaan E, Blauwet C, Webborn N, et al. High precompetition injury rate dominates the injury profile at the Rio 2016 Summer Paralympic Games: a prospective cohort study of 51 198 athlete days. Br J Sports Med 2018 Jan;52(1):24-31. [doi: 10.1136/bjsports-2017-098039] [Medline: 29030389]

7. Hoff J. Training and testing physical capacities for elite soccer players. J Sports Sci 2005 Jun;23(6):573-582. [doi: 10.1080/02640410400021252] [Medline: 16195006]
8. DiStefano LJ, Padua DA, Blackburn JT, Garrett WE, Guskiewicz KM, Marshall SW. Integrated injury prevention program improves balance and vertical jump height in children. J Strength Cond Res 2010 Feb;24(2):332-342. [doi: 10.1519/JSC.0b013e3181cc2225] [Medline: 20072067]
9. Claudino JG, Cronin J, Mezêncio B, McMaster DT, McGuigan M, Tricoli V, et al. The countermovement jump to monitor neuromuscular status: A meta-analysis. J Sci Med Sport 2017 Apr;20(4):397-402. [doi: 10.1016/j.jsams.2016.08.011] [Medline: 27663764]
10. Castagna C, Ganzetti M, Ditroilo M, Giovannelli M, Rocchetti A, Manzi V. Concurrent validity of vertical jump performance assessment systems. J Strength Cond Res 2013 Mar;27(3):761-768. [doi: 10.1519/JSC.0b013e31825dbcc5] [Medline: 22648140]
11. Sattler T, Sekulic D, Esco MR, Mahmutovic I, Hadzic V. Analysis of the association between isokinetic knee strength with offensive and defensive jumping capacity in high-level female volleyball athletes. J Sci Med Sport 2015 Sep;18(5):613-618. [doi: 10.1016/j.jsams.2014.08.002] [Medline: 25218817]
12. Váczi M, Tollár J, Meszler B, Juhász I, Karsai I. Short-term high intensity plyometric training program improves strength, power and agility in male soccer players. J Hum Kinet 2013 Mar;36:17-26 [FREE Full text] [doi: 10.2478/hukin-2013-0002] [Medline: 23717351]
13. Thompson W. Worldwide survey of fitness trends for 2017. ACSM'S Health & Fitness Journal 2016;20(6):8-17. [doi: 10.1249/FIT.0000000000000252]
14. Ahmed OH, Lee H, Marchant H, Jones R, Hall EE. Navigating the new landscape of apps: Overcoming the challenge of poor quality apps in sport and exercise medicine. Br J Sports Med 2018 Oct;52(20):1283-1284. [doi: 10.1136/bjsports-2017-097911] [Medline: 29146753]
15. Carlos-Vivas J, Martin-Martinez JP, Hernandez-Mocholi MA, Perez-Gomez J. Validation of the iPhone app using the force platform to estimate vertical jump height. J Sports Med Phys Fitness 2018 Mar;58(3):227-232. [doi: 10.23736/S0022-4707.16.06664-0] [Medline: 27653154]
16. Gallardo-Fuentes F, Gallardo-Fuentes J, Ramírez-Campillo R, Balsalobre-Fernández C, Martínez C, Caniuqueo A, et al. Intersession and Intrasession Reliability and Validity of the My Jump App for Measuring Different Jump Actions in Trained Male and Female Athletes. J Strength Cond Res 2016 Jul;30(7):2049-2056. [doi: 10.1519/JSC.0000000000001304] [Medline: 27328276]
17. Balsalobre-Fernández C, Glaister M, Lockey RA. The validity and reliability of an iPhone app for measuring vertical jump performance. J Sports Sci 2015;33(15):1574-1579. [doi: 10.1080/02640414.2014.996184] [Medline: 25555023]
18. Driller M, Tavares F, McMaster D, O'Donnell S. Assessing a smartphone application to measure counter-movement jumps in recreational athletes. Int J Sports Sci Coach 2017 Aug 25;12(5):661-664. [doi: 10.1177/1747954117727846]
19. Markovic G, Dizdar D, Jukic I, Cardinale M. Reliability and factorial validity of squat and countermovement jump tests. J Strength Cond Res 2004 Aug;18(3):551-555. [doi: 10.1519/1533-4287(2004)18<551:RAFVOS>2.0.CO;2] [Medline: 15320660]
20. Nogueira FDA, de Freitas V, Nogueira R, Miloski B, Werneck F, Bara-Filho M. Improvement of physical performance, hormonal profile, recovery-stress balance and increase of muscle damage in a specific futsal pre-season planning. Revista Andaluza de Medicina del Deporte 2018 Apr;11(2):63-68. [doi: 10.1016/j.ramd.2015.11.008]
21. Weir JP. Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res 2005 Feb;19(1):231-240. [doi: 10.1519/15184.1] [Medline: 15705040]
22. Cohen J. Statistical Power Analysis for the Behavioral Sciences. New York: Taylor & Francis, Inc; 1988.
23. Carroll KL, Leiser J, Paisley TS. Cerebral palsy: physical activity and sport. Curr Sports Med Rep 2006 Dec;5(6):319-322. [Medline: 17067500]
24. Damiano DL, Abel MF. Functional outcomes of strength training in spastic cerebral palsy. Arch Phys Med Rehabil 1998 Feb;79(2):119-125. [Medline: 9473991]
25. Patatoukas D, Farmakides A, Aggeli V, Fotaki S, Tsibidakis H, Mavrogenis AF, et al. Disability-related injuries in athletes with disabilities. Folia Med (Plovdiv) 2011;53(1):40-46. [Medline: 21644404]
26. Nyland J, Snouse SL, Anderson M, Kelly T, Sterling JC. Soft tissue injuries to USA paralympians at the 1996 summer games. Arch Phys Med Rehabil 2000 Mar;81(3):368-373. [Medline: 10724085]
27. Yingling VR, Castro DA, Duong JT, Malpartida FJ, Usher JR. The reliability of vertical jump tests between the Vertec and phone application. PeerJ 2018;6:e4669 [FREE Full text] [doi: 10.7717/peerj.4669] [Medline: 29692955]
28. Menzel H, Chagas MH, Szmuchrowski LA, Araujo SRS, de AAGP, de JFR. Analysis of lower limb asymmetries by isokinetic and vertical jump tests in soccer players. J Strength Cond Res 2013 May;27(5):1370-1377. [doi: 10.1519/JSC.0b013e318265a3c8] [Medline: 22796999]
29. Eagles AN, Sayers MGL, Bousson M, Lovell DI. Current Methodologies and Implications of Phase Identification of the Vertical Jump: A Systematic Review and Meta-analysis. Sports Med 2015 Sep;45(9):1311-1323. [doi: 10.1007/s40279-015-0350-7] [Medline: 26121950]

30. International Paralympic Committee. Bonn; 2016 Oct 27. International Standards for 2015 Athlete Classification Code published URL: https://www.paralympic.org/news/international-standards-2015-athlete-classification-code-published [accessed 2018-11-27] [WebCite Cache ID 74FY85vKh]
31. Beckman EM, Connick MJ, Tweedy SM. Assessing muscle strength for the purpose of classification in Paralympic sport: A review and recommendations. J Sci Med Sport 2017 Apr;20(4):391-396. [doi: 10.1016/j.jsams.2016.08.010] [Medline: 27692576]

Abbreviations

CMJ: countermovement jump
CP Football: Cerebral Palsy Football
CP: cerebral palsy
ICC: intraclass correlation coefficient
MDC: minimal detectable change
SEM: standard error of measurement
SJ: squat jump

Edited by G Eysenbach; submitted 28.05.18; peer-reviewed by R Raiol, DO Ahmed; comments to author 12.07.18; revised version received 26.07.18; accepted 26.08.18; published 30.01.19.

Please cite as:
Coswig V, Silva ADACE, Barbalho M, Faria FRD, Nogueira CD, Borges M, Buratti JR, Vieira IB, Román FJL, Gorla JI
Assessing the Validity of the MyJump2 App for Measuring Different Jumps in Professional Cerebral Palsy Football Players: An Experimental Study
JMIR Mhealth Uhealth 2019;7(1):e11099
URL: http://mhealth.jmir.org/2019/1/e11099/
doi:10.2196/11099
PMID:30698529

Original Paper

Feasibility and Patient Experience of a Home-Based Rehabilitation Program Driven by a Tablet App and Mobility Monitoring for Patients After a Total Hip Arthroplasty

Jildou Hoogland[1], MSc; Annet Wijnen[1], MSc; Tjerk Munsterman[2], MSc; Carina LE Gerritsma[3], MD, PhD; Baukje Dijkstra[4], MSc; Wierd P Zijlstra[4], MD, PhD; Janneke Annegarn[5], PhD; Francisco Ibarra[6], PhD; Wiebren Zijlstra[7], PhD; Martin Stevens[1], PhD

[1]Department of Orthopedics, University of Groningen, University Medical Center Groningen, Groningen, Netherlands

[2]Department of Physiotherapy, Martini Hospital Groningen, Groningen, Netherlands

[3]Department of Orthopedics, Martini Hospital Groningen, Groningen, Netherlands

[4]Department of Orthopedics, Medical Center Leeuwarden, Leeuwarden, Netherlands

[5]Department of Personal Health, Philips Research, Eindhoven, Netherlands

[6]University of Trento, Trento, Italy

[7]Institute of Movement and Sport Gerontology, German Sport University Cologne, Cologne, Germany

Corresponding Author:
Martin Stevens, PhD
Department of Orthopedics
University of Groningen
University Medical Center Groningen
Hanzeplein 1
Groningen, 9713 GZ
Netherlands
Phone: 31 50 3613271
Fax: 31 50 3614588
Email: m.stevens@umcg.nl

Abstract

Background: Recent developments in technology are promising for providing home-based exercise programs.

Objective: The objective of this study was to evaluate the feasibility and patient experience of a home-based rehabilitation program after total hip arthroplasty (THA) delivered using videos on a tablet personal computer (PC) and a necklace-worn motion sensor to continuously monitor mobility-related activities.

Methods: We enrolled 30 independently living patients aged 18-75 years who had undergone THA as a treatment for primary or secondary osteoarthritis (OA) between December 2015 and February 2017. Patients followed a 12-week exercise program with video instructions on a tablet PC and daily physical activity registration through a motion sensor. Patients were asked to do strengthening and walking exercises at least 5 days a week. There was weekly phone contact with a physiotherapist. Adherence and technical problems were recorded during the intervention. User evaluation was done in week 4 (T1) and at the end of the program (T2).

Results: Overall, 26 patients completed the program. Average adherence for exercising 5 times a week was 92%. Reasons mentioned most often for nonadherence were vacation or a day or weekend off 25% (33/134) and work 15% (20/134). The total number of technical issues was 8. The average score on the user evaluation questionnaire (range 0-5) was 4.6 at T1 and 4.5 at T2. The highest score was for the subscale “coaching” and the lowest for the subscale “sensor.”

Conclusions: A home-based rehabilitation program driven by a tablet app and mobility monitoring seems feasible for THA patients. Adherence was good and patient experience was positive. The novel technology was well accepted. When the home-based rehabilitation program proves to be effective, it could be used as an alternative to formal physiotherapy. However, further research on its effectiveness is needed.

(JMIR Mhealth Uhealth 2019;7(1):e10342) doi:10.2196/10342

KEYWORDS

home-based rehabilitation; mobile phone; osteoarthritis; physiotherapy; total hip arthroplasty

Introduction

Surgical treatment by means of total hip arthroplasty (THA) is most often indicated in end-stage hip osteoarthritis (OA). At present, THA is considered one of the most successful, effective, and cost-effective surgical treatments available. Hence, a total of 29,937 primary THAs were performed in the Netherlands in 2017 [1]. As in other Western countries, there is an increasing tendency in the Netherlands to perform fast-track surgery, after which people leave the hospital within a few days. The downside, however, is a risk of patients being minimally supported in their rehabilitation process during hospital admission and after discharge. At present, postoperative physiotherapy is not always covered by the basic health insurance in the Netherlands [2]. This can ultimately lead to suboptimal recovery [3]. Bandholm and Kehlet emphasized the urge for immediate and intensive postoperative physiotherapy [3]. In addition, Austin et al showed that this physiotherapy need not take place in a formal setting; a home-based program could work as well [4]. Furthermore, Austin et al showed that a home-based rehabilitation program seems to be both safe and efficacious for a majority of patients undergoing THA [4].

In this context, it is important to look at how technical innovations can be supportive of such home-based programs. Recent technological developments such as wearable sensors and tablet use with mobile internet are promising for providing home-based programs [5]. The use of objective activity monitoring with wearable sensors can potentially be helpful in strategies aimed at increasing adherence to home-based rehabilitation programs and daily activity [6]. Furthermore, a home-based program can improve adherence, which is often influenced by aspects such as lack of motivation, the effort and costs of traveling, and a preference for the privacy of the home environment [7].

The use of computers and tablets is rising among older adults in the Netherlands [6]. The ownership of tablets among seniors aged 65-75 years increased from 28% in 2012 to 60% in 2016 [8]. Although home-based rehabilitation programs may be of great importance, research is needed to optimize the programs that are supported by technology. This study, therefore, aimed to evaluate the feasibility and patient experience of a home-based rehabilitation program after THA, delivered using videos on a tablet personal computer (PC) and a necklace-worn motion sensor to continuously monitor mobility-related activities.

Methods

Study Design

A 6-month prospective cohort study was conducted to test the feasibility and patient experience of a home-based rehabilitation program. Patients participated in a 12-week, home-based exercise program after THA, following video instructions on a tablet PC. Physical activity was registered daily through a necklace-worn motion sensor, and patients were contacted weekly by telephone to receive coaching from a physiotherapist. The phone calls were aimed at motivating participants, discussing barriers to exercise and exercise load, and answering questions concerning guidelines in terms of movement and load after surgery. Measurements were taken preoperatively (T0) and at 4 weeks (T1), 12 weeks (T2), and 6 months postoperatively (T3). The study was approved by the Medical Ethical Committee of University Medical Center Groningen (METc2014/399).

Study Population

We included 30 independently living patients aged 18-75 years who had undergone THA as treatment for primary or secondary OA. Patients were waiting for THA at either the Martini Hospital Groningen or the Medical Center Leeuwarden in the Netherlands. Exclusion criteria were as follows: (1) revision surgery; (2) medical conditions that disallow independent living; (3) cognitive impairment; and (4) inability to sufficiently read and understand Dutch. Patients were included from December 2015 to February 2017 and were required to sign a written informed consent form to be able to participate.

Rehabilitation Program

The duration of the program was 12 weeks. Patients started the program within 7 days of the surgery. Patients performed exercises independently at home using the tablet PC for instructions. The program included strengthening and walking exercises based on increasing the muscle force, balance, and functionality. The exercises comprised movements that trained abductors, flexors, and extensors of the affected hip. The content of the program was based on previous research [9,10] and on guidelines from the American Association of Orthopaedic Surgeons. For the rest, the program was designed in line with the most recent guidelines from the Royal Dutch Society for Physical Therapy [11].

Patients were asked to exercise at least 5 days a week, with rest days on Thursday and Sunday. Strengthening exercises were performed 3 times a week. The instructions for the exercises were provided by videos on the tablet PC, which patients had to imitate. The sessions started with exercise bouts of 10 minutes, which progressively went up to 45 minutes during the 12 weeks of the program. The first step-in level of the program consisted of light and easy exercises. Difficulty and exercise duration were increased across levels very gradually. The exercise burden increased by adding more repetitions, more exercises, and longer training time as well as by incorporating the use of ankle weights. Instructions for walking exercises had no video and showed a descriptive message only. Patients started by walking three 5-minute blocks each day, progressing up to a total of 30-minute walking per day (see Multimedia Appendix 1 for a complete overview of the home-based rehabilitation program).

At the end of the week, patients were asked questions on the tablet PC about perceived pain and perceived intensity of the

exercises. A score of self-reported intensity <4 (scale 0-10) was used as an indicator that a patient could train at a higher level. There was weekly telephone support from a physiotherapist. During this phone call, the physiotherapist and the patient evaluated the progress and agreed on whether to train at a higher level. The program consisted of 12 levels, each week intended for a level of increasing difficulty.

During the intervention, the physiotherapist made 3 home visits. On the first visit, participants received an explanation about the exercises and use of the tablet. The second and third visits were, respectively, at weeks 4 and 12 postoperatively and included physical tests and filling out questionnaires.

Technical Apps

Tablet Personal Computer

Patients received exercise instructions through a tablet PC, a Dell Latitude 10 running the Windows 8 operating system. Exercise instructions were provided through a Web-based app. The app provided exercise instructions and gave participants feedback on their training performance. Exercise completion and app use were recorded to track adherence. The app was designed to be as easy as possible so that people with no tablet experience could participate. Internet connection was provided by the subjects' own home Wi-Fi.

The physiotherapist used a coach app that showed daily registration of completion degree or interruption of exercise bouts. Answers on the evaluation questions (about pain and perceived intensity) were also shown at the end of the week. The physiotherapist was able to change the level of the exercises through this app.

Sensor

The necklace-worn sensor (Figure 1) weighed about 30 g and measured 55 mm × 25 mm × 10 mm (Research prototype; Philips Research, Eindhoven, the Netherlands) [6]. The sensor device included a miniature hybrid sensor containing a 3-dimensional microelectromechanical system accelerometer and a barometric pressure sensor. Accelerometry data were sampled at 50 Hz with a range of 8 g; barometric data were sampled at 25 Hz. A micro-SD card was used for storage and exchange of data. Subjects were asked to wear the sensor in the daytime during the 12-week program and connect the sensor to the tablet manually using a USB cable for data transfer and battery charging every night [6].

Evaluation Methods

Patient Characteristics

Preoperative demographic data, height, weight, medical history, and pre- and postoperative complications were recorded. Factors that might have influenced patients' ability to independently perform a home-based program using novel technology were assessed through a questionnaire. Furthermore, questions were asked about the previous and current use of PCs and smartphones.

Adherence

Adherence to the rehabilitation program was evaluated on the basis of the completion of the planned exercises as indicated by watching the exercise videos and reading the instruction messages. Program adherence was considered sufficient when it exceeded 70%. Reasons why patients did not perform the planned exercises were recorded by the physiotherapist during the weekly phone calls.

User Evaluation

User evaluation was performed with a questionnaire adapted from the sensing and action to support mobility in ambient assisted living subject evaluation form [12,13]. The questionnaire contained questions about the user experience, the perceived intensity of the intervention, coaching, wearing of the sensor, and acceptability of the technology. Answer categories ranged on a Likert scale from 0 ("Do not agree at all") to 5 ("Fully agree"). A higher score indicated a more positive opinion. At the end of the questionnaire, patients were able to write down other suggestions or comments. The user evaluation was done at week 4 (T1) and at the end of the program (T2).

Figure 1. The necklace-worn motion sensor.

Technical Problems

Technical issues that interrupted the execution of the program were logged during the program. All phone calls and extra home visits were registered along with the reasons for these calls or visits.

Statistical Analysis

Statistical analyses were performed using IBM SPSS Statistics 22.0 (IBM). Descriptive statistics were used to portray the main characteristics of the research group.

Results

Demographic Characteristics

In total, 9 men and 21 women participated in the study. Mean age was 64 (SD 6.7) years. Table 1 shows an overview of the demographic characteristics. Of all patients, 7 were living alone and the others were living with a partner or with partner and children. Furthermore, 8 patients had undergone THA on the other hip in the past. While 8 patients had back problems, 5 had rheumatic complaints. All patients had previous computer experience, and 25 owned a smartphone.

Adherence rate

A total of 26 patients completed the program. Four patients dropped out in the first 2 weeks: 3 patients dropped out because of severe back pain, preference to visit a regular physical therapist, and reoperation after a fall and the fourth patient performed postoperatively worse than expected; this patient was insecure, needed more direct personal coaching, and went to a regular physical therapist. Because of sustained back pain, 2 patients finished the program 4 weeks before the official end and went to a regular physical therapist. There were no exercise-induced injuries. Of the 26 patients who completed the program, 3 did not participate in the 6-month measurement because of surgery of the other hip (THA and a fracture) and illness.

Table 1. Demographic characteristics of participants.

Characteristic	Participants (N=30), n (%)
Age in years, mean (SD)	64 (6.7)
Gender, n (%)	
Male	9 (30)
Female	21 (70)
Length (cm), mean (SD)	175 (7.2)
Body weight (kg), mean (SD)	79.8 (13.9)
Education, n (%)	
Low	13 (43)
Middle	4 (13)
High	13 (43)
Employed, n (%)	17 (57)
Living situation, n (%)	
Alone	7 (23)
With partner	20 (67)
With partner and children	2 (7)
With children	1 (3)
Computer experience, n (%)	
Daily	25 (83)
Sometimes	5 (17)
Smartphone owners, n (%)	25 (83)
Surgical approach, n (%)	
Posterolateral	22 (73)
Anterior	8 (27)
Previous total hip arthroplasty on the other hip, n (%)	8 (27)

Table 2. Overview of adherence rate, self-reported perceived pain and intensity, and percentage of patients who increased a level that week during the 12-week rehabilitation program.

Week	Adherence total, mean (SD)	Adherence to strengthening exercises, mean (SD)	Adherence to walking exercises, mean (SD)	Self-reported perceived pain[a], mean (SD)	Self-reported perceived intensity[b], mean (SD)	Patients increasing a level (n=26), n (%)
Week 1	96.4 (9.5)	96.5 (10.4)	96.4 (13.1)	4.1 (2.0)	4.6 (2.6)	20 (77)
Week 2	96.7 (8.3)	96.3 (10.6)	97.2 (10.6)	3.6 (1.7)	5.0 (1.9)	18 (69)
Week 3	98.8 (4.3)	98.7 (6.5)	99.0 (4.9)	3.0 (1.9)	3.9 (2.1)	23 (88)
Week 4	96.9 (6.8)	97.5 (9.0)	96.2 (11.6)	2.9 (2.2)	3.0 (2.0)	20 (77)
Week 5	97.3 (6.0)	97.5 (9.0)	97.1 (8.1)	2.2 (2.3)	2.7 (2.3)	25 (96)
Week 6	96.9 (6.8)	96.2 (10.8)	98.1 (6.8)	1.9 (1.9)	2.5 (2.1)	26 (100)
Week 7	96.5 (6.9)	97.5 (9.0)	95.2 (12.3)	2.0 (1.6)	2.2 (1.9)	26 (100)
Week 8	94.6 (9.0)	93.7 (13.3)	96.2 (13.6)	1.9 (2.1)	2.1 (2.0)	22 (85)
Week 9	87.6 (22.2)	84.1 (25.6)	93 (22.3)	1.6 (1.9)	2.3 (1.8)	17 (71)[c]
Week 10	82.8 (28.4)	82.8 (30.5)	83.0 (30.4)	1.9 (1.8)	2.3 (1.8)	20 (83)[c]
Week 11	85.4 (22.5)	83.5 (24.0)	88.0 (24.9)	1.5 (1.1)	2.3 (1.8)	20 (83)[c]
Week 12	69.6 (39.6)	66.7 (40.2)	74.0 (41.6)	1.6 (1.2)	2.1 (1.9)	N/A[d]

[a]When rating perceived pain on a 0-10 scale at the end of the week (0=no pain, 10=worst possible pain).

[b]When rating perceived intensity of the exercises on a 0-10 scale at the end of the week (0=rest, 10=maximal).

[c] Since 2 patients stopped earlier because of sustained back pain, n=24 in these weeks.

[d]N/A: not applicable.

For all patients, average adherence to exercising 5 times a week was 92%. For all weeks, the adherence was sufficient (>70%), except for strengthening exercises on week 12 (Table 2). After week 8, there was a decrease in adherence. Adherence for strengthening and walking exercises was comparable, except for weeks 9 and 12; for both these weeks, adherence to walking exercises was higher than adherence to strengthening exercises. During the intervention, self-reported perceived pain decreased from 4.1 in week 1 to 1.6 in week 12. Self-reported perceived intensity of the exercises decreased from 4.6 in week 1 to 2.1 in week 12. A score of self-reported intensity <4 was used as an indicator that a patient could train at a higher level. These results correspond with the fact that not raising the exercise level at the end of the week mostly occurred in the first 4 weeks of the program.

Table 3 shows the reasons for nonadherence. Participants failed to comply with training due to vacation or a day or weekend off 25% (33/134) of the time. In addition, work 15% (20/134) and internet connectivity problems 10% (13/134) were often mentioned as the reasons for not exercising. Holidays, days off, and work were mentioned mainly in the last 3 weeks of the intervention. Not exercising because of a social activity was mentioned on all weeks, while pain or muscle pain related to the THA was mentioned mainly in the first 2 weeks.

Table 3. Overview of the reasons for nonadherence.

Reasons for nonadherence	Total number of reasons (n=134), n (%)
Holiday or vacation or day or weekend off	33 (25)
Work	20 (15)
Social activity: birthday, family visit, national holiday	20 (15)
(Muscle) pain related to the total hip arthroplasty (THA)	14 (10)
Pain not related to the THA	13 (10)
Internet problems	13 (10)
Unknown	9 (7)
Forgot to do the exercises	6 (4)
App or tablet did not work	3 (2)
No motivation to train	2 (1)
Disease or illness	1 (1)

Table 4. Results of the user evaluation questionnaire.

Subscale[a]	Time point	
	T1[b] (n=26), mean (SD)	T2[c] (n=26), mean (SD)
Rehabilitation program	4.58 (0.66)	4.59 (0.65)
Coaching	4.88 (0.38)	4.85 (0.48)
Sensor	4.11 (1.00)	3.99 (1.07)
Tablet personal computer	4.77 (0.50)	4.74 (0.55)

[a]Answer options varied from "Do not agree at all" (0) to "Fully agree" (5) on a Likert scale. A higher score on the questionnaire indicated a more positive opinion on the intervention.

[b]T1: At 4 weeks into the program.

[c]T2: At the end of the program.

In this study, of the 26 patients who completed the program, 5 (19%) completed all levels of the program, 11 (42%) reached level 11, 5 (19%) reached level 10, 2 (8%) reached level 9, and 3 (12%) reached level 8 of the program. Not raising the exercise level at the end of the week occurred 43 times in total and occurred mostly in the first 2 weeks of the program. Staying on the same level occurred 23 times (23/43, 53%) in weeks 1-4 and 20 times (20/43, 47%) after week 4. Table 2 shows an overview of the increase in level each week.

The total number of technical issues was 8. Of them, 5 issues included errors in the server of the app. These problems were mostly solved within a few hours, so people were able to complete the exercises for that day. Three issues required an extra home visit to be solved—an unstable Wi-Fi connection, a broken tablet PC, and a disconnection of the sensor and the tablet PC.

User Evaluation

The average score on the user evaluation questionnaire (range 0-5) was 4.6 at T1 and 4.5 at T2. The highest score was for the subscale "coaching" and the lowest score was for "sensor" (Table 4). For the subscale "rehabilitation program," the highest scores were given for the statements "The rehabilitation program is effective for improving muscle strength," "The instructions for the exercises were clear," and "I would recommend this rehabilitation program to other patients." Lowest scores (although >4.0) were given for the statements "The rehabilitation program is effective for improving my walking pattern" and "The level of the exercises was adapted to my possibilities."

Overall, 19 patients gave suggestions for improvements or other comments at the end of the user evaluation questionnaire. Among all, 9 patients mentioned that they liked being able to rehabilitate from home (and that they did not have to travel) and felt motivated by the rehabilitation program; 4 patients would have liked an extra home visit in the first few weeks to check the performance of the exercises (mentioned by 2 patients) and the walking pattern (mentioned by 2 patients). Furthermore, 5 patients recommended more diversity in the exercises; 7 patients mentioned that the duration of the program was a bit too long, especially when they felt their recovery was complete and they had started working again, and 8 patients reported that they experienced the daily wearing of the sensor as uncomfortable because the sensor was big (mentioned by 4 patients) and because the cord was irritating (mentioned by 4 patients).

Discussion

Principal Findings

The results of this study provide support for the feasibility of a home-based telemonitored rehabilitation program for patients after a THA. Adherence to the program was good and user evaluation was positive, and there were only 8 technical issues during the intervention.

A total of 30 patients were included in the study, and 26 patients completed the program. Because of pre-existing back pain, 2 patients finished the program 4 weeks before the official end. The back pain was unrelated to the intervention. There were no exercise-induced injuries during the intervention. This indicates that patients after THA can perform a rehabilitation program safely at home. Furthermore, previous studies concluded that unsupervised home exercise is safe for a majority of THA patients [4,14].

Average adherence to exercise 5 times a week was 92%, which was higher than our goal of 70%. However, we must note that after week 8, there was a decrease. Overall, adherence rate for our program is higher than that for similar 12-week programs, such as those by Chang et al and Mikkelsen et al [14,15], who reported an adherence rate of 73% and 77%, respectively; these two home-based rehabilitation programs were not supported by technology. Our study adherence rate was comparable with the 99% rate for the 8-week home-based program combined with weekly institutional exercise sessions used by Steinhilber et al [16]. This suggests that weekly phone contact combined with the use of technology has the potential to replace supervised exercise sessions.

The reasons mentioned most often for nonadherence were vacation (or a day or weekend off) and work. Both reasons were mentioned mainly in the last 3 weeks of the intervention, which explains the decrease in adherence after week 8. Some people even suggested that the program could be shortened. Internet problems concerned 10% (13/134) of the reasons for nonadherence, although this applied only for 2 patients in a short period (6-7 days).

Patients were positive about the program, giving an average score of 4.6 (range 0-5) at T1 and 4.5 at T2 on the user evaluation questionnaire. Patients liked that they could rehabilitate from home (and that they did not have to travel) and felt motivated by the program. The remote support by weekly phone contact with the coach was appreciated by patients. The importance of the weekly phone contact is in line with a previous study reporting that motivation and coaching is an important parameter for home-based exercise performance and enhanced adherence [17].

The rehabilitation program consisted of 12 levels each week, intended for a level of increasing difficulty and exercise duration. Despite the various levels offered, 5 patients suggested more diversity in the exercises. Furthermore, 4 patients would have liked an extra home visit in the first few weeks to check performance of the exercises and their walking pattern. These comments correspond with the lowest scores in the subscale "rehabilitation program" of the evaluation questionnaire for the statements: "The rehabilitation program is effective for improving my walking pattern" and "The level of the exercises was adapted to my possibilities." Of all, 7 patients mentioned that the program duration was a bit too long; 6 of these patients started working again 6-8 weeks postoperatively. It appears difficult to combine the program with work, even though patients could choose for themselves the time of day to exercise. A recommended adjustment is a more individualized program with additional exercise diversity and when necessary extra support, possibly in the form of a home visit, to improve the walking pattern. Another recommendation is adjusting the duration of the program to patients' goal achievement.

Patients were positive about the technology and gave an average score of 4.8 and 4.1 (range 0-5) for the use of the tablet and sensor, respectively. All patients used their own home Wi-Fi. Geraedts et al reported that adherence to their home-based exercise program and dropping out were strongly influenced by the stability of the mobile internet connection [13]. Based on this study and that of Geraedts et al, it can be concluded that Wi-Fi is preferred over mobile internet connection. All patients had previous computer experience and most patients owned a smartphone [13]. This study shows that it is feasible for this patient group to use novel technology in a home-based rehabilitation program.

Austin et al supported unsupervised home exercise as an effective rehabilitation strategy, which is cost effective as well, for most THA patients compared with formal physiotherapy [4]. The study suggests that because of cost-effectiveness, a home-based program should be used as a standard of routine care after THA. However, some patients may benefit more from formal physiotherapy, for instance, some seniors or people with poor preoperative functional status. More research is needed to identify which patient populations benefit more from supervised rehabilitation.

A limitation of the study was the small number of patients, although this was a deliberate choice to test the feasibility of the program for the first time. In addition, patients who had agreed to participate in the study had some computer experience already and were probably more motivated than average patients, which led to some bias. Nonetheless, the wide variety in educational level, age, and living and work situation seem to have provided a representative group.

Conclusions

A home-based rehabilitation program driven by a tablet app and mobility monitoring seems feasible for THA patients. Adherence to the program was good, and patient experience was positive. In addition, the novel technology was accepted well. When the home-based rehabilitation program also proves to be effective, it could be an alternative to formal physiotherapy. However, further research is needed into the effectiveness.

Acknowledgments

This study was supported by a grant from Anna Fonds and University Medical Center Groningen.

Conflicts of Interest

JA is a Philips Research employee. Philips Research provided the necklace-worn sensor for the study.

Multimedia Appendix 1

Content of the home-based rehabilitation program.

[PDF File (Adobe PDF File), 47KB - mhealth_v7i1e10342_app1.pdf]

References

1. LROI Raportage. LROI. S'Hertogenbosch, The Netherlands: Dutch Arthroplasty Register (LROI); 2018 Jul 04. Orthopedische Implantaten in Beeld: Jaarrapportage uit de Landelijke Registratie Orthopedische Implantaten 2018 URL: https://issuu.com/novlroi/docs/magazine_lroi_2018?e=10489651/62934445 [WebCite Cache ID 72OsZryvr]
2. Rijksoverheid. The Hague, The Netherlands: Rijksoverheid Zorgverzekering URL: https://www.rijksoverheid.nl/onderwerpen/zorgverzekering [accessed 2018-03-07] [WebCite Cache ID 6xjj2rWhL]
3. Bandholm T, Kehlet H. Physiotherapy exercise after fast-track total hip and knee arthroplasty: time for reconsideration? Arch Phys Med Rehabil 2012 Jul;93(7):1292-1294. [doi: 10.1016/j.apmr.2012.02.014] [Medline: 22494947]

4. Austin MS, Urbani BT, Fleischman AN, Fernando ND, Purtill JJ, Hozack WJ, et al. Formal Physical Therapy After Total Hip Arthroplasty Is Not Required: A Randomized Controlled Trial. J Bone Joint Surg Am 2017 Apr 19;99(8):648-655. [doi: 10.2106/JBJS.16.00674] [Medline: 28419032]
5. Geraedts H, Zijlstra A, Bulstra SK, Stevens M, Zijlstra W. Effects of remote feedback in home-based physical activity interventions for older adults: a systematic review. Patient Educ Couns 2013 Apr;91(1):14-24. [doi: 10.1016/j.pec.2012.10.018] [Medline: 23194823]
6. Geraedts HAE, Zijlstra W, Zhang W, Bulstra S, Stevens M. Adherence to and effectiveness of an individually tailored home-based exercise program for frail older adults, driven by mobility monitoring: design of a prospective cohort study. BMC Public Health 2014 Jun 07;14:570 [FREE Full text] [doi: 10.1186/1471-2458-14-570] [Medline: 24908049]
7. Franco MR, Tong A, Howard K, Sherrington C, Ferreira PH, Pinto RZ, et al. Older people's perspectives on participation in physical activity: a systematic review and thematic synthesis of qualitative literature. Br J Sports Med 2015 Oct;49(19):1268-1276. [doi: 10.1136/bjsports-2014-094015] [Medline: 25586911]
8. Centraal Bureau voor de Statistiek. The Hague, The Netherlands: Centraal Bureau voor de Statistiek; 2016 Dec 27. 75-plussers sterkst groeiende groep internetters URL: https://www.cbs.nl/nl-nl/nieuws/2016/52/75-plussers-sterkst-groeiende-groep-internetters [accessed 2018-02-23] [WebCite Cache ID 6xRoDTX5N]
9. Galea MP, Levinger P, Lythgo N, Cimoli C, Weller R, Tully E, et al. A targeted home- and center-based exercise program for people after total hip replacement: a randomized clinical trial. Arch Phys Med Rehabil 2008 Aug;89(8):1442-1447. [doi: 10.1016/j.apmr.2007.11.058] [Medline: 18586222]
10. Jan M, Hung J, Lin JC, Wang S, Liu T, Tang P. Effects of a home program on strength, walking speed, and function after total hip replacement. Arch Phys Med Rehabil 2004 Dec;85(12):1943-1951. [Medline: 15605331]
11. Royal Dutch Society for Physical Therapy. Amersfoort, The Netherlands: Royal Dutch Society for Physical Therapy; 2010 Apr. KNGF Guideline for Physical Therapy in patients with Osteoarthritis of the hip and knee: Supplement to the Dutch Journal of Physical Therapy URL: https://www.kngf.nl/binaries/content/assets/kennisplatform/onbeveiligd/guidelines/osteoarthritis_of_the_hip_and_knee_practice_guidelines_2010.pdf [accessed 2018-02-23] [WebCite Cache ID 75gn6OvBh]
12. CORDIS: European Commission. Luxembourgh: CORDIS; 2008 Apr. Sensaction-AAL: Sensing and ACTION to support mobility in ambient assisted living URL: https://cordis.europa.eu/project/rcn/80559_en.html [accessed 2018-02-23] [WebCite Cache ID 6xRnPmnpz]
13. Geraedts HAE, Zijlstra W, Zhang W, Spoorenberg SLW, Báez M, Far IK, et al. A Home-Based Exercise Program Driven by Tablet Application and Mobility Monitoring for Frail Older Adults: Feasibility and Practical Implications. Prev Chronic Dis 2017 Dec 02;14:E12 [FREE Full text] [doi: 10.5888/pcd14.160227] [Medline: 28152361]
14. Chang C, Lin K, Chen W, Jane S, Yeh S, Wang T. Effects of a Home-Based Resistance Training Program on Recovery From Total Hip Replacement Surgery: Feasibility and Pilot Testing. J Nurs Res 2017 Feb;25(1):21-30. [doi: 10.1097/jnr.0000000000000128] [Medline: 26588453]
15. Mikkelsen LR, Mikkelsen SS, Christensen FB. Early, intensified home-based exercise after total hip replacement--a pilot study. Physiother Res Int 2012 Dec;17(4):214-226. [doi: 10.1002/pri.1523] [Medline: 22451440]
16. Steinhilber B, Haupt G, Miller R, Boeer J, Grau S, Janssen P, et al. Feasibility and efficacy of an 8-week progressive home-based strengthening exercise program in patients with osteoarthritis of the hip and/or total hip joint replacement: a preliminary trial. Clin Rheumatol 2012 Mar;31(3):511-519. [doi: 10.1007/s10067-011-1893-0] [Medline: 22086491]
17. Silveira P, van de Langenberg R, van Het Reve E, Daniel F, Casati F, de Bruin ED. Tablet-based strength-balance training to motivate and improve adherence to exercise in independently living older people: a phase II preclinical exploratory trial. J Med Internet Res 2013 Aug 12;15(8):e159 [FREE Full text] [doi: 10.2196/jmir.2579] [Medline: 23939401]

Abbreviations

OA: osteoarthritis
PC: personal computer
THA: total hip arthroplasty

Edited by G Eysenbach; submitted 12.03.18; peer-reviewed by S An, Y Peng, T Jiang; comments to author 05.08.18; revised version received 26.09.18; accepted 11.10.18; published 31.01.19.

Please cite as:
Hoogland J, Wijnen A, Munsterman T, Gerritsma CLE, Dijkstra B, Zijlstra WP, Annegarn J, Ibarra F, Zijlstra W, Stevens M
Feasibility and Patient Experience of a Home-Based Rehabilitation Program Driven by a Tablet App and Mobility Monitoring for Patients After a Total Hip Arthroplasty
JMIR Mhealth Uhealth 2019;7(1):e10342
URL: http://mhealth.jmir.org/2019/1/e10342/
doi:10.2196/10342
PMID:30702438

Original Paper

The App Behavior Change Scale: Creation of a Scale to Assess the Potential of Apps to Promote Behavior Change

Fiona H McKay[1], BSc, MPH, PhD; Sarah Slykerman[1], BHS, MHP, GradDipEd; Matthew Dunn[1], BA, PhD

School of Health and Social Development, Deakin University, Burwood, Australia

Corresponding Author:
Fiona H McKay, BSc, MPH, PhD
School of Health and Social Development
Deakin University
Burwood Hwy
Burwood, 3125
Australia
Phone: 61 392517183
Email: fiona.mckay@deakin.edu.au

Abstract

Background: Using mobile phone apps to promote behavior change is becoming increasingly common. However, there is no clear way to rate apps against their behavior change potential.

Objective: This study aimed to develop a reliable, theory-based scale that can be used to assess the behavior change potential of smartphone apps.

Methods: A systematic review of all studies purporting to investigate app's behavior change potential was conducted. All scales and measures from the identified studies were collected to create an item pool. From this item pool, 3 health promotion exerts created the App Behavior Change Scale (ABACUS). To test the scale, 70 physical activity apps were rated to provide information on reliability.

Results: The systematic review returned 593 papers, the abstracts and titles of all were reviewed, with the full text of 77 papers reviewed; 50 papers met the inclusion criteria. From these 50 papers, 1333 questions were identified. Removing duplicates and unnecessary questions left 130 individual questions, which were then refined into the 21-item scale. The ABACUS demonstrates high percentage agreement among reviewers (over 80%), with 3 questions scoring a Krippendorff alpha that would indicate agreement and a further 7 came close with alphas >.5. The scale overall reported high interrater reliability (2-way mixed interclass coefficient=.92, 95% CI 0.81-0.97) and high internal consistency (Cronbach alpha=.93).

Conclusions: The ABACUS is a reliable tool that can be used to determine the behavior change potential of apps. This instrument fills a gap by allowing the evaluation of a large number of apps to be standardized across a range of health categories.

(JMIR Mhealth Uhealth 2019;7(1):e11130) doi:10.2196/11130

KEYWORDS

apps; smartphone; mobile phone; mobile app; scale development; rating

Introduction

The delivery of psychological and public health interventions through technology is becoming an increasingly common way to prevent illness and promote health. Smartphones and tablets are well positioned to play a role in such interventions as they offer functionalities and opportunities for personalization through the widespread availability of a range of mobile phone apps [1]. Apps play an important role in the management of illness and are a low-cost, easy avenue for the promotion of health and well-being [2-4]. In 2017, there were 325,000 *health* apps across the 2 most common app platforms: Google Play and iTunes [5]. This includes apps that have been developed to assist patients in the management of a range of diseases and conditions, including diabetes mellitus type 1 or 2 [1,6], pain management [7,8], the promotion of increased physical activity [9,10], improve nutrition [11,12], and the promotion of improved mental health [13,14].

Although research investigating mobile phone–based technology over recent years has shown that short message service (SMS) text message–based interventions can have a positive impact on sexual health knowledge [15] and that most health

interventions can benefit from some form of phone-based activity [16], research into the effectiveness of health behavior change through apps is in its infancy, and there is no clear consensus in the research around which specific features of apps can assist in behavior change. Content analyses of apps have identified some features that may promote health behavior change in apps for smoking cessation [17], alcohol reduction [18,19], and physical activity [20,21]. However, most apps only contain a few features that could be considered to have the potential to change behavior [22]. Features that have been found to promote health behavior change include the ability to provide direct advice about behavior change and track behaviors [17] or provide information on the consequences of continuing with the behavior [19]. Conversely, those studies that have found apps to be lacking in health behavior change features have highlighted the absence of individual tailoring such as personalized notifications or the collection of background information, for example, using global position system data to identify when a person might be at a high-risk area for alcohol use [18] or simply asking a user to set a smoking quit date [17].

Studies that report on user outcomes or experiences of apps have had similarly mixed results. One systematic review that investigated the role of apps and other digital media in physical activity and diet as it relates to cancer survivorship found an overall increase in minutes of physical activity with use of the app, but mixed evidence for improved diet, and no improvement for secondary outcomes such as a reduction in anxiety or depression [23]. A recent study investigating the role of apps in improving mental health found that after 30 days of app use, mental well-being improved in those using 1 of 3 mental well-being apps tested and those using 1 of the 3 apps tested showed improvements in depression. None resulted in improvements in anxiety [24]. A systematic review and meta-analysis of studies that employed a smartphone app to increase physical activity found that the use of apps could result in significant changes to body weight and body mass index; however, nonsignificant results were identified in changes to physical activity [25].

Alongside this growing body of interest in the identification of apps that may play a role in behavior change [26,27] is an increasing body of research that seeks to first understand the features of apps that may play a role in behavior change and then to measure and classify these features [28-30]. Common among these studies is an aim to identify features that employ best practice to allow health practitioners to better inform consumers and patients of the apps most suited to their needs. The ability of practitioners to give this advice is predicated on the ability of researchers to effectively classify and evaluate apps suitable for the most common health conditions through a reliable and valid measurement tool.

As described by McKay et al [31], many studies investigating the potential of apps to change behaviors have employed a behavior change taxonomy (either the CALO-RE or 26 or 93 item taxonomy) for the rating and categorization of apps [10,20-22]. The aim of the systematic review undertaken by McKay et al [31] was to investigate ways in which researchers evaluate the potential health behavior change of apps to identify any current best practice approaches. Instruments identified in the review were created to investigate the behavior change potential of Web- and text-based health interventions [32]. The techniques present in these instruments have been identified in a range of studies and then linked back to behavior change potential. Most notable are by Abraham and Michie [33], who suggested a number of behavior change techniques common to many health behavior theories. Michie et al [34] identified 5 techniques present in physical activity and dietary interventions: self-monitoring, intention formation, specific goal setting, review of behavioral goals, and feedback on performance, finding that interventions that included self-monitoring with at least one other technique were responsible for the largest effect size [34]. These findings are supported by other work suggesting that self-monitoring is useful for increasing physical activity and improving diet for those who were overweight with comorbidities [35], with other work suggesting that self-monitoring is one of the strongest predictors of weight loss [36] and can also assist in decreasing alcohol consumption [37].

App-based studies that have employed these taxonomies have found apps to be lacking in the identified characteristics of a good behavior change intervention. For example, in an investigation of 166 apps that encourage medical adherence against 93 behavior change techniques, Morrissey et al [22] found most apps contained between 0 and 7 techniques, with the most common technique identified being *action planning*, where users are able to set a reminder to take medication at a specific time every day, and set prompts or cues, typically through the setting of an alarm. A total of 2 studies investigated physical activity but found few techniques for behavior change. Direito et al [21] found that most apps contained 8 techniques, most frequently providing instruction, setting graded tasks, and employing self-monitoring, whereas Conroy et al [10] identified 4 or fewer techniques in the physical activity apps they reviewed.

As more practitioners begin to recommend apps to patients for a range of health care needs [38,39], it becomes essential that we have a valid and reliable way to evaluate these apps. Although both valid and reliable, the taxonomies of behavior change theory [33] were designed to evaluate the features of text and Web-based interventions [32,40,41], not for the review of apps. For instance, these taxonomies often feature a large number of items that are closely related, and are theoretically important in behavior change theory, but will often only appear once in an app. For example, the behavior change taxonomy used by Morrissey et al [22] includes 93 items, with each item allocated a score of 1 if present and 0 if absent. Many of these items are similar, for example, there are 11 items categorized as *reward* (including material incentives, material rewards, and nonspecific rewards), all of which are classified separately. For most apps, only one of these items would be present, thus although an app may offer rewards and the benefits that they bring to behavior change, they only offer 1 type means that app would receive a low score in that behavior change category. With increasing knowledge and the growing body of research into app-based interventions, there is a clear need for a purpose-designed app rating system to identify the potential for health behavior change. Although there is 1 scale, the Mobile App Rating Scale (MARS), that is able to describe the

functionality of apps, including aesthetics and information shared [42], there is currently no scale that can measure the potential for behavior change.

Over the past 3 years, the 3 authors of this study have been involved in rating and reviewing apps for the Victorian Health Promotion Foundation (VicHealth) Healthy Living Apps project [43]. The VicHealth Healthy Living Apps project is an annual rating activity using the MARS [42] and CALO-RE [40] scales to provide consumers with a guide to which apps may assist them best in promoting health. This project typically sees up to 400 apps rated annually for their functionality and ability to encourage or promote behavior change in 1 of the following 5 categories: healthy eating, physical activity, tobacco prevention, alcohol harm prevention, and mental well-being. These categories have been chosen as they form the key priority areas of VicHealth and, therefore, are those that are investigated in the VicHealth Healthy Living Apps project [43]. This experience has made clear to the authors that a purpose-designed scale to measure the health behavior change potential is needed for any app review that seeks to recommend apps to the public.

This study aims to develop a reliable, theory-based scale that can be used to assess the behavior change potential of smartphone apps.

Methods

Study Design

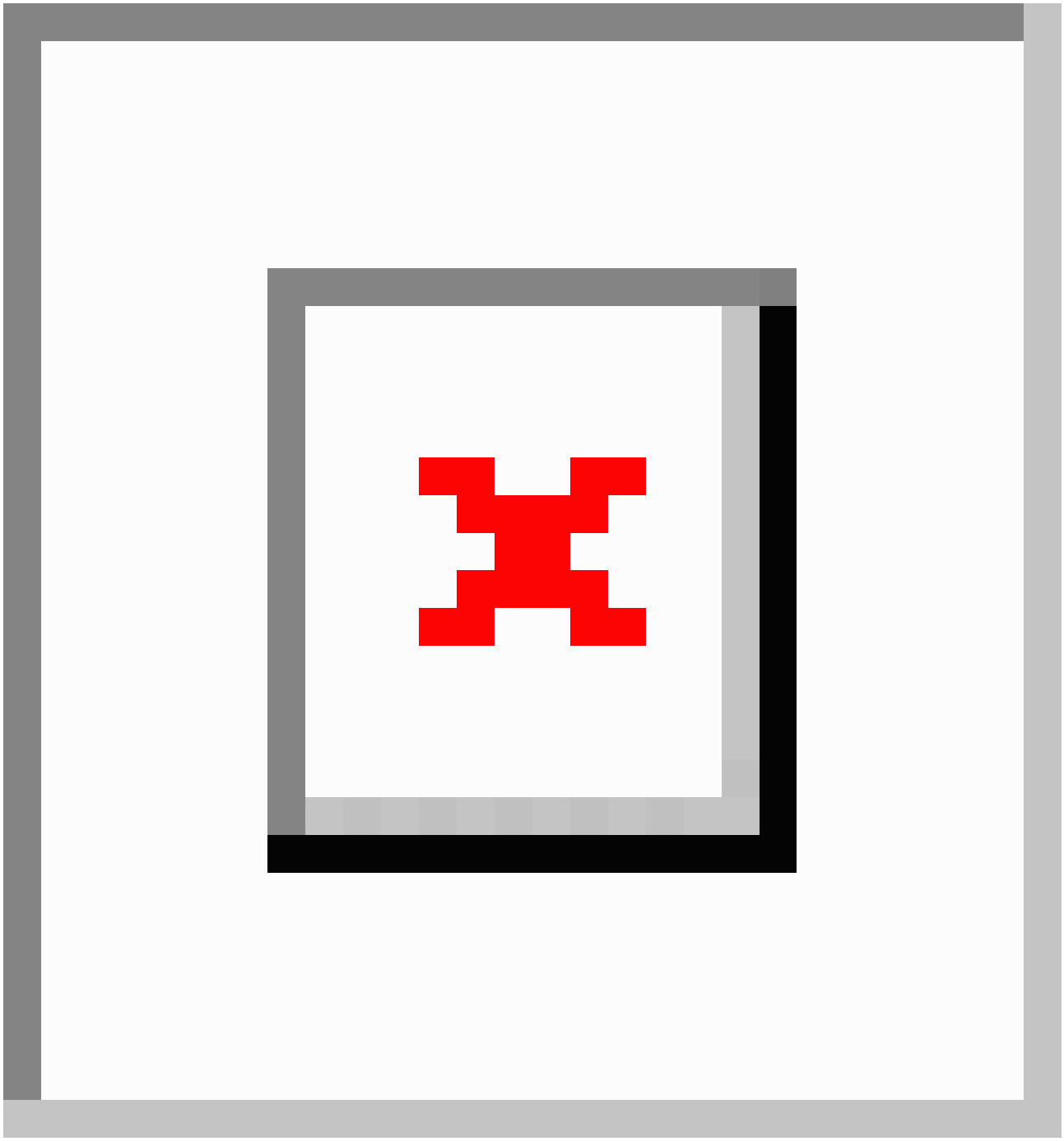
The creation of this scale occurred in 4 phases. Phase 1 included a systematic review to identify all scales that have been used to rate the potential of an app to encourage behavior change. Results from this phase were analyzed and developed into a draft tool. Phases 2 to 4 consisted of series of deductive tests. The results of each round of testing were analyzed and incorporated into the next version of the scale until the team could be confident of reliability and validity of the scale. The final version of the scale was shared with a panel of experts for comment and feedback (see Figure 1 for an overview of the study procedure).

Phase 1: Systematic Review to Develop Initial Item Pool

A systematic search of the literature was conducted to gather all published evidence relating to the various ways that apps have been evaluated for behavior change potential to develop an item pool. This search was based on and extended a previous systematic review [31]. A total of 5 databases (Academic Search Complete, CINAHL Complete, E-Journal, MEDLINE Complete, and PsycINFO) were systematically searched. The search was completed on November 17, 2017, with no temporal limitations placed on the search. The search was limited to studies focusing on mobile phones, smartphones, cell phones, and tablets; used apps; and focused on health behaviors previously investigated in the VicHealth Healthy Living Apps project [43]. Search terms were health, wellbeing, preventative health, smok*, nutrition, alcohol, physical activity, or mental wellbeing.

The inclusion criteria comprised studies that evaluated mobile health apps in English, evaluations or reviews of apps targeted at consumers, alone or in addition to health professionals, and studies that evaluated the effectiveness of mobile health apps.

Figure 1. Study procedure. ICC: interclass coefficient; ABACUS: App Behavior Change Scale.



Excluded studies comprised those that evaluated mobile health apps targeted only at health professionals, formative evaluations of mobile health apps, protocols for evaluations, apps that were not publicly or commercially available, studies that reported primarily on the validation of any mobile health app tool (eg, the MARS), and studies of apps not related to health behavior change. The papers were first screened by title and abstract based on the inclusion and exclusion criteria. The full texts of selected papers were then obtained for further assessment for final inclusion.

Phase 2: Face and Content Validity

The initial version of this scale was pilot-tested with 3 physical activity apps. The pilot testing was conducted by 2 experienced reviewers (FM and SS) and allowed the raters to (1) become familiar with the scale and (2) refine the wording of items and create item descriptors and examples.

Following this pilot, the ABACUS version 1 was used to rate the 3 highest rating apps from each of the 5 categories (15 apps in total) from the VicHealth Healthy Living Apps project [43]. To undertake this testing, the reviewers downloaded and became familiar with each app. Similar to other studies [28,42,44], the authors spent approximately 10 to 15 min testing all app features before rating. After the apps were rated, the raters met to discuss the app and the allocated score as a way to achieve agreement among raters and strengthen the scale. This discussion allowed for an identification of the similarities and differences in rating and, importantly, the strengths and weakness of each question

in the scale, including clarity and specificity. During this process, the raters added and refined descriptors and examples for each item.

Phase 3: Reliability Analysis

The ABACUS version 2 was used by 3 raters (FM, SS, and MD) to review 50 physical activity apps. Physical activity apps were chosen for this phase not only because there are a large number of physical activity apps in the Apple iTunes store, providing a large choice for consumers but also because past reviewing [43] suggests that they represent a wide range in app quality. Apps were downloaded from the app store, and mirroring testing in phase 2, the authors spent approximately 10 min to 15 min reviewing all the features of the app.

Reliability of the scale was assessed using Krippendorff alpha. This allows for rating of ordinal data, can be used with an unlimited number of raters, and has been found to be superior to Cohen kappa [45,46]. Consistent with previous research, an alpha of more than .67 is used to indicate agreement [47], whereas, a negative alpha indicates less agreement than that would be expected by chance and suggests that there may have been inconsistencies in how measures were applied [48]. The internal consistency of the scale was calculated using Cronbach alpha. Interrater reliability was determined by interclass coefficient (ICC) [49]. Percentage agreement was also calculated.

Phase 4: Reliability Analysis 2

To investigate the discrepancies identified in phase 3, the same 3 raters (FM, SS, and MD) rated 5 unrated physical activity apps together. The apps were rated 1 at a time allowing for discussion of the results and for clarification of problem areas, specifically in item descriptions and examples. At the completion of this further moderation activity, an additional 20 apps were independently reviewed against ABACUS version 2, following the same procedure as phase 3.

Results

Phase 1

The search identified 593 unique papers. The abstracts and titles of all papers were reviewed, leaving 77 papers for full-text review. This review resulted in 50 papers that fully met the inclusion criteria and were included in this study (the list of resources is available in Table 1). To determine current best or common practice in app reviewing for behavior change, all scales used in the 50 papers identified were collected. For scales that were not provided as part of the manuscript or as a supplemental material, institution and academic sharing websites (such as Research.net) were searched. If the scale was not able to be located, the authors were emailed and a copy was requested. Only 2 scales [50,51] were unable to be obtained as the author had either moved on from that institution or there was no response to the email.

The scales identified in this systematic review were collated into a single document resulting in 1333 items (see Multimedia Appendix 1), with duplicates and questions present in the MARS removed, leaving 130 individual items (see Multimedia Appendix 1). Moreover, 2 authors (FM and SS) experienced in health promotion and health promoting apps reviewed the item pool. These authors had participated in the VicHealth Healthy Living Apps project [43], and each was experienced in rating hundreds of apps. The item pool was reviewed to identify or create items that were clear and based on previous work by these authors would be present in the highest quality apps [43].

From the 130 items, similar items were collapsed, for example, items that sought to identify avoidance or were collapsed with items that sought to minimize distraction; items that were presented as statements or single words were reworked into questions for ease of use. For example, 1 item that read “discrepancy between current behaviour and goal” was reworked to read “Does the app give the user the ability to quickly and easily understand the difference between current action and future goals?” This process resulted in an initial version of this scale, with 33 items that were categorized into 7 groups: (1) general, (2) goals, (3) feedback and monitoring, (4) knowledge and information, (5) actions, (6) rewards, and (7) environmental factors. These items formed the first version of the scale, the App Behavior Change Scale (ABACUS) version 1.

Table 1. Types and methods of evaluation.

Method	Location	Health condition	Reference
Data usage and user feedback	United Kingdom	Alcohol	Attwood et al [52]
Established evaluation checklist (Abraham and Michie 2008) [33]	New Zealand	Physical activity and dietary	Direito et al [21]
Established evaluation checklist (Abraham and & Michie 2008) [33]	The Netherlands	Health and Fitness	Middelweerd et al [20]
Established evaluation checklist (Abraham and Michie 2008) [33]	United States	Cancer	Vollmer et al [53]
Established evaluation checklist (CALO-RE)	United States	Physical activity	Conroy et al [10]
Established evaluation checklist (MARS[a] and Abroms, 2013 checklist)	Australia	Smoking	Thornton et al [28]
Established evaluation checklist (MARS)	New Zealand	Weight loss and smoking cessation	Patel et al [54]
Established evaluation checklist (MARS)	New Zealand	Travel and dietary behavior associated with health and environmental impact	Sullivan et al [44]
Established evaluation checklist (MARS) and self-developed evaluation checklist based on literature review	United States	Weight management	Bardus et al [55]
Established evaluation checklist (Michie et al) [32]	Ireland	Medication adherence	Morrissey et al [22]
Matched case-control trial	Australia	Physical activity	Kirwan et al [56]
Not discussed	Spain and United Kingdom	Iron-deficiency anemia, hearing loss, migraine, low vision, asthma, diabetes mellitus, osteoarthritis, and unipolar depressive disorders	Martínez-Pérez et al [2]
Self-developed checklist based on diabetes guidelines	United States	Diabetes	Nie et al [58]
Self-developed checklist based on epilepsy guidelines	Australia	Epilepsy	Pandher et al [59]
Self-developed checklist, established evaluation checklist (system usability scale)	United States	Chronic illness	Singh et al [60]
Self-developed evaluation checklist	Unites States	Smoking cessation	Abroms et al [61]
Self-developed evaluation checklist	Germany	Diabetes	Arnhold et al [1]
Self-developed evaluation checklist	United States	Weight management	Azar et al [4]
Self-developed evaluation checklist	Canada	Cancer	Bender et al [62]
Self-developed evaluation checklist	South Korea	Smoking cessation	Choi et al [63]
Self-developed evaluation checklist; user feedback	Norway	Diabetes	Chomutare et al [6]
Self-developed evaluation checklist; user feedback	United States	Alcohol	Cohn et al [50]
Self-developed evaluation checklist	United States	Diabetes and endocrinology	Eng et al [64]
Self-developed evaluation checklist	Unites States	Smoking cessation	Hoeppner et al [17]
Self-developed evaluation checklist	United Kingdom	Asthma	Huckvale et al [3]
Self-developed evaluation checklist	Canada	Headache	Hundert et al [65]
Self-developed evaluation checklist	United Kingdom	Melanoma	Kassianos et al [66]
Self-developed evaluation checklist; user feedback	United States	Hypertension	Kumar et al [67]
Self-developed evaluation checklist	Spain	Heart disease	Martínez-Perez et al [57]
Self-developed evaluation checklist; user feedback	United Kingdom	Breast cancer	Mobasheri et al [68]
Self-developed evaluation checklist	Australia	Bipolar disorder	Nicholas et al [69]
Self-developed evaluation checklist	Italy	Hearing	Paglialonga et al [70]
Self-developed evaluation checklist	United States	Weight-loss	Pagoto et al [71]

Method	Location	Health condition	Reference
Self-developed evaluation checklist; user feedback	United States	Cancer	Pandey et al [72]
Self-developed evaluation checklist; user feedback	Spain	Mindfulness	Plaza et al [73]
Self-developed evaluation checklist	United States	Mental health	Radovic et al [74]
Self-developed evaluation checklist; user feedback	United Kingdom	Pain	Reynoldson et al [8]
Self-developed evaluation checklist	Spain	HIV	Robustillo et al [75]
Self-developed evaluation checklist	United States	Health and wellness	Sama et al [76]
Self-developed evaluation checklist	Canada	Depression	Shen et al [77]
Self-developed evaluation checklist	United Kingdom	Smoking cessation	Ubhi et al [78]
Self-developed evaluation checklist; established evaluation checklist; user feedback	United States	Pediatric obesity	Wearing et al [79]
Self-developed evaluation checklist; user feedback	Australia	Alcohol	Weaver et al [80]
Self-developed evaluation checklist; user feedback	United States	Physical activity	Yang et al [81]
Self-developed evaluation checklist based on literature review	United States	Suicide prevention	Aguirre et al [82]
Self-developed evaluation checklist; established evaluation checklist (MARS)	United States	Pediatric medication adherence	Nguyen et al [83]
User feedback	Ireland	Physical activity	Casey et al [9]
User feedback	United Kingdom	Women's health	Derbyshire and Dancey [84]
User feedback	United States	Smoking	Ferron et al [85]
User feedback	Spain	Type 2 diabetes, obesity, and breastfeeding	García-Gómez et al [51]

[a]MARS: Medication Adherence Rating Scale.

Phase 2

This process resulted in the removal of 9 questions that were deemed to be unclear or were found to be duplicates or unnecessary. For example, in the initial scale, there were 4 separate items that outlined behavior costs, rewards, and encouragement. The authors' experience rating several hundreds of apps over a number of years, combined with the initial round of reviewing and discussion of this scale, allowed for the determination that more than one of these items were unlikely to be in the same app. As a result, these items were collapsed into 1 item: "Does the app provide general encouragement?" Other items were also removed at this point as it was determined that these questions were not relevant to behavior change, for example, a question about whether the app could be used without internet connection and 2 questions about expertise and consistency with national guidelines were collapsed into 1 question: "Was the app created with expertise and/or Does the app provides information that are consistent with national guidelines?"

The resulting scale contained 24 items. Following this, the scale was tested with 15 apps, 3 from each category: physical activity; healthy eating; alcohol; smoking; and mental well-being. Again, this process allowed for a refinement of the scale and resulted in several changes, including clarifying words and descriptors, reordering items, and combining other items, for example, 3 items relating to material, social, and self-reward or incentive were collapsed into a single item: "Does the app provide a material or social reward or incentive?" The authors' experience rating apps lead to the conclusion that it would unlikely that any 1 app would have more than 1 incentive or reward. Phase 2 resulted in the 22-item ABACUS version 2 with questions categorized into the following 4 categories: knowledge and information, goals and planning, feedback and monitoring, and actions.

At this stage, the ABACUS version 2 was sent to 7 external experts for their comment on content. These experts included 3 experts on mental well-being, 1 expert on alcohol and tobacco, 1 on physical activity, 1 on behavioral science, and 1 on health promotion. These experts were able to offer suggestions on language and terminology used, resulting in refinement of terminology and descriptors. For example, one of the reviewers suggested that the descriptor of item 1.4 (Does the app provide instruction on how to perform the behavior?) also includes video instructions (the app is clear in telling the person how to perform a behavior or preparatory behaviors, either verbally, through video, or in written form. Please note, the behavior that is seeking to be changed, not information on how to use the app). This version of the scale is presented in Table 2.

Phase 3

Phase 3 testing was conducted with 50 physical activity apps downloaded from the app store. All apps were rated independently by 3 reviewers against the ABACUS version 2, with ratings entered into Qualtrics to minimize user error. This phase found half of the questions to have high percentage agreement among reviewers (over 80%) with the scale overall reporting moderate interrater reliability (2-way mixed ICC=.69,

95% CI 0.52-0.82) and moderate internal consistency (Cronbach alpha=.71). However, some questions reported very low agreement. For example, question 4.3 "Does the app allow or encourage for practice or rehearsal, in addition to daily activities?" returned only an agreement of 51% with a negative Krippendorff alpha (alpha=−.01). Several other questions showed similarly low scores (see Table 2), and only 1 question achieved an alpha that would indicate agreement. These results prompted an additional round of discussion, and comparison was undertaken.

Phase 4

The initial discussion resulted in the collapsing of 2 *goal* questions into 1 from "Does the app allow for the setting of outcome (long-term) goals?" and "Does the app have the ability to set short and medium-term goals or a plan?" to "Does the app allow for the setting of goals?" Furthermore, a number of descriptors were reworded, and examples were provided for all questions. These changes resulted in ABACUS version 3 containing 21 questions (see Table 3 for final version of the scale).

This round of rating found over 80% of questions to have high percentage agreement among reviewers, with 3 questions scoring a Krippendorff alpha indicating agreement and a further 7 came close with alphas more than .5. The scale overall reported high interrater reliability (2-way mixed ICC=.91, 95% CI 0.81-0.97) and high internal consistency (Cronbach alpha=.93; see Table 4).

Table 2. Percentage agreement and reliability of App Behavior Change Scale version 2.

Item #	Measure	Phase 3 (50 apps)	
		Interrater reliability (Krippendorff alpha)	Percent agreement
1.1	Customize and personalize features	−.010	56
1.2	Consistent with national guidelines or created with expertise	.25	88
1.3	Baseline information	.45	73
1.4	Instruction on how to perform the behavior	.79	91
1.5	Information about the consequences of continuing and/or discontinuing behavior	.21	92
2.1	Willingness for behavior change	−.01	97
2.2	Goal setting	.22	83
2.3	Review goals, update, and change when necessary	.33	75
3.1	Understand the difference between current action and future goals	.60	84
3.2	Self-monitor behavior	.53	81
3.3	Share behaviors with others and/or allow for social comparison	.30	65
3.4	User feedback (in person or automatically)	.12	88
3.5	Export data	.16	77
3.6	Material or social reward or incentive	.19	66
3.7	General encouragement	.23	65
4.1	Reminders and/or prompts or cues for activity	.23	61
4.2	Encourage positive habit formation	.11	71
4.3	Practice or rehearsal, in addition to daily activities	−.01	51
4.4	Opportunity to plan for barriers	−.01	97
4.5	Restructuring the physical or social environment	−.01	97
4.6	Distraction or avoidance	−.02	95

Table 3. Final app behavior change scale, including examples.

Scale: item number and question		Definition	Example or further information	Source of question (from Table 1)
1. Knowledge and information				
1.1	Does the app have the ability to customize and personalize some features?	Elements of the app can be personalized through specific tools or functions that are specific to the individual using the app.	• To select a disease type from among several available and then to follow a specific path or set of tools or systems. • To select to receive emails or texts of a specific nature. • To choose "yes" or "no" to a specific capability of the app would be considered personalization. • To create a personalized exercise plan.	[44,54]
1.2	Was the app created with expertise and/or Does the app provide information that is consistent with national guidelines?	This would be found in the about section or generally in the app.	• Does the app suggest 30 min of exercise each day? • Does it recommend 5 veg and 3 fruit? • Does it seek to build resilience and promote help seeking? • Is there any evidence that the app was created by an expert? (doctor/professional body/university)	[44,54]
1.3	Does the app ask for baseline information?	This includes BMI[a], weight, smoking rate, exercise, or drinking behaviors	• This might be at the set-up phase or in a profile setting.	[28,85]
1.4	Does the app provide instruction on how to perform the behavior?	The app is clear in telling the person how to perform a behavior or preparatory behaviors, either verbally, through video, or in written form. NB: the behavior that is seeking to be changed, not information on how to use the app	• This could include showing person how to use gym equipment, sharing sample plans for action, instruction on suitable clothing, recipes, and general tips.	[20,21,22,81]
1.5	Does the app provide information about the consequences of continuing and/or discontinuing behavior?	The app gives the user information about the consequences of behavior in general, this includes information about the relationship between the behavior and its possible or likely consequences in the general case. This information can be general or personalized.	• Consequences may include health, feelings, or cost consequences.	[22,81]
2. Goals and planning				
2.1	Does the app ask for willingness for behavior change?	Is there a feature during setup where you describe how ready you are for behavior change?	• This may be in the form of a scale of readiness or in a question that asks the user to describe how ready you are.	[17,85]
2.2	Does the app allow for the setting of goals?	The person is encouraged to make a behavioral resolution. The person is encouraged to set a general goal that can be achieved by behavioral means. This includes subgoals or preparatory behaviors and/or specific contexts in which the behavior will be performed. The behavior in this technique will be directly related to or be a necessary condition for the target behavior.	• This is the explicit noting of a goal or choosing a goal from one provided within the app.	[20,21,40,44,54,55,81]

Scale: item number and question		Definition	Example or further information	Source of question (from Table 1)
2.3	Does the app have the ability to review goals, update, and change when necessary?	Involves a review or analysis of the extent to which previously set behavioral goals (regardless of short or long) were achieved.	• This is where a goal can be changed. This allows people to act on previously set goals and then revise or adjust where needed.	[22,40,81]
3. Feedback and monitoring				
3.1	Does the app give the user the ability to quickly and easily understand the difference between current action and future goals?	Allows user to see how they are tracking against a goal and to see the difference between what they want to do and what they are currently doing. This will give some feedback on where they are at and what they need to change to get to where they want to be.	• This could be in the form of a graph or some other visual describing how close the user is to meeting their goals.	[22,40,81]
3.2	Does the app have the ability to allow the user to easily self-monitor behavior?	The app allows for a regular monitoring of the activity.	• Connects with watch that records daily steps that can be reviewed. • Allows for easy logging of exercise or meditation? • Allows for tracking of weight loss. • Allows logging of daily alcoholic drinks or cigarettes.	[20,21]
3.3	Does the app have the ability to share behaviors with others (including social media or forums) and/or allow for social comparison?	The app allows the person to share his or her behaviors on social media or in forums. This could also include a *buddy* system or a leaderboard.	• Share with Facebook or other socials • Tell the user that they are doing x and at this time, other people like them are doing y	[4,20,21,22,85]
3.4	Does the app have the ability to give the user feedback—either from a person or automatically?	The app is able to provide the person with feedback, comments, or data about their own recorded behavior. This might be automatic or could be personal.	• Does the app have a *coach* function?	[22,40,81]
3.5	Does the app have the ability to export data from app?	The app allows for the export of information and progress to an external user.	• Export to a computer or to another user such as a doctor or fitness expert. • Sharing to Facebook does not count.	[65]
3.6	Does the app provide a material or social reward or incentive?	App provides rewards for attempts at achieving a behavioral goal. This might include efforts made toward achieving the behavior or progress made in preparatory steps toward the behavior or in achieving a goal.	• Financial, either in returning money that was not spent on, for example, cigarettes or in paying someone to engage in a specific activity. • Social or public, for example, congratulating the person for each day that he or she meets his or her exercise target.	[22,40,81]
3.7	Does the app provide general encouragement?	The app provides general encouragement and positive reinforcement on actions leading to the goal.	• This could include achievement badges or telling the user that they are a certain percentage closer to their goal.	[22,40,81]
4. Actions				

Scale: item number and question		Definition	Example or further information	Source of question (from Table 1)
4.1	Does the app have reminders and/or prompts or cues for activity?	The app prompts the user to engage in the activity. The app has the ability to give notifications or reminders to cue the behavior.	• This could be like the apple watch reminding you to stand or a meditation app telling you to meditate now.	[20,21]
4.2	Does the app encourage positive habit formation?	The app prompts explicit rehearsal and repetition of the behavior–not just tracking or logging.	• An example of this are the couch to 5 km apps that provide a training schedule.	[21,22,81]
4.3	Does the app allow or encourage for practice or rehearsal, in addition to daily activities?	App does not have a lock on activities or a number that you cannot exceed daily.	• This would include allowing the user to undertake extra activities in a single day.	[20,21]
4.4	Does the app provide opportunity to plan for barriers?	The app encourages the person to think about potential barriers and identify ways of overcoming them.	• Alcohol app might give strategies for a night out that would normally be a big night.	[55]
4.5	Does the app assist with or suggest restructuring the physical or social environment?	The app prompts the person to alter the environment in ways so that it is more supportive of the target behavior.	• Might suggest locking up or throw away or their high-calorie snacks or take their running shoes to work.	[21,22,81]
4.6	Does the app assists with distraction or avoidance?	The app gives suggestions and advice on how the person can avoid situations or distract themselves when trying to reach their goal.	• For example, a smoking cessation app may suggest that the user not drink coffee if this is typically combined with smoking behaviors that they are trying to cease.	[21,22,81]

Table 4. Percentage agreement and reliability of App Behavior Change Scale version 3.

Item #	Measure	Phase 4 (20 apps)	
		Interrater reliability (Krippendorff alpha)	Percent agreement
1.1	Customize and personalize features	.52	83
1.2	Consistent with national guidelines or created with expertise	.73	83
1.3	Baseline information	.79	90
1.4	Instruction on how to perform the behavior	.63	87
1.5	Information about the consequences of continuing and/or discontinuing behavior	−.02	93
2.1	Willingness for behavior change	0	97
2.2	Goal setting	.58	83
2.3	Review goals, update, and change when necessary	.38	80
3.1	Understand the difference between current action and future goals	.34	80
3.2	Self-monitor behavior	.62	83
3.3	Share behaviors with others and/or allow for social comparison	.73	87
3.4	User feedback (in person or automatically)	.26	67
3.5	Export data	.43	87
3.6	Material or social reward or incentive	.15	60
3.7	General encouragement	.54	77
4.1	Reminders and/or prompts or cues for activity	.61	80
4.2	Encourage positive habit formation	.28	63
4.3	Practice or rehearsal, in addition to daily activities	.05	80
4.4	Opportunity to plan for barriers	.31	93
4.5	Restructuring the physical or social environment	.57	93
4.6	Distraction or avoidance	1	100

Discussion

Principal Findings

This study reports on the creation of a scale (ABACUS) to measure the potential behavior change of smartphone apps. After conducting a systematic review to identify all research that has evaluated apps for behavior change, 133 items were identified and later modified after expert review to a final set of 21 items. The items within the scale are grouped into the following 4 categories: knowledge and information, goals and planning, feedback and monitoring, and actions. The ABACUS was reviewed by an expert panel and then tested first against 50 physical activity apps; however, because of concerns relating to moderate internal consistency and interrater reliability, an additional step of moderation was taken. This moderation saw the same raters come together to refine the scale, resulting in improved descriptors and the inclusion of examples for each question. Following this revision, the scale was used to rate an additional 20 apps. This round of ratings resulted in a high internal consistency and interrater reliability. Although previous studies evaluating smartphone apps have focused largely on features available in apps [21] or behavior change techniques through a self-developed evaluation checklist [4,10], the ABACUS provides researchers with a reliable and valid instrument to evaluate apps based on their behavior change potential.

This scale will allow researchers to investigate the behavior change potential of a large number of apps reasonably quickly. This is important, as the fast-moving pace of app technology means that although randomized controlled trials (RCTs) remain important in understanding the impacts of individual apps on behavior [86], it has been suggested that the RCT may not be the most appropriate method to generate evidence around mobile apps [28]. RCTs can take a significant amount of time in planning and design meaning that by the time the RCT is available for publication, the information is no longer current [28]. The scale developed in this research is not a replacement for an RCT but rather will allow researchers and consumers to understand the behavior change potential of an app in the absence of an RCT.

The MARS [42], a 23-item tool included 5 subscales for measuring app quality: engagement, functionality, aesthetics, information, and app subjective quality, with questions such as target age group, ease of navigation, or aesthetics can be used in conjunction with the ABACUS. The MARS is a useful tool in understanding the aesthetic and functional appeal of an app. When used together, the MARS and the ABACUS will allow researchers to provide users with 2 scores for each app: 1 that

measures app quality and 1 that measures potential for behavior change.

This study is only a starting point in the identification and interpretation of the behavior change potential of smartphone apps. This study only reports on the validation and reliability of physical activity apps, and as such, further testing of the scale should be conducted on additional health areas such as smoking, alcohol, and nutrition, as it is possible that different items may be important for these health areas. Furthermore, a more detailed investigation into the relative scores of apps will need to be undertaken. This will allow for an understanding of the importance of the overall score assigned to each app. At present, this scale is best understood as providing a continuous score rather than specific cut-off points. However, this is not to say that with more investigation and testing that clear scores could not provide a consumer with a numerical rating reflecting a behavioral outcome. This study has not purported to demonstrate correlation between an app's score and the health outcome; however, this scale could be used in future along with a more detailed study of individual apps and the behavior change outcomes in using them.

ABACUS has good interrater reliability and is a valid tool for evaluating the potential behavior change in smartphone apps. The validation and reliability testing of ABACUS contributes to the literature by providing a standardized method of evaluating smartphone apps for behavior change.

Limitations

Although this scale shows good reliability and validity, there are several limitations that need to be addressed. The first is that we have not sought to investigate criterion validity. The scale presented in this paper seeks to measure the theoretical behavior change potential of apps; and therefore, we do not seek to investigate the relationship between actual features of apps and behavioral outcomes. This scale has not been designed for this type of activity, so we leave this up to others to identify an appropriate method for such an investigation. Although reducing the numbers of items on the scale facilitates faster rating, there is a risk that removal of duplicate items and streamlining these items into 1 binary response may inflate a score. For example, by collapsing all *goal*-setting activities into 1 item, this scale recognizes apps that have any goals-setting ability, rather than the strength of that ability—a feature found in the behavior change taxonomy. Furthermore, there is a risk that by collapsing items that record starting a positive behavior with stopping a negative behavior, we may be missing a key aspect of behavior change. These decisions were made based on the authors' experience of rating apps with an understanding that a single app will not include both of these features, and as such, in seeking to provide a succinct scale, it makes more sense to only measure 1 outcome. Like other similar studies [42], this study highlights the importance of rater's knowledge of apps when completing such evaluations and with moderating 5 to 10 apps at the beginning of the process as a team is important to ensure a robust score. In addition, similar to other studies, raters in this study spent 10 min to 15 min with the app to become familiar before completing the evaluation. This time spent using the app is consistent with other studies that seek to review apps, as a longer time under review is not realistic [42,87]. Finally, 1 key limitation of this study is that the scale has been validated on physical activity apps. Although this scale seeks to be used in the future for other health behaviors, at this point in time, we are only confident that it can be used to rate the health behavior potential of physical activity apps. Other health behaviors will need to be investigated in future studies.

Conclusions

The ABACUS is a reliable tool that can be used to determine the behavior change potential of apps. This instrument fills a gap by allowing the evaluation of a large number of apps to be standardized across a range of health categories. This scale can be used by teams to rate apps that seek to promote behavior change, allowing for high-quality apps to then be recommended to the general public.

Acknowledgments

This study was partly funded by the Victorian Health Promotion Foundation.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Items from systematic review.

[XLSX File (Microsoft Excel File), 80KB - mhealth_v7i1e11130_app1.xlsx]

References

1. Arnhold M, Quade M, Kirch W. Mobile applications for diabetics: a systematic review and expert-based usability evaluation considering the special requirements of diabetes patients age 50 years or older. J Med Internet Res 2014 Apr 09;16(4):e104 [FREE Full text] [doi: 10.2196/jmir.2968] [Medline: 24718852]
2. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M. Mobile health applications for the most prevalent conditions by the World Health Organization: review and analysis. J Med Internet Res 2013 Jun 14;15(6):e120 [FREE Full text] [doi: 10.2196/jmir.2600] [Medline: 23770578]

3. Huckvale K, Car M, Morrison C, Car J. Apps for asthma self-management: a systematic assessment of content and tools. BMC Med 2012 Nov 22;10:144 [FREE Full text] [doi: 10.1186/1741-7015-10-144] [Medline: 23171675]
4. Azar KM, Lesser LI, Laing BY, Stephens J, Aurora MS, Burke LE, et al. Mobile applications for weight management: theory-based content analysis. Am J Prev Med 2013 Nov;45(5):583-589. [doi: 10.1016/j.amepre.2013.07.005] [Medline: 24139771]
5. Research2Guidance. 325,000 mobile health apps available in 2017 – Android now the leading mHealth platform URL: https://research2guidance.com/325000-mobile-health-apps-available-in-2017/ [accessed 2018-02-07] [WebCite Cache ID 6x2aCVSet]
6. Chomutare T, Fernandez-Luque L, Arsand E, Hartvigsen G. Features of mobile diabetes applications: review of the literature and analysis of current applications compared against evidence-based guidelines. J Med Internet Res 2011 Sep 22;13(3):e65 [FREE Full text] [doi: 10.2196/jmir.1874] [Medline: 21979293]
7. Cardoş RA, Şoflău R, Gherman A, Sucală M, Chiorean A. A mobile intervention for core needle biopsy related pain and anxiety: a usability study. J Evid-Based Psychot 2017 Mar 01;17(1):21-30. [doi: 10.24193/jebp.2017.1.2]
8. Reynoldson C, Stones C, Allsop M, Gardner P, Bennett MI, Closs SJ, et al. Assessing the quality and usability of smartphone apps for pain self-management. Pain Med 2014 Jun;15(6):898-909. [doi: 10.1111/pme.12327] [Medline: 24422990]
9. Casey M, Hayes PS, Glynn F, OLaighin G, Heaney D, Murphy AW, et al. Patients' experiences of using a smartphone application to increase physical activity: the SMART MOVE qualitative study in primary care. Br J Gen Pract 2014 Aug;64(625):e500-e508 [FREE Full text] [doi: 10.3399/bjgp14X680989] [Medline: 25071063]
10. Conroy DE, Yang C, Maher JP. Behavior change techniques in top-ranked mobile apps for physical activity. Am J Prev Med 2014 Jun;46(6):649-652. [doi: 10.1016/j.amepre.2014.01.010] [Medline: 24842742]
11. Lieffers J, Hanning RM. Dietary assessment and self-monitoring with nutrition applications for mobile devices. Can J Diet Pract Res 2012;73(3):e253-e260. [doi: 10.3148/73.3.2012.e253] [Medline: 22958633]
12. Tonkin E, Brimblecombe J, Wycherley T. Characteristics of smartphone applications for nutrition improvement in community settings: a scoping review. Adv Nutr 2017 Mar;8(2):308-322 [FREE Full text] [doi: 10.3945/an.116.013748] [Medline: 28298274]
13. Grist R, Porter J, Stallard P. Mental health mobile apps for preadolescents and adolescents: a systematic review. J Med Internet Res 2017 Dec 25;19(5):e176 [FREE Full text] [doi: 10.2196/jmir.7332] [Medline: 28546138]
14. Lui J, Marcus D, Barry C. Evidence-based apps? A review of mental health mobile applications in a psychotherapy context. Prof Psychol Res Pr 2017;48(3):199-210. [doi: 10.1037/pro0000122]
15. Lim M, Hocking J, Aitken C, Fairley C, Jordan L, Lewis J, et al. Impact of text and email messaging on the sexual health of young people: a randomised controlled trial. J Epidemiol Community Health 2012 Jan;66(1):69-74. [doi: 10.1136/jech.2009.100396] [Medline: 21415232]
16. Hall A, Cole-Lewis H, Bernhardt J. Mobile text messaging for health: a systematic review of reviews. Annu Rev Public Health 2015 Mar 18;36:393-415 [FREE Full text] [doi: 10.1146/annurev-publhealth-031914-122855] [Medline: 25785892]
17. Hoeppner BB, Hoeppner SS, Seaboyer L, Schick MR, Wu GW, Bergman BG, et al. How smart are smartphone apps for smoking cessation? A content analysis. Nicotine Tob Res 2016 May;18(5):1025-1031 [FREE Full text] [doi: 10.1093/ntr/ntv117] [Medline: 26045249]
18. Hoeppner BB, Schick MR, Kelly LM, Hoeppner SS, Bergman B, Kelly JF. There is an app for that - Or is there? A content analysis of publicly available smartphone apps for managing alcohol use. J Subst Abuse Treat 2017 Nov;82:67-73. [doi: 10.1016/j.jsat.2017.09.006] [Medline: 29021117]
19. Crane D, Garnett C, Brown J, West R, Michie S. Behavior change techniques in popular alcohol reduction apps: content analysis. J Med Internet Res 2015 May 14;17(5):e118 [FREE Full text] [doi: 10.2196/jmir.4060] [Medline: 25977135]
20. Middelweerd A, Mollee J, van der Wal CN, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014 Jul 25;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]
21. Direito A, Dale L, Shields E, Dobson R, Whittaker R, Maddison R. Do physical activity and dietary smartphone applications incorporate evidence-based behaviour change techniques? BMC Public Health 2014 Jun 25;14:646 [FREE Full text] [doi: 10.1186/1471-2458-14-646] [Medline: 24965805]
22. Morrissey E, Corbett T, Walsh J, Molloy G. Behavior change techniques in apps for medication adherence: a content analysis. Am J Prev Med 2016 May;50(5):e143-e146. [doi: 10.1016/j.amepre.2015.09.034] [Medline: 26597504]
23. Roberts A, Fisher A, Smith L, Heinrich M, Potts H. Digital health behaviour change interventions targeting physical activity and diet in cancer survivors: a systematic review and meta-analysis. J Cancer Surviv 2017 Dec;11(6):704-719 [FREE Full text] [doi: 10.1007/s11764-017-0632-1] [Medline: 28779220]
24. Bakker D, Kazantzis N, Rickwood D, Rickard N. A randomized controlled trial of three smartphone apps for enhancing public mental health. Behav Res Ther 2018 Oct;109:75-83. [doi: 10.1016/j.brat.2018.08.003] [Medline: 30125790]
25. Flores Mateo G, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015 Nov 10;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]

26. Flaherty S, McCarthy M, Collins A, McAuliffe F. Can existing mobile apps support healthier food purchasing behaviour? Content analysis of nutrition content, behaviour change theory and user quality integration. Public Health Nutr 2018 Dec;21(2):288-298. [doi: 10.1017/S1368980017002889] [Medline: 29081322]
27. Ptakauskaite N, Cox AL, Berthouze N. Knowing what you're doing or knowing what to do: how stress management apps support reflection and behaviour change. In: CHI EA '18 Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems. 2018 Presented at: CHI EA '18; April 21 - 26, 2018; Montreal QC, Canada. [doi: 10.1145/3170427.3188648]
28. Thornton L, Quinn C, Birrell L, Guillaumier A, Shaw B, Forbes E, et al. Free smoking cessation mobile apps available in Australia: a quality review and content analysis. Aust N Z J Public Health 2017 Dec;41(6):625-630. [doi: 10.1111/1753-6405.12688] [Medline: 28749591]
29. Reyes A, Qin P, Brown CA. A standardized review of smartphone applications to promote balance for older adults. Disabil Rehabil 2018 Dec;40(6):690-696. [doi: 10.1080/09638288.2016.1250124] [Medline: 27868438]
30. Weekly T, Walker N, Beck J, Akers S, Weaver M. A review of apps for calming, relaxation, and mindfulness interventions for pediatric palliative care patients. Children (Basel) 2018 Jan 26;5(2) [FREE Full text] [doi: 10.3390/children5020016] [Medline: 29373515]
31. McKay F, Cheng C, Wright A, Shill J, Stephens H, Uccellini M. Evaluating mobile phone applications for health behaviour change: a systematic review. J Telemed Telecare 2018 Jan;24(1):22-30. [doi: 10.1177/1357633X16673538] [Medline: 27760883]
32. Michie S, Richardson M, Johnston M, Abraham C, Francis J, Hardeman W, et al. The behavior change technique taxonomy (v1) of 93 hierarchically clustered techniques: building an international consensus for the reporting of behavior change interventions. Ann Behav Med 2013 Aug;46(1):81-95. [doi: 10.1007/s12160-013-9486-6] [Medline: 23512568]
33. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol 2008 May;27(3):379-387. [doi: 10.1037/0278-6133.27.3.379] [Medline: 18624603]
34. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
35. Dombrowski S, Sniehotta F, Avenell A, Johnston M, MacLennan G, Araújo-Soares V. Identifying active ingredients in complex behavioural interventions for obese adults with obesity-related co-morbidities or additional risk factors for co-morbidities: a systematic review. Health Psychol Rev 2012 Mar;6(1):7-32. [doi: 10.1080/17437199.2010.513298]
36. Burke L, Wang J, Sevick M. Self-monitoring in weight loss: a systematic review of the literature. J Am Diet Assoc 2011 Jan;111(1):92-102 [FREE Full text] [doi: 10.1016/j.jada.2010.10.008] [Medline: 21185970]
37. Michie S, Whittington C, Hamoudi Z, Zarnani F, Tober G, West R. Identification of behaviour change techniques to reduce excessive alcohol consumption. Addiction 2012 Aug;107(8):1431-1440. [doi: 10.1111/j.1360-0443.2012.03845.x] [Medline: 22340523]
38. Cook S, Palmer M, Shuler FD. Smartphone mobile applications to enhance diagnosis of skin cancer: a guide for the rural practitioner. W V Med J 2015;111(5):22-28. [Medline: 26521532]
39. Boyle L, Grainger R, Hall RM, Krebs JD. Use of and beliefs about mobile phone apps for diabetes self-management: surveys of people in a hospital diabetes clinic and diabetes health professionals in New Zealand. JMIR Mhealth Uhealth 2017 Jun 30;5(6):e85 [FREE Full text] [doi: 10.2196/mhealth.7263] [Medline: 28666975]
40. Michie S, Ashford S, Sniehotta FF, Dombrowski SU, Bishop A, French DP. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours: the CALO-RE taxonomy. Psychol Health 2011 Nov;26(11):1479-1498. [doi: 10.1080/08870446.2010.540664] [Medline: 21678185]
41. Michie S, Johnston M, Francis J, Hardeman W, Eccles M. From theory to intervention: mapping theoretically derived behavioural determinants to behaviour change techniques. Appl Psychol 2008;57(4):660-680. [doi: 10.1111/j.1464-0597.2008.00341.x]
42. Stoyanov S, Hides L, Kavanagh D, 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]
43. VicHealth. VicHealth Healthy Living Apps project 2016 URL: https://www.vichealth.vic.gov.au/media-and-resources/vichealth-apps/healthy-living-apps [accessed 2018-02-08] [WebCite Cache ID 6x4GR3CxS]
44. Sullivan R, Marsh S, Halvarsson J, Holdsworth M, Waterlander W, Poelman M, et al. Smartphone apps for measuring human health and climate change co-benefits: a comparison and quality rating of available apps. JMIR Mhealth Uhealth 2016 Dec 19;4(4):e135 [FREE Full text] [doi: 10.2196/mhealth.5931] [Medline: 27993762]
45. Bock MA, Krippendorff K. The Content Analysis Reader. London: SAGE Publications; 2008.
46. Hayes A, Krippendorff K. Answering the call for a standard reliability measure for coding data. Commun Methods Meas 2007 Apr;1(1):77-89. [doi: 10.1080/19312450709336664]
47. Krippendorff K. Content Analysis: An Introduction to Its Methodology. London: Sage Publications; 1980.
48. Krippendorff K. Reliability in content analysis. Human Comm Res 2004 Jul;30(3):411-433. [doi: 10.1111/j.1468-2958.2004.tb00738.x]

49. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979 Mar;86(2):420-428. [Medline: 18839484]
50. Cohn AM, Hunter-Reel D, Hagman BT, Mitchell J. Promoting behavior change from alcohol use through mobile technology: the future of ecological momentary assessment. Alcohol Clin Exp Res 2011 Dec;35(12):2209-2215 [FREE Full text] [doi: 10.1111/j.1530-0277.2011.01571.x] [Medline: 21689119]
51. García-Gómez JM, de la Torre-Díez I, Vicente J, Robles M, López-Coronado M, Rodrigues J. Analysis of mobile health applications for a broad spectrum of consumers: a user experience approach. Health Informatics J 2014 Mar;20(1):74-84. [doi: 10.1177/1460458213479598] [Medline: 24550566]
52. Attwood S, Parke H, Larsen J, Morton K. Using a mobile health application to reduce alcohol consumption: a mixed-methods evaluation of the drinkaware track & calculate units application. BMC Public Health 2017 Dec 17;17(1):394 [FREE Full text] [doi: 10.1186/s12889-017-4358-9] [Medline: 28511698]
53. Vollmer DD, Fair K, Hong YA, Beaudoin CE, Pulczinski J, Ory MG. Apps seeking theories: results of a study on the use of health behavior change theories in cancer survivorship mobile apps. JMIR Mhealth Uhealth 2015;3(1):e31 [FREE Full text] [doi: 10.2196/mhealth.3861] [Medline: 25830810]
54. Patel R, Sulzberger L, Li G, Mair J, Morley H, Shing MN, et al. Smartphone apps for weight loss and smoking cessation: quality ranking of 120 apps. N Z Med J 2015 Sep 04;128(1421):73-76. [Medline: 26370762]
55. Bardus M, van Beurden SB, Smith J, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int J Behav Nutr Phys Act 2016 Mar 10;13:35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
56. Kirwan M, Duncan M, Vandelanotte C, Mummery W. Using smartphone technology to monitor physical activity in the 10,000 Steps program: a matched case-control trial. J Med Internet Res 2012 Apr 20;14(2):e55 [FREE Full text] [doi: 10.2196/jmir.1950] [Medline: 22522112]
57. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M, Herreros-González J. Mobile apps in cardiology: review. JMIR Mhealth Uhealth 2013 Jul 24;1(2):e15 [FREE Full text] [doi: 10.2196/mhealth.2737] [Medline: 25098320]
58. Nie L, Xie B, Yang Y, Shan YM. Characteristics of Chinese m-Health applications for diabetes self-management. Telemed J E Health 2016 Dec;22(7):614-619 [FREE Full text] [doi: 10.1089/tmj.2015.0184] [Medline: 27171016]
59. Pandher PS, Bhullar KK. Smartphone applications for seizure management. Health Informatics J 2016 Dec;22(2):209-220. [doi: 10.1177/1460458214540906] [Medline: 25038202]
60. Singh K, Drouin K, Newmark LP, Lee J, Faxvaag A, Rozenblum R, et al. Many mobile health apps target high-need, high-cost populations, but gaps remain. Health Aff (Millwood) 2016 Dec 01;35(12):2310-2318. [doi: 10.1377/hlthaff.2016.0578] [Medline: 27920321]
61. Abroms LC, Lee Westmaas J, Bontemps-Jones J, Ramani R, Mellerson J. A content analysis of popular smartphone apps for smoking cessation. Am J Prev Med 2013 Dec;45(6):732-736 [FREE Full text] [doi: 10.1016/j.amepre.2013.07.008] [Medline: 24237915]
62. Bender J, Yue R, To M, Deacken L, Jadad A. A lot of action, but not in the right direction: systematic review and content analysis of smartphone applications for the prevention, detection, and management of cancer. J Med Internet Res 2013 Dec 23;15(12):e287 [FREE Full text] [doi: 10.2196/jmir.2661] [Medline: 24366061]
63. Choi J, Noh GY, Park DJ. Smoking cessation apps for smartphones: content analysis with the self-determination theory. J Med Internet Res 2014 Feb 12;16(2):e44 [FREE Full text] [doi: 10.2196/jmir.3061] [Medline: 24521881]
64. Eng DS, Lee JM. The promise and peril of mobile health applications for diabetes and endocrinology. Pediatr Diabetes 2013 Jun;14(4):231-238 [FREE Full text] [doi: 10.1111/pedi.12034] [Medline: 23627878]
65. Hundert A, Huguet A, McGrath P, Stinson J, Wheaton M. Commercially available mobile phone headache diary apps: a systematic review. JMIR Mhealth Uhealth 2014 Aug 19;2(3):e36 [FREE Full text] [doi: 10.2196/mhealth.3452] [Medline: 25138438]
66. Kassianos A, Emery J, Murchie P, Walter F. Smartphone applications for melanoma detection by community, patient and generalist clinician users: a review. Br J Dermatol 2015 Jun;172(6):1507-1518. [doi: 10.1111/bjd.13665] [Medline: 25600815]
67. Kumar N, Khunger M, Gupta A, Garg N. A content analysis of smartphone-based applications for hypertension management. J Am Soc Hypertens 2015 Feb;9(2):130-136. [doi: 10.1016/j.jash.2014.12.001] [Medline: 25660364]
68. Mobasheri MH, Johnston M, King D, Leff D, Thiruchelvam P, Darzi A. Smartphone breast applications - what's the evidence? Breast 2014 Oct;23(5):683-689. [doi: 10.1016/j.breast.2014.07.006] [Medline: 25153432]
69. Nicholas J, Larsen M, Proudfoot J, Christensen H. Mobile apps for bipolar disorder: a systematic review of features and content quality. J Med Internet Res 2015 Aug 17;17(8):e198 [FREE Full text] [doi: 10.2196/jmir.4581] [Medline: 26283290]
70. Paglialonga A, Pinciroli F, Tognola G. The ALFA4Hearing model (At-a-Glance Labeling for Features of Apps for Hearing Health Care) to characterize mobile apps for hearing health care. Am J Audiol 2017 Oct 12;26(3S):408-425. [doi: 10.1044/2017_AJA-16-0132] [Medline: 29049624]
71. Pagoto S, Schneider K, Jojic M, DeBiasse M, Mann D. Evidence-based strategies in weight-loss mobile apps. Am J Prev Med 2013 Nov;45(5):576-582. [doi: 10.1016/j.amepre.2013.04.025] [Medline: 24139770]

72. Pandey A, Hasan S, Dubey D, Sarangi S. Smartphone apps as a source of cancer information: changing trends in health information-seeking behavior. J Cancer Educ 2013 Mar;28(1):138-142. [doi: 10.1007/s13187-012-0446-9] [Medline: 23275239]
73. Plaza I, Demarzo MM, Herrera-Mercadal P, García-Campayo J. Mindfulness-based mobile applications: literature review and analysis of current features. JMIR Mhealth Uhealth 2013 Nov 01;1(2):e24 [FREE Full text] [doi: 10.2196/mhealth.2733] [Medline: 25099314]
74. Radovic A, Vona PL, Santostefano AM, Ciaravino S, Miller E, Stein BD. Smartphone applications for mental health. Cyberpsychol Behav Soc Netw 2016 Jul;19(7):465-470 [FREE Full text] [doi: 10.1089/cyber.2015.0619] [Medline: 27428034]
75. Robustillo Cortés Mde L, Cantudo Cuenca MR, Morillo Verdugo R, Calvo Cidoncha E. High quantity but limited quality in healthcare applications intended for HIV-infected patients. Telemed J E Health 2014 Aug;20(8):729-735. [doi: 10.1089/tmj.2013.0262] [Medline: 24849001]
76. Sama PR, Eapen ZJ, Weinfurt KP, Shah BR, Schulman KA. An evaluation of mobile health application tools. JMIR Mhealth Uhealth 2014 May 01;2(2):e19 [FREE Full text] [doi: 10.2196/mhealth.3088] [Medline: 25099179]
77. Shen N, Levitan M, Johnson A, Bender J, Hamilton-Page M, Jadad A, et al. Finding a depression app: a review and content analysis of the depression app marketplace. JMIR Mhealth Uhealth 2015 Feb 16;3(1):e16 [FREE Full text] [doi: 10.2196/mhealth.3713] [Medline: 25689790]
78. Ubhi HK, Kotz D, Michie S, van Schayck OC, Sheard D, Selladurai A, et al. Comparative analysis of smoking cessation smartphone applications available in 2012 versus 2014. Addict Behav 2016 Jul;58:175-181 [FREE Full text] [doi: 10.1016/j.addbeh.2016.02.026] [Medline: 26950256]
79. Wearing JR, Nollen N, Befort C, Davis AM, Agemy CK. iPhone app adherence to expert-recommended guidelines for pediatric obesity prevention. Child Obes 2014 Apr;10(2):132-144 [FREE Full text] [doi: 10.1089/chi.2013.0084] [Medline: 24655230]
80. Weaver ER, Horyniak DR, Jenkinson R, Dietze P, Lim MS. "Let's get wasted!" and other apps: characteristics, acceptability, and use of alcohol-related smartphone applications. JMIR Mhealth Uhealth 2013 Jun 25;1(1):e9 [FREE Full text] [doi: 10.2196/mhealth.2709] [Medline: 25100681]
81. Yang C, Maher JP, Conroy DE. Implementation of behavior change techniques in mobile applications for physical activity. Am J Prev Med 2015 Apr;48(4):452-455. [doi: 10.1016/j.amepre.2014.10.010] [Medline: 25576494]
82. Aguirre R, McCoy M, Roan M. Development guidelines from a study of suicide prevention mobile spplications (Apps). J Technol Hum Serv 2013 Jul;31(3):269-293. [doi: 10.1080/15228835.2013.814750]
83. Nguyen E, Bugno L, Kandah C, Plevinsky J, Poulopoulos N, Wojtowicz A, et al. Is there a good app for that? Evaluating m-Health apps for strategies that promote pediatric medication adherence. Telemed J E Health 2016 Dec;22(11):929-937. [doi: 10.1089/tmj.2015.0211] [Medline: 27070837]
84. Derbyshire E, Dancey D. Smartphone Medical Applications for Women's Health: What Is the Evidence-Base and Feedback? Int J Telemed Appl 2013;2013 [FREE Full text] [doi: 10.1155/2013/782074] [Medline: 24454354]
85. Ferron J, Brunette M, Geiger P, Marsch L, Adachi-Mejia A, Bartels S. Mobile phone apps for smoking cessation: quality and usability among smokers with psychosis. JMIR Hum Factors 2017 Mar 03;4(1):e7 [FREE Full text] [doi: 10.2196/humanfactors.5933] [Medline: 28258047]
86. Vaz CL, Suthar AG, Pousti BT, Aye SM, Williams KJ, Zhao H. A smartphone app–based lifestyle intervention promotes weight loss—results of a prospective, randomized, controlled clinical trial (RCT). Diabetes 2018;67(Supplement 1). [doi: 10.2337/db18-2070-P]
87. Powell A, Torous J, Chan S, Raynor G, Shwarts E, Shanahan M, et al. Interrater reliability of mHealth app rating measures: analysis of top depression and smoking cessation apps. JMIR Mhealth Uhealth 2016 Feb 10;4(1):e15 [FREE Full text] [doi: 10.2196/mhealth.5176] [Medline: 26863986]

Abbreviations

ABACUS: App Behavior Change Scale
ICC: interclass coefficient
MARS: Mobile App Rating Scale
RCT: randomized controlled trial
VicHealth: Victorian Health Promotion Foundation

Edited by G Eysenbach; submitted 24.05.18; peer-reviewed by K Stawarz, A Crandall, A Nguyen, H Potts; comments to author 04.08.18; revised version received 29.10.18; accepted 29.10.18; published 25.01.19.

Please cite as:
McKay FH, Slykerman S, Dunn M
The App Behavior Change Scale: Creation of a Scale to Assess the Potential of Apps to Promote Behavior Change
JMIR Mhealth Uhealth 2019;7(1):e11130
URL: http://mhealth.jmir.org/2019/1/e11130/
doi:10.2196/11130
PMID:30681967

Review

Usefulness of Wearable Cameras as a Tool to Enhance Chronic Disease Self-Management: Scoping Review

Ralph Maddison[1], BHSc, MSc (Hons), PhD; Susie Cartledge[1,2], BN (Hons), PhD; Michelle Rogerson[1], BSc (Hons), PhD, MAppPsych; Nicole Sylvia Goedhart[3], BSc, MSc; Tarveen Ragbir Singh[1], BSc (Hons), MSc; Christopher Neil[4,5], MBBS, FRACP, PhD; Dinh Phung[6], BSc (Hons), PhD; Kylie Ball[1], PhD, BA (Hons)

[1]Institute for Physical Activity and Nutrition, School of Exercise and Nutrition Sciences, Deakin University, Geelong, Australia

[2]Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia

[3]Athena Institute for Research on Innovation and Communication in Health and Life Sciences, Faculty of Science, VU University, Amsterdam, Netherlands

[4]Department of Medicine, Western Health, University of Melbourne, Melbourne, Australia

[5]Department of Cardiology, Western Health, Melbourne, Australia

[6]Centre for Pattern Recognition and Data Analytics, Deakin University, Geelong, Australia

Corresponding Author:
Ralph Maddison, BHSc, MSc (Hons), PhD
Institute for Physical Activity and Nutrition
School of Exercise and Nutrition Sciences
Deakin University
Locked Bag 20000
Geelong, 3220
Australia
Phone: 61 420232459
Email: ralph.maddison@deakin.edu.au

Abstract

Background: Self-management is a critical component of chronic disease management and can include a host of activities, such as adhering to prescribed medications, undertaking daily care activities, managing dietary intake and body weight, and proactively contacting medical practitioners. The rise of technologies (mobile phones, wearable cameras) for health care use offers potential support for people to better manage their disease in collaboration with their treating health professionals. Wearable cameras can be used to provide rich contextual data and insight into everyday activities and aid in recall. This information can then be used to prompt memory recall or guide the development of interventions to support self-management. Application of wearable cameras to better understand and augment self-management by people with chronic disease has yet to be investigated.

Objective: The objective of our review was to ascertain the scope of the literature on the use of wearable cameras for self-management by people with chronic disease and to determine the potential of wearable cameras to assist people to better manage their disease.

Methods: We conducted a scoping review, which involved a comprehensive electronic literature search of 9 databases in July 2017. The search strategy focused on studies that used wearable cameras to capture one or more modifiable lifestyle risk factors associated with chronic disease or to capture typical self-management behaviors, or studies that involved a chronic disease population. We then categorized and described included studies according to their characteristics (eg, behaviors measured, study design or type, characteristics of the sample).

Results: We identified 31 studies: 25 studies involved primary or secondary data analysis, and 6 were review, discussion, or descriptive articles. Wearable cameras were predominantly used to capture dietary intake, physical activity, activities of daily living, and sedentary behavior. Populations studied were predominantly healthy volunteers, school students, and sports people, with only 1 study examining an intervention using wearable cameras for people with an acquired brain injury. Most studies highlighted technical or ethical issues associated with using wearable cameras, many of which were overcome.

Conclusions: This scoping review highlighted the potential of wearable cameras to capture health-related behaviors and risk factors of chronic disease, such as diet, exercise, and sedentary behaviors. Data collected from wearable cameras can be used as an adjunct to traditional data collection methods such as self-reported diaries in addition to providing valuable contextual

information. While most studies to date have focused on healthy populations, wearable cameras offer promise to better understand self-management of chronic disease and its context.

(JMIR Mhealth Uhealth 2019;7(1):e10371) doi:10.2196/10371

KEYWORDS

eHealth; review; cameras; life-logging; lifestyle behavior; chronic disease

Introduction

Background

Noncommunicable diseases, principally cardiovascular diseases, cancer, diabetes, and chronic respiratory diseases, are the leading causes of death in most developed countries, contributing to 60% of all deaths globally [1,2]. However, improved medical management has extended the life expectancy of people with chronic disease. Therefore, treatment goals for many noncommunicable diseases include mitigating exacerbations of the disease to reduce symptoms and prevent hospitalizations [3]. For people with chronic disease, appropriate self-management is critical to maximize health, treatment benefits, and quality of life [4].

Self-management refers to an individual's engagement in undertaking and managing day-to-day tasks, making and sustaining lifestyle changes, and managing physical symptoms and mental health over the course of an illness [5,6]. This can include a host of activities, such as adhering to prescribed medications, undertaking daily care activities (eg, blood glucose monitoring, self-weighing, rehabilitation exercises, toileting activities), managing body weight (eg, reducing energy intake, increasing physical activity), and managing dietary intake (eg, limiting salt consumption). People with chronic disease need to be supported to engage in and maintain self-management, thus reducing symptoms and empowering them to manage their own health.

Improved self-management has been associated with reduced mortality, hospital admissions, and health care costs [7,8]. For example, a trial of nurse-led education to improve self-management in heart failure demonstrated significant reductions in the relative risk of cardiac (0.59, 95% CI 0.38-0.91) and heart failure-related (0.49, 95% CI 0.27-0.88) hospitalizations [9]. To maximize effective behavioral interventions, efforts must focus on understanding the challenges individuals face in managing the complex demands of their illness and the often multiple and competing conditions. New approaches (with low participant burden and cost) are needed to identify these challenges and effectively tailor interventions to match people's needs.

The rise of technologies for health care use, such as mobile phones and wearable cameras, offers the potential to facilitate self-management for people with chronic disease [8]. Visual "life-logging" is one such technology. It refers to the use of wearable cameras to digitally capture everyday life activities through first-person point-of-view images [10]. Wearable cameras gather data that accurately reflect the participant's real-world experiences and environments [11]. Self-reporting of behaviors is difficult and subject to underreporting (eg, dietary intake [12]), overreporting (eg, physical activity [13]), or simply forgetting activities undertaken or food consumed. Wearable cameras can be used to prompt recall and provide health care practitioners with valuable insight into people's daily behaviors and patterns. This information can assist with prompting subsequent behavior change and developing tailored self-management strategies with patients.

Wearable cameras have been used to assess dietary recall [14] and as an intervention to assist in accurate data collection for a range of activities, such as food purchasing [15], time use [16], sedentary behavior [17], and travel times [18], and to observe behavior changes in early-stage dementia [19]. A recent systematic review to assess the utility of camera images to assist in the assessment of dietary intake found that images can enhance self-report by revealing unreported foods and identify misreporting errors not captured by traditional methods alone [12]. While the feasibility of collecting and analyzing images is well documented across a range of behaviors [14,18], the application of this technology to better understand and augment self-management in people with chronic disease has yet to be investigated and should be considered [20]. For example, wearable cameras could be used for behavior change strategies such as increasing awareness and motivation around specific behaviors [20]. Specifically, wearable cameras could monitor the time and contexts in which individuals take prescribed medications or complete self-monitoring activities [20], such as measuring daily body weight. Using camera-assisted recall methods, this approach could facilitate conversations between patients and health care providers to better tailor self-management strategies.

Objective

The aim of this review was to ascertain in a human population of any age, with or without chronic disease, what is known about the potential use of wearable cameras for assisting with self-management of chronic disease, including self-management practices such as self-weighing and taking medication, as well as capturing lifestyle behaviors associated with chronic disease (eg, physical activity, sedentary behavior, diet, or smoking). To do this, we searched original research articles reporting studies using any methods, as well as review articles and published conference proceedings.

Methods

Rationale for a Scoping Review

Given the novelty of using wearable cameras for enhancing self-management, we considered a scoping review appropriate prior to undertaking a systematic review. Scoping reviews involve systematically searching and selecting literature to map key concepts and summarize a range of evidence to convey the

breadth and depth of a field [21,22]. These reviews can be used to examine the extent, range, and nature of research activities, determine the value in undertaking a full systematic review, summarize and disseminate research findings, or help identify gaps in the research literature [23]. Following current guidelines for conducting scoping reviews [21,23,24], once we identified the research question, we proceeded to (1) identify relevant studies, (2) select studies, and (3) collate, summarize, and report the results.

Identification of Relevant Studies

We conducted a comprehensive electronic literature search in July 2017 in the following databases: PsycINFO, MEDLINE, CINAHL, and SPORTDiscus (all through EBSCO), EMBASE (through OVID), Web of Science, ProQuest, ACM Digital Library, and Cochrane Library. We combined Medical Subject Headings and free terms to search for focused articles. We used a search string including the following search terms and derivatives for each database (see Multimedia Appendix 1 for the full search strategy): (1) wearable camera* OR life-logging OR SenseCam OR Narrative Clip OR GoPro OR Google Glass AND (2) chronic disease OR lifestyle OR lifestyle modification OR rehabilitation OR diet OR physical activity OR medication adherence OR fluid restriction OR smoking.

We also scanned the reference lists of records identified by the search for additional studies that met our inclusion criteria. For the purpose of this review, we were interested in studies that (1) used wearable cameras to capture one or more modifiable lifestyle risk factors associated with chronic disease (eg, physical activity, sedentary behavior, diet, or smoking), (2) used wearable cameras capture typical self-management behaviors (eg, taking medication, self-weighing), or (3) involved a population group with chronic disease (eg, cardiovascular disease, diabetes, respiratory disease). Wearable cameras can capture images actively and passively [25]; in this review, we excluded studies that only used active image capture, were constrained in context (eg, laboratory settings), or collected data for less than 1 day. We also excluded studies using videos due to concerns around battery life constraints (reducing wear time) and the increased difficulty of annotating and coding video images. The search strategy was not limited by study design or year, and it included conference proceedings and full articles but was limited to articles written in English. As a systematic literature review on the effect of wearable cameras on memory disorders had been conducted recently [26], we excluded articles that used wearable cameras in managing forms of dementia, such as Alzheimer disease.

Study Selection

We imported search results from the databases into the reference software package EndNote version X8 (Clarivate Analytics), which automatically removed most duplicates, with the remaining removed manually. Textbox 1 lists the inclusion criteria for the title and abstract screening. Figure 1 presents a flow diagram leading to the included studies.

Textbox 1. Criteria for article inclusion in the scoping review.

Inclusion criteria

Title-level screening:

- If the title of the article contained the following concepts:
  - Wearable cameras (eg, SenseCam, Narrative Clip, GoPro)
  - Life-logging

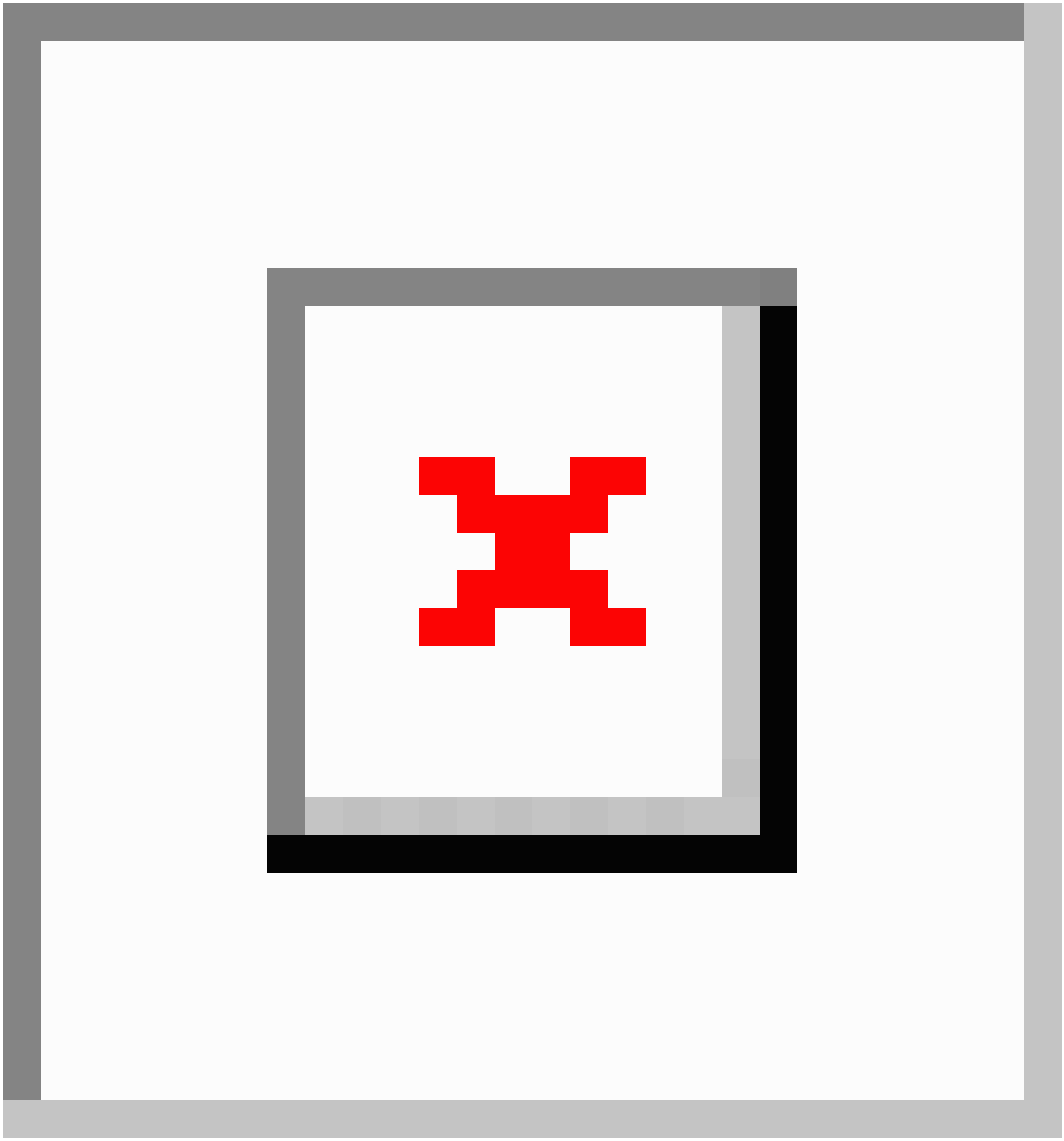
Abstract-level screening:

- If the article
  - Reported the use or effect of the wearable cameras or life-logging on a lifestyle behavior or chronic disease *or* measured lifestyle behavior with wearable cameras or life-logging *and*
  - Was written in English *and*
  - Reported qualitative or quantitative findings *and*
  - Passively captured images *and*
  - Measured free-living activities *and*
  - Used the wearable camera for ≥1 day

Exclusion criteria

- Articles focused on participants with memory disorders

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flowchart depicting the article selection process.



Data Extraction

Data extraction was performed by 4 reviewers (NSG, RM, MR, and SC) using a custom data extraction sheet (see Multimedia Appendix 2), which comprised 3 sections; article overview (eg, authors, title, year, country, journal, study design, study aim, camera device, interval of images taken), measuring lifestyle (eg, lifestyle behavior, aim to measure or to change behavior), and study details (eg, sample size and characteristics, intervention duration, presence of a control group, type of image annotation, data analysis, challenges or issues with cameras).

Summarizing and Reporting the Results

We categorized and described the included studies according to study characteristics, which included the country where the study was undertaken, characteristics of participants (eg, size of sample, age, health status, presence of chronic disease), study design or type, and behaviors measured using wearable cameras (eg, physical activity, dietary intake, sedentary behavior). We also classified studies according to the data collected (primary or secondary) and whether it was a review or discussion article.

Results

Study Characteristics

Multimedia Appendix 3 [14-17,25,27-46] presents the main results of primary and secondary data collection. Table 1 [11,12,47-50] presents study characteristics of review and discussion studies. Of the 31 studies identified, 22 analyzed

primary data and 3 performed a secondary analysis of existing data (Multimedia Appendix 3). These studies were published between 2010 and June 2017, with most primary or secondary data collection studies (22/25, 88%) published since 2013 (Figure 2). Our search strategy also identified 2 discussion articles, 2 reviews, and 2 descriptive articles (Table 1).

Studies undertaking primary or secondary data analysis (n=25) were predominantly feasibility or pilot studies (n=13, 52%) [15,18,25,27-32,34-36,43,46], followed by methodological studies (n=6, 24%) [37,39-41,44] and validation studies (n=4, 16%) [14,33,42,46]. There was 1 randomized controlled trial [38] conducted with acquired brain injury patients where camera images formed part of a health intervention. There was also 1 descriptive study, which described the context of sedentary time in older adults [17]. The majority of studies were conducted in the United States (n=8, 32%) [27,32,35-37,42-44], the United Kingdom (n=6, 24%) [15-17,28,31,45], and Ireland (n=4, 16%) [30,33,39,41]. A total of 3 studies (12%) [29,34,46] were international multicenter studies, while the remaining studies were from New Zealand (n=3, 12%) [14,25,40] and Spain (n=1, 4%) [38].

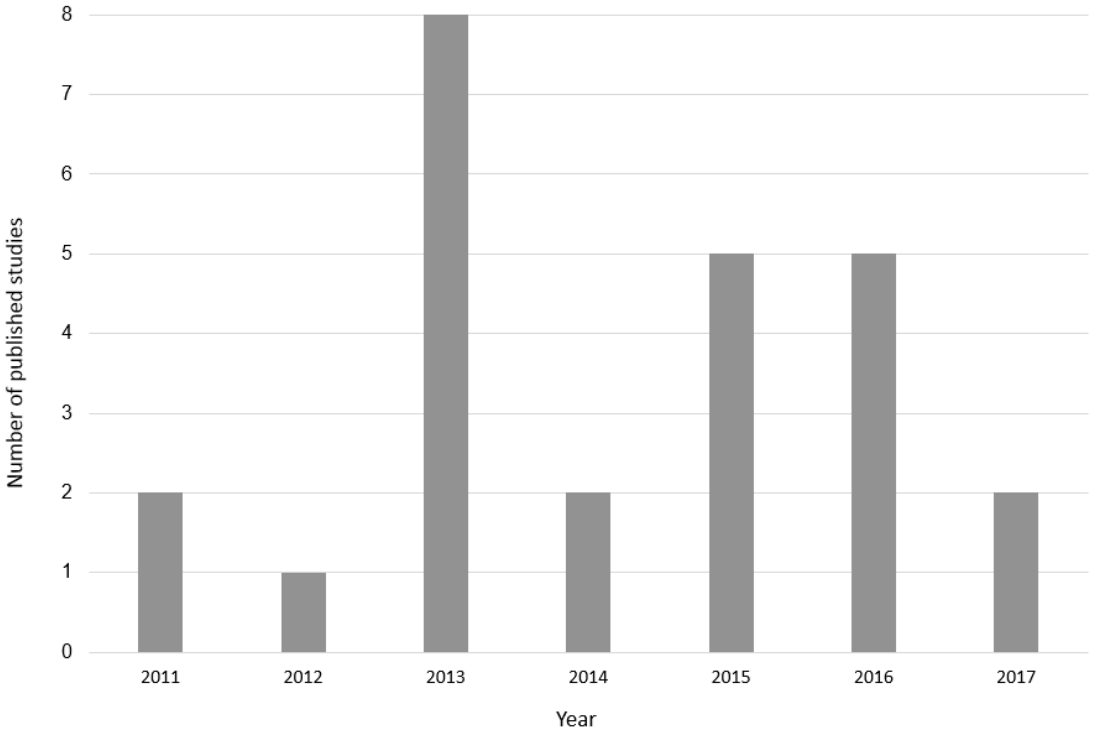
Of the remaining 6 review or discussion studies (Table 1), 2 reviews focused on dietary assessment [14,47]. One descriptive study presented a research program and a new device for recording food intake [48], while the other described a wrist-worn device for measuring physical activity [49]. The final 2 studies were discussion articles examining the utility of using wearable cameras for assessing lifestyle behaviors (eg, sedentary and nutrition behaviors, television viewing), their challenges, and ethical-related issues [11,50].

Feasibility and pilot studies [15,18,25,27-32,34,36,43,46] focused predominantly on assessing the feasibility of using wearable cameras to capture information on dietary intake or food consumption, physical activity, and sedentary behavior.

Table 1. Characteristics of review and discussion studies (n=6).

Study	Design	Camera device	Aim	Findings
Dietary intake				
Boushey, 2016 [47]	Narrative review	Multiple	Overview of image-assisted and image-based methods, including implementation and detail on image-based food records.	Accuracy of dietary assessment was improved. Underreporting was reduced in all included studies.
Gemming, 2015 [12]	Systematic review	Multiple	Examination of studies that evaluated or validated image-assisted methods of assessing dietary energy intake.	Evidence regarding the validity of image-assisted methods of dietary assessment was limited. Self-reported data could be enhanced with images, providing a primary record of dietary intake to obtain valid estimates of energy intake.
Sun, 2010 [48]	Descriptive (research program)	Prototype	Description of emerging science for objective methods of dietary assessment.	A research program to automatically record food intake was described. Hardware (camera, reference lights, accelerometer, microphone, global positioning system, and data processor) had been developed; software (enabling privacy protection, video segmentation, food identification, portion size analysis, and nutrient and calorie determinations) was under development.
Physical activity				
Maekawa, 2012 [49]	Descriptive (device)	Wrist-Sense	Description of the implementation of a wrist-worn sensor device.	The camera supplied information on what object the wearer was holding, which related strongly to the activity the wearer was performing.
Activities of daily living				
Doherty, 2013 [11]	Discussion	Multiple	Assessment of the utility of wearable cameras for objectively measuring lifestyle behaviors.	The use of wearable cameras was considered appropriate to understand lifestyle behaviors.
Loveday, 2016 [50]	Discussion	Multiple	Discussion of the objective measurement of context and illustration of the utility of quantifying context using example data from 3 ongoing studies.	Devices that provide contextual information, such as wearable cameras, location monitors, and proximity sensors, provided researchers with a more comprehensive picture of behavior.

Figure 2. Number of published primary and secondary data studies by year of publication.



Studies classified as validation studies [14,33,42,46] assessed the validity of wearable cameras either compared with or in combination with other measurement approaches, including self-reported diaries and questionnaires, as well as objective measurement techniques such as accelerometers and doubly labelled water. For the methodological studies [35,37,39-41,44], outcomes focused on testing software or analytic approaches to classify and analyze generated images.

Sample Characteristics

Where sample size was relevant and reported (n=24), numbers were generally small (Figure 3): 8 studies (33%) had fewer than 20 participants (minimum of 5) [16,27-29,31,32,38,41], 13 studies (54%) had between 20 and 50 participants [14,15,17,25,30,33,35,37-39,42-44], and only 3 studies (13%) [34,36,46] had more than 50 participants (maximum of 84). Participants were predominantly healthy adults (n=16, 64%) [14,16,17,25,27-29,31,32,34-36,39,41,44,46], with some studies including athletes (n=4, 16%) [30,33,42,43] or school children (n=2, 8%) [15,40]. One study examined physical activity and sedentary time in a sample of older women [37]. No studies identified in our review recruited participants with a chronic disease; however, 1 study included participants with an acquired brain injury [38]. The majority of studies included convenience samples, often recruited through universities; thus, participants were more likely to be well educated with higher socioeconomic status.

Types and Uses of Wearable Cameras

The SenseCam (subsequently called Vicon Revue, then Vicon Autographer) was the most frequently used brand of wearable camera (n=18, 72%). The remaining studies used a similar technology, worn around the neck on a lanyard or on the wrist. In 22 studies (88%), participants wore a camera for 3 to 7 days; in 1 study, participants wore the cameras for 7 weeks [38]. Studies used wearable cameras to measure, observe, or validate specific behaviors related to dietary intake (n=10, 40%) [14,15,25,27-32,48]; physical activity (n=6, 24%) [33-37,49]; a diverse range of daily activities such as travel, work, and exposure to food marketing (n=5, 20%) [16,38-41]; sedentary behavior (n=4, 16%) [17,42-44]; and travel behaviors such as walking, cycling, or motorized transport (n=2, 8%) [45,46].

Our search found only 1 randomized controlled trial [38], which combined the SenseCam and Actiheart device for goal management training over 7 weeks for 16 people with an acquired brain injury. Results showed that goal management training plus the addition of viewing SenseCam images resulted in greater improvements in cognitive skills compared with goal management training alone.

Issues Associated With Wearable Cameras

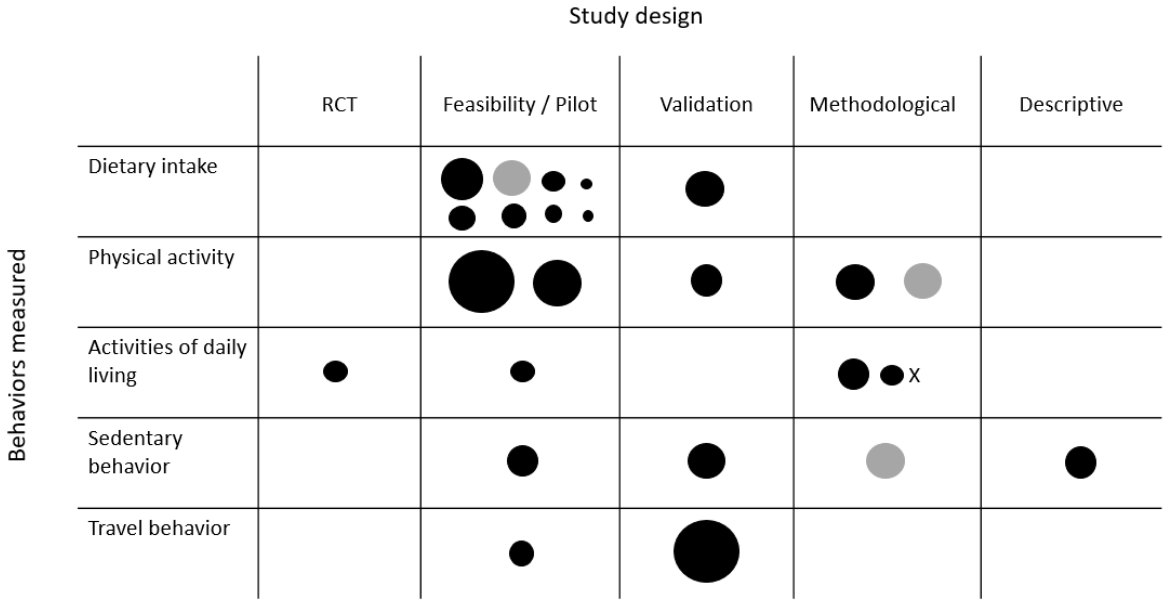
Findings from the studies identified in this review highlighted a variety of technical and personal constraints associated with using this technology. These included person-related factors such as feeling self-conscious while wearing the device, forgetting to put the device on, and privacy and ethical concerns for both the person wearing the device and those whose images were being taken. Technical issues noted included short battery life, challenges in analyzing and classifying very large numbers of images captured by the device, and lack of consistent high-quality images.

Utility of Wearable Cameras for Self-Management

Notwithstanding the technical issues reported in many studies, most highlighted that wearable cameras offered a feasible and acceptable method for measuring specific behaviors, namely food consumption, physical activity, and sedentary behaviors,

and identifying activities of daily living, especially when used alongside more traditional methods of assessment such as logbooks, diaries, and self-recall questionnaires.

Figure 3. Behaviors measured, study design, and sample size of primary and secondary data collection studies (n=25). Studies are represented by black and gray circles proportional to the study sample size. X denotes 1 study that did not report sample size. Black circles represent studies that used primary data. Gray circles represent studies that used secondary data for analysis. RCT: randomized controlled trial.



Based on this review, the most common use of wearable cameras was for prompting recall of dietary intake and to identify unreported items [14,15,27,28,30]. In these studies, participants were asked to recall their dietary intake using 24-hour dietary recall reported questionnaires [14,30] or interview [28] techniques alone and then again in combination with reviewing SenseCam images. All 3 studies showed that viewing the images increased self-reported energy intake by 12.5% [28], 8% (men) and 6% (women) [14], or 10.7% to 17.7% [30] compared with 24-hour recall alone. These results suggested that a more accurate estimate of total energy intake is possible by combining the SenseCam data with conventional food diaries.

Wearable camera data also offered promise for dietary assessment [25,30,31] and to augment understanding of eating moments or episodes by providing spatial and contextual information [25,29]. Data from 1 earlier validation study in 2013 identified that eating moments could be identified with a high degree of accuracy (89.7%) [32].

The next most common use of wearable cameras was to augment measurement of physical activity-related behaviors, including travel behaviors such as cycling and using public transport [45,46]. Kelly et al [45] tested the feasibility of measuring travel to school in 2012, followed by validating travel data against a travel diary in 2014 [46] with 84 adults using a variety of transportation modes. Both studies found that self-reported journey data were accurate at the group level but imprecise at the individual level.

Combined with traditional measurement techniques (eg, accelerometry), wearable cameras were used to provide context (type and location) for physical activity [33,34] and sedentary behaviors (eg, location, activity undertaken while sedentary) [17,43]. However, data analysis in these areas is still in its infancy, with 3 methodological studies describing the challenges of developing algorithms [37], machine learning techniques [35], and classification techniques [44] for processing wearable camera images.

The complexity of identifying activities of daily living using wearable cameras was made evident by the high number of methodological studies (3 of 5 studies) in this category [39-41]. These studies aimed to identify personal traits [39], everyday activities [41], and children's exposure to food marketing and other public health issues (eg, tobacco exposure) [40]. These studies analyzed millions of images (from 1.4 million [40] to 3.5 million [39]) using both manual and software-assisted coding methods.

Using camera images to prompt recall during interviews, 1 study demonstrated the feasibility and acceptability of using wearable cameras to reconstruct daily use of time [16]. We found a single randomized clinical trial, which used daily camera images to enhance self-management by people with an acquired brain injury [38]. This study showed that, in addition to goal management training, reviewing SenseCam images in the intervention arm resulted in health improvements (eg, planning, self-monitoring, error detection), both quantitatively (increased effect sizes between groups) and qualitatively (increased engagement). This was the only study to use wearable cameras as a health intervention and highlighted the potential of extending this approach beyond acquired brain injury to other clinical conditions (such as cardiovascular disease, diabetes, and peritoneal dialysis).

Discussion

Principal Findings

We conducted this scoping review to ascertain what is known about the use of wearable cameras for self-management by people with a chronic disease and to determine the potential for using this approach for helping people to better manage their chronic disease. We identified 31 studies, with most published since 2013, highlighting the infancy of the field. Based on the available evidence, no studies used wearable cameras for self-management activities, but the searched literature demonstrated that wearable cameras have utility for capturing health-related behaviors and lifestyle risk factors of chronic disease, such as dietary intake, physical activity, and sedentary behavior. Thus far, wearable cameras have been predominantly used to augment existing measurement approaches of these behaviors by assisting with recall of activities (eg, eating, sitting at a computer) or have been used to provide contextual information on these behaviors (eg, walking in a park, or associated behaviors such as sitting at a table eating food). These lifestyle behaviors are important for the prevention and management of many chronic diseases but, as we have demonstrated in this review, wearable cameras have yet to be used by people with existing disease. Furthermore, other than Silva and colleagues' [26] study of the use of wearable cameras to assist with memory and recall for those affected by Alzheimer disease, there is a paucity of research on the use of wearable cameras for capturing other disease self-management practices such as taking medication, self-weighing, and undertaking home dialysis.

The majority of included studies focused on young healthy individuals, with a subset of studies investigating wearable camera use with athletes [30,33,42,43] and school children [15,40]. One study [37] investigated walking and sedentary time in older women, a group that may be at higher risk of chronic disease. We could identify only 1 study [38] that used the SenseCam as an intervention tool for people with an acquired brain injury, which highlighted the potential of this approach to improve health outcomes of those with an injury or chronic health condition. At the time of our review, we could not identify any further studies investigating the use of wearable cameras in other clinical populations.

In terms of diet, our findings are consistent with a systematic review (included in this review) [12], which assessed the utility of images captured using both handheld devices and wearable cameras for supporting traditional self-report methods or to provide a primary record of dietary intake. Of the 13 included studies in the review by Gemming et al [12], 10 used handheld devices, while 3 studies used wearable cameras. Findings from the systematic review showed that images enhanced self-report by revealing unreported foods and identifying misreporting error. In addition, when used as a primary record of dietary intake, images provided valid estimates of energy intake [14,30]. However, both image quality and camera position influenced the quality of data collected, which needs to be considered in future studies involving dietary behaviors [14]. Our scoping review also highlighted studies that investigated the use of wearable cameras for capturing contextual factors and other behaviors associated with food consumption and motivations to eat [25,29].

Privacy and ethical issues associated with captured images need to be considered in future research. Doherty et al [11] argued that this method is useful for observing and understanding participant behavior and could be used as a lifestyle behavior change catalyst; however, only 1 study [38] identified in this review used it for this purpose. Given issues of participant and third-party privacy, ethical frameworks have been developed to guide the use of wearable cameras. One such article [51] presented a research checklist to be used for studies with automated wearable cameras. The checklist addressed informed consent, privacy and confidentiality, nonmaleficence, and autonomy of third parties. Future studies using wearable cameras for self-management in clinical populations need to closely consider these privacy and ethical issues.

Technical issues associated with using wearable cameras present numerous challenges, which need to be considered when undertaking future research. These include issues such as participant compliance and adherence with wearing the device due to factors such as limited battery life and the need to recharge devices, not wanting to wear them, and being self-conscious around others. However, the biggest issues that potentially limit the use of these technologies in their current form relate to lack of image clarity, the difficulty of correctly coding images to reflect specific behaviors, and processing vast amounts of image data [11,41,44]. Of the studies included in this review, 4 developed data algorithms and object classification methodology [37,39,41,44]. As highlighted by these studies, using machine learning techniques, such as deep convolutional neural networks, it is possible to undertake automated image recognition [52,53]. Google has a pretrained model for ImageNet, consisting of a collection 10 million images depicting 1000 object categories. Using this approach, it is possible to determine corresponding probabilities of correct identification of images according to specific labels. The precision of this approach and others, while they are useful, remain variable. Manual processing (reviewing and coding) of images is time and resource intensive. Analysis of images from wearable cameras for summary purposes or for creating viable interventions remains challenging. Thus, the utility of wearable cameras with multiple participants over extended periods of times (eg, 6 months) is unclear. Finally, an ongoing challenge is that many of the devices used in the included studies are no longer available (eg, SenseCam/Vicon and Narrative Clip). In this review, we did not include mobile phone–enabled cameras; however, given their ubiquity, they do offer a viable solution for capturing some activities such as dietary intake, but they do not provide the passive data collection afforded by current wearable camera technologies.

Strengths and Limitations

This is, to our knowledge, the first scoping review to examine the use of wearable cameras for self-management. We conducted an extensive search of the literature using 9 databases from both health and sports science. We did not impose limits on study design and therefore included a diverse range of studies from

over 6 countries. On the basis of the number of published studies identified in our original search, we chose not to extend the search to include gray literature. Nor did we perform a quality assessment, as it is suggested that this is not in the remit of scoping reviews [54]. More recently, a consultation phase has been suggested to be included in scoping reviews [24]. This phase involves formally presenting the findings to knowledge users and community members to gain collective experience, expertise, and knowledge on the chosen subject [24]. While recommended, this approach is not mandatory; thus, we did not include a consultation phase as part of this review.

Future Research

Findings from this review suggest that there is interest in the use of wearable cameras to assess lifestyle behaviors, particularly diet, physical activity, and sedentary behavior. In its current form, this method has been used to augment existing measurement techniques through validation, reduction of error associated with recall of behaviors, and provision of rich contextual information. The lessons learned from this research are important if we are to better evaluate these behaviors and empower and support patients with self-management.

There appears to be considerable opportunity to use wearable cameras to specifically assess self-management behaviors and to apply this method to a host of clinical conditions. For example, heart failure is a chronic condition, often with a variable clinical course and frequent exacerbations of symptoms [3]. Appropriate self-management is critical to maximize treatment benefits [4]. To maximize effective behavioral interventions, efforts must focus on understanding the challenges individuals face in managing the complex demands of their illness. Wearable cameras could be used to capture the ecological context in which people manage this disease and identify the extent to which people adhere to key self-management practices (eg, taking medication, daily weighing, fluid restriction, salt consumption in foods). If these behaviors could be correctly identified, then more tailored interventions to match people's needs could be developed. For example, participants with a new diagnosis of heart failure could wear a camera for 1 to 4 weeks. On the patient's return to an outpatient clinic, a nurse specialist or other health professional could, by using software, review images alongside the individual to identify self-management practices and offer suggestions for improvement. Such an approach could easily be applied to other scenarios, such as peritoneal dialysis or stroke rehabilitation. Future research is needed to determine the feasibility of such an approach, including whether a chronic disease population would be able to wear and maintain the device for the duration of the intervention and, therefore, how a wearable camera would fit into their lives and disease management [55]. We would expect that the simple nature of these devices would not add additional burden to people with chronic disease and their caregivers, but future research is required [56].

Conclusion

This scoping review highlighted the use of wearable cameras for the assessment of lifestyle-related behaviors (in particular diet and physical activity) among healthy, adult populations; however, none of the studies specifically focused on self-management behaviors by people with chronic disease or in clinical settings. The advanced capabilities of wearable camera technologies, when considered alongside the gap in the evidence base and early findings of the usefulness of cameras in other populations identified here, all point to the promising potential of this approach and the need for further investigation in clinical populations.

Acknowledgments

This research was conducted as part of a Heart Foundation Vanguard Grant, Australia (101348). KB was supported by a Principal Research Fellowship from the Australian National Health and Medical Research Council (ID 1042442). The contents of this paper are the responsibility of the authors and do not necessarily reflect the views of the funding bodies.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy.

[PDF File (Adobe PDF File), 338KB - mhealth_v7i1e10371_app1.pdf]

Multimedia Appendix 2

Data extraction form.

[PDF File (Adobe PDF File), 316KB - mhealth_v7i1e10371_app2.pdf]

Multimedia Appendix 3

Characteristics of primary and secondary data collection studies.

[PDF File (Adobe PDF File), 362KB - mhealth_v7i1e10371_app3.pdf]

References

1. World Health Organization. Non-communicable diseases factsheet. Geneva, Switzerland: WHO; 2015. URL: http://www.who.int/en/news-room/fact-sheets/detail/noncommunicable-diseases [accessed 2018-11-02] [WebCite Cache ID 73dC3uADD]
2. Abegunde DO, Mathers CD, Adam T, Ortegon M, Strong K. The burden and costs of chronic diseases in low-income and middle-income countries. Lancet 2007 Dec 08;370(9603):1929-1938. [doi: 10.1016/S0140-6736(07)61696-1] [Medline: 18063029]
3. Hobbs FDR, Kenkre JE, Roalfe AK, Davis RC, Hare R, Davies MK. Impact of heart failure and left ventricular systolic dysfunction on quality of life: a cross-sectional study comparing common chronic cardiac and medical disorders and a representative adult population. Eur Heart J 2002 Dec;23(23):1867-1876. [Medline: 12445536]
4. Gardetto NJ. Self-management in heart failure: where have we been and where should we go? J Multidiscip Healthc 2011 Mar 31;4:39-51 [FREE Full text] [doi: 10.2147/JMDH.S8174] [Medline: 21544247]
5. Riegel B, Moser DK, Anker SD, Appel LJ, Dunbar SB, Grady KL, American Heart Association Council on Cardiovascular Nursing, American Heart Association Council on Clinical Cardiology, American Heart Association Council on Nutrition, Physical Activity, and Metabolism, American Heart Association Interdisciplinary Council on Quality of Care and Outcomes Research. State of the science: promoting self-care in persons with heart failure: a scientific statement from the American Heart Association. Circulation 2009 Sep 22;120(12):1141-1163 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.192628] [Medline: 19720935]
6. McManus RJ, Mant J, Franssen M, Nickless A, Schwartz C, Hodgkinson J, TASMINH4 investigators. Efficacy of self-monitored blood pressure, with or without telemonitoring, for titration of antihypertensive medication (TASMINH4): an unmasked randomised controlled trial. Lancet 2018 Mar 10;391(10124):949-959 [FREE Full text] [doi: 10.1016/S0140-6736(18)30309-X] [Medline: 29499873]
7. Savard LA, Thompson DR, Clark AM. A meta-review of evidence on heart failure disease management programs: the challenges of describing and synthesizing evidence on complex interventions. Trials 2011;12:194 [FREE Full text] [doi: 10.1186/1745-6215-12-194] [Medline: 21846340]
8. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
9. Koelling TM, Johnson ML, Cody RJ, Aaronson KD. Discharge education improves clinical outcomes in patients with chronic heart failure. Circulation 2005 Jan 18;111(2):179-185. [doi: 10.1161/01.CIR.0000151811.53450.B8] [Medline: 15642765]
10. Gurrin C, Smeaton AF, Doherty AR. Lifelogging: personal big data. Found Trends Inf Retrieval 2014;8(1):1-125 [FREE Full text]
11. Doherty AR, Hodges SE, King AC, Smeaton AF, Berry E, Moulin CJA, et al. Wearable cameras in health: the state of the art and future possibilities. Am J Prev Med 2013 Mar;44(3):320-323. [doi: 10.1016/j.amepre.2012.11.008] [Medline: 23415132]
12. Gemming L, Utter J, Ni Mhurchu C. Image-assisted dietary assessment: a systematic review of the evidence. J Acad Nutr Diet 2015 Jan;115(1):64-77. [doi: 10.1016/j.jand.2014.09.015] [Medline: 25441955]
13. Welk G. Physical Activity Assessments for Health-Related Research. Champaign, IL: Human Kinetics; 2002.
14. Gemming L, Rush E, Maddison R, Doherty A, Gant N, Utter J, et al. Wearable cameras can reduce dietary under-reporting: doubly labelled water validation of a camera-assisted 24 h recall. Br J Nutr 2015 Jan 28;113(2):284-291. [doi: 10.1017/S0007114514003602] [Medline: 25430667]
15. Cowburn G, Matthews A, Doherty A, Hamilton A, Kelly P, Williams J, et al. Exploring the opportunities for food and drink purchasing and consumption by teenagers during their journeys between home and school: a feasibility study using a novel method. Public Health Nutr 2016 Jan;19(1):93-103. [doi: 10.1017/S1368980015000889] [Medline: 25874731]
16. Kelly P, Thomas E, Doherty A, Harms T, Burke Ó, Gershuny J, et al. Developing a method to test the validity of 24 hour time use diaries using wearable cameras: a feasibility pilot. PLoS One 2015;10(12):e0142198 [FREE Full text] [doi: 10.1371/journal.pone.0142198] [Medline: 26633807]
17. Leask CF, Harvey JA, Skelton DA, Chastin SF. Exploring the context of sedentary behaviour in older adults (what, where, why, when and with whom). Eur Rev Aging Phys Act 2015;12:4 [FREE Full text] [doi: 10.1186/s11556-015-0146-7] [Medline: 26865868]
18. Kelly P, Doherty A, Berry E, Hodges S, Batterham AM, Foster C. Can we use digital life-log images to investigate active and sedentary travel behaviour? Results from a pilot study. Int J Behav Nutr Phys Act 2011;8:44 [FREE Full text] [doi: 10.1186/1479-5868-8-44] [Medline: 21599935]
19. Piasek P, Irving K, Smeaton AF. Case study in SenseCam use as an intervention technology for early-stage dementia. Int J Comput Healthc 2012;1(4):304. [doi: 10.1504/IJCIH.2012.051798]
20. Doherty A, Williamson W, Hillsdon M, Hodges S, Foster C, Kelly P. Influencing health-related behaviour with wearable camerastrategies & ethical considerations. New York, NY: ACM; 2013 Presented at: 4th International SenseCam & Pervasive Imaging Conference; Nov 18-19, 2013; San Diego, CA, USA p. 60-67. [doi: 10.1145/2526667.2526677]

21. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010;5:69 [FREE Full text] [doi: 10.1186/1748-5908-5-69] [Medline: 20854677]
22. Colquhoun HL, Levac D, O'Brien KK, Straus S, Tricco AC, Perrier L, et al. Scoping reviews: time for clarity in definition, methods, and reporting. J Clin Epidemiol 2014 Dec;67(12):1291-1294. [doi: 10.1016/j.jclinepi.2014.03.013] [Medline: 25034198]
23. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol 2005 Feb;8(1):19-32. [doi: 10.1080/1364557032000119616]
24. O'Brien KK, Colquhoun H, Levac D, Baxter L, Tricco AC, Straus S, et al. Advancing scoping study methodology: a web-based survey and consultation of perceptions on terminology, definition and methodological steps. BMC Health Serv Res 2016 Dec 26;16:305 [FREE Full text] [doi: 10.1186/s12913-016-1579-z] [Medline: 27461419]
25. Gemming L, Doherty A, Utter J, Shields E, Ni Mhurchu C. The use of a wearable camera to capture and categorise the environmental and social context of self-identified eating episodes. Appetite 2015 Sep;92:118-125. [doi: 10.1016/j.appet.2015.05.019] [Medline: 26002278]
26. Silva AR, Pinho MS, Macedo L, Moulin CJA. A critical review of the effects of wearable cameras on memory. Neuropsychol Rehabil 2018 Jan;28(1):117-141. [doi: 10.1080/09602011.2015.1128450] [Medline: 26732623]
27. Arab L, Estrin D, Kim DH, Burke J, Goldman J. Feasibility testing of an automated image-capture method to aid dietary recall. Eur J Clin Nutr 2011 Oct;65(10):1156-1162 [FREE Full text] [doi: 10.1038/ejcn.2011.75] [Medline: 21587282]
28. Gemming L, Doherty A, Kelly P, Utter J, Ni Mhurchu C. Feasibility of a SenseCam-assisted 24-h recall to reduce under-reporting of energy intake. Eur J Clin Nutr 2013 Oct;67(10):1095-1099. [doi: 10.1038/ejcn.2013.156] [Medline: 24002044]
29. Ng K, Shipp V, Mortier R, Benford S, Flintham M, Rodden T. Understanding food consumption lifecycles using wearable cameras. Pers Ubiquitous Comput 2015;19(7):1183-1195. [doi: 10.1007%2Fs00779-015-0871-y]
30. O'Loughlin G, Cullen SJ, McGoldrick A, O'Connor S, Blain R, O'Malley S, et al. Using a wearable camera to increase the accuracy of dietary analysis. Am J Prev Med 2013 Mar;44(3):297-301. [doi: 10.1016/j.amepre.2012.11.007] [Medline: 23415128]
31. Pettitt C, Liu J, Kwasnicki RM, Yang G, Preston T, Frost G. A pilot study to determine whether using a lightweight, wearable micro-camera improves dietary assessment accuracy and offers information on macronutrients and eating rate. Br J Nutr 2016 Jan 14;115(1):160-167. [doi: 10.1017/S0007114515004262] [Medline: 26537614]
32. Thomaz E, Parnami A, Essa I, Abowd GD. Feasibility of identifying eating moments from first-person images leveraging human computation. New York, NY: ACM; 2013 Presented at: 4th International SenseCam & Pervasive Imaging Conference; Nov 18-19, 2013; San Diego, CA, USA p. 26-33. [doi: 10.1145/2526667.2526672]
33. O Connor S, McCaffrey N, Whyte E, Moran K. The novel use of a SenseCam and accelerometer to validate training load and training information in a self-recall training diary. J Sports Sci 2016;34(4):303-310. [doi: 10.1080/02640414.2015.1050600] [Medline: 26032327]
34. Doherty AR, Kelly P, Kerr J, Marshall S, Oliver M, Badland H, et al. Using wearable cameras to categorise type and context of accelerometer-identified episodes of physical activity. Int J Behav Nutr Phys Act 2013;10:22 [FREE Full text] [doi: 10.1186/1479-5868-10-22] [Medline: 23406270]
35. Ellis K, Godbole S, Chen J, Marshall S, Lanckriet G, Kerr J. Physical activity recognition in free-living from body-worn sensors. New York, NY: ACM; 2013 Presented at: 4th International SenseCam & Pervasive Imaging Conference; Nov 18-19, 2013; San Diego, CA, USA p. 88-89. [doi: 10.1145/2526667.2526685]
36. Nebeker C, Lagare T, Takemoto M, Lewars B, Crist K, Bloss CS, et al. Engaging research participants to inform the ethical conduct of mobile imaging, pervasive sensing, and location tracking research. Transl Behav Med 2016 Dec;6(4):577-586. [doi: 10.1007/s13142-016-0426-4] [Medline: 27688250]
37. Rosenberg D, Godbole S, Ellis K, Di C, Lacroix A, Natarajan L, et al. Classifiers for accelerometer-measured behaviors in older women. Med Sci Sports Exerc 2017 Dec;49(3):610-616 [FREE Full text] [doi: 10.1249/MSS.0000000000001121] [Medline: 28222058]
38. Cuberos-Urbano G, Caracuel A, Valls-Serrano C, García-Mochón L, Gracey F, Verdejo-García A. A pilot investigation of the potential for incorporating lifelog technology into executive function rehabilitation for enhanced transfer of self-regulation skills to everyday life. Neuropsychol Rehabil 2018 Jun;28(4):589-601. [doi: 10.1080/09602011.2016.1187630] [Medline: 27251578]
39. Doherty AR, Caprani N, Conaire C, Kalnikaite V, Gurrin C, Smeaton AF, et al. Passively recognising human activities through lifelogging. Comput Hum Behav 2011 Sep;27(5):1948-1958. [doi: 10.1016/j.chb.2011.05.002]
40. Signal LN, Smith MB, Barr M, Stanley J, Chambers TJ, Zhou J, et al. Kids'Cam: an objective methodology to study the world in which children live. Am J Prev Med 2017 Sep;53(3):e89-e95. [doi: 10.1016/j.amepre.2017.02.016] [Medline: 28455122]
41. Wang P, Smeaton AF. Using visual lifelogs to automatically characterize everyday activities. Inf Sci 2013 May;230:147-161. [doi: 10.1016/j.ins.2012.12.028]

42. Kerr J, Marshall SJ, Godbole S, Chen J, Legge A, Doherty AR, et al. Using the SenseCam to improve classifications of sedentary behavior in free-living settings. Am J Prev Med 2013 Mar;44(3):290-296. [doi: 10.1016/j.amepre.2012.11.004] [Medline: 23415127]
43. Marinac C, Merchant G, Godbole S, Chen J, Kerr J, Clark B, et al. The feasibility of using SenseCams to measure the typecontext of daily sedentary behaviors. New York, NY: ACM; 2013 Presented at: 4th International SenseCam & Pervasive Imaging Conference; Nov 18-19, 2013; San Diego, CA, USA p. 42-49. [doi: 10.1145/2526667.2526674]
44. Moghimi M, Wu W, Chen J, Godbole S, Marshall S, Kerr J, et al. Analyzing sedentary behavior in life-logging images. 2014 Presented at: 2014 IEEE International Conference on Image Processing; Oct 27-30, 2014; Paris, France p. 1011-1015. [doi: 10.1109/ICIP.2014.7025202]
45. Kelly P, Doherty AR, Hamilton A, Matthews A, Batterham AM, Nelson M, et al. Evaluating the feasibility of measuring travel to school using a wearable camera. Am J Prev Med 2012 Nov;43(5):546-550 [FREE Full text] [doi: 10.1016/j.amepre.2012.07.027] [Medline: 23079179]
46. Kelly P, Doherty A, Mizdrak A, Marshall S, Kerr J, Legge A, et al. High group level validity but high random error of a self-report travel diary, as assessed by wearable cameras. J Transport Health 2014 Sep;1(3):190-201. [doi: 10.1016/j.jth.2014.04.003]
47. Boushey CJ, Spoden M, Zhu FM, Delp EJ, Kerr DA. New mobile methods for dietary assessment: review of image-assisted and image-based dietary assessment methods. Proc Nutr Soc 2016 Dec 12:1-12. [doi: 10.1017/S0029665116002913] [Medline: 27938425]
48. Sun M, Fernstrom JD, Jia W, Hackworth SA, Yao N, Li Y, et al. A wearable electronic system for objective dietary assessment. J Am Diet Assoc 2010 Jan;110(1):45-47 [FREE Full text] [doi: 10.1016/j.jada.2009.10.013] [Medline: 20102825]
49. Maekawa T, Kishino Y, Yanagisawa Y, Sakurai Y. WristSense: wrist-worn sensor device with camera for daily activity recognition. 2012 Presented at: 2012 IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops); Mar 19-23, 2012; Lugano, Switzerland p. 510-512. [doi: 10.1109/PerComW.2012.6197551]
50. Loveday A, Sherar LB, Sanders JP, Sanderson PW, Esliger DW. Novel technology to help understand the context of physical activity and sedentary behaviour. Physiol Meas 2016 Dec;37(10):1834-1851. [doi: 10.1088/0967-3334/37/10/1834] [Medline: 27654030]
51. Kelly P, Marshall SJ, Badland H, Kerr J, Oliver M, Doherty AR, et al. An ethical framework for automated, wearable cameras in health behavior research. Am J Prev Med 2013 Mar;44(3):314-319. [doi: 10.1016/j.amepre.2012.11.006] [Medline: 23415131]
52. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. 2012 Presented at: 25th International Conference on Neural Information Processing Systems; Dec 3-6, 2012; Lake Tahoe, NV, USA p. 1097-1105.
53. Biswas D, Su H, Wang C, Blankenship J, Stevanovic A. An automatic car counting system using OverFeat framework. Sensors (Basel) 2017 Jun 30;17(7) [FREE Full text] [doi: 10.3390/s17071535] [Medline: 28665360]
54. Peters MDJ, Godfrey CM, Khalil H, McInerney P, Parker D, Soares CB. Guidance for conducting systematic scoping reviews. Int J Evid Based Healthc 2015 Sep;13(3):141-146. [doi: 10.1097/XEB.0000000000000050] [Medline: 26134548]
55. Lazar A, Koehler C, Tanenbaum J, Nguyen DH. Why we use and abandon smart devices. New York, NY: ACM; 2015 Presented at: 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing; Sep 7-11, 2015; Osaka, Japan p. 635-646. [doi: 10.1145/2750858.2804288]
56. May CR, Eton DT, Boehmer K, Gallacher K, Hunt K, MacDonald S, et al. Rethinking the patient: using burden of treatment theory to understand the changing dynamics of illness. BMC Health Serv Res 2014;14:281 [FREE Full text] [doi: 10.1186/1472-6963-14-281] [Medline: 24969758]

Edited by G Eysenbach; submitted 13.03.18; peer-reviewed by A Doherty, IN Gomez, P Wark, M Wolters; comments to author 18.06.18; revised version received 13.08.18; accepted 30.08.18; published 03.01.19.

Please cite as:
Maddison R, Cartledge S, Rogerson M, Goedhart NS, Ragbir Singh T, Neil C, Phung D, Ball K
Usefulness of Wearable Cameras as a Tool to Enhance Chronic Disease Self-Management: Scoping Review
JMIR Mhealth Uhealth 2019;7(1):e10371
URL: https://mhealth.jmir.org/2019/1/e10371/
doi:10.2196/10371
PMID:30609985

Original Paper

Using Passive Smartphone Sensing for Improved Risk Stratification of Patients With Depression and Diabetes: Cross-Sectional Observational Study

Archana Sarda[1], MBBS, MD; Suresh Munuswamy[2], MBBS, MPH, PhD; Shubhankar Sarda[3], BSc; Vinod Subramanian[3], MBA

[1]Sarda Centre for Diabetes and Selfcare, Aurangabad, India

[2]DST Health Informatics Rapid Design Lab, Indian Institute of Public Health-Hyderabad, Hyderabad, India

[3]Touchkin eServices Private Limited, Bangalore, India

Corresponding Author:
Shubhankar Sarda, BSc
Touchkin eServices Private Limited
1st Floor, Manjusha, No 532
16th Cross, 2nd Main Road, 2nd Stage, Indira Nagar
Bangalore, 560038
India
Phone: 91 9762665119
Email: shubhankar@touchkin.com

Abstract

Background: Research studies are establishing the use of smartphone sensing to measure mental well-being. Smartphone sensor information captures behavioral patterns, and its analysis helps reveal well-being changes. Depression in diabetes goes highly underdiagnosed and underreported. The comorbidity has been associated with increased mortality and worse clinical outcomes, including poor glycemic control and self-management. Clinical-only intervention has been found to have a very modest effect on diabetes management among people with depression. Smartphone technologies could play a significant role in complementing comorbid care.

Objective: This study aimed to analyze the association between smartphone-sensing parameters and symptoms of depression and to explore an approach to risk-stratify people with diabetes.

Methods: A cross-sectional observational study *(Project SHADO—Analyzing Social and Health Attributes through Daily Digital Observation)* was conducted on 47 participants with diabetes. The study's smartphone-sensing app passively collected data regarding activity, mobility, sleep, and communication from each participant. Self-reported symptoms of depression using a validated Patient Health Questionnaire-9 (PHQ-9) were collected once every 2 weeks from all participants. A *descriptive analysis* was performed to understand the representation of the participants. A *univariate analysis* was performed on each derived sensing variable to compare behavioral changes between depression states—*those with self-reported major depression (PHQ-9>9)* and *those with none (PHQ-9≤9)*. A *classification predictive modeling*, using supervised machine-learning methods, was explored using derived sensing variables as input to construct and compare classifiers that could risk-stratify people with diabetes based on symptoms of depression.

Results: A noticeably high prevalence of self-reported depression (30 out of 47 participants, 63%) was found among the participants. Between depression states, a significant difference was found for *average activity rates (daytime)* between participant-day instances with symptoms of major depression (mean 16.06 [SD 14.90]) and those with none (mean 18.79 [SD 16.72]), P=.005. For *average number of people called (calls made and received)*, a significant difference was found between participant-day instances with symptoms of major depression (mean 5.08 [SD 3.83]) and those with none (mean 8.59 [SD 7.05]), P<.001. These results suggest that participants with diabetes and symptoms of major depression exhibited lower activity through the day and maintained contact with fewer people. Using all the derived sensing variables, the extreme gradient boosting machine-learning classifier provided the best performance with an average cross-validation accuracy of 79.07% (95% CI 74%-84%) and test accuracy of 81.05% to classify symptoms of depression.

Conclusions: Participants with diabetes and self-reported symptoms of major depression were observed to show lower levels of social contact and lower activity levels during the day. Although findings must be reproduced in a broader randomized controlled

study, this study shows promise in the use of predictive modeling for early detection of symptoms of depression in people with diabetes using smartphone-sensing information.

(JMIR Mhealth Uhealth 2019;7(1):e11041) doi:10.2196/11041

KEYWORDS

depression; diabetes; mental health; comorbidity; passive sensing; smartphone; classification; machine learning; mHealth; risk assessment

Introduction

Background

There exists growing evidence regarding the bidirectional association between diabetes (type 2) and depression [1,2]. A meta-analysis of published studies on adults [3] reported depression to be 2 to 3 times more common in people with diabetes (both types) than in those without, with the odds of depression significantly higher in women than in men with diabetes. An estimated 8% to 35% of people across ages with diabetes mellitus (both type 1 and 2) also suffer from depression [4]. Depression increases the risk of nonadherence to medical treatment by 27% to 30% [5-7], which is a significant problem in diabetes self-care. Comorbidity (diabetes and depression) has also been associated with increased health care costs. Individuals (US adults) with diabetes who also had depression were found to be 2 to 4.5 times more expensive to treat than those with diabetes alone (US $247 million compared with US $55 million in 2001, dollars after adjusting for differences in age, sex, race, ethnicity, health insurance, and comorbidity) [8,9]. Depression is known to be associated with abnormalities in metabolism of biologics (eg, increased counterregulatory hormone release and action, changes in glucose transport function, and increased immuno-inflammatory activation) [10]. Depression might also increase the risk of developing type 2 diabetes with an increase in insulin resistance and reduction of glucose uptake in adults [11]. Comorbidity of depression and diabetes is associated with a high likelihood of complications [12,13], lower quality of life [14], increased mortality [15], poor management and control [16,17], and poor disease outcomes through decreased physical activity [18,19] as reported in studies covering a diverse population across age groups.

In an analysis of worldwide studies [20], it was found that primary care physicians fail to correctly diagnose between 30% and 50% of patients who present with a depressive disorder. The same study also pointed to poor recognition rates of depression symptoms among both men and women aged less than 40 years. A retrospective study [21] among a population-based sample of primary care patients with diabetes (both type 1 and type 2, across age groups) within a US-based health maintenance organization revealed that depression is identified only half of the time (approximately 51%). The same study also pointed that only 31% of patients with comorbid diabetes and depression received adequate antidepressant treatment and only 6.7% received 4 or more psychotherapy sessions during a 12-month period. Clinical-only interventions seem to have very modest effects in diabetes management of patients with depression [22-24]. The American Diabetes Association recommends that patients with diabetes be screened for psychosocial and psychological problems or disorders, such as depression [4,25]. However, this appears to happen rarely [26].

The number of global smartphone users is expected to surpass 2.3 billion by 2017 [27]. Smartphones carry sensors such as accelerometer, global positioning system (GPS), and ambient light sensors that capture data and that could provide information on someone's behavior. In this context, the smartphone could be the most ubiquitous data collection device today. It also presents with huge privacy and security concerns. Developed nations have higher smartphone penetration, and the ownership rates in emerging and developing nations have been rising at an extraordinary rate [28]. Passive data from smartphone sensors have been known to detect patterns of behavior in people with depression [11,29-32]. Research has been establishing the link between smartphone-sensing data and its application in overall well-being [33-36] and depression [37-39]. Smartphones for social sensing [40,41], in monitoring and possibly as an intervention in mental health [42,43], has the advantage of ubiquity, discretion, and low cost. Comorbidity poses a large economic burden; hence, there exists a need for effective screening and treatment. To use limited resources efficiently, risk stratification is important to target appropriate intervention for people with diabetes and depression.

Previous Work

Many studies have explored how smartphone-sensing data can be used as a predictor for depression and mental health [44-47], but very few studies have applied passive sensing to predict symptoms of depression among people with diabetes, thereby enabling improved risk stratification. There are recent studies that have attempted to predict depression among patients with diabetes using longitudinal patient records or data from clinical trials or surveys, but not using sensing data as an indicator [48-50].

Study Objective

The aim of the study was (1) to identify behaviors derived from smartphone-sensing data that are significant with symptoms of major depression compared with those with no symptoms among patients with diabetes and (2) to evaluate a risk stratification approach for early detection of symptoms of depression in patients with diabetes using smartphone-sensing parameters. To the best of our knowledge, this study was the first such implementation of using automatically captured smartphone-sensing data such as activity, communication, mobility, and sleep to screen for symptoms of depression among primary care patients with diabetes.

Methods

About the Study

The pilot study named *Project SHADO (Analyzing Social and Health Attributes through Daily Digital Observation)* was conducted in 2016 on a cross section of participants with diabetes located in periurban India and owning low-cost smartphones. The institutional Ethics Committee at the Public Health Foundation of India's Indian Institute of Public Health at Hyderabad, India, approved the study (Approval number IIPHH/TRCIEC/073/2016).

Study Design

The study design was conducted in association with a diabetes clinic situated in Aurangabad, a city in the state of Maharashtra, India. A cross-sectional observational study was designed to be conducted on a sample of patients undergoing diabetes treatment at the clinic. The study did not require any intervention or change in treatment or lifestyle for the participants. It did not involve a control group.

The period of the study was originally 14 weeks and later extended to 20 weeks to collect sufficient smartphone-sensing data. The study app that was used passively and anonymously collected data regarding activity, mobility, sleep, and communication from each participant. The actual conversation from the call was never collected. For identifying symptoms of depression among enrolled participants, a globally validated screening tool, Patient Health Questionnaire-9 (PHQ-9) was used [51,52]. The PHQ-9 survey was made available in both English and the local language (Marathi). The language-modified version of PHQ-9 had been validated in other studies [53,54]. The PHQ-9 English and Marathi language questionnaire are available in Multimedia Appendices 1 and 2, respectively. For the participants to feel comfortable keeping data services enabled on their smartphones, they were provided with a 1 GB data recharge per month to cover for usage costs. No other incentive was provided to the participants. The care providers at the clinic administered the study. Administrator guidelines were set that ensured effective participant enrollment and onboarding process. A Web-based patient administration system was used to manage participant details. The clinical staff were oriented about the study, familiarized about the study app, and trained on the administrator system and guidelines. Effective monitoring and support was established to manage any issues that could occur during the study. The conceptual framework of the research design is shown in Multimedia Appendix 3. The study was designed to have no financial burden on the participant, nor any drug or device hazard.

Study Participants

A list of 100 periurban patients undergoing treatment for diabetes and who satisfied the study inclusion criteria (see Textbox 1) were contacted for their interest and participation in the study. Patients who did not meet the inclusion criteria, who had mobility restrictions, who were bedridden, or who had serious comorbid conditions including disabilities and visual or hearing impairment were excluded.

Overall, 47 out of the 100 patients provided their consent and were enrolled for the study. Participation in the study was voluntary, and as per the informed consent provided, a participant could decide not to participate or could withdraw from the study at any time without having to provide any reasons or justifications. All study participants followed a formal onboarding process where they were provided with information about the study; were educated about the privacy, security, and consent process; had the study app setup; provided explicit consent; and completed the initial PHQ-9 survey. The first self-reported depression score was assessed in person at the clinic followed by collection over telephone once every 2 weeks during the study period.

Study App

The study used a smartphone-sensing app ("app") developed by Touchkin. The app assisted family members to care for their loved ones remotely and nonintrusively and to check on their well-being. The app's machine learning (ML) platform helped detect probable well-being changes by using activity rates, communication levels, sleep patterns, and mobility information collected from the user's smartphone sensors.

Data Collection

Data were deidentified before use for research purposes.

Sociodemographic Data

The participant's sociodemographic information such as gender, marital status, occupation, age, education, and family particulars were captured at enrollment. The level of control over the existing diabetes condition for each participant was assessed by the diabetologist. The participant's level of control was assessed based on their existing condition, lifestyle, and medication adherence history. Participants were classified as having low, moderate, or high control over their condition.

Passive Sensing Data

Smartphone-sensing data were captured by the app automatically every 2 minutes and stored on-device using a read or write memory card. The app was designed to capture only hashed identifiers, and the collected data were secured and anonymized on-device before being transferred to the storage servers for an aggregate analysis. All transmissions were in encrypted form using the HTTPS secure sockets layer protocol. On the server side, these files were merged, parsed, and synchronized by Python-based postprocessing infrastructure and stored in not only SQL–based servers. The servers and the data thereof were access restricted, allowing only the engineering lead to retrieve the minimal needed data for research. The raw passive sensing data were processed and daily values derived for the sensing variables. Running of the ML models on the entire deidentified dataset was performed securely on the cloud with the research analyst getting to view only the performance results.

Textbox 1. Study inclusion and exclusion criteria.

Inclusion criteria:

- Those who were 1+ year on diabetes treatment, with recent 6 months of consulting the diabetologist
- Those who had at least one clinic visit per month
- Those who were 18 years and above
- Those who owned a smartphone with Android operating system
- Those with normal mobility, with no known debilitating comorbidities
- Those willing to participate for the duration of the study
- Those willing to carry a smartphone at all times
- Those willing to download the smartphone-sensing app

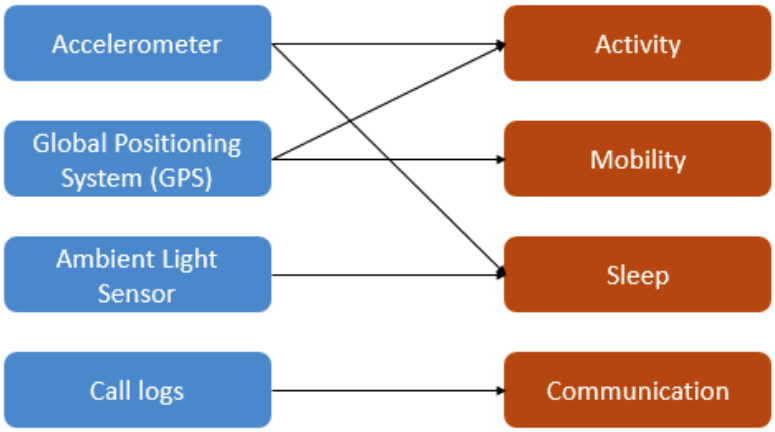
Exclusion criteria:

- Those with known debilitating comorbidities
- Those with known mobility restrictions
- Those who were bedridden
- Those with disabilities
- Those with visual or hearing impairment
- Those who owned a non-Android smartphone (eg, iOS)

The social interaction data were captured from 3 main sensors and the call logs. These were the accelerometer, the GPS, and the ambient light sensor. All these sensors reflect pairwise communication and face-to-face proximity, intensity and nature of social ties, the dynamics of network, and amount of light in the background. To ensure that no loss of sensing data occurred in the event of network drops, they were stored in the participant's smartphone for up to 3 days.

A total of 53 sensing variables were derived from the activity, mobility, sleep, and communication data collected from the smartphone sensors. Figure 1 outlines the sensor-feature map. Activity variables were derived based on periods where the participant was found to be active. This was measured by calculating the number of times the relative gravity values, derived from accelerometer readings, exceeded the stationary threshold range (as defined from 0.8 to 1.2). Relative gravity is a measure of acceleration experienced by the mobile device with reference to earth's gravity. Mobility variables were derived based on the number of locations and the distance traveled (in meters). Sleep variables were based on relative gravity and the number of screen-ons. Call-related variables were based on the number of total calls (made and received), missed calls, and call duration (in minutes). Call-related variables do not include participant texting, as this was not considered in the scope of data capture. The values for these variables were derived as day-wise aggregates from the raw-sensing data collected for each participant. Details on each of the 53 derived sensing variables can be found in Multimedia Appendix 4.

Figure 1. Sensor-feature map.



Screening Instrument Data

PHQ-9 consists of a validated 9-item depression screening tool, with each item having 4 options (scored 0-3) as responses namely: “not at all,” “several days,” “more than half the days,” and “nearly every day.” PHQ-9 helps screen for the presence and severity of depression with a maximum total score of 27. The values for each of the 53 derived sensing variables were aggregated day wise, whereas the screening data (PHQ-9 scores) were collected 14 days apart for each participant. Therefore, it was decided to impute the average of 2 consecutive PHQ-9 surveys submitted by a participant as the score for all the participant-day instances occurring between the 2 screening dates for that participant.

Data Quality

Of the 47 participants, only 1 participant did not complete the study. The final analysis was performed on data from 46 participants, of which 29 were men and 17 were women. Over the course of the study, a total of 2694 *participant-day* records (instances) were collected. This formed the dataset for subsequent analysis.

All PHQ-9 surveys had reminders for the participants as well as administrators to follow up. The opportunity to miss capture of sensing data could arise from many reasons, and mostly these are random in nature. Some include participants changing phone settings, forgetting to carry phones or there are errors during capture, storage, and manipulation. This resulted in dataset instances having values missing for some of the derived sensing variables. Data quality for smartphone-sensing data was an area of focus during the data gathering and data cleaning phase. Typically, low-end smartphone devices present several challenges when sending out passive sensing data, namely, frequency of data transmission varies, unexpected shutdown occurs for extended periods during the day, and data values (eg, accelerometer readings) sometimes do not match higher-end devices. Within the study, data quality measures were implemented to solve 2 issues: (1) data integrity: how to conduct reliable sampling of data to ensure there was minimum loss in continuity of data feeds and (2) data accuracy: how to ensure the validity of the data being collected, so that we could be confident that this correctly represented participant context (see Multimedia Appendix 5 for the measures taken to ensure data integrity and accuracy).

Monitoring and Support

During the study period, active support was provided to ensure minimal dropout and to ensure priority resolution of issues and monitoring of data quality. The administrators monitored the Web administration system for alerts and reached out to participants as required. Alerts included unusual smartphone usage or when 2-week surveys were due and other reasons (eg, not receiving sensing data).

Smartphone Battery and Memory Optimization

The app’s technical and proprietary data collection methods ensured that the participant’s smartphone battery impact was kept low. The app occupied less than 10 MB of storage space on a typical Android smartphone and consumed less than 2% of total battery. This was lower than that consumed by other apps usually installed on a smartphone and as measured over a 24-hour period. A recent study [55] had pointed out the importance of reporting battery performance as it plays a major role in sensing data collection and quality.

Data Analysis

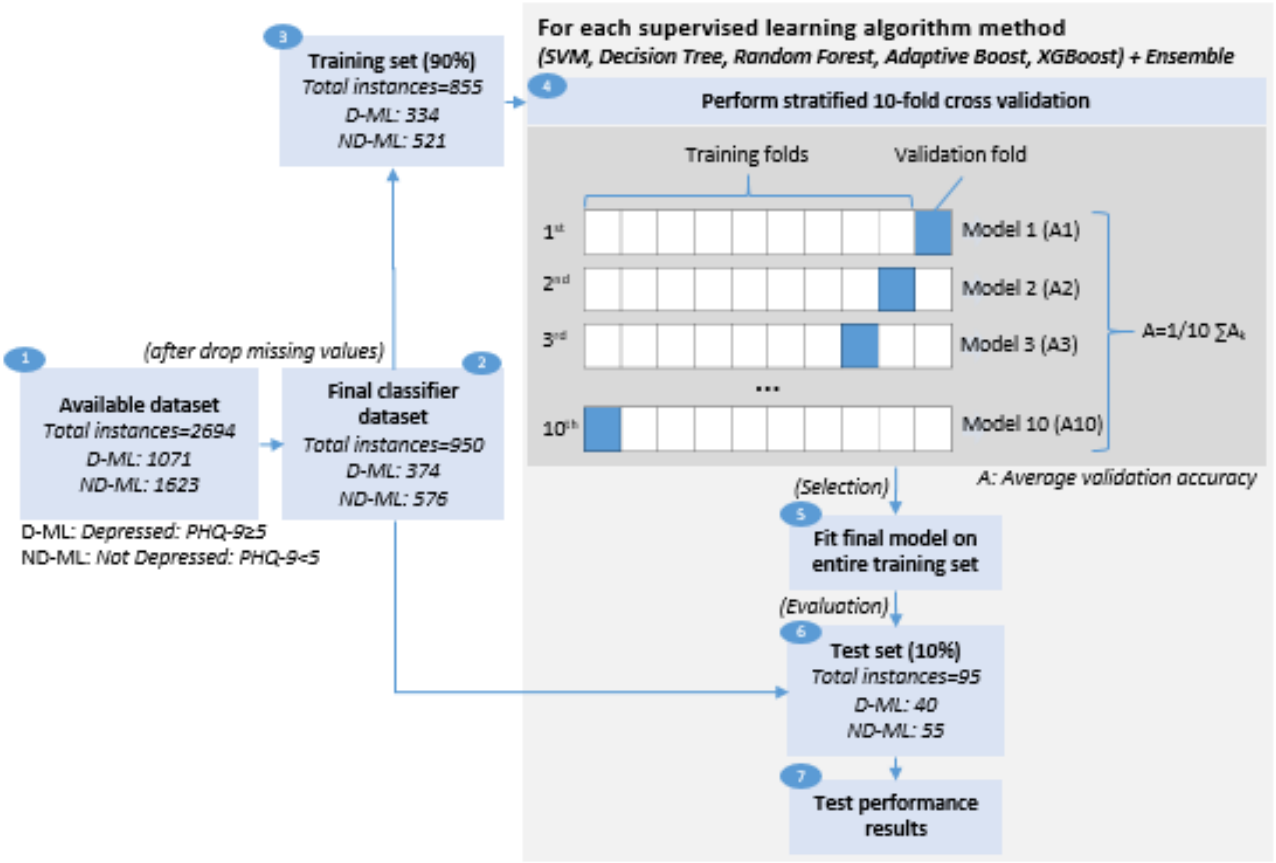
Analysis was performed in 3 parts: descriptive, univariate, and classification modeling.

Descriptive analysis was performed to understand the representation of the participants based on their sociodemographic factors, clinical presentations, and mental well-being.

Univariate analysis was performed to understand whether there were observable differences in behavior between a set of instances tagged with symptoms of major depression (depressed—D) and those with none (not depressed—ND). The D class included instances with PHQ-9 greater than 9 (moderate to severe severity) scores, and the ND class included instances with PHQ-9 less than or equal to 9 (none to mild severity) scores. The PHQ-9 cutoff for major depression was decided based on published studies [56]. An independent *t* test was conducted to compare the 2 classes, considering the test’s robustness with large number of instances [57]. Therefore, it was safe to assume that the data in each class were normally distributed for each derived sensing variable. It was also observed that the number of instances in the 2 classes for each derived sensing variable was significantly unbalanced. It was, therefore, decided that the unequal variance independent *t* test [58] be applied (with and without outliers) to compare the 2 classes. It was observed that any imputation of missing values in sensing variables would potentially introduce bias, and hence no treatment was affected on the missing values.

Classification modeling was performed with the objective to explore, compare, and identify the best performing classifier method to build a risk stratification model for early detection of symptoms of depression in participants with diabetes. A total of 5 supervised ML methods and their ensemble were explored. Tree-based supervised ML methods were mostly considered, given their robustness to multicollinearity, outliers, and missing values. These include support vector machine (SVM), decision tree (DT), random forest (RF), adaptive boosting (AdaBoost), and extreme gradient boosting (XGBoost). The radial basis function (RBF) kernel used in SVM methods is known to handle large feature sets and their nonlinear interactions. The other 4 methods were tree-based methods that included the basic DT along with boosting trees (AdaBoost and XGBoost) and bagging trees (RF). Both bagging and boosting trees combine several DTs to reduce error and improve classification performance. Boosting trees help reduce bias, whereas bagging trees help reduce variance. A voting ensemble was also trained that combined each of the 5 methods to check for improved classification performance. Each of the 5 methods provides a class (D or ND) prediction (vote) for each participant-day instance, whereas the voting ensemble counts these votes and tags the majority class voted for that instance.

Figure 2. Classification modeling (train-validate-test) approach. PHQ-9: Patient Health Questionnaire-9; SVM: support vector machine; XGBoost: extreme gradient boosting.



Open-source Python software on Jupyter Notebook was used for modeling. A lower PHQ-9 cutoff of 5 was considered for the modeling to broaden the scope and include instances with mild symptoms of depression.

Mild symptoms or subclinical depression in people with diabetes has been found to be common, associated with high levels of diabetes-related distress, psychological distress, and lower quality of life and also a risk indicator of major depression [59-61]. Those instances with self-reported symptoms of depression (PHQ-9≥5—mild to severe) were grouped under the depressed-machine learning (D-ML) class and those with none (PHQ-9<5) under the not depressed-machine learning (ND-ML) class. Instances that contained missing values in any of the 53 derived sensing variables were entirely dropped. The dataset so obtained was divided into a training set (90%) and test set (10%), and a stratified K-fold cross-validation (CV) with K equal to 10 folds was performed on the training set for each of the 5 classifier methods (see Figure 2 for the modeling approach [train-validate-test] followed). The CV has been widely used to compare different classifier methods. A major advantage of using the 10-fold CV approach is that every data instance gets to be in a validation set exactly once and gets to be in a training set 9 times, leading to lower variance in the resulting estimate. Stratification ensures that the ratio of both classes (D-ML and ND-ML) is equally represented in each fold. A CV splits the training set into K equal parts. A model is trained on K−1 parts and gets validated on the remaining part. This process leads to development of K models for each method. The average performance of K models is then compared for each method. A nested CV was not opted for because of computational challenges. The final model for each method was then trained on the entire training set and tested with the unseen test set. Test performance results were then compared to identify the best ML method. Classification performance in terms of accuracy, specificity, sensitivity, and precision along with the confusion matrix was used to compare the ML methods. For this study, a lower number of false negatives (wrongly classifies symptoms of depression to be absent) was important as a wrong classification would lead to patients with symptoms of depression being missed out for priority diabetes care. Therefore, apart from high accuracy, a high recall with a reasonably high precision formed the basis to compare and select the appropriate method to build risk stratification models.

Results

Descriptive Analysis

At the start of the study, 1 out of the 47 participants had a known diagnosis for depression. A noticeably high percentage of participants (30/46), excluding 1 with a known diagnosis, self-reported symptoms of depression on the PHQ-9 survey during the study (Figure 3). Overall, 6 out of 31 reported severe symptoms of depression, including suicidal tendency. These severe cases were referred to psychologists promptly. The diabetologist ensured regular follow-up of their referred patients

for their depression condition as per the care protocol established at the diabetes clinic.

At the beginning, 30 men and 17 women participated in the study. Of them, 1 male participant dropped out making it 29 men and 17 women at the end of study. As seen in Table 1, participants had a mean age of 35 years (SD 12). In total, 60% (28/46) of the participants were in the age group of 21 to 40. Of them, 58% (27/46) of the participants were married. A majority of the participants were office goers at 69% (32/46), whereas 17% (8/46) were students. Of them, 58% (27/46) of the participants held a bachelor's or master's degree, whereas 34% (16/46) had completed schooling. Some of the clinical characteristics of the participants included an almost equal mix of diabetes condition (type 1/type 2) and 63% (29/46) had moderate level of control over their diabetes condition.

Univariate Analysis

All the 2694 instances (participant-day records) were included for this analysis. The results with and without outliers are summarized in Tables 2 and 3.

- A significant difference was observed in the *average activity rates in the morning hours (from 6:00 am until 11:59 am)* among those with symptoms of major depression (mean 13.70 [SD 14.04]) compared with those with none (mean 18.48 [SD 18.44]), *P*<.001. A significant difference was also observed in *average activity rates in the remaining part of the day (from noon until 4:00 pm)* among those with symptoms of major depression (mean 16.06 [SD 14.90]) than those with none (mean 18.79 [SD 16.72]), *P*=.005. These results suggested that those with symptoms of major depression exhibited lower and irregular activity rates through the day as compared with those with none.
- A significant difference was observed in the *number of screen-on times at night (from midnight until 6:00 am)* among those with symptoms of major depression (mean 6.70 [SD 9.33]) compared with those with none (mean 3.16 [SD 8.91]), *P*<.001. The results suggested that those with symptoms of major depression possibly had an impacted sleep quality due to higher screen-ons.
- A significant difference was observed in the *average total number of calls (made and received)* among those with symptoms of major depression (mean 12.61 [SD 9.15]) compared with those with none (mean 22.28 [SD 50.76]), *P*<.001. A significant difference was also observed in the *average number of people called* among those with symptoms of major depression (mean 5.08 [SD 3.83]) compared with those with none (mean 8.59 [SD 7.05]), *P*<.001. The results suggested that those with symptoms of major depression maintained contact with fewer people and attended fewer calls.
- Mobility variables showed limited to no statistical significance at 95% and 99% CI, respectively, between the 2 depression states and hence have not been reported.

Univariate trends over the weeks also showed that those with symptoms of major depression (D) exhibited irregular and lower daytime average activity rates (Figure 4) compared with those with none (ND).

Trends over the week showed that those with symptoms of major depression (D) had irregular and higher average number of screen-ons at nighttime (Figure 5) than those with none (ND).

Trends over the week also showed that those with symptoms of major depression (D) withdrew socially with lower average number of calls, lower average number of people contacted, and lower average duration per call (Figure 6) than those with none (ND).

Figure 3. Prevalence of depression.

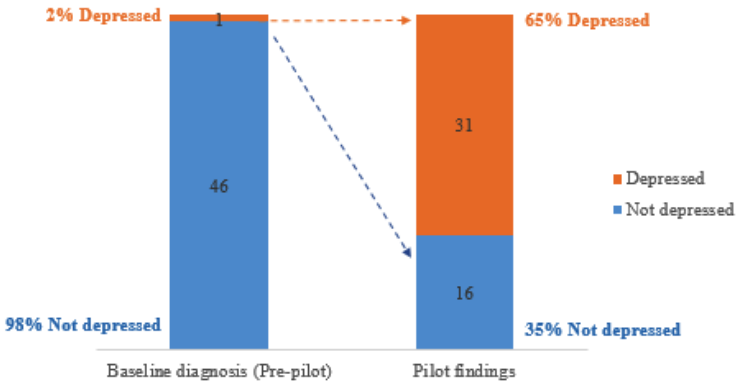


Table 1. Participant demographics (N=46).

Participant characteristics	Statistic, n (%)
Age (years)	
15-20	4 (10)
21-30	14 (30)
31-40	15 (32)
41-50	5 (11)
51+	8 (17)
Gender	
Men	29 (64)
Women	17 (36)
Marital status	
Single	18 (40)
Married	28 (60)
Education	
Grade 10-12	15 (34)
Bachelor’s degree	19 (41)
Master’s degree	9 (19)
Vocational education	3 (6)
Occupation	
Student	8 (17)
Home	6 (13)
Office	32 (70)
Family	
Living alone	4 (9)
Living with family	42 (91)
Chronic condition	
Diabetes type 1	21 (45)
Diabetes type 2	25 (55)
Patient location	
Outstation	14 (30)
In city	32 (70)
Level of control over diabetes condition	
Low	11 (23)
High	7 (15)
Moderate	28 (62)

Table 2. Univariate analysis results (with outliers).

Key smartphone-sensing variables	Depressed (PHQ-9[a]>9)		Not depressed (PHQ-9≤9)		*P* value (with unequal variance)
	n1[b] (%)	Mean (SD)	n2[c] (%)	Mean (SD)	
Activity rate[d] (am[e])	194 (11)	13.70 (14.04)	1598 (89)	18.48 (18.44)	<.001
Activity rate (day[f])	228 (12)	16.06 (14.91)	1761 (88)	18.79 (16.72)	.005
Screen-on[g] (night[h])	130 (9)	6.70 (9.33)	1301(91)	3.16 (8.91)	<.001
Calls (made and received)	262 (11)	12.61 (9.15)	2057 (89)	22.28 (50.76)	<.001
People called	262 (11)	5.08 (3.83)	2057 (89)	8.59 (7.05)	<.001
Call duration (minutes)	262 (11)	18.95 (19.32)	2057 (89)	37.59 (174.88)	<.001

[a]PHQ: Patient Health Questionnaire.

[b]n1: Number of instances with values for depressed.

[c]n2: Number of instances with values for not depressed.

[d]Total number of *active* polled every 2 min. *Active:* where relative gravity values exceed the stationary threshold range (0.8-1.2).

[e]From 6:00 am until 11:59 am.

[f]From noon until 4:00 pm.

[g]Total number of *Screen-on* polled every 2 min. *Screen on:* where the user had their mobile screen switched on and unlocked.

[h]From midnight until 6:00 am.

Table 3. Univariate analysis results (without outliers).

Key smartphone-sensing variables	Depressed (PHQ[a]-9>9)		Not depressed (PHQ-9≤9)		*P* value (with unequal variance)
	n1[b] (%)	Mean (SD)	n2[c] (%)	Mean (SD)	
Activity rate[d] (am[e])	183 (11)	11.06 (7.93)	1507 (89)	14.87 (10.80)	<.001
Activity rate (day[f])	214 (11)	12.95 (6.96)	1754 (89)	18.61 (16.49)	<.001
Screen-on[g] (night[h])	120 (9)	4.58 (5.54)	1156 (91)	1.32 (1.69)	<.001
Calls (made and received)	254 (12)	11.69 (7.58)	1933 (88)	16.02 (11.54)	<.001
People called	240 (11)	4.22 (2.37)	1965 (89)	7.59 (5.34)	<.001
Call duration (minutes)	246 (11)	15.24 (12.52)	1917 (89)	21.36 (19.21)	<.001

[a]PHQ: Patient Health Questionnaire.

[b]n1: Number of instances with values for depressed.

[c]n2: Number of instances with values for not depressed.

[d]Total number of *active* polled every 2 min. *Active:* where relative gravity values exceed the stationary threshold range (0.8-1.2).

[e]From 6:00 am until 11:59 am.

[f]From noon until 4:00 pm.

[g]Total number of *Screen-on* polled every 2 min. *Screen on:* where the user had their mobile screen switched on and unlocked.

[h]From midnight until 6:00 am.

Figure 4. Week-wise trend of average activity rates (day). D: depressed; ND: not depressed.

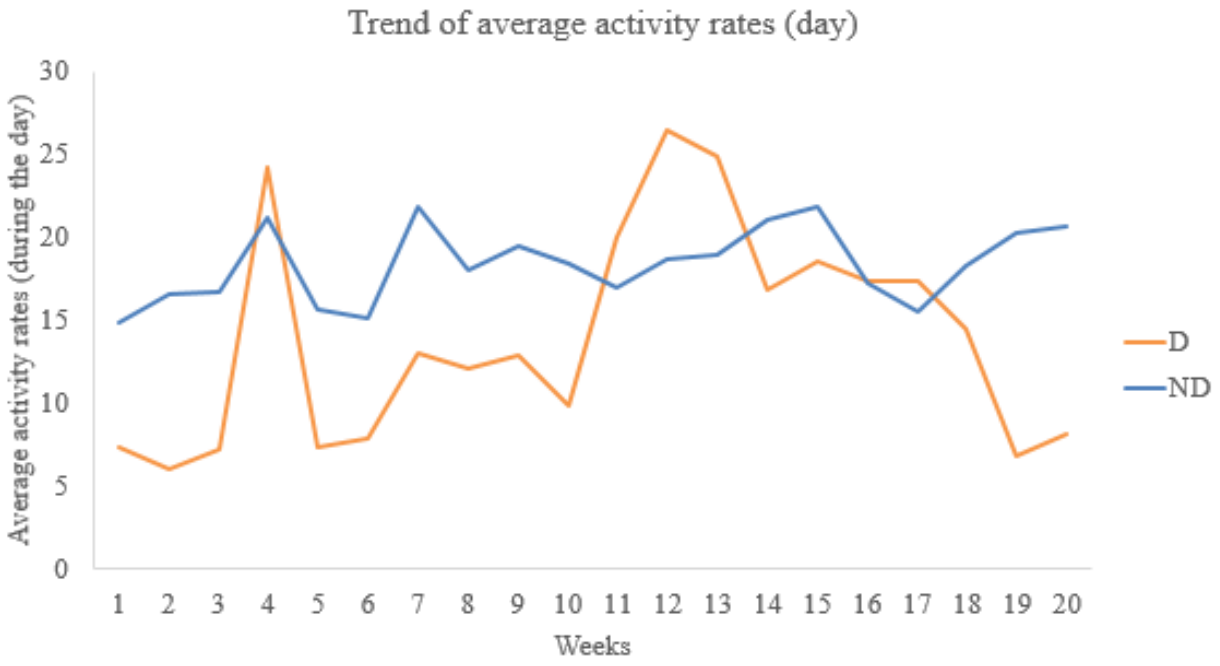


Figure 5. Week-wise trend of average screen-ons (night). D: depressed; ND: not depressed.

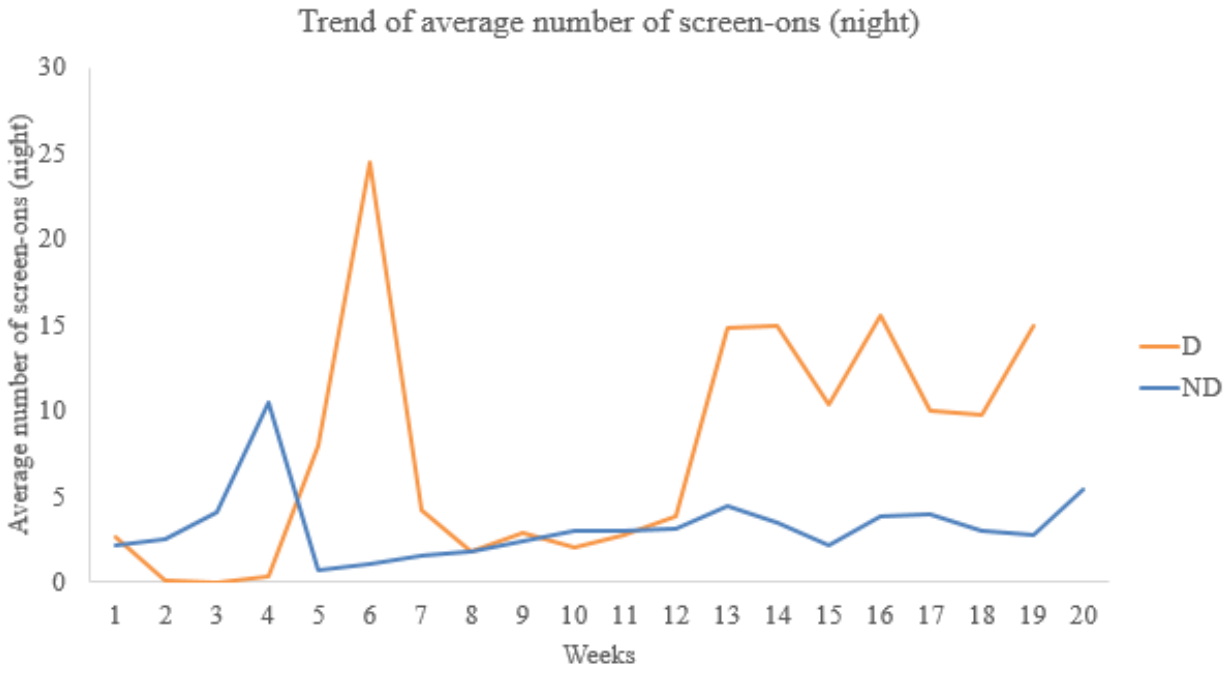
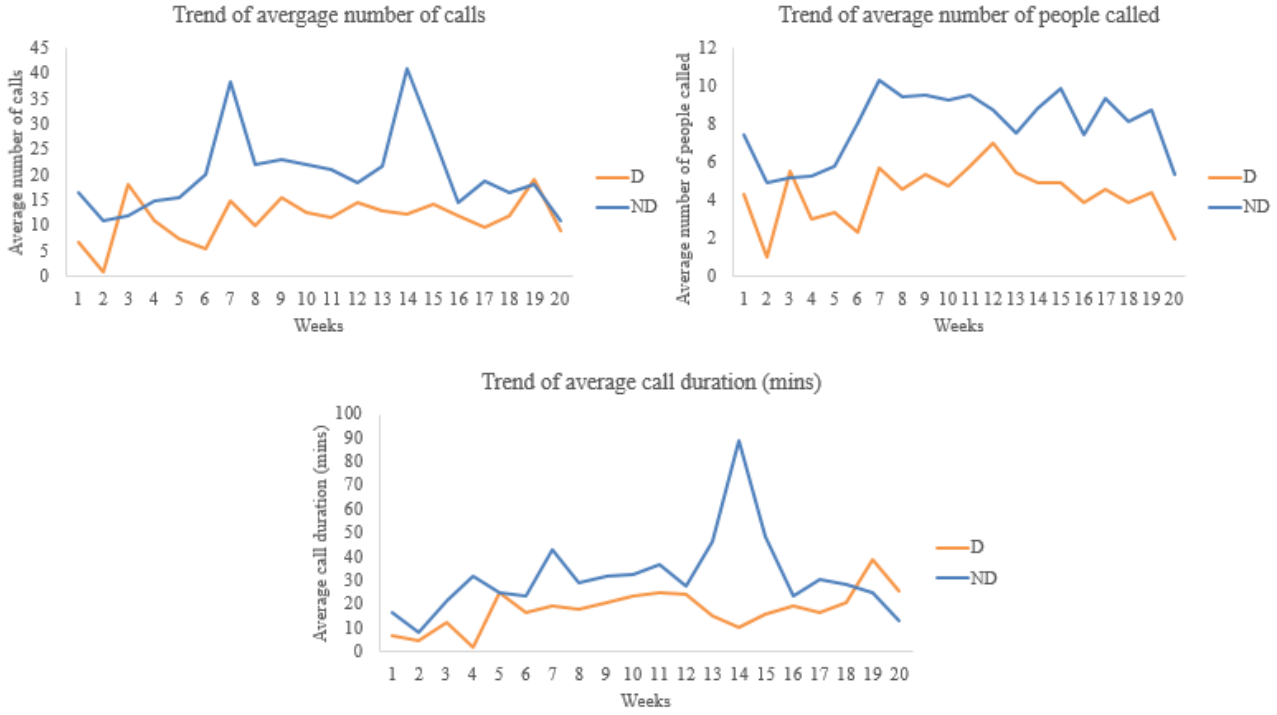


Figure 6. Week-wise trends of average calls-people-duration. D: depressed; ND: not depressed.



Classification Analysis

All participant-day instances with missing values in 1 or more of the 53 derived sensing variables were removed, which resulted in 950 out of 2694 instances available for analysis. A 90:10 (training:test) split resulted in 855 instances in the training set and 95 instances in the test set. A stratified 10-fold CV was performed on this training set before testing on an unseen test set.

Collinearity Check

Very low correlation (*r* values ranging from −.15 to .13) was observed between the self-reported depression state *(PHQ-9 cutoff 5)* and each of the 53 derived sensing variables as measured by Pearson correlation. High pair-wise correlation or collinearity (>80%) was observed among some of the derived sensing variables (Multimedia Appendix 6). Mobility and communication-based variables showed higher collinearity. This was primarily due to the association between the variables and their subset, for example, “total calls at peak” and “total calls at off-peak” variables were subsets of the “total calls” variable. Activity and sleep-based variables showed lower collinearity. This was again due to the association between variables such as “total activity rates,” “screen-on times,” and their “time of day” subsets. For example, “total activity by day/night/eve/am” were all subsets of the “total activity” variable. As each derived sensing variable, of its own, provided rich behavioral context, it was decided to retain all the 53 variables and their subsets as inputs into the modeling.

Model Development

In total, 5 ML methods (SVM, DT, RF, AdaBoost, and XGBoost) and their ensemble were trained and compared for performance. Accuracy provided the fraction of correctly classified samples of both classes (D-ML and ND-ML). XGBoost and RF performed the best in terms of accuracy. They reported an average cross-validated accuracy of 79.1% (95% CI 74% to 84%) and 78.3% (95% CI, 71% to 85%), respectively, and a higher test accuracy of 81.1% and 80.0%, respectively (see Table 4) . Both the methods also reported a higher recall of 75.0% and 70.0%, respectively, and a reasonable precision of 78.9% and 80.0%, respectively, as compared with other methods. Recall can be interpreted as “Of the participant-day instances that were actually symptoms of depression, what proportion was classified as having symptoms of depression.” Precision can be interpreted as “Of the participant-day instances that were classified as symptoms of depression, what proportion actually had symptoms of depression.”

Table 4. Classification performance.

Performance	SVM[a] (RBF[b])	Decision tree – single	Random forest	Extreme gradient boosting	Adaptive boost	Voting ensemble
Accuracy						
Average cross-validation accuracy, % (95% CI)	73.8 (67-81)	69.1 (60-78)	78.3 (71-85)	79.1 (74-84)	74.3 (67-81)	75.3 (68-82)
Test accuracy, %	80.0	66.3	80.0	81.1	73.7	77.9
Precision (test), %	86.2	60.5	80.0	78.9	75.9	80.6
Sensitivity and recall (test), %	62.5	57.5	70.0	75.0	55.0	62.5
Specificity (test), %	92.7	72.7	87.3	85.5	87.3	89.1
Confusion matrix: training counts						
True positive	228	304	324	330	241	321
True negative	487	503	521	520	485	519
False positive	34	18	0	1	36	2
False negative	106	30	10	4	93	13
Confusion matrix: test counts						
True positive	25	23	28	30	22	25
True negative	51	40	48	47	48	49
False positive	4	15	7	8	7	6
False negative	15	17	12	10	18	15

[a]SVM: support vector machine.

[b]RBF: radial basis function.

Discussion

Principal Findings

A noticeably high prevalence of self-reported symptoms of depression (63%) was observed in this study as compared with the 8% to 35% normally reported in other studies [4]. This could be attributable to the single study site and the characteristics of the recruited participants. A detailed analysis was not within the scope of study.

Low correlation was observed between self-reported symptoms of depression and each of the derived sensing variables. This contrasts from the results observed in the Dartmouth Student Life study [41] where results indicated a strong correlation between automatic sensing data (derived for sleep, conversation, and location) and PHQ-9 scores. The difference can be attributed to the use of different sets of derived sensing variables and possibly due to missing values in the sensing variables. Again, a detailed analysis on correlation was not within the scope of study. The study did show, at 95% and 99% significance, lower levels of social contact (total calls—made and received), higher phone access at night (number of screen-ons), and lower daytime activity (activity rates during the day) among those with self-reported symptoms of depression (PHQ-9>9).

In total, 1744 out of 2694 participant-day instances were removed from the dataset as they contained missing values in 1 or more of the derived sensing variables. Between large numbers of training instances available for modeling and avoidance of any bias being introduced in the dataset because of imputation of missing values in the derived sensing variables, a decision was taken in favor of the latter. The XGBoost and RF methods were able to classify each participant-day instance with a test accuracy of 81.1% and 80.0%, respectively, and with a sensitivity and recall of 75.0% and 70.0%, respectively. From among the recently published passive sensing studies, only 1 study [38] was found comparable with the approach followed in our study. In that, they used a smartphone app to collect sensing data, used a PHQ-9 self-report scale but with a cutoff of 11 to separate participants into 2 classes, and also built binary classifiers with a leave-one-out CV approach to predict symptoms of depression. That study used 2 classification methods, namely RF and SVM, which resulted in an accuracy of 61% and 59%, respectively, and a sensitivity and recall of 62% and 72%, respectively. Although the results obtained from the *Project SHADO* study did show better performance, it would not be appropriate to make a direct comparison given the different study design adopted by both papers. However, both studies did show a performance superior to a random classification. The classifier results show promise and points to a need to develop high-performing ML models, including evaluation of unsupervised learning approaches as ML and smartphone-sensing technologies advance. A high classifier performance allows for an early and improved detection of symptoms of depression among patients with diabetes. This enables the primary care physician to use the results of the classifier as one of the several biomarkers for high-risk classification and prioritization of patients with diabetes and provide for personalized and empathetic care.

Limitations

Key limitations of the study included a single study site, small participant size, and the nonrandomized–based approach. The average of 2 consecutive PHQ-9 scores reported by a participant was assumed to be the depression symptom of the participant for the days and instances between the 2 screening time points. This could induce an error in outcome on days or instances where a participant was to exhibit a different mood or symptom. There is a need to investigate a better approach to capture daily symptoms of depression for each participant instead of the imputation approach taken for this study. PHQ-9, although a validated scale to screen for symptoms of depression, is not a tool to firmly diagnose depression [62]. Therefore, participants with high PHQ-9 scores need not necessarily have depression and vice versa. The study was also limited by missing values in derived sensing variables. The study design should include a plan for specific follow-up with participants, without influencing the participants, to help reduce possibilities of missing values in sensing information that might be introduced due to the participant's smartphone usage behavior. Although several measures to monitor and ensure data quality were used, limitations do exist in the data collection methods. Data collection methods can be improved further to manage the variability of mobile devices and how they respond to the data collection code set. The approach and observations from this pilot study are at best preliminary and a larger, randomized control–based study would help in validation of the findings.

Conclusions

Smartphone sensor–enabled daily digital observation of health and social attributes is a promising new approach with significant potential for management of comorbid conditions. Although the findings need to be replicated with a larger multisite randomized control study, this observational study has opened up the possibility of understanding the real-world everyday mental well-being and social attributes of people with diabetes in a clinical setting. Supplementing the smartphone-sensing data with clinical records from each visit along with daily behavioral information aggregated from a smartphone-based conversational chatbot app [63] would help further the risk stratification objectives. It is equally important to be sensitive and treat passive sensing data as sensitive health information and ensure adequate privacy and security controls are in place before wider use.

Acknowledgments

Touchkin funded the pilot study and the publication fees for the study. Touchkin provided access to their smartphone app product along with technical support during the study period. Participant recruitment and study administration were managed by clinical staff and administrators at Sarda Centre for Diabetes and Selfcare. The study acknowledges the following for their support during the study: Dr Banshi Saboo, MD, FIACM, FICN, FACE, MNAMS, a diabetologist associated with Dia Care & Hormone Clinic, Diabetology, Ahmedabad, India; Dr Girija Chetty, Associate Professor, PhD, associated with the University of Canberra, Computer Science, Canberra, Australia; Stephen Raj and Priyadarsini Suresh, associated with Public Health Foundation of India-Indian Institute of Public Health, Health Informatics, Hyderabad, India. The study acknowledges Soundarya Murthy, content editor at Touchkin, for her paper reviews and Himanshu Panwar, artificial intelligence engineer at Touchkin, along with Bindhya Rajendran and Aayush Agarwal, analytics interns at Touchkin, for their support in building and validating the machine learning methods.

Authors' Contributions

Study concept and design was provided by AS with support from SM. Ethics committee approval for the study was provided by SM. Study administration was provided by AS, SS, and the staff at diabetes clinic. Data collection, secure storage, and data aggregation were managed by SS. Research and data analysis were done by VS, SM, and SS. VS helped in drafting the paper. Data interpretation was done by all authors. The paper was reviewed by all authors.

Conflicts of Interest

AS (first author) is also an advisor to Touchkin and without any fiduciary relation. SM (second author) is a program coordinator with Public Health Foundation of India and holds no conflict of interest. SS (third author) is an engineering lead and a paid employee of Touchkin. VS (fourth author) is an independent research consultant at Touchkin and draws a consulting fee.

Multimedia Appendix 1

Patient Health Questionnaire- 9 English language version.

[PDF File (Adobe PDF File), 39KB - mhealth_v7i1e11041_app1.pdf]

Multimedia Appendix 2

Patient Health Questionnaire- 9 Marathi language version.

[PDF File (Adobe PDF File), 109KB - mhealth_v7i1e11041_app2.pdf]

Multimedia Appendix 3

Conceptual framework of study research design.

[PNG File, 20KB - mhealth_v7i1e11041_app3.png]

Multimedia Appendix 4

Derived smartphone-sensing variable description.

[PDF File (Adobe PDF File), 68KB - mhealth_v7i1e11041_app4.pdf]

Multimedia Appendix 5

Measures taken to ensure sensing data integrity and accuracy.

[PDF File (Adobe PDF File), 23KB - mhealth_v7i1e11041_app5.pdf]

Multimedia Appendix 6

Pearson correlation of derived sensing variables and depression state.

[PNG File, 163KB - mhealth_v7i1e11041_app6.png]

References

1. Golden SH, Lazo M, Carnethon M, Bertoni AG, Schreiner PJ, Diez Roux AV, et al. Examining a bidirectional association between depressive symptoms and diabetes. J Am Med Assoc 2008 Jun 18;299(23):2751-2759 [FREE Full text] [doi: 10.1001/jama.299.23.2751] [Medline: 18560002]
2. Nouwen A, Winkley K, Twisk J, Lloyd CE, Peyrot M, Ismail K, European Depression in Diabetes (EDID) Research Consortium. Type 2 diabetes mellitus as a risk factor for the onset of depression: a systematic review and meta-analysis. Diabetologia 2010 Dec;53(12):2480-2486 [FREE Full text] [doi: 10.1007/s00125-010-1874-x] [Medline: 20711716]
3. Anderson RJ, Freedland KE, Clouse RE, Lustman PJ. The prevalence of comorbid depression in adults with diabetes: a meta-analysis. Diabetes Care 2001 Jun;24(6):1069-1078 [FREE Full text] [Medline: 11375373]
4. Andreoulakis E, Hyphantis T, Kandylis D, Iacovides A. Depression in diabetes mellitus: a comprehensive review. Hippokratia 2012 Jul;16(3):205-214 [FREE Full text] [Medline: 23935284]
5. Martin LR, Williams SL, Haskard KB, Dimatteo MR. The challenge of patient adherence. Ther Clin Risk Manag 2005 Sep;1(3):189-199 [FREE Full text] [Medline: 18360559]
6. Gonzalez JS, Peyrot M, McCarl LA, Collins EM, Serpa L, Mimiaga MJ, et al. Depression and diabetes treatment nonadherence: a meta-analysis. Diabetes Care 2008 Dec;31(12):2398-2403 [FREE Full text] [doi: 10.2337/dc08-1341] [Medline: 19033420]
7. Dirmaier J, Watzke B, Koch U, Schulz H, Lehnert H, Pieper L, et al. Diabetes in primary care: prospective associations between depression, nonadherence and glycemic control. Psychother Psychosom 2010;79(3):172-178. [doi: 10.1159/000296135] [Medline: 20234147]
8. Chapman DP, Perry GS, Strine TW. The vital link between chronic disease and depressive disorders. Prev Chronic Dis 2005 Jan;2(1):A14 [FREE Full text] [Medline: 15670467]
9. Egede LE, Walker RJ, Bishu K, Dismuke CE. Trends in costs of depression in adults with diabetes in the United States: Medical Expenditure Panel Survey, 2004-2011. J Gen Intern Med 2016 Jun;31(6):615-622. [doi: 10.1007/s11606-016-3650-1] [Medline: 26969312]
10. Musselman DL, Betan E, Larsen H, Phillips LS. Relationship of depression to diabetes types 1 and 2: epidemiology, biology, and treatment. Biol Psychiatry 2003 Aug 01;54(3):317-329. [Medline: 12893107]
11. Talbot F, Nouwen A. A review of the relationship between depression and diabetes in adults: is there a link? Diabetes Care 2000 Oct;23(10):1556-1562. [Medline: 11023152]
12. Aguilera A, Lyles CR. The case for jointly targeting diabetes and depression among vulnerable patients using digital technology. JMIR Diabetes 2017 Jan 17;2(1):e1 [FREE Full text] [doi: 10.2196/diabetes.6916] [Medline: 30291080]
13. Colon E, Giachello A, McIver L, Pacheco G, Vela L. Diabetes and depression in the Hispanic/Latino community. Clin Diabetes 2013 Jan 17;31(1):43-45. [doi: 10.2337/diaclin.31.1.43]
14. Schram MT, Baan CA, Pouwer F. Depression and quality of life in patients with diabetes: a systematic review from the European Depression in Diabetes (EDID) Research Consortium. Curr Diabetes Rev 2009 May;5(2):112-119 [FREE Full text] [Medline: 19442096]
15. Katon WJ, Rutter C, Simon G, Lin EH, Ludman E, Ciechanowski P, et al. The association of comorbid depression with mortality in patients with type 2 diabetes. Diabetes Care 2005 Nov;28(11):2668-2672 [FREE Full text] [Medline: 16249537]

16. Lustman PJ, Clouse RE. Depression in diabetic patients: the relationship between mood and glycemic control. J Diabetes Complications 2005;19(2):113-122. [doi: 10.1016/j.jdiacomp.2004.01.002] [Medline: 15745842]
17. McKellar JD, Humphreys K, Piette JD. Depression increases diabetes symptoms by complicating patients' self-care adherence. Diabetes Educ 2004;30(3):485-492. [doi: 10.1177/014572170403000320] [Medline: 15208846]
18. Koopmans B, Pouwer F, de Bie RA, van Rooij ES, Leusink GL, Pop VJ. Depressive symptoms are associated with physical inactivity in patients with type 2 diabetes. The DIAZOB Primary Care Diabetes study. Fam Pract 2009 Jun;26(3):171-173. [doi: 10.1093/fampra/cmp016] [Medline: 19321598]
19. von Korff M, Katon W, Lin EH, Simon G, Ludman E, Oliver M, et al. Potentially modifiable factors associated with disability among people with diabetes. Psychosom Med 2005;67(2):233-240. [doi: 10.1097/01.psy.0000155662.82621.50] [Medline: 15784788]
20. Wittchen H, Mühlig S, Beesdo K. Mental disorders in primary care. Dialogues Clin Neurosci 2003 Jun;5(2):115-128 [FREE Full text] [Medline: 22034245]
21. Katon WJ, Simon G, Russo J, von Korff M, Lin EH, Ludman E, et al. Quality of depression care in a population-based sample of patients with diabetes and major depression. Med Care 2004 Dec;42(12):1222-1229. [Medline: 15550802]
22. Cherrington A, Ayala GX, Sleath B, Corbie-Smith G. Examining knowledge, attitudes, and beliefs about depression among Latino adults with type 2 diabetes. Diabetes Educ 2006;32(4):603-613. [doi: 10.1177/0145721706290836] [Medline: 16873598]
23. Rubin RR, Knowler WC, Ma Y, Marrero DG, Edelstein SL, Walker EA, Diabetes Prevention Program Research Group. Depression symptoms and antidepressant medicine use in Diabetes Prevention Program participants. Diabetes Care 2005 Apr;28(4):830-837 [FREE Full text] [Medline: 15793181]
24. Rubin RR, Ciechanowski P, Egede LE, Lin EH, Lustman PJ. Recognizing and treating depression in patients with diabetes. Curr Diab Rep 2004 Apr;4(2):119-125. [Medline: 15035972]
25. American Diabetes Association. Standards of medical care in diabetes-2015 abridged for primary care providers. Clin Diabetes 2015 Apr;33(2):97-111 [FREE Full text] [doi: 10.2337/diaclin.33.2.97] [Medline: 25897193]
26. Akincigil A, Matthews EB. National rates and patterns of depression screening in primary care: results from 2012 and 2013. Psychiatr Serv 2017 Jul 01;68(7):660-666. [doi: 10.1176/appi.ps.201600096] [Medline: 28196461]
27. eMarketer. 2014 Dec 11. 2 Billion Consumers Worldwide to Get Smart(phones) by 2016 URL: https://www.emarketer.com/Article/2-Billion-Consumers-Worldwide-Smartphones-by-2016/1011694 [accessed 2018-01-15] [WebCite Cache ID 6wTzEPULT]
28. Poushter J. Pew Research Center. 2016 Feb 22. Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies URL: http://assets.pewresearch.org/wp-content/uploads/sites/2/2016/02/pew_research_center_global_technology_report_final_february_22__2016.pdf [accessed 2018-01-15] [WebCite Cache ID 6wU1u4wUn]
29. Harari GM, Lane ND, Wang R, Crosier BS, Campbell AT, Gosling SD. Using smartphones to collect behavioral data in psychological science: opportunities, practical considerations, and challenges. Perspect Psychol Sci 2016 Nov;11(6):838-854. [doi: 10.1177/1745691616650285] [Medline: 27899727]
30. Doryab A, Min J, Wiese J, Zimmerman J, Hong J. Association for the Advancement of Artificial Intelligence. North America: AAAI Workshops; 2014 Jun. Detection of behavior change in people with depression URL: https://www.aaai.org/ocs/index.php/WS/AAAIW14/paper/view/8850/8303 [accessed 2018-12-10] [WebCite Cache ID 74Yaewm7M]
31. Glynn LG, Hayes PS, Casey M, Glynn F, Alvarez-Iglesias A, Newell J, et al. Effectiveness of a smartphone application to promote physical activity in primary care: the SMART MOVE randomised controlled trial. Br J Gen Pract 2014 Jul;64(624):e384-e391. [doi: 10.3399/bjgp14X680461] [Medline: 24982490]
32. Canzian L, Musolesi M. Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2015 Presented at: UbiComp'15; September 07-11, 2015; Osaka, Japan p. 1293-1304. [doi: 10.1145/2750858.2805845]
33. Cornet VP, Holden RJ. Systematic review of smartphone-based passive sensing for health and wellbeing. J Biomed Inform 2018 Jan;77:120-132. [doi: 10.1016/j.jbi.2017.12.008] [Medline: 29248628]
34. DeMasi O, Feygin S, Dembo A, Aguilera A, Recht B. Well-being tracking via smartphone-measured activity and sleep: cohort study. JMIR Mhealth Uhealth 2017 Oct 05;5(10):e137 [FREE Full text] [doi: 10.2196/mhealth.7820] [Medline: 28982643]
35. Jaques N, Taylor S, Azaria A, Ghandeharioun A, Sano A, Picard R. Predicting students' happiness from physiology, phone, mobility, and behavioral data. In: International Conference on Affective Computing and Intelligent Interaction and workshops: [proceedings]. 2015 Sep Presented at: ACII 2015; September 21-24, 2015; Xi'an, China p. 222-228 URL: http://europepmc.org/articles/PMC5431070 [doi: 10.1109/ACII.2015.7344575]
36. Rabbi M, Ali S, Choudhury T, Berke E. Passive and in-situ assessment of mental and physical well-being using mobile sensors. In: Proceedings of the 13th international conference on Ubiquitous computing. 2011 Presented at: UbiComp'11; September 17 - 21, 2011; Beijing, China p. 385-394 URL: http://europepmc.org/abstract/MED/25285324 [doi: 10.1145/2030112.2030164]

37. Place S, Blanch-Hartigan D, Rubin C, Gorrostieta C, Mead C, Kane J, 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]
38. Wahle F, Kowatsch T, Fleisch E, Rufer M, Weidt S. Mobile sensing and support for people with depression: a pilot trial in the wild. JMIR Mhealth Uhealth 2016 Sep 21;4(3):e111 [FREE Full text] [doi: 10.2196/mhealth.5960] [Medline: 27655245]
39. Ben-Zeev D, Scherer EA, Wang R, Xie H, Campbell AT. Next-generation psychiatric assessment: using smartphone sensors to monitor behavior and mental health. Psychiatr Rehabil J 2015 Sep;38(3):218-226 [FREE Full text] [doi: 10.1037/prj0000130] [Medline: 25844912]
40. Madan A, Cebrian M, Lazer D, Pentland A. Social sensing for epidemiological behaviour change. In: Proceedings of the 12th ACM international conference on Ubiquitous computing. 2010 Presented at: UbiComp'10; September 26 - 29, 2010; Copenhagen, Denmark p. 291-300. [doi: 10.1145/1864349.1864394]
41. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S, et al. Student Life: assessing mental health, academic performance and behavioural trends of college students using smartphones. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing. 2014 Presented at: UbiComp'14; September 13 - 17, 2014; Seattle, Washington p. 3-14. [doi: 10.1145/2632048.2632054]
42. Torous J, Friedman R, Keshavan M. Smartphone ownership and interest in mobile applications to monitor symptoms of mental health conditions. JMIR Mhealth Uhealth 2014;2(1):e2 [FREE Full text] [doi: 10.2196/mhealth.2994] [Medline: 25098314]
43. Osmani V. Smartphones in mental health: detecting depressive and manic episodes. IEEE Pervasive Comput 2015;14(3):10-13. [doi: 10.1109/MPRV.2015.54]
44. Hung GC, Yang PC, Chang CC, Chiang JH, Chen YY. Predicting negative emotions based on mobile phone usage patterns: an exploratory study. JMIR Res Protoc 2016 Aug 10;5(3):e160 [FREE Full text] [doi: 10.2196/resprot.5551] [Medline: 27511748]
45. Saeb S, Lattie EG, Schueller SM, Kording KP, Mohr DC. The relationship between mobile phone location sensor data and depressive symptom severity. PeerJ 2016;4:e2537 [FREE Full text] [doi: 10.7717/peerj.2537] [Medline: 28344895]
46. Saeb S, Zhang M, Karr CJ, Schueller SM, Corden ME, Kording KP, et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study. J Med Internet Res 2015;17(7):e175 [FREE Full text] [doi: 10.2196/jmir.4273] [Medline: 26180009]
47. Saeb S, Zhang M, Kwasny MM, Karr CJ, Kording K, Mohr DC. The Relationship between Clinical, Momentary, and Sensor-based Assessment of Depression. In: Proceedings of the 9th International Conference on Pervasive Computing Technologies for Healthcare. 2015 Aug Presented at: PervasiveHealth'15; May 20-23, 2015; Istanbul, Turkey p. 229-232 URL: http://europepmc.org/abstract/MED/26640739 [doi: 10.4108/icst.pervasivehealth.2015.259034]
48. Jin H, Wu S, Vidyanti I, di Capua P, Wu B. Predicting depression among patients with diabetes using longitudinal data. A multilevel regression model. Methods Inf Med 2015;54(6):553-559. [doi: 10.3414/ME14-02-0009] [Medline: 26577265]
49. Jin H, Wu S, di Capua P. Development of a clinical forecasting model to predict comorbid depression among diabetes patients and an application in depression screening policy making. Prev Chronic Dis 2015 Sep 03;12:E142 [FREE Full text] [doi: 10.5888/pcd12.150047] [Medline: 26334714]
50. Khater D, Omar M. Frequency and risk factors of depression in type 1 diabetes in a developing country. J Pediatr Endocrinol Metab 2017 Aug 28;30(9):917-922. [doi: 10.1515/jpem-2016-0414] [Medline: 28771435]
51. Kocalevent R, Hinz A, Brähler E. Standardization of the depression screener Patient Health Questionnaire (PHQ-9) in the general population. Gen Hosp Psychiatry 2013;35(5):551-555. [doi: 10.1016/j.genhosppsych.2013.04.006] [Medline: 23664569]
52. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med 2001 Sep;16(9):606-613 [FREE Full text] [Medline: 11556941]
53. Shidhaye R, Gangale S, Patel V. Prevalence and treatment coverage for depression: a population-based survey in Vidarbha, India. Soc Psychiatry Psychiatr Epidemiol 2016 Dec;51(7):993-1003 [FREE Full text] [doi: 10.1007/s00127-016-1220-9] [Medline: 27106852]
54. Kochhar PH, Rajadhyaksha SS, Suvarna VR. Translation and validation of brief patient health questionnaire against DSM IV as a tool to diagnose major depressive disorder in Indian patients. J Postgrad Med 2007;53(2):102-107 [FREE Full text] [Medline: 17495375]
55. Boonstra TW, Nicholas J, Wong QJ, Shaw F, Townsend S, Christensen H. Using mobile phone sensor technology for mental health research: integrated analysis to identify hidden challenges and potential solutions. J Med Internet Res 2018 Jul 30;20(7):e10131 [FREE Full text] [doi: 10.2196/10131] [Medline: 30061092]
56. Manea L, Gilbody S, McMillan D. Optimal cut-off score for diagnosing depression with the Patient Health Questionnaire (PHQ-9): a meta-analysis. CMAJ 2012 Feb 21;184(3):E191-E196 [FREE Full text] [doi: 10.1503/cmaj.110829] [Medline: 22184363]
57. Rochon J, Gondan M, Kieser M. To test or not to test: preliminary assessment of normality when comparing two independent samples. BMC Med Res Methodol 2012 Jun 19;12:81 [FREE Full text] [doi: 10.1186/1471-2288-12-81] [Medline: 22712852]

58. Ruxton GD. The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test. Behav Ecol 2006 Jul 01;17(4):688-690 [FREE Full text] [doi: 10.1093/beheco/ark016]
59. Hermanns N, Kulzer B, Krichbaum M, Kubiak T, Haak T. How to screen for depression and emotional problems in patients with diabetes: comparison of screening characteristics of depression questionnaires, measurement of diabetes-specific emotional problems and standard clinical assessment. Diabetologia 2006 Mar;49(3):469-477. [doi: 10.1007/s00125-005-0094-2] [Medline: 16432706]
60. Cuijpers P, Smit F. Subthreshold depression as a risk indicator for major depressive disorder: a systematic review of prospective studies. Acta Psychiatr Scand 2004 May;109(5):325-331. [doi: 10.1111/j.1600-0447.2004.00301.x] [Medline: 15049768]
61. Hermanns N, Schmitt A, Gahr A, Herder C, Nowotny B, Roden M, et al. The effect of a Diabetes-Specific Cognitive Behavioral Treatment Program (DIAMOS) for patients with diabetes and subclinical depression: results of a randomized controlled trial. Diabetes Care 2015 Apr;38(4):551-560. [doi: 10.2337/dc14-1416] [Medline: 25605812]
62. Mitchell AJ, Yadegarfar M, Gill J, Stubbs B. Case finding and screening clinical utility of the Patient Health Questionnaire (PHQ-9 and PHQ-2) for depression in primary care: a diagnostic meta-analysis of 40 studies. BJPsych Open 2016 Mar;2(2):127-138 [FREE Full text] [doi: 10.1192/bjpo.bp.115.001685] [Medline: 27703765]
63. Inkster B, Sarda S, Subramanian V. An empathy-driven, conversational artificial intelligence agent (Wysa) for digital mental well-being: real-world data evaluation mixed-methods study. JMIR Mhealth Uhealth 2018 Nov 23;6(11):e12106 [FREE Full text] [doi: 10.2196/12106] [Medline: 30470676]

Abbreviations

AdaBoost: adaptive boosting
CV: cross-validation
D: depressed
D-ML: depressed-machine learning
DT: decision tree
GPS: global positioning system
ML: machine learning
ND: not depressed
ND-ML: not depressed-machine learning
PHQ-9: Patient Health Questionnaire-9
RBF: radial basis function
RF: random forest
Project SHADO: Analyzing Social and Health Attributes through Daily Digital Observation
SVM: support vector machine
XGBoost: extreme gradient boosting

Edited by G Eysenbach; submitted 14.05.18; peer-reviewed by P Soderlund, F Fatehi, Z Huang; comments to author 24.09.18; revised version received 19.11.18; accepted 25.11.18; published 29.01.19.

Please cite as:
Sarda A, Munuswamy S, Sarda S, Subramanian V
Using Passive Smartphone Sensing for Improved Risk Stratification of Patients With Depression and Diabetes: Cross-Sectional Observational Study
JMIR Mhealth Uhealth 2019;7(1):e11041
URL: http://mhealth.jmir.org/2019/1/e11041/
doi:10.2196/11041
PMID:30694197

Original Paper

Impact of Personal Health Records and Wearables on Health Outcomes and Patient Response: Three-Arm Randomized Controlled Trial

Jeong-Whun Kim[1*], MD, PhD; Borim Ryu[2*], MS; Seoyoon Cho[2], MS; Eunyoung Heo[2], MS; Yoojung Kim[3], MS; Joongseek Lee[3], MS; Se Young Jung[2,4], MD, MPH; Sooyoung Yoo[2], PhD

[1]Department of Otorhinolaryngology, Seoul National University Bundang Hospital, Seongnam, Republic of Korea

[2]Office of eHealth Research and Businesses, Seoul National University Bundang Hospital, Seongnam, Republic of Korea

[3]Graduate School of Convergence Science and Technology, Seoul National University, Suwon, Republic of Korea

[4]Department of Family Medicine, Seoul National University Bundang Hospital, Seongnam, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Sooyoung Yoo, PhD
Office of eHealth Research and Businesses
Seoul National University Bundang Hospital
82 173rd Street
Gumi-ro, Bundang-gu
Seongnam, 13620
Republic of Korea
Phone: 82 317878980
Fax: 82 317874054
Email: yoosoo0@snubh.org

Abstract

Background: Although using the technologies for a variety of chronic health conditions such as personal health record (PHR) is reported to be acceptable and useful, there is a lack of evidence on the associations between the use of the technologies and the change of health outcome and patients' response to a digital health app.

Objective: This study aimed to examine the impact of the use of PHR and wearables on health outcome improvement and sustained use of the health app that can be associated with patient engagement.

Methods: We developed an Android-based mobile phone app and used a wristband-type activity tracker (Samsung Charm) to collect data on health-related daily activities from individual patients. Dietary record, daily step counts, sleep log, subjective stress amount, blood pressure, and weight values were recorded. We conducted a prospective randomized clinical trial across 4 weeks on those diagnosed with obstructive sleep apnea (OSA) who had visited the outpatient clinic of Seoul National University Bundang Hospital. The trial randomly assigned 60 patients to 3 subgroups including 2 intervention groups: (1) mobile app and wearable device users (n=20), (2) mobile app–only users (n=20), and (3) controls (n=20). The primary outcome measure was weight change. Body weights before and after the trial were recorded and analyzed during clinic visits. Changes in OSA–related respiratory parameters such as respiratory disturbance, apnea-hypopnea, and oxygenation desaturation indexes and snoring comprised the secondary outcome and were analyzed for each participant.

Results: We collected the individual data for each group during the trial, specifically anthropometric measurement and laboratory test results for health outcomes, and the app usage logs for patient response were collected and analyzed. The body weight showed a significant reduction in the 2 intervention groups after intervention, and the mobile app–only group showed more weight loss compared with the controls (P=.01). There were no significant changes in sleep-related health outcomes. From a patient response point of view, the average daily step counts (8165 steps) from the app plus wearable group were significantly higher than those (6034 steps) from the app-only group because they collected step count data from different devices (P=.02). The average rate of data collection was not different in physical activity (P=.99), food intake (P=.98), sleep (P=.95), stress (P=.70), and weight (P=.90) in the app plus wearable and app-only groups, respectively.

Conclusions: We tried to integrate PHR data that allow clinicians and patients to share lifelog data with the clinical workflow to support lifestyle interventions. Our results suggest that a PHR–based intervention may be successful in losing body weight and improvement in lifestyle behavior.

Trial Registration: ClinicalTrials.gov NCT03200223; https://clinicaltrials.gov/ct2/show/NCT03200223 (Archived by WebCite at http://www.webcitation.org/74baZmnCX).

(JMIR Mhealth Uhealth 2019;7(1):e12070) doi:10.2196/12070

KEYWORDS

personal health record; lifestyle; sleep apnea, obstructive; delivery of health care; electronic health record; mobile health

Introduction

Lifelog data or patient-generated health data, considered important for precision medicine initiative implementation, form the next frontier in patient engagement and customized health care [1,2]. Through the availability of numerous devices and compatible mobile apps, patients can collect their own health-related lifestyle data, which can be aggregated with their clinical data into their own personal health record (PHR) [3-6]. A lifelog is a detailed chronicle of a person's life involving large amounts of data. In recent years, the data are usually captured automatically by wearable technology or mobile devices. People who keep lifelogs about themselves are known as lifeloggers [7]. The purpose of lifelogging is to help users collect data for self-monitoring and reflection [8]. Technology has nearly reached the stage when all information, interesting or otherwise, generated in a lifetime by a single person can be assembled and queried relatively efficiently, creating a need for personal information management [9]. Lifelogging can be passive—one stores the by-products of the life one would have lived anyway—or active—one surrounds oneself with sensors and information capture tools to create as rich a picture of one's life as possible [10].

Clinical wearables can be defined as health technology that can be worn by the patient. The wearables contain sensors and use a wireless connection to pass data to a smartphone or similar devices. Wearables are being used in the health care industry to help health care practitioners collect, analyze, and leverage patient data for clinical trials while also significantly improving patient care and overall quality of life [11]. Recent technology advancements in health care have the potential to close the communication and information gap between patients and providers. This has created a mandate for more interactive, demand-responsive mobile health (mHealth) tools that empower consumers to actively manage their own health [12]. Currently, there is an increasing awareness of the health care system's responsibility to provide easily accessible ways for patients to be engaged in their own care by creating effective partnerships that lead to the patient's ability to make competent and well-informed decisions [13]. Although an electronic PHR tethered to an electronic health record (EHR), also known as a patient portal, is currently recognized as a promising mechanism to support greater patient engagement, questions remain about how health care leaders, policy makers, and designers can encourage adoption by both providers and patients and what factors might contribute to sustained utilization [14]. In addition to user-specific characteristics (eg, age, sex, and diagnosis), studies should be conducted on the modifiable factors that affect use duration, to facilitate activities that promote continued use [15].

In our previous study [16], we showed that patients with chronic diseases are more likely to use a PHR system that is integrated into a comprehensive EHR. We also showed that patients with more chronic diseases tend to use PHR more actively, employing the self-administration function. Furthermore, our other previous clinical trial study [17] primarily aimed to demonstrate the development of an EHR-tethered PHR system in which a comprehensive EHR system that can retrieve data from a wearable device has been operated successfully for over 12 years as well as the efficacy of such a system paired with a lifelog data–driven intervention modality [18].

Our clinical trial focused on patients with obstructive sleep apnea (OSA), which is caused by complete or partial obstruction of the upper airway because OSA is closely related to individual lifestyle [19]. Obesity is the most important cause of OSA, and smoking and alcohol consumption are also important causes of OSA. Therefore, a better lifestyle can be a cornerstone of treatment for both obesity and OSA. The aim of the clinical trial was to help OSA patients to correct obesity through a healthier lifestyle and to obtain a better quality of sleep. Smartphone platforms (mHealth systems) are being considered as an innovative solution, thanks to the integration of the essential sensors to obtain clinically relevant parameters in the same device or in combination with wireless wearable devices [20,21]. A recent study from Cardiogram and the University of California, San Francisco, suggested that the Apple Watch can be used to test for OSA. This study indicates that wearable devices could provide an accessible, low-cost approach to evaluate OSA [22]. From the perspective of weight loss, there have been several studies trying to verify the effects of technology-based interventions. There were studies comparing in-person behavioral weight loss intervention with a technology-based system over a 3-month period in overweight adults [23,24]. A recent study tried a long-term observation with wearables during a 24-month study period [25]. The study stated that effective long-term treatments are needed to address the obesity epidemic and that numerous wearable technologies are unclear if these are effective in improving weight loss.

Our first clinical study was a preliminary study aimed at observing the weight loss impact for obesity patients of conventional care versus EHR-integrated PHR–based care after system development [17,23]. In particular, as a continuous development from previous studies, we aimed to reveal the

effectiveness of a PHR–based health care app with 3 arms: mobile app and wearables, mobile app alone, and control. Furthermore, we examined the actual mobile lifelogs for each lifestyle categories: activity step counts, meals, sleep logs, and stress.

The primary objective of this more sophisticated 3-arm clinical trial design was to (1) explore the effect of a wearable device and mobile PHR app based on the patient's weight loss and sleep-related health outcome for a research reproducibility perspective, (2) observe patient response as a proxy measure for patient engagement in EHR-integrated PHR use, and (3) study health app usability based on patients' responses to automated lifestyle comments.

Methods

Design of Patient-Friendly Mobile Personal Health Record App

We designed a new mobile PHR app, *MyHealthKeeper* (Samsung Electronics Co, Seoul, South Korea), to be compatible with health data collection platforms of private companies and to be linked to our hospital EHR system. The user interfaces (Multimedia Appendix 1) were designed to be more patient-friendly when patients collect health-related lifestyle data on their own. A more sophisticated *MyHealthKeeper* PHR app to collect health-related lifestyle data was developed and tested in 2 different experimental patient groups (app and wearable user vs app-only user). The lifestyle data we tried to collect included weight, step counts as physical activity, food intake, sleep hours, and subjective strength of daily stress. The *MyHealthKeeper* app is largely composed of (1) logging according to the nature of the information and (2) screens registered in an accordion format, which can be used to navigate through the tabs. The app is composed of several subpages for recording daily meal, sleep log, stress, blood pressure, weight value, and synced activity step counts. When the user clicks the OK tab in the recording step, the objectives and guides, records, and daily span are presented, and the recording and modification functions are provided when each tab is clicked. The verification phase provides clinical advice that gives confirmation through weekly assessments. This clinician feedback comment is implemented on each page to improve communication between doctors and patients. The patient's response to automated comment logs as well as other usage logs was collected and analyzed for each group.

In particular, we designed our app to be compatible with the Samsung Health platform, which is one of the biggest worldwide mobile platforms, and to be linked to our EHR-tethered PHR system. This is the first approach to be aligned with a private company platform and a tertiary general university hospital PHR in Korea. Therefore, we could develop flexible data collection techniques that can leverage user lifestyle logs in both devices and smartphones from users (walking steps in this study). Samsung Health (originally S Health) is a free app developed by Samsung that tracks various aspects of daily life contributing to well-being such as physical activity, diet, and sleep [24]. The app was installed by default only on some smartphones of the brand. It could also be downloaded from the Samsung Galaxy Apps store. The app is obviously compatible with Samsung fitness trackers and smartwatches on Samsung phones. Users get health data from their preferred tools, reducing the burden of collection, and Samsung Charm (Samsung Electronics Co, Seoul, South Korea), which was developed by Samsung Electronics and was used to collect daily activity data in this study.

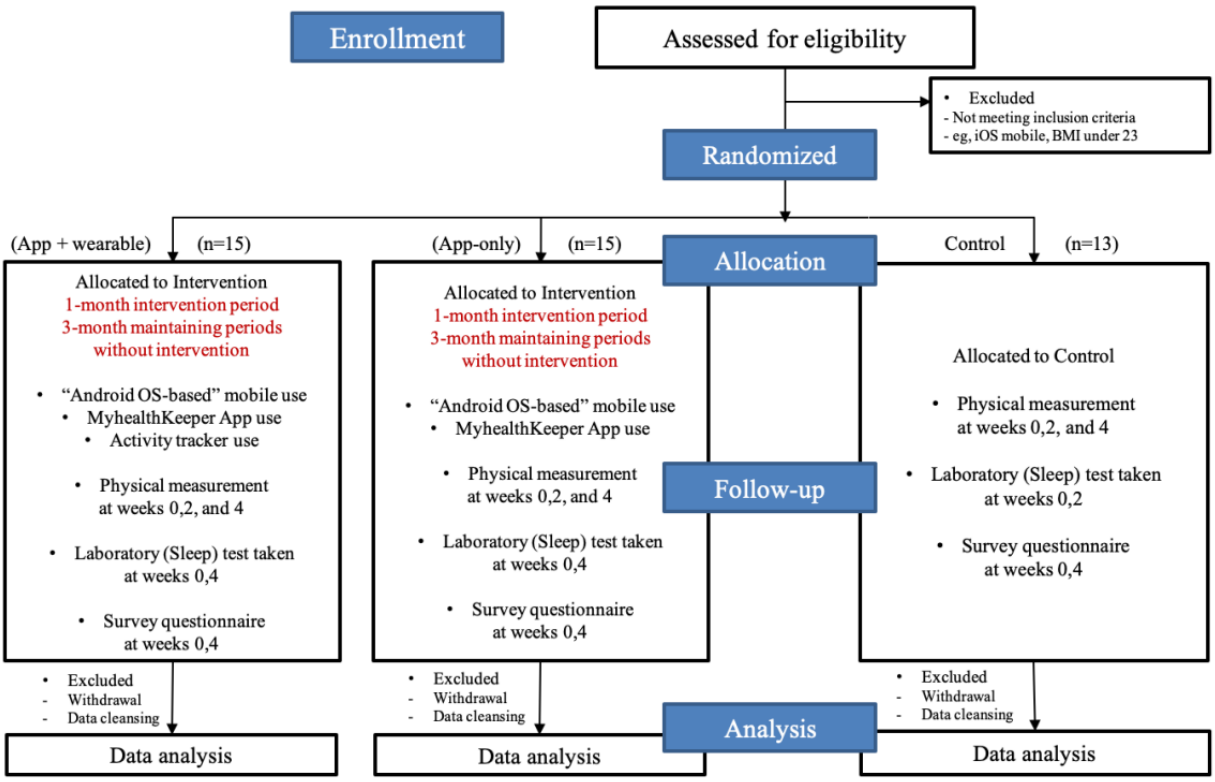
Design of a Clinical Trial

We conducted a prospective randomized clinical trial. Patients who visited the outpatient clinic of Seoul National University Bundang Hospital (SNUBH) were recruited from July 1 to August 31, 2017. We set the following inclusion criteria for enrollment in the trial: (1) diagnosis with OSA; (2) no cardiopulmonary disease, cancer, or other acute diseases; (3) body mass index (BMI) over 23 kg/m^2; and (4) provision of prior consent to comply with self-management. We excluded patients who would not be able to use a mobile app and a wearable device and those who were pregnant. Figure 1 demonstrates the overall clinical trial design in this study. The participants had to follow the instructions of the app and wearable for 4 weeks, but they could use the app and wearable after 4 weeks on their own.

The participants were randomized into 3 groups: (1) app plus wearable group, (2) app-only group, and (3) control group. The app plus wearable group used the PHR app and Samsung Charm wearable activity tracker band. The step counts of the app-only group were collected from the mobile phone itself and delivered to the Samsung Health platform. The app plus wearable and app-only groups were instructed to use the apps as per the guideline. On the contrary, the control group was managed ad libitum. The control group did not receive the PHR app or the wearable device, and they did not use any other intervention than the verbal advice to lose weight during their visit to the clinician. All the participants in the 3 groups underwent sleep WatchPAT tests twice at weeks 0 and 4 to measure an objective respiratory parameter (apnea-hypopnea index per hour) during sleep. WatchPAT is a Food and Drugs Administration–approved portable diagnostic device that uniquely uses finger-based physiology and innovative technology to enable simple and accurate OSA testing while avoiding the complexity and discomfort associated with traditional airflow-based systems [25]. Body weight and height were also recorded twice at weeks 0 and 4.

This study was approved by the SNUBH Institutional Review Board (B-1504-296-302) and registered at the United States National Institutes of Health clinical trial registry (ClinicalTrails.gov registration number: NCT03200223).

Figure 1. Clinical trial study design. BMI: body mass index; OS: operating system.



Personal Health Record–Based Patient Lifelog Data Collection

All study participants completed a paper-based survey at the first and last day of the trial. Subjective respiratory parameters during sleep in the survey questions included snoring frequency, snoring intensity, sleep apnea witnessed frequency, and daytime sleepiness. Snoring frequency, sleep apnea witnessed frequency, and daytime sleepiness were asked to answer by days in a week in the questionnaire (eg, How many days of sleep apnea a week is found by you or your family?). Snoring intensity, sleep apnea severity, and daytime sleepiness severity were asked to answer by 0- to 10-point scale in the questionnaire (eg, How loud is the snoring sound?; degree from 0 to 10). Furthermore, objective WatchPAT test was performed to examine deep and light sleep amount (percentage), Peripheral arterial tonometry Rapid Eye Movement (pREM; percentage), sleep latency (minutes), sleep efficiency, and number of wakes.

Meanwhile, the *MyHealthKeeper* app is designed and developed to collect manually recorded lifelogs and apply them to individual health management. Lifelog data collected by the app are categorized as weight, stress, meal, and sleep logs. Weight value, subjective stress, sleep time, and sleep satisfaction logs should be recorded once a day, whereas meal logs are recorded 3 times a day (breakfast, lunch, and dinner). Each meal record includes the input time and the amount on a 5-point scale. Physical activity is concerned with daily step counts, which are easily synced and delivered from the Samsung Health platform. The control group did not receive the lifestyle modification app and the wearable device. They received conventional care pertaining to lifestyle modification for achieving weight loss goals during the 4-week study period.

With *MyhealthKeeper* app, we calculated the average value of collected data from each lifelog such as physical activity, food intake, sleep time, daily stress, and weight. We simply combined all data logging scores for each patient. For example, if a user recorded every lifelog in a day, total score of the day is 5. Therefore, we could derive a star-shaped pentagon plot with 5 types of lifestyle log data. The mean area value for each plot is a range of 0 to 1 by every user. On the basis of this area value score, patients could be divided into 2 subgroups according to the median score. We defined a compliant subgroup with a score over 0.4 and noncompliant subgroup with a score under 0.4.

Clinical Study Outcome Measure

The primary outcome of the trial was collection of lifestyle data and weight change. Any decrease in body weight during the study period (4 weeks) was defined as successful weight reduction. The secondary outcomes were changes in the subjective and objective respiratory parameters during sleep.

Body weights before and after PHR–based clinical intervention were recorded and analyzed. The BMI of each participant is defined as the body mass divided by the square of the body height and is expressed in units of kg/m^2; the difference between BMI before and after the study was analyzed at the end of the study period. Any decrease in body weight during the study period (4 weeks) was defined as successful weight reduction. It is very important that the measurement be taken using the same method and in the same conditions to ensure uniformity between participants and in the same participant over time. In our study, a skilled nurse helped to measure the patient's body weight in the hospital health checkup center using the conventional health checkup process (place and dress).

Statistical Analysis Method

Results are presented as means (SD). Differences in various parameters between the PHR–based intervention group and the control group were analyzed using the chi-square test as appropriate. Paired *t* test was used to examine changes in primary or secondary outcomes in the groups. All statistical analyses were performed using R version 3.0.2 developed by R Core Team (R Foundation for Statistical Computing, Vienna 2013), and a *P* value of <.05 was considered statistically significant.

Results

Architecture of New MyhealthKeeper Linked to Samsung Health and Hospital Electronic Health Record

Our new PHR app, *MyhealthKeeper*, was developed to be compatible with the Samsung Health platform, which is one of the biggest worldwide mobile platforms, and to be linked to our hospital EHR system, Bestcare (ezCaretech Co, Seoul, South Korea; Figure 2).

The *MyHealthKeeper* interface was designed to work in the accordion format, which can be used to navigate through the tabs. The app was composed of several subpages for recording daily meal, sleep log, stress, blood pressure, weight value, and synced activity step counts. Weight value, subjective stress, sleep time, and sleep satisfaction logs were recorded once a day, whereas meal logs were recorded 3 times a day (breakfast, lunch, and dinner). Each meal record includes the input time and the amount on a 5-point scale. Daily step counts were collected through Samsung Health platform. Clinician feedback comments were implemented on each page to improve communication between doctors and patients (Figure 3). The patient's response to the comments as well as other subpage usage logs were collected and analyzed for each group.

General Characteristics of Clinical Trial Participants

Table 1 shows general characteristics of participants. A total of 60 patients (51 males and 9 females) were enrolled in this study. In addition, 43 patients (43 males) finished the study, whereas 17 patients (8 males and 9 females) were excluded because of withdrawal or incomplete follow-up sleep study. The analysis used a per-protocol methodology for the withdrawn patients. The mean age in app plus wearable, app-only, and control groups was 45.3, 41.5, and 40.5, respectively (*P*=.45). The BMI was 29.1, 30.8 and 28.7 in the app plus wearable, app-only, and control groups, respectively (*P*=.15). There were also no significant differences in demographics and smoking and drinking behaviors among the 3 groups.

Figure 2. MyHealthKeeper personal health record development.

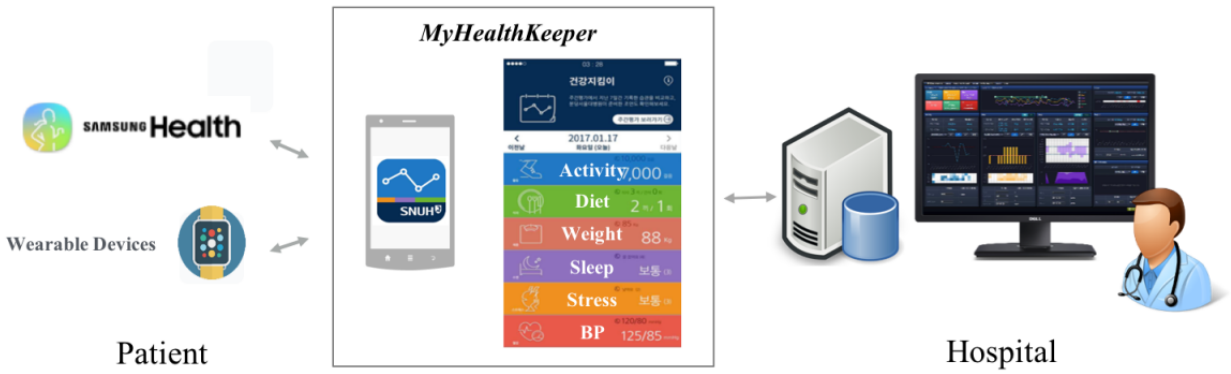


Figure 3. Clinician's comment delivery on individual lifestyle.

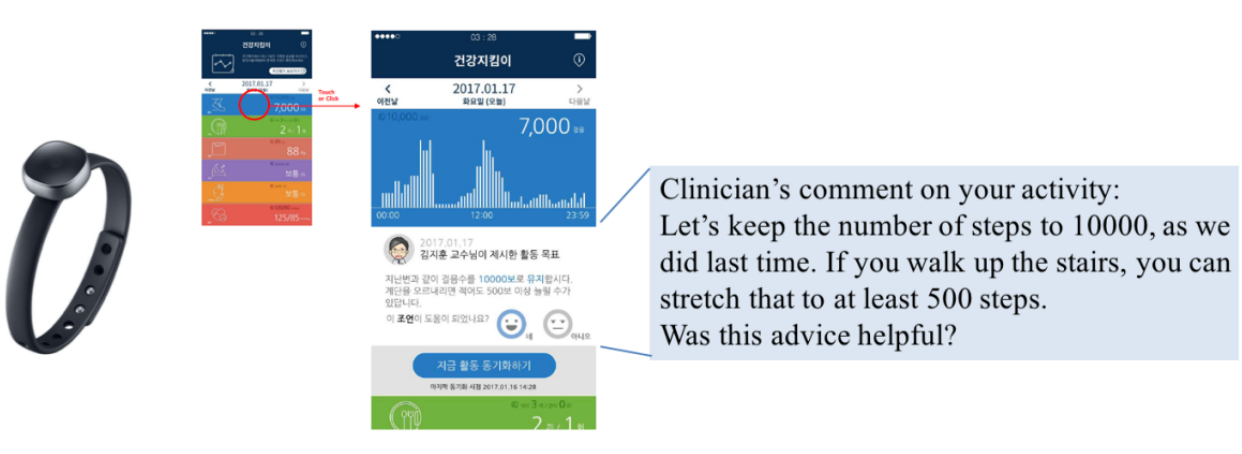


Table 1. General characteristics of study participants.

Characteristics	App + wearable (n=15)	App-only (n=15)	Control (n=13)	*P* value
Age (year), mean (SD)	45.3 (9.5)	41.5 (11.8)	40.5 (7.4)	.45
Weight (kg), mean (SD)	84.3 (5.6)	93.7 (22.3)	84.3 (12.1)	.13
Body mass index (kg/m^2), mean (SD)	29.1 (2.8)	30.8 (6.0)	28.7 (2.8)	.15
Height (cm), mean (SD)	170.4 (6.7)	174.0 (5.7)	171 (7.0)	.42
College or higher, n	13	12	13	.37
White collar, n	9	7	10	.36
Married, n	13	11	12	.44
Smokers, n	5	4	3	.77
Drink twice or more per week, n	5	6	6	.86

Changes in Health Outcome: Analysis of the Primary and Secondary Outcomes

Our clinical trial study results revealed a significant change in weight loss in both intervention groups (both wearable device users and PHR app–only users). Moreover, 2 PHR intervention group participants who used the *MyHealthKeeper* mobile app every day showed significantly larger changes in weight and BMI than those in the control group (Table 2; average: 1.4 kg and 2 kg; 95% CI: 0.9-1.9; *P*<.001) In this study, there were no statistically significant changes in subjective improvement of respiratory system test results (Table 3 and Figure 4).

Table 2. Primary outcome changes of each group.

Characteristics	App + wearable (n=15)			App-only (n=15)			Control (n=13)		
	Pre, mean (SD)	Post, mean (SD)	*P* value	Pre, mean (SD)	Post, mean (SD)	*P* value	Pre, mean (SD)	Post, mean (SD)	*P* value
Weight (kg)	84.3 (5.6)	82.9 (5.6)	0.02	93.7 (22.3)	91.7 (22.9)	0.003	84.3 (12.1)	83.9 (12.6)	0.3
Body mass index (kg/m^2)	29.1 (2.8)	28.7 (3.0)	0.02	30.8 (6.0)	30.1 (6.2)	0.002	28.7 (2.8)	28.6 (3.0)	0.34

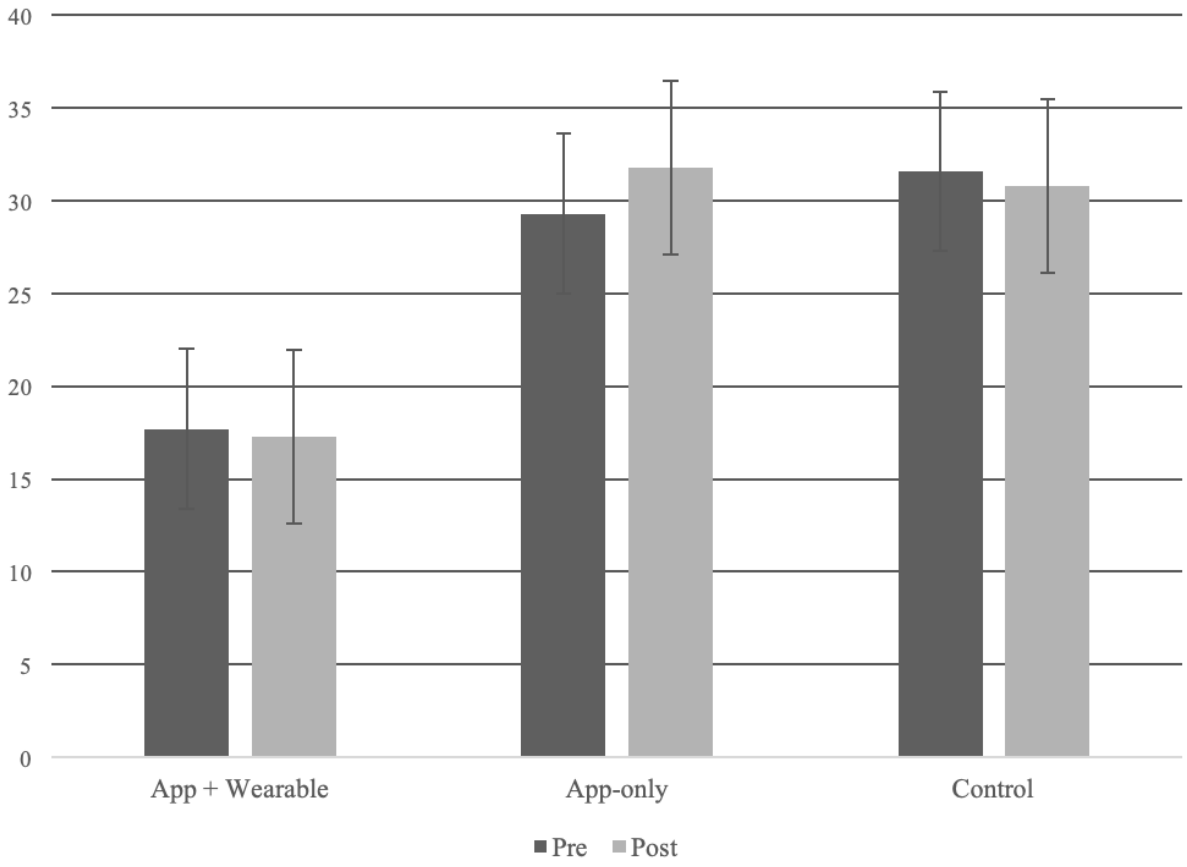
Table 3. Subjective improvement of respiratory system.

Characteristics	App + wearable (n=15)			App-only (n=15)			Control (n=13)			*P* value[a]
	Pre, mean (SD)	Post, mean (SD)	*P* value[b]	Pre, mean (SD)	Post, mean (SD)	*P* value[b]	Pre, mean (SD)	Post, mean (SD)	*P* value[b]	
Snoring frequency (days/week)	5.9 (1.5)	4.7 (2.3)	0.11	5.4 (1.9)	4.2 (2.7)	0.04	6.2 (1.3)	4.3 (2.7)	0.02	0.79
Apnea witnessed (days/week)	3.7 (2.7)	2.9 (2.8)	0.36	3.9 (2.6)	3 (2.7)	0.25	4.2 (2.7)	2.8 (2.7)	0.11	0.92
Daytime sleepiness (days/week)	5.1 (1.9)	3.6 (2.2)	0.01	4.2 (2.4)	3.4 (2.4)	0.17	3.7 (2.7)	2.5 (2.4)	0.22	0.62

[a]*P* value: pre- and postcomparison among the 3 groups.

[b]*P* value: pre- and postcomparison within each group.

Figure 4. Pre- and postchanges in apnea-hypopnea index profile.



Patients' Response: Comparison of Step Counts Between App Plus Wearable and App-Only Groups

The average daily step counts (8165 steps) from the app plus wearable group were significantly higher than those (6034 steps) from the app-only group because they collected step-count data from different devices (*P*=.02). However, the diurnal hourly pattern of step counts was almost similar (Figure 5). Due to the nature of the collection devices, such as wrist wearable or mobile phone, the gait information was slightly different.

We assumed that each lifestyle data record is patient engagement reflection as patient response to PHR system and the clinician. The data collection record rate was calculated according to the records of 1 day. Study participants were moderately required to record daily weight, stress, snack, meal, and sleep logs. For daily diet records, which can be logged as breakfast, lunch, and dinner, if any of the 3 records were recorded, we processed as 1 record for that day.

The average rate of data collection was not different in physical activity (49.82% vs 49.96%; *P*=.99), food intake (32.67% vs 32.82%; *P*=.98), sleep (32.01% vs 32.45%; *P*=.95), stress (30.11% vs 27.33%; *P*=.70), and weight (32.82% vs 31.87%; *P*=.90) in the app plus wearable and app-only groups, respectively.

We combined all the lifestyle data and data logging scores for each patient and derived a star-shaped pentagon plot with 5 types of lifestyle log data. The mean area value for each plot is a range of 0 to 1 by every user (Table 4). According to the composite lifestyle scores, we calculated the star-shaped pentagon plot area ranging from 0 to 1 (Table 4), and the plots were drawn as shown in Figure 6. Patients could be divided into 2 subgroups: compliant versus noncompliant, based on the average median value as 0.4 (Table 4), star plot square measure over 0.4 grouped as compliant. The percentage of compliant patients was 63.64% and 36.36% in the app plus wearable and app-only groups, respectively (*P*=.26). Figure 7 demonstrates lifelog records per each group: red color stands for recorded days by app plus wearable users, whereas blue color stands for recorded days by app-only users, and gray color is used to maintain period days without intervention.

Figure 5. Changes in daily steps per group.

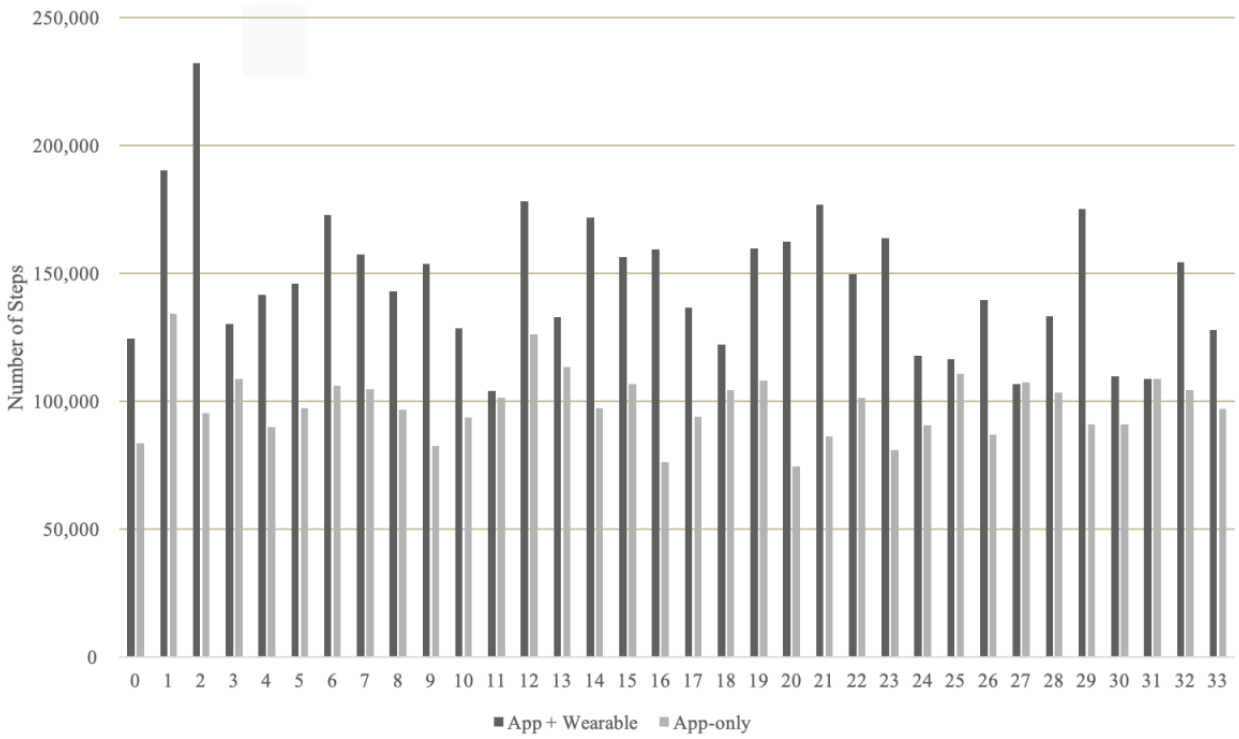


Table 4. Star-shaped pentagon plot area calculation with meal, snack, stress, and sleep log records. The area value was normalized between 0 to 1. Each intervention group participant's minimum, first quartile, median, mean, third quartile, and maximum value was described below.

Indices	Minimum	First quartile	Median	Mean	Third quartile	Maximum
App + wearable (n=15)	0.06	0.17	0.37	0.33	0.45	0.66
App-only (n=15)	0.01	0.09	0.14	0.25	0.4	0.67

Figure 6. Star plot based on recorded lifelog per each group.

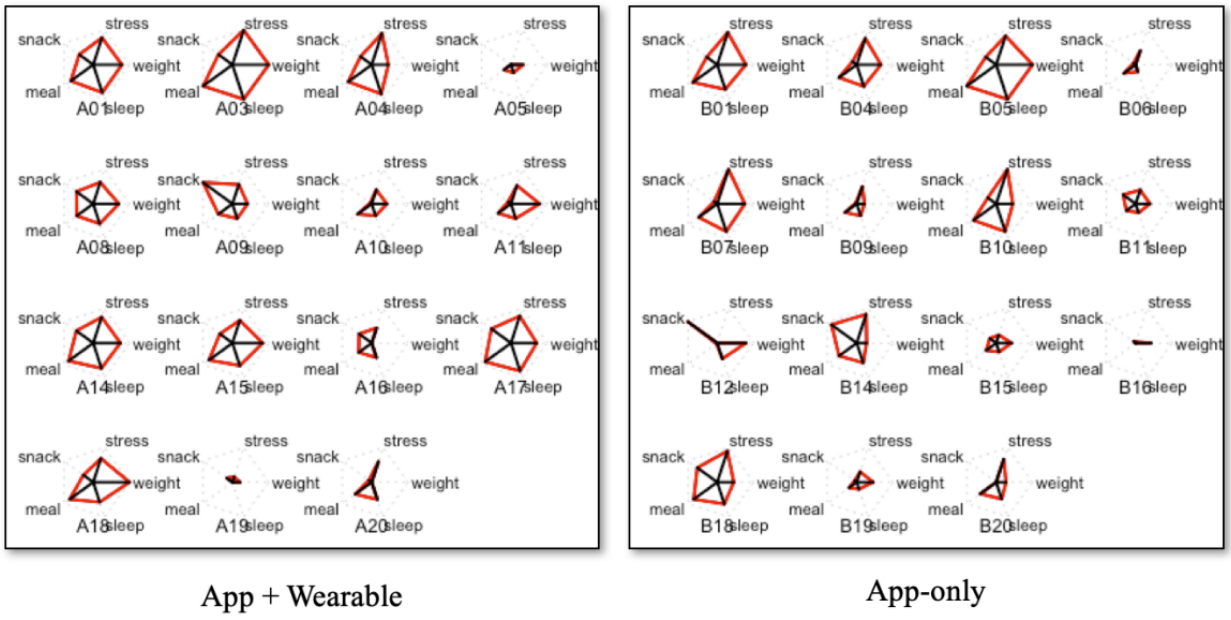
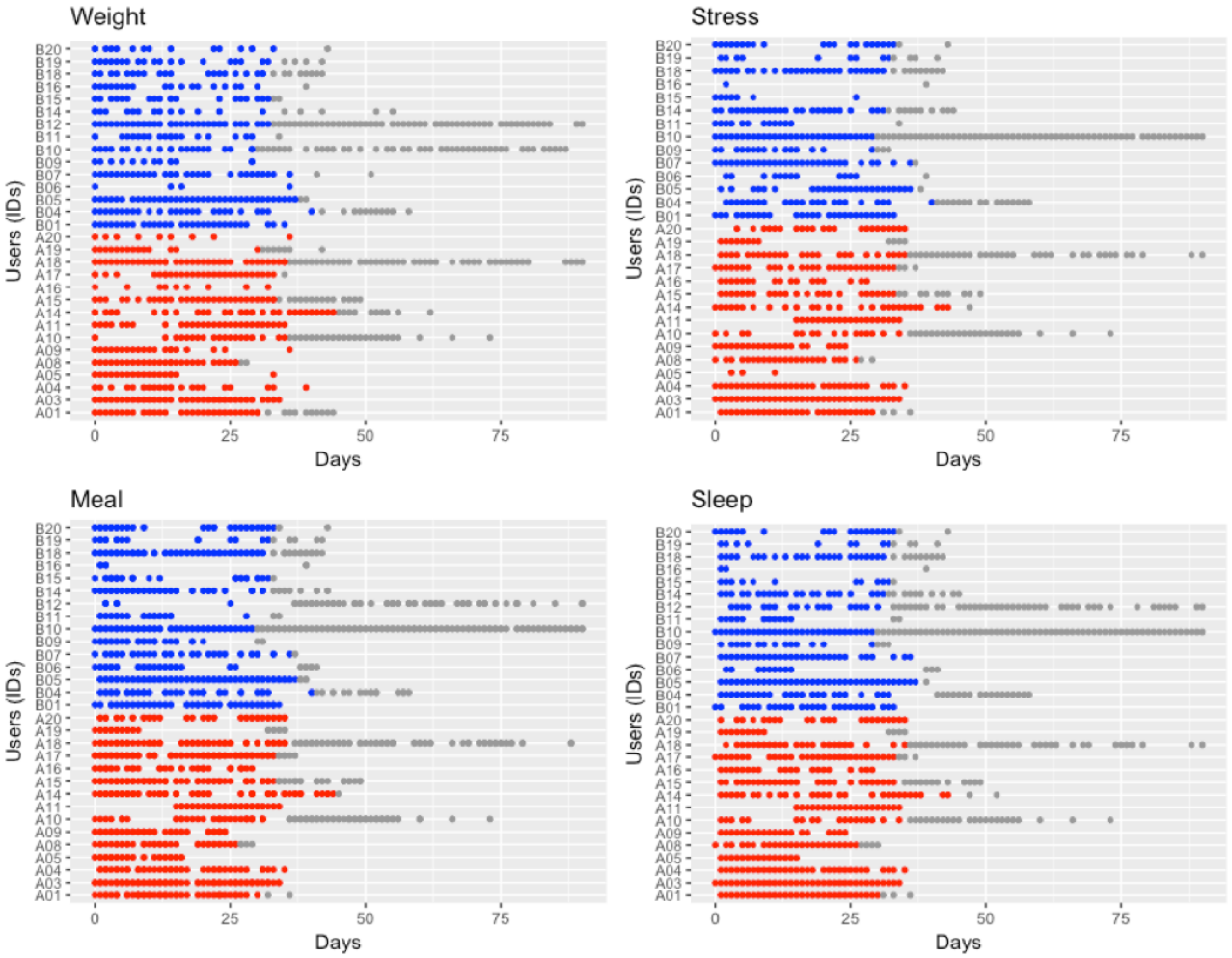


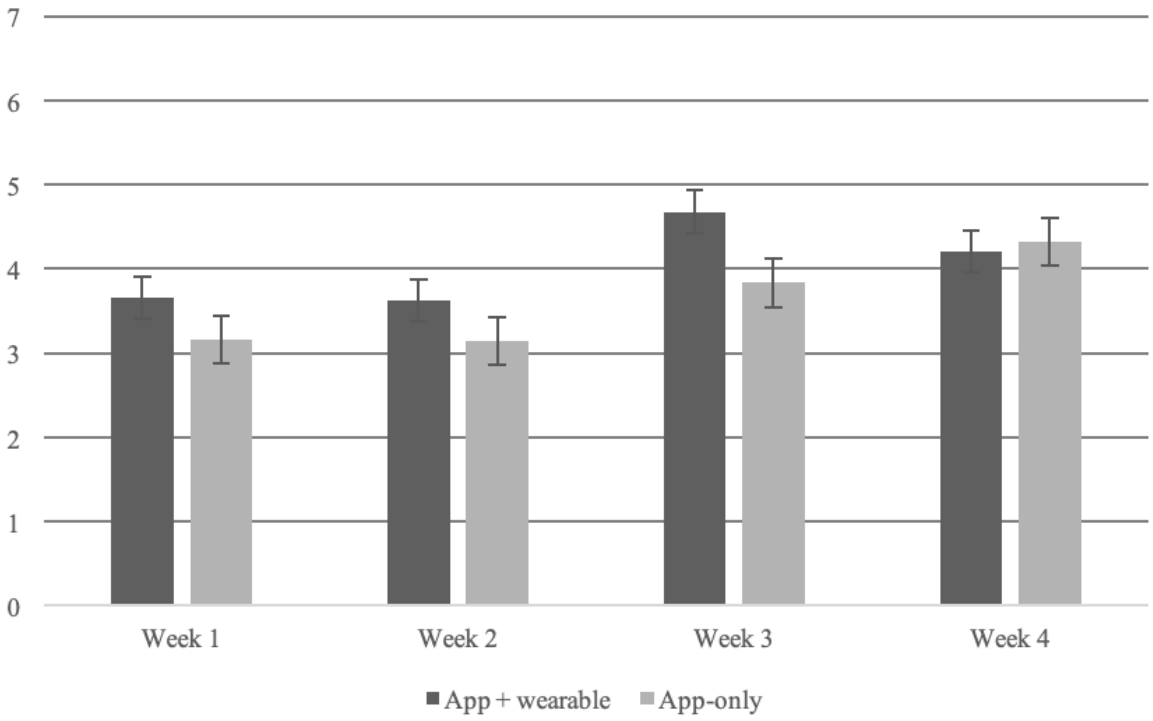
Figure 7. The number of recorded lifelog data for weight, stress, meal, and sleep logs (red: recorded days by app plus wearable users; blue: recorded days by app-only users; and gray: maintain period days without intervention).



Patients' Response: Usability Reflecting Patients' Response to Automated Clinician Comments on Daily Lifestyle

To study health app usability based on patients' responses to automated lifestyle comments, we analyzed usage logs for clinician comment feedback, which is depicted in Figure 3. Figure 8 demonstrates patients' response to clinician feedback comment per each group as a satisfaction score. When a patient responded satisfied, the response was scored as 1, and 0 otherwise. App plus wearable users tend to answer more than app-only users in detail. Furthermore, these 2 intervention group participants showed increased answers between both weeks 2 and 3. At this point, patients visited to see their clinician and anthropometric measurement was done in the hospital, these actions may influenced to motivate PHR usage.

Figure 8. Patient response to clinician feedback comment.



Discussion

Principal Findings

The mobile phone–based PHR system in this study has a flexible and unique platform architecture to collect patient-generated health data and deliver patient response to EHR for the following reasons. First, the user can use PHR service in this study only if he or she has a smartphone, regardless of whether he or she has a wearable device or not. It was available to record daily health logs with Android OS app, as one of the nation's largest market share, and to connect the data with the most popular Samsung Health platform. To the best of our knowledge, this is the first study on the system that links EHR-tethered PHR with conventional health platform such as Samsung Health. Second, the 3-arm clinical trial design, which targeted to observe the impact of wearable device and mobile app, revealed that the reduction of weight was because of the use of our PHR system, not the use of the wearable. Given that the majority of mobile phones already have a health platform function such as activity tracking, using the mobile phone health platform itself might be enough for PHR health care service without wearables. We suggest that a long-term prospective study including a larger number of participants be warranted to show whether patients require a higher precision in step count to make weight loss decisions. Third, with EHR-tethered PHR system, clinicians are available to show their patient's daily lifestyle and to deliver coaching feedback. Conventionally, clinicians just ask the patients to recall their lifelog or lifestyle to identify their lifestyles. Therefore, it seems to be quite inaccurate. The patient recall basis system is not enough for both patients and clinicians to pay much attention to the importance of lifestyle. However, the EHR-tethered PHR system may allow doctors and patients to check the patients' lifelog information together on EHR screen and to prescribe lifestyle for patients. Therefore, clinicians can review these data on the PHR module interface on the EHR and provide health-related lifestyle management feedback to the patients during the patient's visit to the clinic. From the experience using this system, patients are more likely to be interested in their lifelog data because the lifelog can be objectively summarized on the EHR and shared with their own clinicians.

Wearable biosensors are noninvasive devices used to acquire, transmit, process, store, and retrieve health-related data [26]. Biosensors have been integrated into various platforms, including watches, wristbands, skin patches, shoes, belts, textiles, and smartphones. Patients have the option to share data obtained by biosensors with their providers or social networks to support clinical treatment decisions and disease self-management [27]. In our previous study, we aimed to demonstrate the development of an EHR-tethered PHR app called *MyHealthKeeper*, which can retrieve data from a wearable device and deliver them to a hospital EHR system, and to study the effectiveness of PHR data–driven clinical intervention with clinical trial results [17]. We gathered this patient-generated lifestyle-related health information with a mobile app and activity tracking device and transferred it to a PHR data server to create a summary view based on the practical needs of the clinicians. Our first clinical study was a preliminary study aimed at observing the weight loss impact for obesity patients of

conventional care versus EHR-integrated PHR–based care after system development.

The related previous study examined the patient usage logs with the 5-year use of a mobile PHR system distributed by a tertiary hospital in South Korea [15]. On the basis of actual mobile PHR usage data, they investigated the usage pattern and characteristics of the users of patient-generated health data services. This was the first approach to analyze long-term usage data in mobile PHR research [15]. In addition to user-specific characteristics (eg, age, sex, and diagnosis), the study suggested that more research should be conducted on the modifiable factors that affect use duration, to facilitate activities that promote continued use.

In this study, we aimed to observe the effects of a PHR–based health care app with 3 arms: the mobile app and wearables, mobile app alone, and control in patients with sleep apnea. We also aimed at examining the impact of the use of PHR and wearables on health outcome improvement and sustained use of the health app that can be associated with patient engagement. As a primary outcome, we observed OSA patient's weight loss, BMI, and other sleep-related parameters to examine health outcome changes. Moreover, 2 intervention group patients showed weight loss and BMI change during the trial period. We also collected and analyzed lifelogs, which are dependent on health behaviors influenced by patient engagement (meal habits, weight control, diet, and exercise). Patient response was analyzed by mobile app records. To analyze usability and patient response, we assumed patient engagement by proxy through feedback logs from automated clinician comments on each lifestyle, which reflects patient activation.

This study is a unique approach rather than other studies based on the following characteristics. First, we collected 2 types of intervention group users' (app + wearable vs app-only) actual usage data to investigate the difference. Although we did not find a statistical difference in sleep-related health outcome between the groups, this study design is thoroughly concerned with lifestyle-related disease and can be further examined as a long-term observation. Second, our analysis found the continuous usage pattern, including without intervention period. This intervention-free observation reflects the actual user pattern in mobile PHR app without consciousness.

Limitations

This study could not provide a longitudinal observation of the EHR-tethered PHR system because of the practical constraints. Due to the short clinical trial period and the small number of study participants, it was difficult to determine a causal relationship, and the study did not provide information about the precise improvement in the health outcomes of PHR users. Nevertheless, we tried to observe the effect of PHR system for OSA patients, which is closely related to personal lifestyle-related sleep factor. Furthermore, with this integrated PHR system, we also expect longitudinal follow-up and continuous patient engagement in future studies. We hope to derive and apply many PHR features of an EHR-tethered PHR system for a variety of lifestyle-related disease management in further studies based on this study protocol.

Acknowledgments

This study was supported by the Creative Industrial Technology Development Program (10053249, Development of Personalized Healthcare System Exploiting User Life-Log and Open Government Data for Business Service Model Proof on Whole Life Cycle Care) funded by the Ministry of Trade, Industry and Energy (MOTIE, Korea). This study was approved by the SNUBH Institutional Review Board (B-1706/405-301) and was registered at the United States National Institutes of Health clinical trial registry (ClinicalTrails.gov registration number: NCT03200223).

Authors' Contributions

JWK and BR analyzed the data and drafted the manuscript as the first authors. SC contributed to data analysis, and EH and SYJ helped in conducting the clinical trial and contributed to data discussions. YK and JL contributed the user experience (UX)-based PHR interface design development. SY supervised the overall study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Supplementary image describes MyHealthKeeper mobile interfaces.

[PNG File, 590KB - mhealth_v7i1e12070_app1.png]

Multimedia Appendix 2

CONSORT-EHEALTH checklist (V 1.6.2).

[PDF File (Adobe PDF File), 98KB - mhealth_v7i1e12070_app2.pdf]

References

1. Camara MA, Castillo Y, Blanco-Almazan D, Estrada L, Jane R. Lifelog Health. Patient Generated Health Data URL: https://lifeloghealth.com/patient-generated-health-data/ [accessed 2018-12-06] [WebCite Cache ID 74STJoEzj]
2. Lifelog Health. PGHD Leads to Precision Medicine URL: https://lifeloghealth.com/pghd-leads-to-precision-medicine/ [accessed 2018-12-06] [WebCite Cache ID 74STuwAwV]
3. Caligtan CA, Dykes PC. Electronic health records and personal health records. Semin Oncol Nurs 2011 Aug;27(3):218-228. [doi: 10.1016/j.soncn.2011.04.007] [Medline: 21783013]
4. Roehrs A, da Costa CA, Righi RD, de Oliveira KS. Personal health records: a systematic literature review. J Med Internet Res 2017 Jan 6;19(1):e13 [FREE Full text] [doi: 10.2196/jmir.5876] [Medline: 28062391]
5. Heart T, Ben-Assuli O, Shabtai I. A review of PHR, EMR and EHR integration: a more personalized healthcare and public health policy. Health Policy Technol 2017 Mar;6(1):20-25. [doi: 10.1016/j.hlpt.2016.08.002]
6. Archer N, Fevrier-Thomas U, Lokker C, McKibbon KA, Straus SE. Personal health records: a scoping review. J Am Med Inform Assoc 2011;18(4):515-522 [FREE Full text] [doi: 10.1136/amiajnl-2011-000105] [Medline: 21672914]
7. Wikipedia. Lifelog URL: https://en.wikipedia.org/wiki/Lifelog [accessed 2018-12-06] [WebCite Cache ID 74SUKRQ9r]
8. Arvola M, Blomkvist J, Wahlman F. Lifelogging in user experience research: supporting recall and improving data richness. Des J 2017 Sep 6;20(Supp 1):S3954-S3965. [doi: 10.1080/14606925.2017.1352898]
9. Jones W. Keeping found things found: The study and practice of personal information management. Burlington, Massachusetts, United States: Morgan Kaufmann; 2010:448.
10. O'Hara K, Tuffield MM, Shadbolt N. Lifelogging: privacy and empowerment with memories for life. Identity Inf Soc 2008 Dec;1(1):155-172. [doi: 10.1007/s12394-009-0008-4]
11. Zomorodi B. HIT Consultant. 2018 Sep 4. How do Clinical Wearables Impact Patient Care and Quality of Life? URL: https://hitconsultant.net/2018/04/09/clinical-wearables-revolutionizing-care/ [accessed 2018-12-06] [WebCite Cache ID 74SWTKSts]
12. Liquid State. 2018 Mar 2. The State of Patient Engagement: 4 Trends to Watch in 2018 URL: https://liquid-state.com/state-patient-engagement-trends-2018/ [accessed 2018-12-06] [WebCite Cache ID 74SWZ6WwQ]
13. Kupchunas WR. Personal health record: new opportunity for patient education. Orthop Nurs 2007;26(3):185-191. [doi: 10.1097/01.NOR.0000276971.86937.c4] [Medline: 17538475]
14. Irizarry T, DeVito Dabbs A, Curran CR. Patient portals and patient engagement: a state of the science review. J Med Internet Res 2015 Jun 23;17(6):e148 [FREE Full text] [doi: 10.2196/jmir.4255] [Medline: 26104044]
15. Park YR, Lee Y, Kim JY, Kim J, Kim HR, Kim Y, et al. Managing patient-generated health data through mobile personal health records: analysis of usage data. J Med Internet Res Mhealth Uhealth 2018 Apr 9;6(4):e89 [FREE Full text] [doi: 10.2196/mhealth.9620] [Medline: 29631989]
16. Jung SY, Lee K, Hwang H, Yoo S, Baek HY, Kim J. Support for sustainable use of personal health records: understanding the needs of users as a first step towards patient-driven mobile health. J Med Internet Res Mhealth Uhealth 2017 Feb 23;5(2):e19 [FREE Full text] [doi: 10.2196/mhealth.6021] [Medline: 28232300]
17. Ryu B, Kim N, Heo E, Yoo S, Lee K, Hwang H, et al. Impact of an electronic health record-integrated personal health record on patient participation in health care: development and randomized controlled trial of MyHealthKeeper. J Med Internet Res 2017 Dec 7;19(12):e401 [FREE Full text] [doi: 10.2196/jmir.8867] [Medline: 29217503]
18. Yoo S, Lee KH, Lee HJ, Ha K, Lim C, Chin HJ, et al. Seoul National University Bundang Hospital's electronic system for total care. Healthc Inform Res 2012 Jun;18(2):145-152 [FREE Full text] [doi: 10.4258/hir.2012.18.2.145] [Medline: 22844650]
19. Wikipedia. Obstructive sleep apnea URL: https://en.wikipedia.org/wiki/Obstructive_sleep_apnea [accessed 2018-12-06] [WebCite Cache ID 74SWnsbDz]
20. Camara MA, Castillo Y, Blanco-Almazan D, Estrada L, Jane R. mHealth tools for monitoring Obstructive Sleep Apnea patients at home: Proof-of-concept. 2017 Jul 11 Presented at: 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC); July 11-15, 2017; Seogwipo, South Korea.
21. Villanueva JA, Suarez MC, Garmendia O, Lugo V, Ruiz C, Montserrat JM. The role of telemedicine and mobile health in the monitoring of sleep-breathing disorders: improving patient outcomes. Smart Homecare Technol TeleHealth 2017 Feb 3;4:1-11. [doi: 10.2147/SHTT.S108048]
22. Ballinger B, Hsieh J, Singh A, Sohoni N, Wang J, Tison GH, et al. DeepHeart: semi-supervised sequence learning for cardiovascular risk prediction. 2018 Presented at: Thirty-Second AAAI Conference on Artificial Intelligence; February 2-7, 2018; Louisiana, USA URL: https://dx.doi.org/10.1155/2016/8094275 [doi: 10.1155/2016/8094275]
23. Kim Y, Ji S, Lee H, Kim JW, Yoo S, Lee J. "My Doctor is Keeping an Eye on Me!": Exploring the Clinical Applicability of a Mobile Food Logger. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. 2016 May 07 Presented at: CHI '16; May 07-12, 2016; San Jose, California, USA p. 5620-5631. [doi: 10.1145/2858036.2858145]
24. Wikipedia. Samsung Health URL: https://en.wikipedia.org/wiki/Samsung_Health [accessed 2018-12-06] [WebCite Cache ID 74SXDehAe]
25. Itamar Medical Ltd. WatchPAT - Simple Sleep Health Management URL: https://www.itamar-medical.com/watchpat-main/ [accessed 2018-12-06] [WebCite Cache ID 74SXLCMV2]

26. De Sanctis A, Russo S, Craciun MF, Alexeev A, Barnes MD, Nagareddy VK, et al. New routes to the functionalization patterning and manufacture of graphene-based materials for biomedical applications. Interface Focus 2018 Jun 6;8(3) [FREE Full text] [doi: 10.1098/rsfs.2017.0057]
27. Noah B, Keller MS, Mosadeghi S, Stein L, Johl S, Delshad S, et al. Impact of remote patient monitoring on clinical outcomes: an updated meta-analysis of randomized controlled trials. NPJ Digit Med 2018 Jan 15;1. [doi: 10.1038/s41746-017-0002-4]

Abbreviations

BMI: body mass index
EHR: electronic health record
mHealth: mobile health
OSA: obstructive sleep apnea
PHR: personal health record
SNUBH: Seoul National University Bundang Hospital

Edited by G Eysenbach; submitted 31.08.18; peer-reviewed by JH Lee, SY Shin, H Feldman; comments to author 21.09.18; revised version received 14.11.18; accepted 22.11.18; published 04.01.19.

Please cite as:
Kim JW, Ryu B, Cho S, Heo E, Kim Y, Lee J, Jung SY, Yoo S
Impact of Personal Health Records and Wearables on Health Outcomes and Patient Response: Three-Arm Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(1):e12070
URL: https://mhealth.jmir.org/2019/1/e12070/
doi:10.2196/12070
PMID:30609978

Short Paper

How Well iPhones Measure Steps in Free-Living Conditions: Cross-Sectional Validation Study

Shiho Amagasa[1], MMS; Masamitsu Kamada[2,3], PhD; Hiroyuki Sasai[4], PhD; Noritoshi Fukushima[1], PhD, MD; Hiroyuki Kikuchi[1], PhD; I-Min Lee[5,6], MBBS, MPH, ScD; Shigeru Inoue[1], PhD, MD

[1]Department of Preventive Medicine and Public Health, Tokyo Medical University, Shinjuku-ku, Japan

[2]Department of Social and Behavioral Sciences, Harvard T.H. Chan School of Public Health, Boston, MA, United States

[3]Department of Health Sociology and Health Education, School of Public Health, Graduate School of Medicine, The University of Tokyo, Bunkyo-ku, Japan

[4]Department of Life Sciences, Graduate School of Arts and Sciences, The University of Tokyo, Meguro-ku, Japan

[5]Division of Preventive Medicine, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, United States

[6]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, United States

Corresponding Author:
Shigeru Inoue, PhD, MD
Department of Preventive Medicine and Public Health
Tokyo Medical University
6-1-1 Shinjuku
Shinjuku-ku, 160-8402
Japan
Phone: 81 3 3351 6141 ext 237
Email: inoue@tokyo-med.ac.jp

Abstract

Background: Smartphones have great potential for monitoring physical activity. Although a previous laboratory-based study reported that smartphone apps were accurate for tracking step counts, little evidence on their accuracy in free-living conditions currently exists.

Objective: We aimed to investigate the accuracy of step counts measured using iPhone in the real world.

Methods: We recruited a convenience sample of 54 adults (mean age 31 [SD 10] years) who owned an iPhone and analyzed data collected in 2016 and 2017. Step count was simultaneously measured using a validated pedometer (Kenz Lifecorder) and the iPhone. Participants were asked to carry and use their own iPhones as they typically would while wearing a pedometer on the waist for 7 consecutive days during waking hours. To assess the agreement between the two measurements, we calculated Spearman correlation coefficients and prepared a Bland-Altman plot.

Results: The mean step count measured using the iPhone was 9253 (3787) steps per day, significantly lower by 12% (1277/10,530) than that measured using the pedometer, 10,530 (3490) steps per day (P<.001). The Spearman correlation coefficient between devices was 0.78 (P<.001). The largest underestimation of steps by the iPhone was observed among those who reported to have seldom carried their iPhones (seldom carry: mean −3036, SD 2990, steps/day; sometimes carry: mean −1424, SD 2619, steps/day; and almost always carry: mean −929, SD 1443, steps/day; P for linear trend=.08).

Conclusions: Smartphones may be of practical use to individuals, clinicians, and researchers for monitoring physical activity. However, their data on step counts should be interpreted cautiously because of the possibility of underestimation due to noncarrying time.

(JMIR Mhealth Uhealth 2019;7(1):e10418) doi:10.2196/10418

KEYWORDS

mobile phone; step count; physical activity; pedometer; epidemiology; population; validation; free-living conditions

Introduction

Monitoring daily physical activity using smartphones may have a great potential for public health applications [1]. Althoff et al [1] described how step-determined physical activity is distributed using a large-scale database consisting of 68 million days from 717,527 people in 111 countries, automatically measured using iPhones. However, little evidence exists on their measurement accuracy [2-5]. It is unclear how accurately step counts can be tracked via built-in algorithms of smartphones in free-living conditions because the smartphones may not be "tethered" to an individual at all times. For example, Hekler et al [3] examined the validity of physical activity measurement by a custom app of Android phones against an accelerometer in free-living conditions and showed that smartphones appear to be acceptable for estimating physical activity time. However, participants were instructed to carry their smartphones and wear the accelerometers at the same time during waking hours. In another study, Duncan et al [5] assessed various iPhone models in free-living conditions, but they did not fully account for the frequency and location of iPhone carrying. In the real world, individuals vary considerably regarding how much they carry their smartphones with them. Therefore, we aimed to assess the accuracy of step counts measured using smartphones in free-living conditions, under typical conditions where the smartphones may not always be carried by the individuals, using the default installment of a step counter app on the iPhone, against a pedometer.

Methods

Study Sample

We recruited a convenience sample of 54 healthy adults (mean age 31, SD 10, years; 48%, 26/54, men) who owned an iPhone 5S, 6, 6S, 6plus, SE, or 7 (Apple Inc, California, United States) through direct outreach and flyers at a university in 2016 and 2017. Each participant received a 3000 Japanese Yen (US $25) gift card for participating in the study. Ethical approval was granted by Tokyo Medical University Ethics Committee.

Measures

Daily step count was measured using both a validated pedometer, Kenz Lifecorder Ex (Suzuken Co, Ltd, Nagoya, Japan) [6,7], and an iPhone. Schneider et al, in their validation study using 13 pedometer models, have reported that Kenz Lifecorder Ex is suitable for most research purposes (compared to the criterion pedometer, Yamax SW-200), with an observed mean difference in the step count of −703 (SD 1537) steps per day [7]. We used the Health app preinstalled on the iPhone to measure steps using iPhone. Participants were asked to carry their own iPhones as usual and wear a pedometer on their waist for 7 consecutive days during waking hours. A self-reported questionnaire evaluated sociodemographic and health-related factors, as well as how (in their pockets or bags) and how often (almost always, sometimes, seldom) participants carried their iPhones. A record was deemed valid if the pedometer was worn for ≥10 hours a day [8,9] for at least 3 days [10].

Statistical Analysis

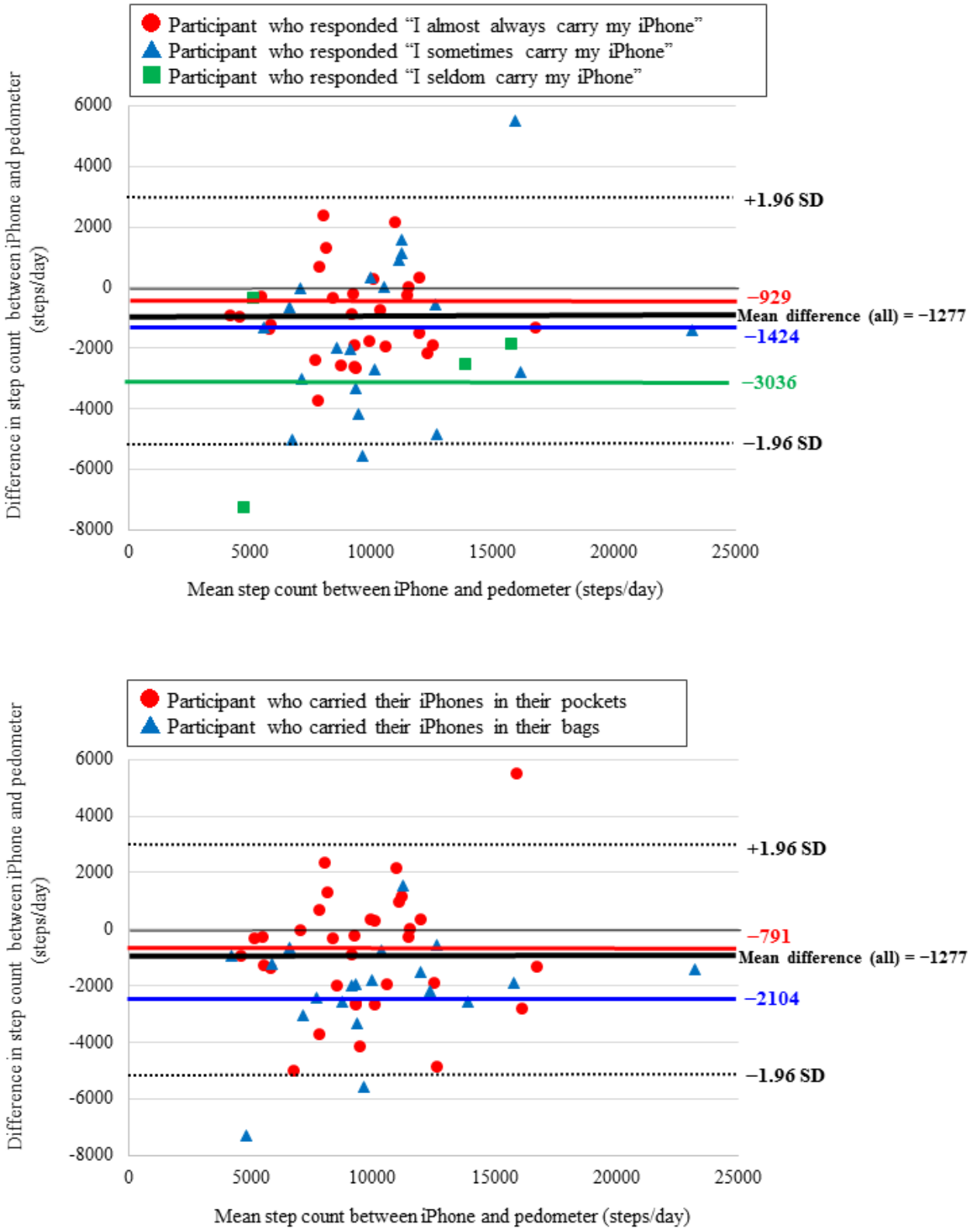
The mean and SD of the step count for each device was obtained. We calculated Spearman correlation coefficients, intraclass correlation coefficient (ICC), and weighted kappa using a classification matrix. The difference in the step count between device measurements was calculated by subtracting the step count of the pedometer from that of the iPhone. A paired *t* test was performed to determine whether the differences between step counts were statistically significant. We performed a 2-sample *t* test and linear regression analysis to detect differences according to iPhone carrying locations and linear trend for frequency, respectively. An ordinal scale was used when the trend tests were run. A Bland-Altman plot was created to assess the agreement between the two measurements [11]. In sensitivity analysis, we included data only from participants with ≥13 hours of pedometer wear time [12,13]. Analyses were conducted in 2017 using IBM SPSS Statistics version 21 (IBM Corp).

Results

The mean step count measured using the iPhones was 9253 (SD 3787) steps per day; this was significantly lower than that measured using the pedometer, 10,530 (SD 3490) steps per day (mean relative difference 12% [SD 21%]; $P<.001$). Spearman correlation coefficient between the devices was 0.78 ($P<.001$), and ICC was 0.88 (95% CI 0.79-0.93; $P<.001$). When categorized into quartiles based on step count, the pedometer and iPhone classified participants into the same quartile 54% (29/54) of the time, resulting in a weighted kappa coefficient of 0.69. The Bland-Altman plot revealed a mean difference in step count of −1277 (SD 2122) steps per day, with no significant proportional bias (Figure 1).

In the first graph in Figure 1, the thick black line shows mean difference among overall sample; dotted black lines show mean (SD 1.96); red line shows mean difference among those who almost always carry their iPhone; blue line shows mean difference among those who sometimes carry their iPhone; and green line shows mean difference among those who seldom carry their iPhone. A negative difference value means the step count measured using the iPhone was lower than that measured using the pedometer (ie, underestimated). There was no significant proportional bias between the two methods ($r=0.06$). In the second graph, the thick black line shows mean difference among overall sample; dotted black lines show mean (SD 1.96); red line shows mean difference among those who carry their iPhone in their pockets; and blue line shows mean difference among those who carry their iPhone in their bags. A negative difference value means the step count measure using the iPhone was lower than that measured using the pedometer (ie, underestimated).

Figure 1. Comparison of daily steps measured using iPhone and pedometer in free-living conditions (N=54).



We then assessed whether step counts from smartphones may be sensitive to how frequently participants carried their iPhones with them (Figure 1). The largest underestimation of steps using the iPhones against the pedometer was observed among those who reported to have seldom carried their iPhones, with borderline statistical significance (seldom carry: −3036, SD 2990, steps/day; sometimes carry: −1424, SD 2619, steps/day; and almost always carry: −929, SD 1443, steps/day; *P* for linear trend=.08). Sensitivity analyses restricting the analyses to participants with ≥13 hours of pedometer wear time also yielded

similar findings that were statistically significant (seldom carry: −3036, SD 2990, steps/day; sometimes carry: −1721, SD 2095, steps/day; and almost always carry: −1032, SD 1401, steps/day; *P* for linear trend=0.03). Additionally, step counts were more underestimated among participants who typically carried their iPhones in their bags (−2104, SD 1844, steps/day) than among those carrying the smartphones in their pockets (−791, SD 2149, steps/day; *P*=.02; Figure 1), although the tests for interaction of iPhone carrying location and frequency with the differences in step counts were not significant, possibly due to small sample sizes in the subgroups. There was no significant interaction of iPhone carrying location and frequency with differences in step count between the pedometer and iPhone.

When stratified by gender, difference in the step count between device measurements was larger among women than among men (−1847, SD 1880, steps/day vs −664, SD 2231, steps/day; *P*=.04; Table 1). Most (18/28, 64%) of the women carried their iPhones in their bags rather than in their pockets, whereas almost all (24/26, 92%) of the men carried them in their pockets (Table 2).

Table 1. Gender differences in daily steps measured using iPhone.

Characteristics	Men	Women	*P* value
Age in years, mean (SD)	30 (10)	32 (10)	.39[a]
Steps measured using, mean (SD)			
Pedometer	9864 (3094)	11,149 (3770)	.18[a]
iPhone	9200 (3332)	9302 (4227)	.92[a]
Differences between the two measurements (iPhone−pedometer steps), mean (SD)	−664 (2231)	−1847 (1880)	*.04* [a,b]
Usage of iPhone model, n (%)			.87[c]
5S	9 (35)	11 (39)	
6	8 (31)	5 (18)	
6S	6 (23)	8 (29)	
SE	1 (4)	2 (7)	
7	2 (8)	2 (7)	

[a]*P* value was calculated using *t* test.

[b]Italicized values indicate statistically significant differences.

[c]*P* value was calculated using Fisher Exact test.

Table 2. Gender differences in frequency and location of carrying an iPhone.

Characteristics	Men		Women		*P* value
	n (%)	Mean (SD)	n (%)	Mean (SD)	
Frequency of carrying an iPhone					.30[a]
Almost always	17 (65)	−439 (1647)	12 (43)	−1623 (679)	*.03* [b,c]
Sometimes	8 (31)	−1181 (3336)	13 (46)	−1573 (2208)	.75[b]
Seldom	1 (4)	N/A[d]	3 (11)	−3928 (2939)	N/A
Location of carrying an iPhone					*<.001* [a]
In the pocket	24 (92)	−573 (2289)	10 (36)	−1314 (1765)	.37[b]
In a bag	2 (8)	−1757 (1151)	18 (64)	−2143 (1925)	.79[b]

[a]*P* value was calculated using Fisher Exact test.

[b]*P* value was calculated using *t* test.

[c]Italicized values indicate statistically significant differences.

[d]N/A: not applicable.

Discussion

Principal Findings

We found that step counts measured using a pedometer or iPhone correlated moderately well under free-living conditions. In contrast to a previous laboratory-based study where only a small difference in the mean step count between iPhone apps and direct observation was found [4], we found that iPhone underestimated average step count by 12% (1277/10,530) compared to a pedometer. These findings were similar to that of previous study where step counts measured using iPhone were underestimated by 1340 steps per day in free-living conditions [5]. Furthermore, the level of underestimation depended on how often participants typically carried the phone with them, as well as different carrying locations of the phone. To improve the accuracy of step counts measured using iPhones, carrying a phone as frequently as possible appears important.

With the growing popularity of smartphones [14], step counting apps make objective tracking of physical activity available to a tremendous number of people [1]. Smartphones may be of practical use to researchers for monitoring step-determined physical activity and for health promotion. Furthermore, clinicians can obtain a patient's daily physical activity data immediately in clinical practice. However, investigators and clinicians also should be aware of the potential for underestimation of step counts using smartphones especially when the interest is in its between-individual variation, including country-level comparisons. For example, a previous study of step-determined physical activity for free-living individuals measured using an iPhone app identified inactive subpopulations such as women [1]. The finding that women took fewer steps than men regardless of age groups may partly be attributable to the phone carrying habits and location of phone carrying among women. In particular, women's clothing, such as dresses, rarely have pockets large enough to fit a smartphone, and in our study, most women carried their iPhones in their bags rather than in their pockets.

The mean bias of step counts measured using iPhone slightly exceeded the ±10% "acceptable" difference range used in previous free-living studies [7,15]. In addition, limits of agreement ranged from −5436 to 2882 steps per day for all participants (−3757 to 1899 steps/day among those who almost always carried an iPhone). However, this difference is comparable to that observed for other pedometers that are considered acceptable for research purpose [7,15].

Limitations

We investigated only healthy, young Japanese adults who were more active than the general population [16] and owned an iPhone; it is unclear whether our results are applicable to other individuals and other smartphone apps. In this study, there might have been an underestimation of differences in step counts between Kenz Lifecorder Ex and the iPhone. Although previous studies have found Kenz Lifecorder Ex to be acceptable compared to gold standard pedometers, the former may slightly underestimate step counts in free-living conditions [7]. Thus, the inherent technical measurement error of the pedometer used in this study is a limitation.

Conclusions

We found that step count measured using a pedometer and iPhone correlated moderately well in free-living conditions. Smartphones can be of practical use to individuals, clinicians, and researchers for monitoring physical activity and for health promotion. However, their data on step counts should be interpreted cautiously because of the possibility of underestimation due to noncarrying time and carrying locations, as well as gender differences.

Acknowledgments

This study was supported by a Grant-in-Aid for Scientific Research (16H03249) from the Japan Ministry of Education, Culture, Sports, Science and Technology, and the Uehara Memorial Foundation. No financial disclosures were reported by the authors of this paper.

Authors' Contributions

SA collected the data and wrote the first draft of this manuscript. MK, HS, NF, HK, IML, and SI participated in the interpretation of the data, revised the draft versions of the manuscript, and provided critical comments during the process. All authors contributed to the writing of the manuscript and approved the final version.

Conflicts of Interest

None declared.

References

1. Althoff T, Sosič R, Hicks JL, King AC, Delp SL, Leskovec J. Large-scale physical activity data reveal worldwide activity inequality. Nature 2017 Dec 20;547(7663):336-339 [FREE Full text] [doi: 10.1038/nature23018] [Medline: 28693034]
2. Orr K, Howe HS, Omran J, Smith KA, Palmateer TM, Ma AE, et al. Validity of smartphone pedometer applications. BMC Res Notes 2015 Nov 30;8:733 [FREE Full text] [doi: 10.1186/s13104-015-1705-8] [Medline: 26621351]

3. Hekler EB, Buman MP, Grieco L, Rosenberger M, Winter SJ, Haskell W, et al. Validation of Physical Activity Tracking via Android Smartphones Compared to ActiGraph Accelerometer: Laboratory-Based and Free-Living Validation Studies. JMIR Mhealth Uhealth 2015 Apr 15;3(2):e36 [FREE Full text] [doi: 10.2196/mhealth.3505] [Medline: 25881662]
4. Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. JAMA 2015 Feb 10;313(6):625-626. [doi: 10.1001/jama.2014.17841] [Medline: 25668268]
5. Duncan MJ, Wunderlich K, Zhao Y, Faulkner G. Walk this way: validity evidence of iphone health application step count in laboratory and free-living conditions. J Sports Sci 2018 Aug;36(15):1695-1704. [doi: 10.1080/02640414.2017.1409855] [Medline: 29179653]
6. Crouter SE, Schneider PL, Karabulut M, Bassett DR. Validity of 10 electronic pedometers for measuring steps, distance, and energy cost. Med Sci Sports Exerc 2003 Aug;35(8):1455-1460. [doi: 10.1249/01.MSS.0000078932.61440.A2] [Medline: 12900704]
7. Schneider PL, Crouter SE, Bassett DR. Pedometer measures of free-living physical activity: comparison of 13 models. Med Sci Sports Exerc 2004 Feb;36(2):331-335. [doi: 10.1249/01.MSS.0000113486.60548.E9] [Medline: 14767259]
8. Tudor-Locke C, Camhi SM, Troiano RP. A catalog of rules, variables, and definitions applied to accelerometer data in the National Health and Nutrition Examination Survey, 2003-2006. Prev Chronic Dis 2012;9:E113 [FREE Full text] [Medline: 22698174]
9. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]
10. Matthews CE, Ainsworth BE, Thompson RW, Bassett DR. Sources of variance in daily physical activity levels as measured by an accelerometer. Med Sci Sports Exerc 2002 Aug;34(8):1376-1381. [Medline: 12165695]
11. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986 Feb 08;1(8476):307-310. [Medline: 2868172]
12. Herrmann SD, Barreira TV, Kang M, Ainsworth BE. How many hours are enough? Accelerometer wear time may provide bias in daily activity estimates. J Phys Act Health 2013 Jul;10(5):742-749. [Medline: 23036822]
13. Herrmann SD, Barreira TV, Kang M, Ainsworth BE. Impact of accelerometer wear time on physical activity data: a NHANES semisimulation data approach. Br J Sports Med 2014 Feb;48(3):278-282. [doi: 10.1136/bjsports-2012-091410] [Medline: 22936409]
14. Anthes E. Mental health: There's an app for that. Nature 2016 Apr 07;532(7597):20-23. [doi: 10.1038/532020a] [Medline: 27078548]
15. Barriera TV, Tudor-Locke C, Champagne CM, Broyles ST, Johnson WD, Katzmarzyk PT. Comparison of GT3X accelerometer and YAMAX pedometer steps/day in a free-living sample of overweight and obese adults. J Phys Act Health 2013 Feb;10(2):263-270. [Medline: 22821951]
16. Ministry of Health Labour and Welfare. 2016. National Health and Nutrition Survey URL: https://www.mhlw.go.jp/file/04-Houdouhappyou-10904750-Kenkoukyoku-Gantaisakukenkouzoushinka/kekkagaiyou_7.pdf [accessed 2018-11-20] [WebCite Cache ID 744WgYxMg]

Abbreviations

ICC: intraclass correlation coefficient

Edited by G Eysenbach; submitted 19.03.18; peer-reviewed by W Tian, J Goris, M Duncan; comments to author 17.08.18; revised version received 09.10.18; accepted 14.10.18; published 09.01.19.

Please cite as:
Amagasa S, Kamada M, Sasai H, Fukushima N, Kikuchi H, Lee IM, Inoue S
How Well iPhones Measure Steps in Free-Living Conditions: Cross-Sectional Validation Study
JMIR Mhealth Uhealth 2019;7(1):e10418
URL: http://mhealth.jmir.org/2019/1/e10418/
doi:10.2196/10418
PMID:30626569

Original Paper

Valuing Mobile Health: An Open-Ended Contingent Valuation Survey of a National Digital Health Program

Camilla Somers[1], PhD; Eleanor Grieve[1], MPH; Marilyn Lennon[2], PhD; Matt-Mouley Bouamrane[2], PhD; Frances S Mair[3], PhD; Emma McIntosh[1], PhD

[1]Health Economics and Health Technology Assessment, University of Glasgow, Glasgow, United Kingdom

[2]Digital Health and Wellness Group, Department of Computer and Information Sciences, University of Strathclyde, Glasgow, United Kingdom

[3]General Practice and Primary Care, University of Glasgow, Glasgow, United Kingdom

Corresponding Author:
Frances S Mair, PhD
General Practice and Primary Care
University of Glasgow
1 Horselethill Road
Glasgow, G12 9LX
United Kingdom
Phone: 44 1413308317
Fax: 44 1413305018
Email: Frances.Mair@glasgow.ac.uk

Abstract

Background: Changing population demographics and technology developments have resulted in growing interest in the potential of consumer-facing digital health. In the United Kingdom, a £37 million (US $49 million) national digital health program *delivering assisted living lifestyles at scale* (dallas) aimed to deploy such technologies at scale. However, little is known about how consumers *value* such digital health opportunities.

Objective: This study explored consumers' perspectives on the potential value of digital health technologies, particularly mobile health (mHealth), to promote well-being by examining their *willingness-to-pay* (WTP) for such health solutions.

Methods: A contingent valuation study involving a UK-wide survey that asked participants to report open-ended absolute and marginal WTP or willingness-to-accept for the gain or loss of a hypothetical mHealth app, *Healthy Connections*.

Results: A UK-representative cohort (n=1697) and a dallas-like (representative of dallas intervention communities) cohort (n=305) were surveyed. Positive absolute and marginal WTP valuations of the app were identified across both cohorts (absolute WTP: UK-representative cohort £196 or US $258 and dallas-like cohort £162 or US $214; marginal WTP: UK-representative cohort £160 or US $211 and dallas-like cohort £151 or US $199). Among both cohorts, there was a high prevalence of zeros for both the absolute WTP (UK-representative cohort: 467/1697, 27.52% and dallas-like cohort: 95/305, 31.15%) and marginal WTP (UK-representative cohort: 487/1697, 28.70% and dallas-like cohort: 99/305, 32.5%). In both cohorts, better general health, previous amount spent on health apps (UK-representative cohort 0.64, 95% CI 0.27 to 1.01; dallas-like cohort: 1.27, 95% CI 0.32 to 2.23), and age had a significant ($P>.00$) association with WTP (UK-representative cohort: −0.1, 95% CI −0.02 to −0.01; dallas-like cohort: −0.02, 95% CI −0.03 to −0.01), with younger participants willing to pay more for the app. In the UK-representative cohort, as expected, higher WTP was positively associated with income up to £30,000 or US $39,642 (0.21, 95% CI 0.14 to 0.4) and increased spending on existing phone and internet services (0.52, 95% CI 0.30 to 0.74). The amount spent on existing health apps was shown to be a positive indicator of WTP across cohorts, although the effect was marginal (UK-representative cohort 0.01, 95% CI 0.01 to 0.01; dallas-like cohort 0.01, 95% CI 0.01 to 0.02).

Conclusions: This study demonstrates that consumers value mHealth solutions that promote well-being, social connectivity, and health care control, but it is not universally embraced. For mHealth to achieve its potential, apps need to be tailored to user accessibility and health needs, and more understanding of what hinders frequent users of digital technologies and those with long-term conditions is required. This novel application of WTP in a digital health context demonstrates an economic argument for investing in upskilling the population to promote access and expedite uptake and utilization of such digital health and well-being apps.

(JMIR Mhealth Uhealth 2019;7(1):e3) doi:10.2196/mhealth.9990

KEYWORDS

mHealth; public health; delivery of health care; public health systems research

Introduction

Background

Globally, more than 50% of the world population owns a mobile device, rising to nearly 90% in the developed world [1]. Digital technology use is becoming integrated into our daily lives and clearly has potential to promote physical and psychological well-being [2]. Digital health is seen as having the potential to transform health care [3] at a time when changing population demographics and rising levels of chronic illness and multimorbidity (the presence of 2 or more long-term conditions) make change imperative [4]. However, this opportunity presents a number of challenges as developers must tackle a current underuse of readily available digital health innovations and there is a need for more evidence to aid understanding of *what* is of value to users [5]. In recent years, the United Kingdom has prioritized developing a digital health strategy to be implemented nationally [3,6]. A key driving force behind digital health is the need to move to more cost-effective health care delivery models, with the National Institute for Health and Care Excellence (NICE) announcing plans to develop a new digital health apps evaluation system to respond to the recent growth in digital health. Digital health, particularly mobile health (mHealth), refers to *raising awareness of health information via mobile or wireless devices* and has the potential to provide an alternative, less resource-intensive delivery of health to a changing population [7,8]. In their most recent communications report, Ofcom declared the United Kingdom to be a *smartphone society*, with more than 60% of the population owning a smartphone [9]. The number of mHealth apps continues to grow at an ever-increasing pace, with as many as 325,000 health apps available in 2017 and 78,000 new health apps added to major app stores in the last year [10]. However, existing studies have demonstrated the complex and highly variable nature of implementing successful well-being digital technologies and tools [11-15].

mHealth technology should be flexible and accessible for both users and practitioners [7]. Using digital devices also enables users to create platforms for support and self-management, providing opportunities for wider aspects of one's well-being to be improved. Looking beyond health to include nonhealth aspects of quality of life, such as one's sense of empowerment and ability to participate in community activities, has increasingly become a focal point of health service interventions [16,17]. Furthermore, NICE has emphasized the need for successful community engagement initiatives by health services to produce positive health gains and tackle health inequalities [18,19]. However, the personalization of digital health technologies and their focus on seeking to improve multiple broader aspects of health pose a number of challenges for economic evaluations in determining the value of their delivery and outcomes [20]. Issues include the need for wider measurement of costs and benefits as well as the handling of development costs [21].

In this paper, we present a contingent valuation (CV) willingness-to-pay (WTP) study for a hypothetical mHealth app that would deliver against 6 well-being outcomes alongside any other health services or treatments. The study examines the value the public places on improving broader well-being outcomes with mHealth. This was part of a wider program *delivering assisted living lifestyles at scale* (dallas). The dallas program launched in 2012 and funded by Innovate UK established 4 multiagency communities across the United Kingdom, who were to show "how independent living technologies, services and systems can be used to promote wellbeing, and provide integrated top quality health and care, enabling people to live independently" [22]. These communities worked in collaboration with a number of stakeholders, including health care services, industry, third sector voluntary organizations, and academic and government bodies, to explore how digital health can be delivered successfully for preventative care and to promote well-being across the United Kingdom [14]. Further details of the communities and their associated partnerships have previously been reported [14,23].

Existing Studies

Research on WTP for specific treatment or disease management using digital health technologies has been conducted but is still in its infancy. In Ireland, they have examined women's valuation of an integrated app and stand-alone app for postoperative monitoring post cesarean section [24]. WTP levels were considerably smaller than anticipated, and this was attributed to the participant's experiences of paying small amounts for mobile phone apps previously [24]. In Bangladesh, a country where health care is provided on a fee for service basis, WTP for mobile phone short message service text messaging to promote diabetes self-management was explored [25]. The researchers found that participants were generally willing to pay for the service and that those males with higher household income and higher levels of education reported higher WTP levels. However, research on WTP for mHealth apps looking at improving broader lifestyle well-being outcomes is currently an understudied area. This study seeks to build on previous studies such as that by Callan and O'Shea [18], which focused on determining societal values for different telecare solutions for older people. Their study demonstrates that there is a preference for developing supportive technologies, which seek to keep older people in their community, and that above telecare for physical or cognitive care needs, strongest preferences were for telecare that sought to improve user's social connections. This potential for mHealth as an individual's own *tailored* health service is further emphasized by researchers such as Klasnja and Pratt [24] who argue that if delivered in a sensitive and appropriate manner, mHealth could be effective in managing both specific diseases and general health while also enabling communities to support one another. This would allow virtual

networks or communities of users with similar goals or location to connect to one another [26]. Reviews of mHealth interventions have demonstrated that few evaluations have captured data that allow for consideration of economic outcomes and overall effectiveness and cost-effectiveness of interventions [27-31]. The lack of standardization in the delivery of mHealth programs means that currently full societal outcomes are not being captured and decision makers cannot make fully informed decisions when comparing the cost-effectiveness of different programs [32]. The potential to understand the value the public places on aspects of broader well-being, lifestyle, or other measures of individual autonomy is important and a much-needed advance in evaluations of these types of person-centered digital health and wellness products and services. Indeed, guidelines for producing high-quality evidence of digital health programs have emphasized the need for appropriate analytical methodology that can capture these noncost-related outcomes [30,33].

Broader Lifestyle Outcomes (The 6Cs)

As the name of dallas explicitly highlights, the program from the outset had an emphasis on making a positive impact on citizens' *lifestyles*, moving away from a purely medical model. As the dallas communities' implementation plans included specific targets on recruitment number, the program funder, Innovate UK, took steps to ensure that impact on lifestyles could also be objectively measured as part of the broader dallas initiatives' deployment. For this purpose, it proposed the use of 6 key concepts that could demonstrate commonality of purpose for the broader program regardless of the details of each of the many communities' interventions. These key concepts were called the *6Cs*, namely, connectedness, control, choice, collaboration, community, and contribution. To achieve some degree of consensus on how these key concepts could be applied to each community implementation plan, a workshop was organized at the outset of the program by Innovate UK in June 2012 (Birmingham). Key representatives from the program funder, the 4 dallas communities, and the program evaluation team (ie, University of Glasgow) attended this workshop. During the workshop, a series of focus groups was undertaken to develop and iteratively refine a detailed mapping of how the 6Cs applied to each community's specific implementation plan [13,14,34]. Table 1 [22] also demonstrates how mHealth app features could foster and improve a user's sense of each of the 6Cs.

Table 1. Innovate UK 6Cs

Concept	Definition	Example of concept as a mobile health app feature
Connectedness	Connections and networking between individuals through real or virtual interaction	Call and messaging features to connect directly to friends or family, local health care, and other users with similar health conditions or goals in their network
Control	Individuals' ability to control their own health care and well-being	Ability to personalize profile and create health goals and details of their health status. Can control who can see aspects of their health status they wish to share and reflect on what is happening in their lives
Choice	Choice in terms of products, services, and systems available to suit needs	Being provided with a suite of alternative apps to manage symptoms at home
Collaboration	Organizations and communities collaborating together to develop and deliver products, systems, and services	Can share health data with others and contribute to forums to raise issues and share experiences
Community	Individuals that are part of a community rather than living in isolation, connected to others with shared needs, interests, and aims	Can share to and link with Web-based and local communities through social media and can gain information about local community resources that might be helpful for individuals or their caregivers
Contribution	Individuals' ability to contribute to their local community	By selecting their *home* location and their interest areas, individuals can receive alerts about local happenings and can also organize their own events or groups

Methods

Contingent Valuation

CV is a form of stated preference methodology used to estimate welfare gains or losses. CV allows researchers to value nonmarket commodities [35]. In the absence of a market for a good, such as that occurring in publicly funded health care systems, surveys can be used to directly ask participants to report their WTP or willingness-to-accept (WTA) the gain or loss of a hypothetical good or service. Values elicited are then regarded as a value indicator and measure of the demand for the good [36]. This allows a direct valuation for the 6Cs, which could be used within a cost-benefit analysis (CBA). The application of WTP methodology can provide insights into what people value (or not) in future digital health services and, therefore, inform both commercial endeavors to provide what the market wants and will pay for and also the planning of health and care services in a future where health care will most certainly be supported by digital products. In this study, the approach provides an indication of people's valuation of a change in the 6Cs.

The study design was a self-complete, stated preference, open-ended WTP survey embedded within a questionnaire, which also asked respondents to self-report sociodemographic information, their general health status, and details of any existing health conditions as well as report their current app-

and digital- device and services ownership and usage. Data were collected through the use of Web-based survey panels accessed through the survey host, ResearchNow. In exchange for completing surveys, members were offered e-currency (points). For a 10 minute survey, they receive approximately £0.50 (US $0.66). Panelists accrue this as e-currency and can exchange it for goods.

Sample

Data were collected from 2 cohorts of participants. First, ResearchNow contacted UK-based panel members to create a representative sample based on age, gender, and income demographics. Second, a subsample whose characteristics mirror those of dallas communities (a *dallas-like* sample) gave the opportunity to generate a WTP estimation for those citizens currently being targeted by dallas and similar National Health Service (NHS) initiatives.

Following guidance on optimal sample sizes for CV open-ended questions, it was predicted that a sample size of no less than 400 was required [37]. To undertake subgroup analyses by including a cohort *dallas-like* sample and to take into consideration the prevalence of multimorbidity in the UK population, advice was sought from a statistician and existing literature and the sample size was increased to approximately 2000 [4,38].

Survey Design

Contextual Information

Pivotal to the success and accuracy of valuations derived from CV studies is the development of realistic, plausible scenarios, which are then presented to individuals. Poorly designed, cognitively burdensome surveys, which respondents find unrealistic, can generate biased responses and can undermine the reliability of their WTP or WTA estimates [35,39]. Before completing the WTP task, respondents were presented with contextual information (see Multimedia Appendix 1).

Respondents were then presented with a hypothetical mHealth app called *Healthy Connections* that was designed to describe the broader lifestyle and well-being outcomes (the 6Cs) that were embedded as part of the dallas program, as described in Table 1.

Willingness-to-Pay Questions

A key consideration of a stated preference WTP study is the type of hypothetical payment *vehicle* used to generate monetary values. The payment vehicle must be realistic to avoid provoking a rejection of the task [40]. For the purpose of this study, a monthly subscription fee was used. Both absolute and marginal WTP questions were included [41]. An open-ended WTP question confirmed the participants' absolute WTP for access to the app and their marginal WTP. The absolute WTP question was framed with participants asked to consider their WTP in relation to what they currently pay to stay connected to others (ie, mobile broadband charges) and for health benefits (ie, mHealth apps or gym memberships). This was to ensure that WTP for the physical mHealth features was similar to that for the current mHealth and health service markets. However, the research team acknowledged that such framing could introduce bias into the WTP results by asking respondents to state a WTP linked to their current spending on similar health or digital services and would not fully capture the respondents' valuation of the health benefits of an improvement in their sense of the 6Cs. Therefore, the marginal WTP question asked participants to consider the maximum they would be willing to pay for improved levels of 6Cs from their current 6Cs' situation. Capturing both these results allows for the researchers to understand the value placed on the health improvement expected and the value for the product or service needed to produce these.

Sociodemographic and Economic Characteristics

Our hypothesis was that general health (complete physical, mental, and social well-being) and experiences of living with long-term health conditions were likely to affect valuations for the 6Cs and that participants' familiarity with mobile technology and mHealth apps may lead to a higher WTP for the *Healthy Connections* app. Respondents were asked to rate their overall general health and well-being from *excellent* to *poor*. When referring to long-term conditions, the examples of asthma, diabetes, cancer, psoriasis, lung disease, heart disease, and depression were provided to respondents to demonstrate the diversity of conditions they should consider when describing their own health. In addition, we hypothesized that younger users could have a higher WTP (more risk taking and more familiar with newer technologies); however, we acknowledged that this had to be balanced with the likelihood that their incomes will likely be lower. Finally, we expected an income effect, with those with higher incomes and with more disposable income reporting a higher WTP.

To examine these possible influences, questions on health (self-reported general health, long-term health conditions, and medication history); ownership of, and accessibility to, technologies (computers, smartphones, internet, previous health apps' history, and total monthly spending on technology); age; and total annual income were included in the survey.

Validity Testing: Pilot Survey

To test the face validity of the survey and the suitability of the open-ended question format, a soft pilot survey was conducted (n=52) before the main Web-based survey. From these results, we were able to test the validity of our survey and whether our open-ended WTP question format was suitable and understood by participants. No respondents were reported to have struggled with the task or were unable to complete.

Analysis

Stata/12SE software (Stata Corp)was used to analyze the data [42]. To estimate a demand function for the 6Cs and the mean WTP, linear regression analyses were used. The open-ended WTP was used as the continuous, dependent variable. Socioeconomic characteristics of the participants were used as *predictor* independent variables. This allowed for the opportunity to test and profile WTP. Furthermore, the pilot study data demonstrated the wide range of WTP responses and prevalence of zero responses. Zero valuations are common in this form of study as the good or service in question is a UK health app and would be part of the suite of NHS services, which are all free at point of use (covered by taxation), and thus, there

would be an assumption that this mHealth app should not differ. Indeed, all apps on the NHS digital library are free for use. Thus, to reduce the large skew in the results and learning from the pilot study, the WTP values were converted into natural logarithms (LN) before running regressions with the main survey data. It should be noted that before taking the natural log, a value of 1 was added to WTP values to avoid the problem of 0 values. Thus, in each of the models presented, the dependent variable used was LN(WTP)=log(1+WTP). The same calculation was conducted for the marginal WTP values.

Ethics Approval and Consent to Participate

The survey and project received confirmation of University of Glasgow ethics approval (July 29, 2015).

Results

Survey Cohorts

Throughout September to October 2015, a total of 2002 respondents were surveyed for the 2 cohorts (UK general population and *dallas-like* cohorts). The general UK-representative cohort consisted of 1697 respondents. On the basis of the UK general population, 48.67% (826/1697) of the cohort were male and 51.33% (871/1697) were female. The average age of respondents was 47 years, ranging from 18 to 89 years. The majority of respondents (1421/1697, 83.73%) were from the United Kingdom. Furthermore, 68.18% (1157/1697) of the sample were in a relationship, whereas 61.52% (1044/1697) had children. Moreover, 18.70% (317/1697) of respondents had a total household income of less than £14,999 (US $19,820), 60.00% (1017/1697) earned £15,000 to £49,999 (US $19,821 to US $66,069), and the remaining 21.40% (363/1697) earned more than £50,000 (US $66,070). In addition, 62.90% (1066/1697) had no long-term health conditions, whereas 37.20% (631/1697) had long-term health conditions. In total, 14.50% (246/1697) were smokers, and 47.55% (807/1697) stated they took medications regularly. The dallas-like cohort consisted of 305 respondents. Of 305 respondents, 27.9% (85/305) were male and 72.1% (220/305) were female. The cohort had an average age of 48 years, with an age range of 16 to 86 years. Similar to the UK general population cohort, 67.2% (205/305) were in a relationship and 63.3% (193/305) had children. Overall, 15.1% (46/305) of respondents had a total household income of less than £14,999 (US $19,820), 60.0% (183/305) earned £15,000 to £49,999 (US $19,821 to US $66,069), and the remaining 25.0% (76/305) earned more than £50,000 (US $66,070). Moreover, 65.2% (199/305) had no long-term health conditions, whereas 34.8% (106/305) had long-term health conditions. In addition, 14.4% (44/305) of the cohort were smokers, and 43.6% (133/305) took medications regularly. The full details of the 2 cohorts can be found in Multimedia Appendix 2.

Absolute and Marginal Willingness-to-Pay

Summary statistics of both cohorts' respondents' absolute WTP and marginal WTP are shown in Table 2. When compared with the WTP figures of the UK general population sample, *dallas-like* respondents reported lower mean WTP and marginal WTP than that estimated in the general population survey, whereas both samples' marginal WTP estimates had a range of £600 (US $793). Furthermore, among both cohorts, there was a high prevalence of zeros for both the absolute WTP (UK general population sample: 467/1697, 27.52% and dallas-like cohort: 95/305, 31.15%) and marginal WTP (UK-representative cohort: 487/1697, 28.70% and dallas-like cohort: 99/305, 32.5%) estimates.

Results of linear regressions conducted are shown in Table 3. The results illustrate that for the general UK population cohort, respondents who felt they *disagree*, were *neutral*, or *agree* to the statement that they feel connected to health care providers were more likely to pay more ($P<.05$) for the *optimal* scenario presented to them than the reference group (*strongly disagree*). Furthermore, feeling connected to social care services or providers was shown to act as a predictor of higher WTP. The dallas-like cohort demonstrated that the only potential predictor was the sense of *control* responses. Higher levels of control over health management acted as an inverse indicator of WTP as respondents (relative to the reference level of *strongly disagree*) were more likely to pay less for the improvement provided by *Healthy Connections*.

Table 2. Absolute and marginal willingness-to-pay.

Descriptive statistics	General UK population (n=1697)		Dallas-like respondents (n=305)	
	Absolute WTP[a] (£/month)	Marginal WTP (£/month)	Absolute WTP (£/month)	Marginal WTP (£/month)
Mean	16.3 (US $21.5)	13.3 (US $17.6)	13.5 (US $17.8)	12.6 (US $16.7)
Median	5 (US $6.6)	5 (US $6.6)	5 (US $6.6)	5 (US $6.6)
Mode	0	0	0	0
Range	£0-900 (US $0-1189)	£0-600 (US $0-793)	£0-600 (US$0-793)	£0-600 (US $0-793)

[a]WTP: willingness-to-pay.

Table 3. Linear regression results for marginal willingness-to-pay and respondents' current 6Cs levels (adjusted for age, total household income, and gender).

Variable	General UK population (n=1697)			Dallas-like cohort (n=305)		
	Coefficient	*P* value	95% CI	Coefficient	*P* value	95% CI
Connections (I feel connected with/to...)						
My friends and family						
Strongly disagree	__[a]	—	—	—	—	—
Disagree	−0.2	.59	−0.82 to 0.47	−0.16	.85	−1.79 to 1.48
Neutral	−0.29	.34	−0.89 to 0.30	−0.44	.53	−1.79 to 0.92
Agree	−0.36	.22	−0.94 to 0.21	−0.44	.52	−1.79 to 0.91
Strongly agree	−0.36	.22	−0.94 to 0.22	−0.47	.5	−1.83 to 0.89
Health care services and/or providers						
Strongly disagree	—	—	—	—	—	—
Disagree	0.46	.05	0.00 to 0.93	0.33	.59	−0.88 to 1.54
Neutral	0.44	.06	−0.01 to 0.88	0.19	.75	−0.98 to 1.36
Agree	0.46	.05	0.01 to 0.92	0.06	.93	−1.14 to 1.26
Strongly agree	0.42	.1	−0.09 to 0.92	0	1	−1.40 to 1.41
Social care services and/or providers						
Strongly disagree	—	—	—	—	—	—
Disagree	0.06	.59	−0.17 to 0.30	−0.06	.82	−0.60 to 0.48
Neutral	0.3	.01	0.08 to 0.53	0.36	.16	−0.15 to 0.87
Agree	0.89	0	0.61 to 1.17	0.5	.14	−0.16 to 1.16
Strongly agree	0.98	0	0.55 to 1.42	0.77	.2	−0.42 to 1.95
I feel I make a contribution in my community						
Strongly disagree	—	—	—	—	—	—
Disagree	−0.08	.68	−0.45 to 0.29	0.11	.85	−1.00 to 1.22
Neutral	0.1	.59	−0.27 to 0.47	0.22	.69	−0.87 to 1.32
Agree	0.2	.31	−0.19 to 0.58	0.45	.44	−0.69 to 1.58
Strongly agree	0.17	.45	−0.27 to 0.61	−0.06	.93	−1.35 to 1.23
I feel I have control in how I manage my health and well-being						
Strongly disagree	—	—	—	—	—	—
Disagree	−0.06	.89	−0.83 to 0.72	−1.67	.11	−3.70 to 0.36
Neutral	0.06	.88	−0.70 to 0.81	−2.48	.01	−4.31 to −0.65
Agree	−0.19	.62	−0.94 to 0.56	−2.62	.01	−4.44 to −0.79
Strongly agree	−0.32	.41	−1.10 to −0.45	−2.35	.01	−4.22 to −0.47
I feel I have a choice in how I manage my health and well-being						
Strongly disagree	—	—	—	—	—	—
Disagree	0.44	.25	−0.32 to 1.19	0.86	.44	−1.35 to 3.08
Neutral	0.42	.27	−0.33 to 1.17	1.39	.2	−0.73 to 3.51
Agree	0.61	.11	−0.13 to 1.35	1.88	.08	−0.24 to 4.00
Strongly agree	0.65	.09	−0.11 to 1.41	1.64	.13	−0.47 to 3.74
I feel that I am part of my community						
Strongly disagree	—	—	—	—	—	—
Disagree	0.06	.78	−0.34 to 0.46	0.24	.7	−0.99 to 1.46

Variable	General UK population (n=1697)			Dallas-like cohort (n=305)		
	Coefficient	*P* value	95% CI	Coefficient	*P* value	95% CI
Neutral	0.04	.86	−0.36 to 0.43	0.32	.59	−0.85 to 1.49
Agree	0.24	.25	−0.17 to 0.65	0.55	.37	−0.65 to 1.76
Strongly agree	0.24	.31	−0.23 to 0.72	1.4	.06	−0.03 to 2.82

[a]Standard linear regression conducted and, therefore, coefficients show the difference between the variable category and "Strongly Disagree" as reference category. Strongly disagree *P* values are not applicable.

Sociodemographic and Economic Characteristics

Both cohorts indicated that respondents' age has a significant (*P*<.05) relationship with WTP (UK population cohort: −0.1, 95% CI −0.02 to −0.01; dallas-like cohort: −0.02, 95% CI −0.03 to −0.01), illustrating that younger respondents will pay more for the health connections app. In the general UK population cohort, relative to the reference level group (≤£14,999/US $19,819), £15,000 to £29,999 (US $19,821 to $39,641), income level acts as a significant, positive predictor of higher WTP (0.21, 95% CI 0.14 to 0.4). This is the theoretically expected result. However, this trend is not shown in the income earning brackets of £30,000 to £49,999 (US $39,642 to $66,069) or ≥£50,000 (≥US $66,070), and no relationship between income and WTP is estimated in the dallas-like cohort. Gender differences were statistically significant only in the dallas-like cohort where females had a lower WTP relative to the male reference level (−0.35, 95% CI −0.69 to −0.01). For both cohorts, general health was a positive predictor of WTP, with those respondents who describe themselves in better health being more likely to spend more for the *Healthy Connections* app.

However, only in the general UK population sample, there was a statistically significant positive relationship between regularly taking medication and higher WTP (0.16, 95% CI −0.01 to 0.32). This trend was not statistically significant in the dallas-like cohort, and neither cohort illustrated that long-term illness was a factor influencing WTP. These results suggest that individuals who are currently in better health value the mHealth app the most. The full analysis can be found in Multimedia Appendix 3.

In the general UK population cohort, higher WTP values were positively associated with current total monthly payments on phone, internet, and additional features (ie, app subscriptions), with respondents who reported they currently spent more on these services monthly stating larger WTP for the *Healthy Connections* app. Respondents who described themselves as having the internet yet never use it had a significant (*P*<.05), positive relationship with WTP (1.18, 95% CI 0.34 to 2.01) and were more likely to pay a higher amount than the reference group who have no access to the internet at home. In addition, those who have access to the internet at home and use it regularly demonstrated a negative association with WTP (−0.5, 95% CI −0.93 to −0.07) relative to the reference group. Finally, owning a computer but rarely using it acted as a statistically significant predictor of an inverse WTP (−0.5, 95% CI −0.95 to −0.05), paying less than those who do not own a computer. Results from the dallas-like cohort highlighted that owning a computer or smartphone, having regular access to the internet, and the total monthly payment for phones and internet usage (and additional features) were not indicators for higher WTP. For both cohorts, previous amount spent on health apps acted as a significant positive predictor of WTP, yet the effect was minimal (0.01). These linear regression results on familiarity and accessibility to mHealth and technology demonstrate that aside from the UK population cohort's positive association between current payments for phone, internet, and additional features and higher WTP, having access to a computer and internet is not a clear indicator of higher value and WTP for mHealth and was shown to be a negative indicator of WTP in the UK population cohort (see Multimedia Appendix 3 for further details).

Discussion

Principal Findings

Drawing on data from 2 cohorts, we have demonstrated that both the general UK population and a cohort whose characteristics are similar to those already receiving a large-scale digital health program valued both the access to improved broader well-being (6Cs) and the development of an mHealth app such as *Healthy Connections*. This WTP study revealed a positive valuation of the 6Cs of £196 (US $258) per annum for the general UK population cohort's absolute WTP values and a value of £160 (US $211) for the marginal WTP (ie, to move participants from their current 6Cs' position to the highest level of 6Cs). In addition, the dallas-like cohort's absolute WTP valued the 6Cs mHealth app at £162 (US $214) and a value of £151 (US $200) for the marginal WTP. By incorporating questions about both these forms of WTP, we were able to evidence positive valuations for both the possibility of the improvement in their sense of each of the 6Cs' lifestyle components from their current 6Cs' experience (marginal WTP) and also for the value for the app itself (absolute WTP). Therefore, the study's results lend themselves to a wider evidence base than just mHealth apps and solutions and can demonstrate that investment in other activities or services, which seek to foster improvement in 6Cs lifestyle components, may also be a worthwhile investment in resource allocation.

Furthermore, the study illustrates that for the general UK population cohort, this WTP was positively affected by participants' existing sense of connection to social care services and having current connections to health care services or staff. Conversely, dallas-like respondents who felt they already had a sense of control in their health and well-being management demonstrated an inverse relationship to WTP. Such sensitivity to individual needs and preferences may represent a costly or

time-consuming development process, yet these results further evidence the challenges associated with obtaining consistent, homogenous preferences from WTP surveys of digital health programs.

In addition, the valuations are based on the understanding that the *Healthy Connections* mHealth app was a generalizable (not disease specific) mHealth service suitable for the whole population. The research team envisaged that the 6Cs lifestyle components were aspects of health and well-being that could be valuable for all users, not just those currently suffering from an illness. The results highlighted that for both cohorts, better self-reported health was positively associated with WTP, and long-term illness was not a factor that influenced respondents' WTP, whereas regular medication was associated with higher WTP in the general UK population cohort. The lack of clarity on the relationship between a person's health, health behaviors, and WTP for an app such as *Healthy Connections* in the results suggest that the true value of an app such as *Healthy Connections* could be investigated further with a more detailed focus on types of health (physical and mental well-being) and disease types. The strength of this study is that it shows there is an inherent value for the 6Cs for multiple types of users with differing health needs and status and, thus, provides initial evidence of the need for further investigation of the role of mHealth to improve lifestyle. Examination of sociodemographic and economic factors and familiarity with mHealth technology demonstrated some user traits that may help inform future development of similar mHealth apps. Age was shown to act as a predictor of higher WTP in both cohorts, with younger respondents being willing to pay more for the app. Female respondents were shown to have a lower WTP than their male counterparts in the dallas-like cohort. Beyond higher current spending on digital devices and health apps as indicators of higher WTP, no clear trends were shown across internet, computer, or smartphone access and use. In fact, although the dallas-like cohort results showed no statistically significant trends, the general UK population results showed owning a home computer and using it rarely and being regular users of internet as negatively associated with WTP. This variability in the results highlights a clear need for more research on how *type* of digital platform or accessibility options may impact the success of mHealth apps and the investment in upskilling of users required. Importantly, it suggests that it is incorrect to assume that levels of access to smartphones or the internet can be used to reliably predict uptake of digital health services. Surprisingly, despite other cost indicators acting as indicators of WTP, an increase in total household income was not shown to have the expected significant trend on WTP. There was an increase in WTP, relative to the reference level of *less than £14,999* (US $19,819); however, this was not significant beyond £15,000 to 29,999 (US $19,821 to $39,641). Such confounders suggest that although our study demonstrates that there is clear evidence to support the rationale for developing mHealth as a new supporting method for health care delivery, inherently their use or appropriateness may not be solely reliant on income but perhaps existing familiarity and acceptance for these forms of health-related technologies as a norm or part of daily routine.

Limitations

A limitation of the dallas evaluation is that impact on health and social care resource use was not captured. We can, however, compare our WTP results with both the cost of a dallas-type product and also the cost of the dallas program. The cost of an app can range from free to £1, £10s, and £100s [43], dependent on the type of app. The dallas program included costs of recruiting and reaching users and interoperability costs (ie, enabling work to integrate the apps with health records and social care systems). Further research would enable these WTP results to be used in a CBA framework [20]. To do this, longer-term follow-up would be required to capture impact on health and social care resource use and any potential cost-savings, for example, an attributable reduction in hospital admissions in addition to the cost of an app itself.

The open-ended WTP approach is typically associated with large values, skewed data, and zeros [44]. We have found this to be the case in this study; however, through the decision to capture data on both absolute and marginal WTP, we were able to mitigate the effect of anchoring bias. The study was able to determine value of both the development of mHealth apps and of users' improving their sense of the 6Cs [45].

Another limitation of this study is the UK context, an environment in which there is free universal access to health care. WTP might be quite different in a fee-paying environment, for example, the United States, where use of mHealth apps to avoid attending traditional health care professionals might be valued differently.

Researchers such as Klasnja and Pratt [24] have highlighted how advancements in mHealth technology could, if delivered in a sensitive and appropriate manner, not only be effective for solely specific disease management or general health improvement but could also leverage social networks and communities to support one another. This would allow virtual networks or communities of users with similar goals or location to connect to one another.

Conclusions

This study demonstrates that although consumers value mHealth solutions that promote well-being, social connectivity, and health care control, mHealth is not universally embraced, and more research is needed to understand the relationship between health status of the potential user and how to tailor an app such as *Healthy Connections* to suit their needs. Furthermore, the study evidences that accessibility and use of smartphones, internet, or computers do not equate to WTP for mHealth apps. For mHealth to achieve its potential, apps need to be tailored to the accessibility and health needs of the user and more understanding of what hinders the use or acceptability of mHealth apps to even the most frequent users of multiple digital technologies is required. A key challenge is how to engage people with long-term conditions to encourage uptake of mHealth apps. This novel application of WTP in a digital health context presents a compelling economic argument for further research and future investment in both improving the accessibility and, where necessary, upskilling the population to

promote access and expedite uptake and utilization of such digital health and well-being apps.

Acknowledgments

The authors wish to express their thanks to the funding body Innovate UK, the 4 dallas communities, and the dallas evaluation advisory board. The authors would also like to acknowledge the wider dallas evaluation team, including Andrew Briggs, Alison Devlin, Catherine O'Donnell, Siobhan O'Connor, Ruth Agbakoba, Sarah Barry, and Stephen Brewster. The dallas program was funded by Innovate UK (formerly known as the UK Technology Strategy Board). The views expressed in this paper are those of the authors' and not necessarily those of Innovate UK or any of the dallas funding bodies.

Authors' Contributions

MMB, FSM, ML, and EG as part of the wider dallas team conceived and designed the dallas evaluation framework and secured funding for this study. EM, EG, and FSM conceived the WTP embedded study. CS, EM, and EG designed the WTP study. CS led the WTP data collection and analysis, with EG and EM contributing to the data collection and analysis and interpretation of results. CS and EG drafted the paper, and EM, ML, and FM contributed to the redrafting. All authors approved the final draft for submission.

Conflicts of Interest

None declared.

Multimedia Appendix 1

The willingness-to-pay survey used in the study.

[PDF File (Adobe PDF File), 140KB - mhealth_v7i1e3_app1.pdf]

Multimedia Appendix 2

Sociodemographic information about both cohorts.

[PDF File (Adobe PDF File), 87KB - mhealth_v7i1e3_app2.pdf]

Multimedia Appendix 3

Linear regression results of cohorts' sociodemographic and economic characteristics and willingness-to-pay and familiarity and accessibility to mobile health technology and willingness-to-pay.

[PDF File (Adobe PDF File), 56KB - mhealth_v7i1e3_app3.pdf]

References

1. Poushter J. Pew Research Center. Washington, DC: Pew Internet and American Life Project; 2016 Feb 22. Smartphone ownership and Internet usage continues to climb in emerging economies URL: http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/ [accessed 2018-11-20] [WebCite Cache ID 744UwlNOn]
2. Loiselle CG, Ahmed S. Is Connected Health Contributing to a Healthier Population? J Med Internet Res 2017 Nov 10;19(11):e386 [FREE Full text] [doi: 10.2196/jmir.8309] [Medline: 29127077]
3. Hostetter M, Klein S, McCarthy D. The Commonwealth Fund.: The Commonwealth Fund; 2014 Oct 01. Taking Digital Health to the Next Level: Promoting Technologies That Empower Consumers and Drive Health System Transformation URL: https://www.commonwealthfund.org/publications/fund-reports/2014/oct/taking-digital-health-next-level [accessed 2018-11-20]
4. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. The Lancet 2012 Jul;380(9836):37-43. [doi: 10.1016/S0140-6736(12)60240-2]
5. Hesse BW, Greenberg AJ, Rutten LJF. The role of Internet resources in clinical oncology: promises and challenges. Nat Rev Clin Oncol 2016 Dec;13(12):767-776. [doi: 10.1038/nrclinonc.2016.78] [Medline: 27273045]
6. Scottish Government. Edinburgh: Scottish Government; 2011 Mar 03. Scotland's Digital Future: A Strategy for Scotland URL: https://www2.gov.scot/Topics/ArtsCultureSport/arts/Archive/DigitalAmbition/DigitalAmbitionScotland [accessed 2018-10-17]
7. Bert F, Giacometti M, Gualano MR, Siliquini R. Smartphones and health promotion: a review of the evidence. J Med Syst 2014 Jan;38(1):9995. [doi: 10.1007/s10916-013-9995-7] [Medline: 24346929]

8. EuroHealth. European observatory on health systems and policies. 2013. Building resilient and innovative health systems URL: http://www.euro.who.int/en/about-us/partners/observatory/publications/eurohealth/previous-issues/back-issues-volumes-12-20-2006-2014/building-resilient-and-innovative-health-systems [accessed 2018-10-16]
9. Ofcom. 2015 Aug 17. The Communications Market Report URL: https://www.ofcom.org.uk/research-and-data/multi-sector-research/cmr [accessed 2018-11-20] [WebCite Cache ID 744Vqpner]
10. Jahns R. Research 2 Guidance. 2017. The market for mHealth app services will reach $26 billion by 2017 URL: https://research2guidance.com/the-market-for-mhealth-app-services-will-reach-26-billion-by-2017/ [accessed 2018-11-20] [WebCite Cache ID 744W8AeXa]
11. May C, Gask L, Atkinson T, Ellis N, Mair F, Esmail A. Resisting and promoting new technologies in clinical practice: the case of telepsychiatry. Soc Sci Med 2001 Jun;52(12):1889-1901. [Medline: 11352414]
12. Greenhalgh T, Hinder S, Stramer K, Bratan T, Russell J. Adoption, non-adoption, and abandonment of a personal electronic health record: case study of HealthSpace. Br Med J 2010 Nov 16;341(nov16 1):c5814-c5814 [FREE Full text] [doi: 10.1136/bmj.c5814] [Medline: 21081595]
13. Lennon MR, Bouamrane M, Devlin AM, O'Connor S, O'Donnell C, Chetty U, et al. Readiness for delivering digital health at scale: lessons from a longitudinal qualitative evaluation of a national digital health innovation program in the United Kingdom. J Med Internet Res 2017 Dec 16;19(2):e42 [FREE Full text] [doi: 10.2196/jmir.6900] [Medline: 28209558]
14. Devlin AM, McGee-Lennon M, O'Donnell CA, Bouamrane M, Agbakoba R, O'Connor S, et al. Delivering digital health and well-being at scale: lessons learned during the implementation of the dallas program in the United Kingdom. J Am Med Inform Assoc 2016 Jan 08;23(1):48-59 [FREE Full text] [doi: 10.1093/jamia/ocv097] [Medline: 26254480]
15. Mair FS, May C, O'Donnell C, Finch T, Sullivan F, Murray E. Factors that promote or inhibit the implementation of e-health systems: an explanatory systematic review. Bull World Health Organ 2012 May 01;90(5):357-364 [FREE Full text] [doi: 10.2471/BLT.11.099424] [Medline: 22589569]
16. Lorgelly PK, Lawson KD, Fenwick EA, Briggs AH. Outcome measurement in economic evaluations of public health interventions: a role for the capability approach? Int J Environ Res Public Health 2010 Dec 06;7(5):2274-2289 [FREE Full text] [doi: 10.3390/ijerph7052274] [Medline: 20623024]
17. Baba C, Kearns A, McIntosh E, Tannahill C, Lewsey J. Is empowerment a route to improving mental health and wellbeing in an urban regeneration (UR) context? Urban Studies. Urban Stud 2017 Feb 26;54(7):1619-1637 [FREE Full text] [doi: 10.1177/0042098016632435]
18. Callan A, O'Shea E. Willingness to pay for telecare programmes to support independent living: results from a contingent valuation study. Soc Sci Med 2015 Jan;124:94-102 [FREE Full text] [doi: 10.1016/j.socscimed.2014.11.002]
19. NICE. 2016 Mar. Community engagement to Improve health URL: https://www.nice.org.uk/guidance/ng44 [accessed 2018-11-20] [WebCite Cache ID 744WqQJOr]
20. Wildman J, McMeekin P, Grieve E, Briggs A. Economic evaluation of integrated new technologies for health and social care: suggestions for policy makers, users and evaluators. Soc Sci Med 2016 Dec;169:141-148. [doi: 10.1016/j.socscimed.2016.09.033] [Medline: 27721138]
21. McNamee P, Murray E, Kelly M, Bojke L, Chilcott J, Fischer A, et al. Designing and undertaking a health economics study of digital health interventions. Am J Prev Med 2016 Dec;51(5):852-860 [FREE Full text] [doi: 10.1016/j.amepre.2016.05.007] [Medline: 27745685]
22. Bentley A, King K, Radevsky A, Rainsford C. Winning Moves. 2013. Final report: dallas baseline economic evaluation URL: https://www.winningmoves.com/publications/final-report-dallas-delivering-assisted-living-lifestyles-scale-baseline-economic [accessed 2018-10-11]
23. McGee-Lennon M, Bouamrane M, Barry S, Grieve E, Latina D, Watson N. Evaluating the Delivery of Assisted Living Lifestyles at Scale (dallas). 2012 Sep 13 Presented at: The 26th BCS Conference on Human Computer interaction Workshop; 2012; Birmingham URL: http://ewic.bcs.org/content/ConWebDoc/48790
24. Fawsitt CG, Bourke J, Murphy A, McElroy B, Lutomski JE, Murphy R, et al. A cost-benefit analysis of two alternative models of maternity care in Ireland. Appl Health Econ Health Policy 2017 Dec;15(6):785-794 [FREE Full text] [doi: 10.1007/s40258-017-0344-8] [Medline: 28828573]
25. Shariful Islam SM, Lechner A, Ferrari U, Seissler J, Holle R, Niessen LW. Mobile phone use and willingness to pay for SMS for diabetes in Bangladesh. J Public Health (Oxf) 2016 Mar;38(1):163-169. [doi: 10.1093/pubmed/fdv009] [Medline: 25687131]
26. O'Connor S, Hanlon P, O'Donnell CA, Garcia S, Glanville J, Mair FS. Understanding factors affecting patient and public engagement and recruitment to digital health interventions: a systematic review of qualitative studies. BMC Med Inform Decis Mak 2016 Dec 15;16(1):120 [FREE Full text] [doi: 10.1186/s12911-016-0359-3] [Medline: 27630020]
27. Bergmo TS. Can economic evaluation in telemedicine be trusted? A systematic review of the literature. Cost Eff Resour Alloc 2009 Oct 24;7:18 [FREE Full text] [doi: 10.1186/1478-7547-7-18] [Medline: 19852828]
28. Bensink M, Hailey D, Wootton R. A systematic review of successes and failures in home telehealth: preliminary results. J Telemed Telecare 2016 Dec 02;12(3_suppl):8-16. [doi: 10.1258/135763306779380174]

29. Darkins A, Ryan P, Kobb R, Foster L, Edmonson E, Wakefield B, et al. Care coordination/home telehealth: the systematic implementation of health informatics, home telehealth, and disease management to support the care of veteran patients with chronic conditions. Telemed J E Health 2008 Dec;14(10):1118-1126. [doi: 10.1089/tmj.2008.0021] [Medline: 19119835]
30. Polisena J, Coyle D, Coyle K, McGill S. Home telehealth for chronic disease management: a systematic review and an analysis of economic evaluations. Int J Technol Assess Health Care 2009 Jul;25(3):339-349. [doi: 10.1017/S0266462309990201] [Medline: 19619353]
31. Rojas SV, Gagnon M. A systematic review of the key indicators for assessing telehomecare cost-effectiveness. Telemed J E Health 2008 Nov;14(9):896-904 [FREE Full text] [doi: 10.1089/tmj.2008.0009] [Medline: 19035798]
32. Dávalos ME, French MT, Burdick AE, Simmons SC. Economic evaluation of telemedicine: review of the literature and research guidelines for benefit-cost analysis. Telemed J E Health 2009 Dec;15(10):933-948. [doi: 10.1089/tmj.2009.0067] [Medline: 19954346]
33. Davies A, Newman S. The Kings Fund. 2011 Feb. Evaluating telecare and telehealth interventions: WSDAN briefing paper URL: https://www.kingsfund.org.uk/sites/default/files/Evaluating-telecare-telehealth-interventions-Feb2011.pdf [accessed 2018-11-20] [WebCite Cache ID 744XHiooW]
34. McGee-Lennon M, Bouamrane M, Grieve E, O'Donnell C, O'Connor S, Agbakoba R. A Flexible Toolkit for Evaluating Person-Centred Digital Health and Wellness at Scale. In: Advances in Human Factors and Ergonomics in Healthcare. Switzerland: Springer International Publishing; 2017:105-118.
35. Frew E. Benefit assessment for cost-benefit analysis studies in health care: a guide to carrying out a stated preference willingness to pay survey in health care. In: Applied Methods of Cost-Benefit Analysis in Health Care. 2 ed. Oxford: Oxford University Press; 2011:119-138.
36. Drummond M, Sculpher M, Claxton K, Stoddart G, Torrance G. Methods for the Economic Evaluation of Health Care Programmes. Oxford: Oxford University Press; 2015.
37. Vaughan W, Darling A. Inter-American Development Bank. 2000. The Optimal Sample Size for Contingent Valuation Surveys: Applications to Project Analysis URL: https://publications.iadb.org/bitstream/handle/11319/4837/The%20Optimal%20Sample%20Size%20for%20Contingent%20Valuation%20Surveys%3a%20Applications%20to%20Project%20Analysis.pdf?sequence=1&isAllowed=y [accessed 2018-09-11]
38. Lawson KD, Mercer SW, Wyke S, Grieve E, Guthrie B, Watt GC, et al. Double trouble: the impact of multimorbidity and deprivation on preference-weighted health related quality of life a cross sectional analysis of the Scottish Health Survey. Int J Equity Health 2013 Aug 20;12:67 [FREE Full text] [doi: 10.1186/1475-9276-12-67] [Medline: 23962150]
39. Smith RD. Construction of the contingent valuation market in health care: a critical assessment. Health Econ 2003 Aug;12(8):609-628. [doi: 10.1002/hec.755] [Medline: 12898660]
40. Gyrd-Hansen D. Using the stated preference technique for eliciting valuations: the role of the payment vehicle. Pharmacoeconomics 2013 Oct;31(10):853-861. [doi: 10.1007/s40273-013-0086-x] [Medline: 24009048]
41. McIntosh E. Applied Methods of Cost-Benefit Analysis in Health Care. Oxford: Oxford Univ Press; 2010.
42. StataCorp. Stata. Texas: StataCorp LP; 2011. Stata Statistical Software: Release URL: https://www.stata.com/stata12/
43. Tonnel A. YouTube. 2015. NICE assessment of NHS and social care apps URL: https://www.youtube.com/watch?v=xkglgo-UQ4E [accessed 2018-11-20] [WebCite Cache ID 744eWaNEg]
44. Mitchell R, Carson R. Using Surveys To Value Public Goods. New York: RFF Press; 2000.
45. Pearce D, Ozdemiroglu E. Department for Transport, Local Government and the Regions. Rotherham: Department for Transport, Local Government and the Regions; 2002. Economic valuation with stated preference techniques URL: https://webarchive.nationalarchives.gov.uk/20120919162306/http:/www.communities.gov.uk/documents/corporate/pdf/146871.pdf [accessed 2018-09-01]

Abbreviations

CBA: cost-benefit analysis
CV: contingent valuation
dallas: delivering assisted living lifestyles at scale
LN: natural logarithms
mHealth: mobile health
NHS: National Health Services
NICE: National Institute for Health and Care Excellence
WTA: willingness-to-accept
WTP: willingness-to-pay

Edited by G Eysenbach; submitted 30.01.18; peer-reviewed by LM Aiello, S Della Rosa; comments to author 17.08.18; revised version received 12.10.18; accepted 12.10.18; published 17.01.19.

Please cite as:
Somers C, Grieve E, Lennon M, Bouamrane MM, Mair FS, McIntosh E
Valuing Mobile Health: An Open-Ended Contingent Valuation Survey of a National Digital Health Program
JMIR Mhealth Uhealth 2019;7(1):e3
URL: http://mhealth.jmir.org/2019/1/e3/
doi:10.2196/mhealth.9990
PMID:30664488

Corrigenda and Addenda

Figure Correction: Resting and Postexercise Heart Rate Detection From Fingertip and Facial Photoplethysmography Using a Smartphone Camera: A Validation Study

Bryan P Yan[1], MBBS, FRACP, FRCP (Edin), FACC; Christy KY Chan[1], BSc (Hons); Christien KH Li[2]; Olivia TL To[1], RN, MSc; William HS Lai[1], BSc (Hons); Gary Tse[1,3], PhD, FACC; Yukkee C Poh[4], PhD; Ming-Zher Poh[4], PhD

[1]Division of Cardiology, Department of Medicine and Therapeutics, The Chinese University of Hong Kong and Prince of Wales Hospital, Hong Kong, China (Hong Kong)

[2]Faculty of Medicine, Newcastle University, Newcastle upon Tyne, United Kingdom

[3]Li Ka Shing Institute of Health Sciences, Faculty of Medicine, The Chinese University of Hong Kong, Hong Kong, China (Hong Kong)

[4]Cardiio Inc, Cambridge, MA, United States

Corresponding Author:
Bryan P Yan, MBBS, FRACP, FRCP (Edin), FACC
Division of Cardiology
Department of Medicine and Therapeutics
The Chinese University of Hong Kong and Prince of Wales Hospital
9/F, Division of Cardiology, Department of Medicine and Therapeutics, Clinical Sciences Building
Prince of Wales Hospital, Shatin, N.T.
Hong Kong,
China (Hong Kong)
Phone: 852 2632 3142
Fax: 852 2637 3852
Email: bryan.yan@cuhk.edu.hk

Related Article:

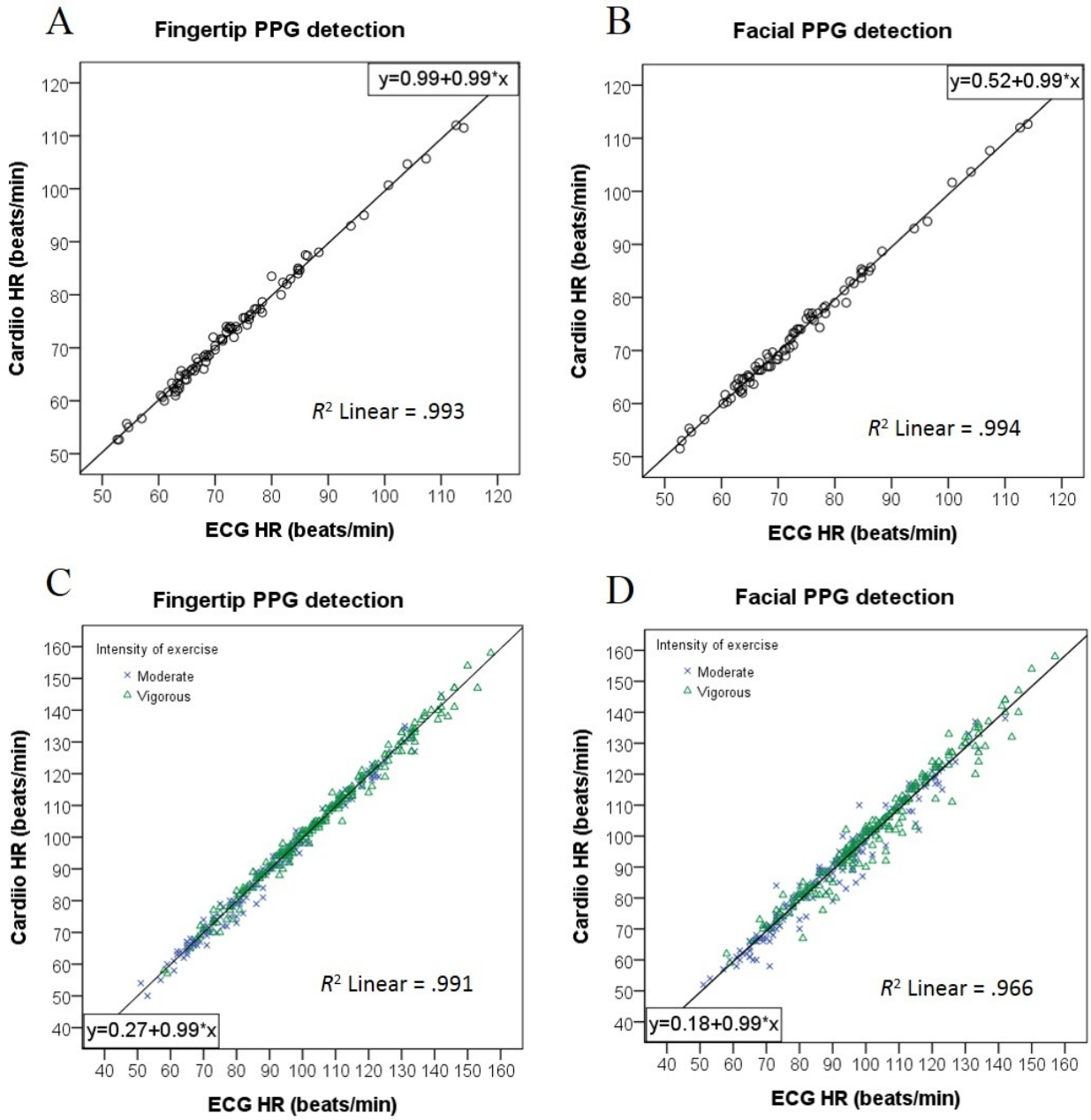
Correction of: http://mhealth.jmir.org/2017/3/e33/

(JMIR Mhealth Uhealth 2019;7(1):e11616) doi:10.2196/11616

In "Resting and Postexercise Heart Rate Detection From Fingertip and Facial Photoplethysmography Using a Smartphone Camera: A Validation Study" (JMIR Mhealth Uhealth 2017;5(3):e33), there was an error in Figure 2. The scatter plot of Figure 2C (Postexercise HR from fingertip PPG signals) was a duplicate of Figure 2D (Postexercise HR from facial PPG signals). The scatter plot of Figure 2C has been updated with a correct value of R^2=0.991. The equation of line (y=0.27+0.99*x) remains unchanged.

The correction will appear in the online version of the paper on the JMIR website on January 3, 2019, 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 also has been resubmitted to those repositories.

Figure 2. Scatter plots comparing measurements of heart rate (HR) estimated from the Cardiio smartphone phone app photoplethysmographic (PPG) signals and from a reference electrocardiogram (ECG). *P*<.001 for all correlations. (A) Resting estimated HR from fingertip PPG signals. (B) Resting estimated HR from facial PPG signals. (C) Postexercise HR from fingertip PPG signals. (D) Postexercise HR from facial PPG signals.



Edited by G Eysenbach; submitted 17.07.18; this is a non–peer-reviewed article;accepted 09.10.18; published 03.01.19.

Please cite as:
Yan BP, Chan CKY, Li CKH, To OTL, Lai WHS, Tse G, Poh YC, Poh MZ
Figure Correction: Resting and Postexercise Heart Rate Detection From Fingertip and Facial Photoplethysmography Using a Smartphone Camera: A Validation Study
JMIR Mhealth Uhealth 2019;7(1):e11616
URL: http://mhealth.jmir.org/2019/1/e11616/
doi:10.2196/11616
PMID:30609980

Corrigenda and Addenda

Correction: Medical-Grade Physical Activity Monitoring for Measuring Step Count and Moderate-to-Vigorous Physical Activity: Validity and Reliability Study

Myles William O'Brien[1,2], BKinH; William Robert Wojcik[1], BKinH; Jonathon Richard Fowles[1], PhD

[1]Centre of Lifestyle Studies, School of Kinesiology, Acadia University, Wolfville, NS, Canada

[2]Division of Kinesiology, Dalhousie University, Halifax, NS, Canada

Corresponding Author:
Jonathon Richard Fowles, PhD
Centre of Lifestyle Studies, School of Kinesiology
Acadia University
550 Main Street
Wolfville, NS,
Canada
Phone: 1 9025851560
Fax: 1 902 585 1702
Email: jonathon.fowles@acadiau.ca

Related Article:

Correction of: http://mhealth.jmir.org/2018/9/e10706/

(JMIR Mhealth Uhealth 2019;7(1):e12576) doi:10.2196/12576

The Authors of "Medical-Grade Physical Activity Monitoring for Measuring Step Count and Moderate-to-Vigorous Physical Activity: Validity and Reliability Study" (JMIR Mhealth Uhealth 2018;6(9):e10706) mistakenly represented the Yamax Digiwalker in the Discussion section. Unlike some of Yamax's newer devices (ie, Yamax EX-510), the Yamax Digiwalker is a spring-levered pedometer and not a piezoelectric pedometer.

Thus, the following sentence has been removed from the Discussion:

> *Similar to the PiezoRx, the Yamax also uses a piezoelectric sensor, which is consistent with this study.*

The correction will appear in the online version of the paper on the JMIR website on January 3, 2019, 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 also has been resubmitted to those repositories.

Edited by G Eysenbach; submitted 22.10.18; this is a non–peer-reviewed article;accepted 25.10.18; published 03.01.19.

Please cite as:
O'Brien MW, Wojcik WR, Fowles JR
Correction: Medical-Grade Physical Activity Monitoring for Measuring Step Count and Moderate-to-Vigorous Physical Activity: Validity and Reliability Study
JMIR Mhealth Uhealth 2019;7(1):e12576
URL: https://mhealth.jmir.org/2019/1/e12576/
doi:10.2196/12576
PMID:30609987

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