JMIR mHealth and uHealth

Contents

Original Papers

Short Paper

Reviews

Corrigenda and Addenda

Correction of: Sleep Quality Prediction From Wearable Data Using Deep Learning (e130)

Aarti Sathyanarayana, Shafiq Joty, Luis Fernandez-Luque, Ferda Ofli, Jaideep Srivastava, Ahmed Elmagarmid, Teresa Arora, Shahrad Taheri.
2   5   4

Original Paper

Formative Work to Develop a Tailored HIV Testing Smartphone App for Diverse, At-Risk, HIV-Negative Men Who Have Sex With Men: A Focus Group Study

Jason W Mitchell[1], MPH, PhD (Public Health); Maria Beatriz Torres[2], PhD; Jennifer Joe[3], MPH; Thu Danh[3], MPH; Bobbi Gass[3], BA; Keith J Horvath[3], PhD

[1]Office of Public Health Studies, University of Hawai'i at Mānoa, Honolulu, HI, United States
[2]Communication Studies, Gustavus Adolphus College, Saint Peter, MN, United States
[3]School of Public Health, University of Minnesota, Minneapolis, MN, United States

Corresponding Author:
Jason W Mitchell, MPH, PhD (Public Health)
Office of Public Health Studies
University of Hawai'i at Mānoa
1960 East-West Road, Biomed D104AA
Honolulu, HI, 96822
United States
Phone: 1 808 956 3342
Fax: 1 808 956 3342
Email: jasonmit@hawaii.edu

Abstract

Background: Although gay, bisexual, and other men who have sex with men (MSM) are disproportionately affected by human immunodeficiency virus (HIV) infection, few test for HIV at regular intervals. Smartphone apps may be an ideal tool to increase regular testing among MSM. However, the success of apps to encourage regular testing among MSM will depend on how frequently the apps are downloaded, whether they continue to be used over months or years, and the degree to which such apps are tailored to the needs of this population.

Objective: The primary objectives of this study were to answer the following questions. (1) What features and functions of smartphone apps do MSM believe are associated with downloading apps to their mobile phones? (2) What features and functions of smartphone apps are most likely to influence MSM's sustained use of apps over time? (3) What features and functions do MSM prefer in an HIV testing smartphone app?

Methods: We conducted focus groups (n=7, with a total of 34 participants) with a racially and ethnically diverse group of sexually active HIV-negative MSM (mean age 32 years; 11/34 men, 33%, tested for HIV ≥10 months ago) in the United States in Miami, Florida and Minneapolis, Minnesota. Focus groups were digitally recorded, transcribed verbatim, and deidentified for analysis. We used a constant comparison method (ie, grounded theory coding) to examine and reexamine the themes that emerged from the focus groups.

Results: Men reported cost, security, and efficiency as their primary reasons influencing whether they download an app. Usefulness and perceived necessity, as well as peer and posted reviews, affected whether they downloaded and used the app over time. Factors that influenced whether they keep and continue to use an app over time included reliability, ease of use, and frequency of updates. Poor performance and functionality and lack of use were the primary reasons why men would delete an app from their phone. Participants also shared their preferences for an app to encourage regular HIV testing by providing feedback on test reminders, tailored testing interval recommendations, HIV test locator, and monitoring of personal sexual behaviors.

Conclusions: Mobile apps for HIV prevention have proliferated, despite relatively little formative research to understand best practices for their development and implementation. The findings of this study suggest key design characteristics that should be used to guide development of an HIV testing app to promote regular HIV testing for MSM. The features and functions identified in this and prior research, as well as existing theories of behavior change, should be used to guide mobile app development in this critical area.

XSL•FO
RenderX

KEYWORDS

smartphone apps; mHealth; HIV testing; HIV-negative men who have sex with men; men who have sex with men; MSM

Introduction

Approximately 1.2 million people are living with human immunodeficiency virus (HIV) in the United States, and 1 in 8 individuals is unaware of their infection [1]. Despite advances in antiretroviral therapy and steady prevention efforts, gay, bisexual, and other men who have sex with men (MSM) continue to be disproportionately affected by HIV. In 2014, MSM accounted for 82% of diagnoses of HIV infections among males, despite representing only 2% of the US population [2,3].

Given the HIV burden among MSM, the Centers for Disease Control and Prevention recommends that all sexually active MSM test for HIV at least annually, with more frequent testing for those who engage in high-risk behaviors (eg, condomless anal sex or drug use during sex) [4]. Studies show that nearly all MSM in the United States have tested for HIV in their lifetime, and approximately two-thirds have done so in the past year [5]. A study of young MSM residing in 5 US cities (Baltimore, Los Angeles, Miami, New York City, and San Francisco) found that 62% of participants had been tested in 2011, with increases in rates of HIV testing since the mid-1990s [6]. However, consistent and repeated testing for HIV (ie, testing for HIV at regular intervals) is needed to reduce the onward transmission of HIV associated with not knowing one's status and, if HIV-positive, to reap the benefits of prompt antiretroviral therapy. Repeated HIV testing appears to be less common among MSM. A study showed that only half of sexually active HIV-negative MSM in concordant primary relationships tested for HIV at least annually (21% tested 2 or more times a year and 29% tested annually), while the remainder of the sample tested less frequently [7]. In the same study, 20% of men never tested for HIV while in their current relationship. Interventions encouraging repeated HIV testing among MSM are needed to address this ongoing need.

Mobile technologies have expanded dramatically in recent years, mirroring increased rates of mobile phone ownership. Ownership of mobile phones with advanced capabilities (referred to here as smartphones), such as those that allow access to the Internet and use apps, grew from 35% in 2011 to 64% in 2015 [8]. Smartphone ownership is particularly high among young adults (18- to 29-year-olds; 85%), and is higher among black (70%) and Hispanic (71%) US adults than among their white peers (61%) [8]. MSM were early adopters of technology [9], including the use of smartphone apps to sexually and socially connect with other MSM [10]. Because mobile device ownership and use has steadily risen over the years 2011-2015 [11], the use of apps and other mobile-based interventions is promising.

Although HIV prevention and treatment technologies that leverage technology are widespread [12,13], further development and testing of smartphone-based app interventions targeting MSM is needed. This need is relevant because MSM—particularly black and Latino MSM—remain disproportionately affected by HIV compared with their white counterparts. However, a review of available HIV and AIDS smartphone apps on the Google Play and Apple stores as of May 2015 found that only 7% of 285 available apps specifically targeted MSM [14]. Schnall and colleagues [15] applied the Information Systems Research (ISR) framework [16] to develop a smartphone app for HIV prevention that meets the needs of high-risk MSM. The ISR framework consists of 3 interrelated cycles: (1) a relevance cycle, (2) a rigor cycle, and (3) a design cycle. First, Schnall and colleagues [15] conducted focus groups with high-risk MSM to identify which features and functions were relevant for HIV prevention, which included (self-) information management, staying healthy, HIV testing, a chat/communication function, and resources. Through the rigor cycle of ISR, Schnall and colleagues reviewed mobile app use for HIV prevention with MSM and highlighted that the development and evaluation of smartphone apps for this purpose have not been well documented. The design cycle of ISR included the development phase of an HIV prevention app by incorporating findings obtained from the relevance and rigor cycles and eliciting feedback about the app from members of the target population (ie, high-risk MSM).

Focus groups have been used in several recent studies to increase the relevance of mobile apps tailored to high-risk MSM. First, for instance, Goldenberg and colleagues [17] recruited MSM (n=38) residing in Atlanta, Seattle, and rural US regions to obtain data about their preferences for an HIV prevention app. Across groups, men reported that HIV prevention smartphone apps should (1) have an educational component to guide their decisions for which test is best for them and prevention options; (2) be interactive and engaging with personalized feedback about their own sexual behaviors; (3) provide a social networking component with other MSM; (4) use language that is simple and understandable to the community; and (5) address privacy concerns by ensuring that the app is from a credible source and by having secure messaging features [17]. Second, Aliabadi and colleagues [18] used the information, motivation, and behavioral skills model to guide focus group discussions with high-risk MSM to better understand their preferences for an HIV prevention app. Key informational (eg, HIV testing and support group information), motivational (eg, addressing sexual encounters in which men intend to use condoms, but do not), and behavioral skill (eg, negotiating safer sex, understanding signs of HIV infection) needs were identified as critical content for their HIV prevention app [18].

Similar to the studies [15,17,18] described above, in our study we conducted focus groups with at-risk MSM to inform the subsequent development of an HIV-related smartphone app. However, this study expanded on the findings of prior studies in several important ways. First, our focus groups explored MSM's use of and attitudes toward smartphone apps that they currently had on their mobile phones to better understand what features and functions they perceived to be associated with regular use of apps over time. Understanding why men may continue to use (or not use) certain apps over time may provide critical insights into how to design a sustainable mobile HIV testing app for MSM. Second, we recruited men living in Miami,

Florida and Minneapolis, Minnesota to assess whether the findings in the earlier studies [15,17,18] were applicable to MSM living in other regions of the United States.

The overarching goal of this study was to elicit feedback about smartphone app use from HIV-negative MSM to apply these lessons toward the development of an engaging and sustainable HIV testing smartphone app. We established 3 primary objectives of these focus groups to meet our goal. First, we sought to understand what features and functions of smartphone apps MSM believed were associated with downloading apps to their mobile phones. Next, we asked MSM to reflect on what features and functions of smartphone apps they believed influenced them to sustain their use of apps on their phone over time. Finally, similar to prior studies [15,17-19], we asked men to describe what features and functions they would prefer to have in an HIV testing smartphone app.

Methods

Participants

We conducted 5 focus groups, 3 in Miami, Florida and 2 in Minneapolis, Minnesota, in January and February 2015. We recruited participants for the focus groups through targeted advertisements placed on Facebook, as well as by flyers placed at local community-based and AIDS service organizations and by word of mouth. Inclusion criteria for the study were self-reported and were (1) being a man 18 years of age or older, (2) being HIV-negative or having an unknown serostatus, (3) having had anal sex with another man in the past year, (4) owning a mobile phone with smartphone features (ie, global positioning system [GPS] technology, short message service, Internet browser capabilities, apps), and (5) being an English speaker. A total of 34 participants, 17 from Miami and 17 from Minneapolis, participated in the study.

Procedures

The University of Minnesota and University of Miami institutional review boards approved all study procedures. The focus group questionnaire was developed by the research team and included questions, in a semistructured format, that explored participants' experiences with smartphones and apps. Specifically, we asked participants to describe which apps they currently had on their smartphone. Then, men were asked to reflect on what general features and functions of smartphone apps they believed to influence their decisions to download, initiate use, and continue to use the app over time. Follow-up prompts were used to gain more insight into men's decisions into smartphone app use if a particular topic (eg, usefulness, enjoyment, ease of use, security, cost, and peer influence) did not spontaneously emerge. Finally, participants were also asked to provide input about what features and functions that they would like to have in a smartphone app to encourage testing for HIV and other sexually transmitted infections.

All recruitment materials included a link to the study website where interested persons were welcomed and asked to complete eligibility screener items, from which we obtained sociodemographic and most recent HIV testing behavior data. We asked those who met eligibility criteria to provide consent.

Focus groups were conducted in confidential settings. To maintain confidentiality and promote truthful answers to the focus group questions, participants were encouraged to use and refer to each other by their first name only. Focus group discussions lasted from 90 to 105 minutes. All focus groups were digitally recorded, transcribed verbatim, and deidentified for analysis.

Data Analysis

A team of 3 independent research associates coded the focus group data. Several meetings were conducted to (1) train associates in qualitative coding and (2) ensure that associates could accurately and consistently identify emerging codes and patterns in a sample transcript, while following the same iterative process. Training was led by 1 of the associates (MBT) with extensive experience in qualitative coding. Analysis of the collected data followed the process of "interrogating, sorting and synthetizing interviews" [20]. Coding helps identify and categorize the meanings expressed by interview participants.

The first part of coding started by naming words, sentences, and paragraphs (codes). Then, the researcher grouped codes with similar meanings into larger categories of meaning. Codes and categories were refined, added, or eliminated as the heuristic process continued [21]. A constant comparison method (ie, grounded theory coding) was employed [22], with the focus group interviewees' statements being continually examined and reexamined in terms of the themes revealed, points of consistency and of difference, and answers to the research questions. This process allowed us "to make implicit views, actions and processes more visible" [20].

To start, team members read the focus group transcripts several times. Working separately, each associate established a first cycle of coding. They identified codes (or themes) emerging from the data. The codes were noted in an Excel (Microsoft Corporation) spreadsheet, along with the representative interview quotes that referred to that theme. Associates then proceeded to cluster these themes and codes into categories or patterns. Next, the team met to revise and refine their coding scheme to prepare for the second cycle of coding. The Excel files of the 3 coders were compiled and compared to determine areas of agreement and disagreements. Areas of disagreement ranged from using different labels to name a code or a pattern, to some codes or patterns not being identified by all 3 associates. We also examined quotes that were classified differently by coders. Disagreements were discussed and resolved to reach full consensus. In that process, associates collectively refined some codes and patterns that represented similar meanings, making sure that all possible meanings were accurately captured and recorded. Finally, a master file was created to reflect the agreements reached about the themes.

Results

As Table 1 shows, participants were, on average, 32 years of age; the sample's age ranged between 18 and 56 years. The study sample was racially and ethnically diverse: 7 of 17 men (41%) in Minneapolis and 4 of 17 men (23%) in Miami self-reported as nonwhite, and 7 of 17 men (41%) in Miami

self-reported as Hispanic. With respect to participant's most recent HIV test, 13 of 17 men (76%) in Minneapolis had been tested within 6 months prior to study enrollment compared with 7 of 17 men (41%) in Miami. Of the 32 men who responded to the question about what type of smartphone they owned (not shown), 16 (50%) owned an iPhone, 14 (44%) owned an Android-based phone, and 2 (6%) owned another type of phone (eg, Microsoft-based phone).

As Table 2 shows, we organized the themes that emerged from the focus groups into 5 main categories, aligning with the general structure of the interview: (1) reasons to download an app, (2) reasons associated with downloading and using the app over time, (3) reasons associated with keeping and using the app over time, (4) reasons associated with deleting an app, and (5) preferences for features and functionality in an HIV testing app. Some themes that emerged from the focus group data were unique to participants' downloading an app, keeping and using an app over time, or deleting an app. In contrast, a few themes transcended and appeared to influence whether participants downloaded and used the app over time, which are presented simultaneously below. In addition, themes that emerged from the data did not differ by location; the information and opinions shared by the men were similar among those in Miami and those in Minneapolis.

Table 1. Focus group sociodemographic characteristics.

Characteristics	Total (N=34)	Miami (n=17)	Minneapolis (n=17)
Age in years (mean)	32	34	31
Race, n (%)			
White	23 (68)	13 (76)	10 (59)
Black/African American	5 (15)	3 (18)	2 (12)
Asian	2 (6)	1 (6)	1 (6)
Native Hawaiian/Pacific Islander	1 (3)	0	1 (6)
American Indian	1 (3)	0	1 (6)
Other	2 (6)	0	2 (12)
Ethnicity, n (%)			
Hispanic	9 (27)	7 (41)	2 (12)
Non-Hispanic	25 (73)	10 (59)	15 (88)
Most recent HIV[a] test, n (%)			
1-3 months ago	12 (35)	4 (24)	8 (47)
4-6 months ago	8 (24)	3 (18)	5 (30)
7-9 months ago	3 (9)	2 (12)	1 (6)
10-12 months ago	5 (15)	4 (24)	1 (6)
>1 year ago	3 (9)	2 (12)	1 (6)
≥5 years ago	3 (9)	2 (12)	1 (6)

[a]HIV: human immunodeficiency virus.

Table 2. Themes, definitions, and participant endorsements of reasons to download, continue to use, and delete apps (N=34).

Theme	Definition	Participant endorsement, n (%)
Reasons to download an app		
Cost	How much participants would spend on an app and whether cost would deter them from downloading it.	29 (85%) 18 preferred free 11 would pay
Security	How secure an app is in terms of it having access to or protecting information.	25 (74%) 13 concerned 12 not concerned
Efficiency	Discussion of whether the app enabled them to save time and added convenience in their life.	18 (53%)
Reasons associated with downloading and using the app over time		
Usefulness and perceived necessity	Perception of the app to provide a certain function that helps to fill a certain need.	24 (71%)
Influence by peers and posted reviews	Downloading and sustained use of certain apps because of reviews, rating, and word of mouth from peers.	Influence by others: 24 (71%) 17 yes, influenced 3 not influenced 4 sometimes Influence by reviews: 21 (62%) 15 yes, influenced 3 not influenced 3 sometimes
Reasons associated with keeping *and* using the app over time		
Reliability	Discussion of whether the app is working properly and reliably compared with other apps.	4 (12%)
Ease of use	The need for an app to be simple and easy to use to be kept.	13 (38%)
Updates	Frequency at which the app would be updated.	16 (47%)
Reasons associated with deleting an app		
Poor performance and functionality	App that does not work or needs too frequent updating, or has too many crashes.	11 (32%)
Boredom and lack of use	Apps not being relevant anymore.	8 (24%)
Preferences for HIV[a] testing App features and functionality		
HIV test reminders	Discussion of opinions about receiving reminders to get tested and preferences about format, frequency, and customization of those reminders.	34 (100%) 8 format 17 frequency 9 customization
Recommended HIV testing intervals with dates	Discussion of receiving personalized, recommended testing intervals with specific dates of when to be tested next.	17 (50%)
Details about HIV testing locations and HIV test locator	Sharing of opinions about wanting to know nearby locations to test and information about the testing sites.	23 (68%)
Monitoring personal behaviors	Sharing of opinions about monitoring their own sexual behaviors.	21 (62%) 16 in favor 1 optional 4 against

[a]HIV: human immunodeficiency virus.

XSL•FO
RenderX

Reasons to Download an App

Cost

The majority of participants preferred apps that were free:

It has to be free. I'm sorry. I'm on a budget. I can't afford all of these apps, especially the really good ones. [33 years old, African American, Miami]

However, a few indicated they did not mind paying a small amount (US $1-2) for an app, particularly if they thought it would be useful:

I've noticed that a big thing for me is barrier to entry, being cost primarily. If it's a paid app I'm a lot less likely to download it to see what it is. If it's a paid app, I want to know that it's something, A) I'm going to use and B) I'm going to enjoy. If there's a free version, even though it has ads and all that stuff, I'll try that first. If I like it then I'll pay for it. Having a cost associated with it, if it's something that I'm not sure about, I definitely kind of stray away from that. [24 years old, white, Minneapolis]

Security

Participants' attitudes varied about the importance of the app being secure in terms of it having access to or protecting their information. For some, security was important to them:

I would say security is a big thing for me. It's got to be secure for the information that I'm allowing it to have or giving it, especially if it's a payment option or purchasing something. Also with the information that it utilizes from either other apps or from information I input, I want that to be secure. I make sure that it's got security features that it's a trusted app so before I download an app, it will run through the security feature on my phone [34 years old, white, Miami]

Other men expressed less concern, although their concerns were heightened by the type and amount of personal information stored in the app:

I generally don't worry about security all that much unless I'm entering a decent amount of personal information. As far as people seeing what's on my phone, I really don't [care]...if they really want to know what's there then I'll show them. It's not a big deal. But my banking information, stuff like that, social security, those I'm quite protective of. [43 years old, white, Minneapolis]

In contrast, others had little to no concerns about whether the app was secure:

I, to be honest, it doesn't. That's not something I really worry about at all. I guess I just don't think about it. [28 years old, biracial Hispanic, Minneapolis]

Efficiency

Participants shared that they were more likely to download an app if it enabled them to save time and added convenience in their life. For instance:

If an app gives me the ability to do something...more quickly...then I'll download that app. [27 years old, biracial, Minneapolis]

If it's going to add a lot of convenience to something [then I'm more likely to download it]. [19 years old, white, Miami]

Reasons Associated With Both Downloading and Using an App Over Time

Usefulness and Perceived Necessity

Another key influential factor was whether men perceived the app to help fill a particular need in their life; that is, the men would often ask themselves if the app performed a particular function that would be useful for them:

For myself, I tend to look at utility apps, productivity apps, banking apps, chat apps, or text—things like that that are very utility based; absolutely. That's the deciding factor. It always come down to, is it going to be useful for me or is this just going to be another app. [37 years old, white, Miami]

Men also questioned the need for the app, what the app could do for them (ie, usefulness), and the importance of having it over a given time period:

It's really based on, for me, the importance of the app and what it can do for me today, tomorrow, and next week. [37 years old, African American, Miami]

Similarly, men perceiving a need for the app for a particular purpose also contributed to their rationale for continuing to use the app over time:

I'll continue to use an app if it still meets the criteria that cause me to download it in the first place. If that app continues to be something that's useful then I'll keep it. There are apps on my phone that I might use twice a year, but if I know I'll probably use this, then I'll keep it. [27 years old, biracial, Minneapolis]

Influence by Peers and Posted Reviews

Peer influence and reviews posted by others who used a particular app were influential to men in deciding whether to download an app (or not) onto their smartphone. For some participants, peer influence was a primary influencer:

I think number one reason I would download something is because I heard about it from someone, or people talking about it [app]. Just hearing about something a lot of. Even if I don't really know what is it, download and just check it out and delete if it it's nothing that I need. [28 years old, biracial Hispanic, Minneapolis]

And although this individual's peers influenced his decision to download an app, he also read the reviews posted by others about the app:

I usually read five reviews before I download it [app] and it's deterred me from downloading a few apps before. [28 years old, biracial Hispanic, Minneapolis]

Participants also expressed how peers both positively and negatively influenced whether they continued to use the app over time. For example:

> *My friends, I have a group of friends, we all have an iPhone and we're all in a group chat. We send different stuff to each other if we like it, if we don't. If somebody says something wrong with it, of course, end of the day, it's your opinion but if two or three people say it's a problem…then usually we end up removing it collectively because it's not really of use.* [37 years old, African American, Miami]

In contrast, others voiced that peers had absolutely no influence on their decision to continue to use the app:

> *No, not at all. I'm my own person. If I like it and it's worth it for me, then I'll continue to [use it].* [56 years old, white Hispanic, Miami]

Other Reasons Associated With Keeping and Using an App Over Time

Other reasons associated with men keeping and using their apps over time (ie, 3 months or longer) pertained to whether they perceived the app to be reliable or easy to use, and how often the app was updated.

Reliability

Men expressed their expectations of the app needing to reliably work when they use it:

> *I would say ease of access and reliability. Huge thing. If it's crashing every two seconds then it starts becoming, is this something that I really, really need or send a crash report and just hold off or whatever.* [34 years old, white, Miami]

Ease of Use

Men also identified that intuitive features and functions were important for them to continue using an app over time:

> *I think simplicity is really the one thing that attracts me the most to an app. If it's not too complicated and it serves my need, then I will definitely continue to use it. As soon as it starts to introduce a lot of features that I don't really need…I will probably start to think about downloading some other apps that are simple to use and that can still serve my purpose.* [23 years old, Asian, Miami]]

Updates

Similar to the influence of peers about downloading and sustained use of an app, the frequency at which an app is updated was also voiced in both a positive and negative frame, which was dependent on personal preference. One participant noted that

> *I think daily would be annoying.* [34 years old, white, Miami]

This was in contrast to another participant, who stated that

> *It's not going to bother me too much, if it's a few times a week or something [like that].* [28 years old, biracial Hispanic, Minneapolis]

Reasons Associated With Deleting an App

The primary reasons why participants deleted apps on their smartphone were largely contributed by their expectations not being met about the app.

Poor Performance and Functionality

Some of those reasons pertained to the app performing poorly, as noted by these 2 participants:

> *If it keeps on shutting down on me, I'm going to stop [and delete it].* [26 years old, white Hispanic, Miami]

> *If it's [a] horrible user interface and it's clunky and I can't figure it out within the first two minutes of downloading the app, I'm just going to delete it.* [24 years old, white, Minneapolis]

Boredom and Lack of Use

Some men noted that being bored with an app and not using it for a while were reasons leading them to delete it:

> *If I'm not using it at all, I have a rule with myself. If I haven't used it in two months I'll delete it.* [18 years old, white, Minneapolis]

Boredom with the app was keenly expressed by this participant:

> *One day I'll just wake up and say, I'm going to take a break from this one and I just [delete it]…when I get tired of it or bored.* [56 years old, white Hispanic, Miami]

Preferences for HIV Testing App Features and Functionality

The latter part of the focus group discussion pertained to exploring participants' attitudes about potential components for a future HIV testing app. The specific components explored were HIV test reminders, recommended testing intervals with dates, details about testing locations and a HIV test locator, and monitoring personal behaviors.

HIV Test Reminders

The discussion about HIV test reminders centered on 3 primary components: format, frequency, and customization. For format, a few participants preferred receiving the reminders via text, while others indicated they wanted the notifications sent through the app:

> *I would say just have it send it through the app itself. It will pop up with the little thing that says…* [49 years old, white Hispanic, Miami]

The frequency in which men wanted to receive the HIV test reminders varied and ranged from daily, weekly, and monthly to other options, including being able to set it up themselves. One participant shared that

> *I would say a base for me I would say a monthly reminder would be kind of nice. Anything sooner than*

that, unless I preference it sooner, would be irritating. [34 years old, white, Miami]

Another voiced that the frequency should depend on his current sex life and related behaviors:

It would depend, if I'm in a monogamous relationship, because I don't need a message. I don't need them. I'm not going outside that relationship. If I'm playing around or if there's consensual nonmonogamy, then, yeah, I think it would be beneficial if not every 60 days, maybe every 90 just to be a little extra safe. That's what I was doing when I was way more active, it's like every three months. [37 years old, white, Miami]

In general, though, many preferred being able to control or customize how and specifically when they receive the messages:

What I would most prefer to be able to do as far as having a scheduled reminder would be just tell it to customize when it would notify me. [19 years old, white, Miami]

Recommended HIV Testing Intervals With Predetermined Dates

Universally, participants liked this component idea for the testing app. Specifically, some liked to be informed about when to test for HIV:

I really, really like that idea. It would make me even more likely to actually use an app like this. It would give me the tools to know when would be a smart interval to get tested on. [19 years old, white, Miami]

Others thought this idea was nice because it made their test decision easy for them:

That's easy and to me, easy is a good thing. [28 years old, white, Minneapolis]

Details About HIV Testing Locations and Testing Locator

Participants thought the app must have information beyond where to just get tested for HIV locally:

Just the location is definitely important, how you're going to get there is definitely important, transportationwise. Information about each testing site is also important, maybe you can add a review sort of thing to each testing site. [23 years old, Asian, Miami]

Others also expressed that this component of the testing app should use the GPS that smartphones have:

Maybe the ability to maybe use, like utilize location services and find where is a place to get tested near me. [27 years old, biracial, Minneapolis]

Monitoring Personal Behaviors

Participants' opinions varied about including a feature that would allow monitoring of their sexual behaviors in the proposed testing app. Some men liked the idea to help them learn about their own behaviors over time with respect to health and prevention:

For me, it would be fine to have that information there. It's healthy, it's educational at the same time, to remind you what you're doing, or what you're not. I would use it. Definitely, yes, my friends will [use it]. [26 years old, white Hispanic, Miami]

Other men indicated they would most likely not use this feature yet understood why it could be appealing to others:

I don't know if I would honestly use something like that. In the way of putting in my conquests, I guess, into an app. I don't feel like I would have a use or case for that necessarily. I get the appeal of it and I can see the use or case but in my personal use and case, I don't see myself using it. [24 years old, white, Minneapolis]

In addition, others thought this feature should be optional:

Optional—It's a lot like the calorie counting programs. You're only going to do it if you want to and you're interested in your own progress, but again, that's something that can put your life into perspective. If you're actually following it and you see, "Wow, I need to calm down. I'm putting myself at risk," that's a good thing, but it should be optional, not something that you're forced to do but just something that if you want to do it, you can just to keep an eye on yourself. [34 years old, white Hispanic, Miami]

Discussion

Principal Findings and Recommendations

Findings from this study identified features and functions that MSM identified as influencing them to download and continually use an app over time on their smartphone, as well as some key components of what they would want in an app that aimed to promote regular testing for HIV. Results from this study showed both unique and overlapping themes related to app preferences (ie, general preferences and HIV prevention content) when compared with findings from focus group conducted in prior studies with MSM [15,17-19] Below, we discuss major findings and how they are similar to and unique compared with these existing studies.

Men in our study noted that cost, security, and efficiency were important reasons that influenced whether they downloaded an app onto their smartphone. Perceiving the app to be useful, one of necessity, and positively reviewed by others were additional reasons men described as being important to influencing not only whether they download an app, but also whether to continue to use it over time. Goldenberg and colleagues [17,19] noted that MSM in their studies also prioritized security, the usefulness of the app, and ease of navigation as important factors. Additionally, others have noted that peers are an important influence on app use [15]. Uniquely to this study, we found that men preferred free or very low-cost apps, which may be important for widespread dissemination of HIV testing and prevention apps. Taken together, this and prior studies suggest

XSL•FO

RenderX

that developmental work of future HIV prevention app interventions should consider how best to incorporate these critical considerations into the app development process. For example, in addition to usual usability testing, formative work during the alpha and beta testing phases of HIV-related apps may benefit from asking target users to reflect on whether the app helps them to more efficiently address their HIV prevention needs, which features and functions of the app are most useful and necessary to help them reduce their risk for HIV, and whether they perceive the app to be secure for maintaining their private information. The results of our study suggest that future HIV prevention apps must meet these basic requirements for MSM to trust and use an app for any length of time.

Men in our study believed that, in order for them to keep an app and continually use it over longer periods of time, the app must work properly (ie, be reliable), be easy to use, and not be constantly updated. Ease of use has been described as an important consideration for continued app use in at least one prior study [17]. In addition, our results draw attention to the need to ensure that HIV testing and prevention apps work consistently as intended and do not overburden MSM with frequent updates. Relatedly, men in this sample were likely to delete an app from their phone that did not work properly, or they became bored with the app and sought another app that better met their needs. Many of these reasons for using or deleting apps from their phone are intuitive and emphasize the expectations that MSM hold about mobile technologies. For these reasons, and in addition to pilot testing to ensure the functions of an app are working correctly, formative work should ensure that men report the app as being easy to use. Along these lines, researchers may want to consider developing HIV prevention apps that include navigation functions (eg, drop-down menus, home navigation buttons) found in other frequently used apps to help with MSM's perceptions of ease of use. Such formative research may benefit from using the ISR framework described by Schnall et al [15] or the iterative, community-driven process described by Goldenberg et al [19].

As HIV testing continues to be an important avenue to prevent onward transmission of HIV among MSM, mobile apps aimed at increasing the frequency of HIV testing are needed. To best tailor mobile apps to MSM, we asked men in the focus groups to reflect on their preferences for the following features and functions of an HIV testing smartphone app: HIV test reminders, recommended testing intervals with dates, details about testing locations with an HIV test locator, and monitoring personal sexual behaviors. Overall, men especially appreciated the idea of receiving a reminder to test for HIV, particularly if it was tailored to them and helped reduce the burden of finding locations to test for HIV locally and elsewhere by capitalizing on and integrating with the GPS features found in smartphones. Many of the preferred features and functions of an HIV testing app described by men in this study reflect those in prior studies, especially personalizing the app and providing GPS-enabled HIV testing information [15,17,19]. However, participants' opinions varied about certain aspects of the proposed HIV testing app, particularly about the self-monitoring of their personal sexual behaviors and how they would receive updates from the app. The same ambivalence about monitoring personal

sexual encounters was expressed by men in at least one prior study [19]. These variations in opinions highlight the need to provide options within the HIV testing app that allow users to select delivery option(s) and how often they would like to be notified by the app. Additionally, the self-monitoring of one's sexual behaviors may aid in increasing HIV testing rates among MSM if this component is viewed as educational, rather than merely a log of their sexual activities. Behavioral self-monitoring aligns with the concept of self-regulation from social cognitive theory [23], which may be used as a theoretical guide to develop future HIV testing apps.

With respect to the development of mobile apps to encourage MSM to test for HIV at regular intervals, researchers should also consider how best to align individuals' risk behaviors for HIV with the current recommended HIV testing intervals of 3, 6, and 12 months. A variety of options exist for this alignment and will depend on the components of the specific HIV testing app. Moreover, newer options for HIV prevention now exist for MSM, which may positively help affect their frequency of HIV testing. The advent of oral preexposure prophylaxis (PrEP)—a regimen of antiretrovirals taken by those who are HIV-negative to prevent the acquisition of HIV—offers a promising means of prevention. Recent clinical trials have demonstrated the safety of PrEP [24-27] and significant efficacy for HIV prevention. To be most effective, the use of PrEP requires MSM to be tested for HIV every 3 months and to be seen by their medical provider; thus, uptake of PrEP would also help encourage MSM to test for HIV at regular intervals (eg, every 3 months).

Limitations

This study has important limitations to acknowledge. The results from these focus groups are not intended to be generalizable to all English-speaking, HIV-negative MSM who are smartphone owners in the United States. In addition, participants lived in 2 large urban cities and, therefore, app preferences by other samples of MSM in the Unites States, including those in rural locales and other locations in the South, may differ from this study's sample. Recall bias may have also affected MSM's ability to accurately identify features and functions that influenced them to download and use their smartphone app over time. Due to the qualitative, cross-sectional nature of the work conducted, causal inference is not possible; as such, longitudinal research methods should be used in future studies to more accurately illuminate the reasons that motivate MSM to download and use apps on their smartphones. Furthermore, future research should assess smartphone app preferences among a more geographically diverse sample of MSM to determine whether similar attitudes are expressed. These results are meant to be first steps in more fully understanding the needs of HIV-negative MSM with respect to smartphone apps to encourage regular HIV testing.

Conclusions

This study highlights findings obtained from formative work about the features and functions of smartphone apps that HIV-negative MSM consider when deciding to download mobile apps to their phone and continue to use or not use them over time. Our findings also reflect what components participants

would want in an app that promotes regular HIV testing. Themes captured from this study reiterate the importance of formative work to help enhance uptake and sustained use of smartphone apps for HIV prevention, including the promotion of regular testing for HIV. Given the complexity and potential barriers of encouraging HIV-negative MSM to test more regularly for HIV, future smartphone app interventions should also be guided by this and other formative research [15,17-19] and theories of behavior change [28].

Conflicts of Interest

None declared.

References

1. Centers for Disease Control and Prevention. HIV in the United States: at a glance. 2015. URL: http://www.cdc.gov/hiv/statistics/overview/ataglance.html [accessed 2016-09-16] [WebCite Cache ID 6kZo6TEEI]

2. Centers for Disease Control and Prevention. HIV surveillance report, 2014; vol 26. 2015 Nov. URL: http://www.cdc.gov/hiv/pdf/library/reports/surveillance/cdc-hiv-surveillance-report-us.pdf [accessed 2016-09-16] [WebCite Cache ID 6kZoujjOQ]

3. Centers for Disease Control and Prevention. HIV among gay and bisexual men. 2015. URL: http://www.cdc.gov/hiv/group/msm/ [accessed 2016-09-16] [WebCite Cache ID 6kZp3uaBB]

4. Branson BM, Handsfield HH, Lampe MA, Janssen RS, Taylor AW, Lyss SB, Centers for Disease ControlPrevention (CDC). Revised recommendations for HIV testing of adults, adolescents, and pregnant women in health-care settings. MMWR Recomm Rep 2006 Sep 22;55(RR-14):1-17; quiz CE1 [FREE Full text] [Medline: 16988643]

5. Centers for Disease Control and Prevention. HIV risk, prevention, and testing behaviors: National HIV Behavioral Surveillance System: men who have sex with men, 20 U.S. Cities, 2011. HIV Surveillance Special Report, Number 8. 2014. URL: http://www.cdc.gov/hiv/pdf/HSSR_8_NHBS_MSM_PDF-03.pdf [accessed 2016-09-16] [WebCite Cache ID 6kZpKggM3]

6. Oster AM, Johnson CH, Le BC, Balaji AB, Finlayson TJ, Lansky A, et al. Trends in HIV prevalence and HIV testing among young MSM: five United States cities, 1994-2011. AIDS Behav 2014 Apr;18 Suppl 3:S237-S247. [doi: 10.1007/s10461-013-0566-1] [Medline: 23955658]

7. Mitchell JW, Horvath KJ. Factors associated with regular HIV testing among a sample of US MSM with HIV-negative main partners. J Acquir Immune Defic Syndr 2013 Dec 1;64(4):417-423 [FREE Full text] [doi: 10.1097/QAI.0b013e3182a6c8d9] [Medline: 23933766]

8. Smith A, Page D. Washington, DC U.S. smartphone use in 2015. Pew Research Center. Washington, DC; 2015 Apr 01. URL: http://www.pewinternet.org/files/2015/03/PI_Smartphones_0401151.pdf [accessed 2016-09-16] [WebCite Cache ID 6kZpbBx3d]

9. Bolding G, Davis M, Sherr L, Hart G, Elford J. Use of gay Internet sites and views about online health promotion among men who have sex with men. AIDS Care 2004 Nov;16(8):993-1001. [doi: 10.1080/09540120412331292453] [Medline: 15511731]

10. Landovitz RJ, Tseng C, Weissman M, Haymer M, Mendenhall B, Rogers K, et al. Epidemiology, sexual risk behavior, and HIV prevention practices of men who have sex with men using GRINDR in Los Angeles, California. J Urban Health 2013 Aug;90(4):729-739 [FREE Full text] [doi: 10.1007/s11524-012-9766-7] [Medline: 22983721]

11. Anderson M. Technology device ownership: 2015. Pew Internet and American Life Project. Washington, DC: Pew Research Center; 2015 Oct 29. URL: http://www.pewinternet.org/files/2015/10/PI_2015-10-29_device-ownership_FINAL.pdf [WebCite Cache ID 6kZpkIYGj]

12. Simoni JM, Kutner BA, Horvath KJ. Opportunities and challenges of digital technology for HIV treatment and prevention. Curr HIV/AIDS Rep 2015 Sep 28;12(4):437-440. [doi: 10.1007/s11904-015-0289-1] [Medline: 26412082]

13. Schnall R, Travers J, Rojas M, Carballo-Diéguez A. eHealth interventions for HIV prevention in high-risk men who have sex with men: a systematic review. J Med Internet Res 2014;16(5):e134 [FREE Full text] [doi: 10.2196/jmir.3393] [Medline: 24862459]

14. Sullivan PS, Jones J, Kishore N, Stephenson R. The roles of technology in primary HIV prevention for men who have sex with men. Curr HIV/AIDS Rep 2015 Dec;12(4):481-488. [doi: 10.1007/s11904-015-0293-5] [Medline: 26519083]

15. 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]

16. Hevner A. A three cycle view of design science research. Scand J Inf Syst 2007;19(2):4.

17. Goldenberg T, McDougal SJ, Sullivan PS, Stekler JD, Stephenson R. Preferences for a mobile HIV prevention app for men who have sex with men. JMIR Mhealth Uhealth 2014;2(4):e47 [FREE Full text] [doi: 10.2196/mhealth.3745] [Medline: 25355249]

18. Aliabadi N, Carballo-Dieguez A, Bakken S, Rojas M, Brown W, Carry M, et al. Using the information-motivation-behavioral skills model to guide the development of an HIV prevention smartphone application for high-risk MSM. AIDS Educ Prev 2015 Dec;27(6):522-537. [doi: 10.1521/aeap.2015.27.6.522] [Medline: 26595265]

19. Goldenberg T, McDougal SJ, Sullivan PS, Stekler JD, Stephenson R. Building a mobile HIV prevention app for men who have sex with men: an iterative and community-driven process. JMIR Public Health Surveill 2015;1(2):e18 [FREE Full text] [doi: 10.2196/publichealth.4449] [Medline: 27227136]

20. Charmaz K. Constructing Grounded Theory. 2nd edition. Thousands Oak, CA: Sage Publications; 2014.

21. Saldaña J. The Coding Manual for Qualitative Researchers. Thousands Oak, CA: Sage Publications; 2013.

22. Glaser B, Strauss A. The Discovery of Grounded Theory: Strategies for Qualitative Research. New York, NY: Aldine De Gruyter; 1967.

23. Bandura A. Social cognitive theory of self-regulation. Organ Behav Hum Decis Processes 1991;50(2):248-287.

24. Grant RM, Lama JR, Anderson PL, McMahan V, Liu AY, Vargas L, et al. Preexposure chemoprophylaxis for HIV prevention in men who have sex with men. N Engl J Med 2010 Dec 30;363(27):2587-2599 [FREE Full text] [doi: 10.1056/NEJMoa1011205] [Medline: 21091279]

25. Grohskopf LA, Chillag KL, Gvetadze R, Liu AY, Thompson M, Mayer KH, et al. Randomized trial of clinical safety of daily oral tenofovir disoproxil fumarate among HIV-uninfected men who have sex with men in the United States. J Acquir Immune Defic Syndr 2013 Sep 1;64(1):79-86. [doi: 10.1097/QAI.0b013e31828ece33] [Medline: 23466649]

26. McCormack S, Dunn D. Pragmatic open-label randomised trial of preexposure prophylaxis: the PROUD study. Presented at: Conference on Retroviruses and Opportunistic Infections (CROI); 2015 February 23-26; Seattle, WA, USA.

27. Molina J, Capitant C, Spire B, Pialoux G, Chidiac C, Charreau I, ANRS Ipergay Study Group. On demand PrEP with oral TDF-FTC in MSM: results of the ANRS Ipergay trial. Presented at: Conference on Retroviruses and Opportunistic Infections (CROI); 2015 February 23-26; Seattle, WA, USA.

28. Tomlinson M, Rotheram-Borus MJ, Swartz L, Tsai AC. Scaling up mHealth: where is the evidence? PLoS Med 2013;10(2):e1001382 [FREE Full text] [doi: 10.1371/journal.pmed.1001382] [Medline: 23424286]

Abbreviations

GPS: global positioning system
HIV: human immunodeficiency virus
ISR: Information Systems Research
MSM: men who have sex with men
PrEP: preexposure prophylaxis

Original Paper

Breadth of Coverage, Ease of Use, and Quality of Mobile Point-of-Care Tool Information Summaries: An Evaluation

Emily Johnson[1*], MLIS; Vamsi K Emani[2*], MBBS, MD; Jinma Ren[3], PhD

[1]University of Illinois at Chicago Library, Library of the Health Sciences - Peoria, University of Illinois at Chicago, Peoria, IL, United States

[2]Department of Internal Medicine, University of Illinois College of Medicine at Peoria, Peoria, IL, United States

[3]Center for Outcomes Research, Department of Medicine, University of Illinois College of Medicine at Peoria, Peoria, IL, United States

[*]these authors contributed equally

Corresponding Author:
Emily Johnson, MLIS
University of Illinois at Chicago Library
Library of the Health Sciences - Peoria
University of Illinois at Chicago
One Illini Dr.
PO Box 1649
Peoria, IL,
United States
Phone: 1 309 671 8491
Fax: 1 312 996 9584
Email: emj11@uic.edu

Abstract

Background: With advances in mobile technology, accessibility of clinical resources at the point of care has increased.

Objective: The objective of this research was to identify if six selected mobile point-of-care tools meet the needs of clinicians in internal medicine. Point-of-care tools were evaluated for breadth of coverage, ease of use, and quality.

Methods: Six point-of-care tools were evaluated utilizing four different devices (two smartphones and two tablets). Breadth of coverage was measured using select *International Classification of Diseases, Ninth Revision*, codes if information on summary, etiology, pathophysiology, clinical manifestations, diagnosis, treatment, and prognosis was provided. Quality measures included treatment and diagnostic inline references and individual and application time stamping. Ease of use covered search within topic, table of contents, scrolling, affordance, connectivity, and personal accounts. Analysis of variance based on the rank of score was used.

Results: Breadth of coverage was similar among Medscape (mean 6.88), Uptodate (mean 6.51), DynaMedPlus (mean 6.46), and EvidencePlus (mean 6.41) ($P>.05$) with DynaMed (mean 5.53) and Epocrates (mean 6.12) scoring significantly lower ($P<.05$). Ease of use had DynaMedPlus with the highest score, and EvidencePlus was lowest (6.0 vs 4.0, respectively, $P<.05$). For quality, reviewers rated the same score (4.00) for all tools except for Medscape, which was rated lower ($P<.05$).

Conclusions: For breadth of coverage, most point-of-care tools were similar with the exception of DynaMed. For ease of use, only UpToDate and DynaMedPlus allow for search within a topic. All point-of-care tools have remote access with the exception of UpToDate and Essential Evidence Plus. All tools except Medscape covered criteria for quality evaluation. Overall, there was no significant difference between the point-of-care tools with regard to coverage on common topics used by internal medicine clinicians. Selection of point-of-care tools is highly dependent on individual preference based on ease of use and cost of the application.

XSL•FO
RenderX

Introduction

State of Mobile Use in Public and Health Care

For the last decade, a surge of technology has had an increasing influence on human life. Over two-thirds of American adults now own a smartphone of some kind and mobile now represents almost two out of three digital media minutes time [1,2]. The field of medicine is not an exception in this adoption of new technology, changing the practice from electronic health records to robotic surgeries.

A tool that is becoming an integral part of medical practice is the mobile device, a telecommunication device running an operating system with the ability to customize by accessing apps, word processing, and Internet connectivity. Common terms for the mobile device include smartphone and tablet, frequently adopting the branded name of the device. Use of smartphones by physicians has plateaued to 80% to 85%, and adoption of tablets has risen to 76% in 2014, according the group Manhattan Research [3]. In 2014, another survey reported over half of US hospitals use mobile devices in their facilities either by offering device programs in their institution or using a bring-your-own-device model [4].

Through the prevalent use of the mobile device, clinical professions in the health sciences now have immediate access to essential information for patient treatment, which has shown improvement in decision making and reduced medical errors, improved communication between medical staff, and enhanced telemedicine capability [5,6].

Definition of a Point-of-Care Tool

In the everyday practice of medicine, physicians are increasingly relying on access to electronic medical resources to make clinical decisions, replacing the condensed pocket reference texts. Physicians often utilize online point-of-care tools, a resource with immediate access to filtered summaries providing current recommendations for the management of medical problems [7,8]. These resources are available globally and many are accessible via a personal subscription or an institutional license on a desktop or laptop computer.

With the emergence of mobile devices, the market of mobile health applications has grown to over 100,000 apps in the two leading operating systems, iOS and Android [9]. Point-of-care tools were no exception to this expansion of this market, including many entering the marketplace since 2008. Medical reference and diagnostic medical apps make up about 18% of the mobile health (mHealth) app market share [9]. Point-of-care tools are increasingly accessible via mobile native apps, Web apps, or both.

Evaluation of Point-of-Care Tools on Mobile Devices

A literature review was conducted to identify articles evaluating mobile point-of-care tools with a focus on breadth of coverage, quality of evidence, and mobile ease of use design [8,10-18]. Studies were identified evaluating the quality and breadth of content of the full online Web versions of these point-of-care tools. Prorok et al and Banzi et al both evaluated the quality and breadth of coverage, determining no single point-of-care tool was ideal and clinicians should not rely on one single product [8,12].

The evaluation of mobile medical apps is in its infancy, with much of the evidence stemming from evaluation from handheld personal digital assistant devices [6,19-23]. Ease of use and mobile interface design evaluation criteria are often too general, complex, or specific for health-related mobile resources [22]. The Healthcare Information and Management Systems Society (HIMSS) has compiled guidelines to evaluate mHealth, including criteria for efficiency and platform optimization, but did not include methods of evaluation of information [24]. Stoyanov et al published the Mobile App Rating Scale for consumer health applications during the data collection of this study and covers similar criteria regarding ease of use on navigation and gestural design but did not have the detailed criteria for rating clinical point-of-care tools [22].

The Goal of This Research Study

The research has, to date, focused on evaluating the breadth of content and quality measures and timeliness of the online versions of point-of-care tools and the user experience within mobile application design; no study has reviewed multiple mobile point-of-care tools and devices using the same evaluation criteria. With increased usage of mobile technology, there is a need to evaluate measures of quality and breadth of content of mobile versions of point-of-care tools. In addition, this research study will include an evaluation of ease of use which may affect the way these point-of-care tool appls are used and could impact patient care.

Objective

The objective is to evaluate if six selected mobile point-of-care tools meet the defined criteria for breadth of coverage, ease of use, and quality on the different mobile devices and operating systems.

Methods

Selection of Point-of-Care Tools

Point-of-care tools were created for use by health care providers and focused on providing answers to clinical questions. Selection for evaluation used the following criteria: English language only, availability in the United States, and accessibility via a mobile device using either a website or an app. If the point-of-care tool was an app, it needed to be accessible on both iOS and Android devices, and it needed to identify as an evidence-based summary product with topics of internal medicine.

The researchers reviewed past evaluations of point-of-care tools within the literature [8,10-18] and reviewed the app store categories *medical* and *medicine* for Android and iOS devices; if there was no app, researchers verified there was a mobile-enabled website. Based on the defined criteria, the researchers selected six point-of-care tools: DynaMed, DynaMedPlus, Epocrates, Essential Evidence Plus, Medscape, and UpToDate. While DynaMed and DynaMedPlus are created by the same vendor, they provided unique content summaries

and user interfaces, so both tools were included in the evaluation.

Accessibility of these resources was either provided by free access (Medscape), paid individual subscription (Epocrates and DynaMedPlus), or made available with a paid institutional subscription from the University of Illinois at Chicago University Library (DynaMed, Essential Evidence Plus, and UptoDate) for complete access of the content.

Selection of Devices

Four different mobile devices were used in the analysis: two smartphones, the iPhone 5S and Moto X second generation, and two tablets, the iPad Mini 2 and Samsung Galaxy tablet. One investigator utilized the iOS operating system on the iPhone and iPad and the other investigator worked with the Android operating system on the Moto X and Samsung tablet. The selections of these devices were based on availability to the researchers and popularity in the medical field [9].

Interrater Reliability

Prior to the full assessment of the criteria, the interrater reliability was examined in a separate dataset, in which two reviewers rated four point-of-care tools, (ACP SmartMedicine, DynaMed, Essential Evidence Plus, and UptoDate) for three *International Classification of Diseases, Ninth Revision (ICD-9)*, diagnosis codes for breadth of coverage, ease of use, and quality on four mobile devices independently.

Breadth of Coverage Assessment

To evaluate the breadth of coverage, ICD-9 billing codes were culled from the setting of a teaching medical center with 609 beds and a level I trauma center in the state of Illinois. A total of 30 codes for the most billed from January to July 2015 were selected from the inpatient and outpatient settings. From the 30 codes, specialties other than general internal medicine were excluded and any duplicated codes utilized both in inpatient and outpatient treatment were removed, bringing the count to 17 total codes used in the evaluation (Multimedia Appendix 1).

The definition of breath of coverage was modified from prior studies [8,12]; each medical topic needed to include information consisting of a summary of the topic, etiology, pathophysiology, clinical manifestations, diagnosis, treatment, and prognosis (Multimedia Appendix 2). These topics cover the most consistently needed information for a patient encounter. If any of the information was incomplete or missing from the medical topic, it was deemed to be uncovered [8,12].

Two researchers independently entered text searches of the ICD-9 diagnosis code into the search engine of each point-of-care tool on the four devices. If searches did not return relevant results, synonyms for the code were utilized. When content was retrieved, the researcher reviewed the content to meet the breadth of coverage topics. Any disagreements on search terms and content topics was resolved by consensus.

Quality Assessment

Methodology to assess quality of point-of-care tools was adapted from a previous study by Prorok et al8 and modified for a mobile point-of-care tool investigation. Quality measures included

inline references for treatment and inline references for diagnostic recommendations (Multimedia Appendix 3). Time stamping of individual content and the app platform were included in the quality measures to evaluate the currency of the content and application. The app version number was included within the time stamping of the app platform criteria. The app's update logs were able to verify the date of that version of the app. The two researchers independently verified the quality measures in all of the point-of-care tools on the four devices, and any disagreements were resolved by consensus.

The other editorial quality measures set by Prorok, such as reviewing policies on finding new evidence, grading recommendations, and updating materials, were surveyed in the initial screening of the mobile point-of-care tools and were consistent with the previous evaluation. The investigators focused on the quality measures that could have presented differently on the mobile platform.

Ease of Use Assessment

The scale defined to report on ease of use of the point-of-care tools on a mobile device was developed utilizing design aspects by HIMSS and Nielsen Norman Group user experience criteria [24,25]. The ease of use for each point-of-care tool application was based on the following defined factors:

- Search within topic: user can use a search feature within a topic to find information.
- Table of contents: topic contains appropriate table of contents for easy navigation.
- Scrolling: appropriate scrolling patterns were utilized within the topic for mobile applications.
- Affordance: it is made clear what items can be selected, tapped, or swiped to connect with other content.
- Connectivity: content is available via Wi-Fi or mobile data connection.
- Personal account: required personal account is easy to log in and modify settings for the point-of-care tool.

The two researchers independently verified the ease of use measures in all of the point-of-care tools on the four devices, and any disagreements were resolved by consensus.

The final version of the assessment criteria, definitions, and rating system are available with this paper (see Multimedia Appendix 4).

Statistical Analysis

Statistical analysis was performed by SAS version 9.4 (SAS Institute Inc). Mean and standard deviation as well as histogram were applied to depict the scores of point-of-care tools. The mean for breadth of coverage was calculated using the average total score of each point-of-care tool. The means for ease of use and the quality measurement was the average total score for each mobile device. A statistical significance level of .05 was set for all hypothesis tests.

The interrater reliability was statistically calculated by using both Pearson correlation coefficient and kappa coefficient for each domain (breadth, ease of use, and quality) between two reviewers in order to determine how consistent their ratings were prior to the full dataset assessment.

XSL•FO
RenderX

Later, in the whole dataset, the researchers generated new rank scores based on the rating scores for each domain because the distribution of rating scores was extremely skewed. A general linear regression model was used to compare the rank scores among the six point-of-care tools, controlling for type of mobile devices and diagnosis codes. Bonferroni-adjusted P values were provided when multiple comparisons were performed (See Multimedia Appendix 5 for calculations). For clearer understanding of the results, a multiple comparison model was used. If point-of-care tools are grouped under the same letter, it indicates that the differences are not statistically significant. If the point-of-care tools are not grouped, the comparisons for that criteria are statistically significant.

Results

Main Findings

Each of the six selected point-of-care tools was evaluated on the breadth of coverage, quality of evidence, and ease of use on the four mobile devices. After discussion and calibration of the survey instruments through an interrater review, data were collected and evaluated between May and September 2015.

Interrater Reliability

The two reviewers had high interrater reliability overall. The correlation coefficients were $r=.97$ ($P<.001$), $r=.88$ ($P<.001$), and $r=.41$ ($P=.07$) in the domains of breadth, ease of use, and quality, respectively. The kappa coefficients were .84 (95% CI 0.77-0.91), .76 (95% CI 0.56-0.96), and .29 (95% CI 0.02-0.55) in the domains of breath, ease of use, and quality, respectively.

Breadth of Coverage Measure

On all devices, the breadth of coverage of selected ICD-9 codes by each point-of-care tool was scored (n=68). The result represents a mean of the scores for each tool on all the devices. The mean is average total score of breath of coverage of each tool.

The mean scores of breadth of coverage were very close among Medscape (6.88), DynaMedPlus (6.47), and Essential Evidence Plus (6.41) with P values of >.99 (Medscape vs DynaMedPlus), 0.259 (Medscape vs Essential Evidence Plus), and >.99 (DynaMedPlus vs Essential Evidence Plus). As shown in Table 1, DynaMed had lowest score of breadth among the tools ($P<.01$), and Epocrates also had a relative lower score of breadth compared to the top four tools ($P<.01$).

The researchers did not find any statistical difference in the total score attained by each point-of-care tool on different electronic devices with different operating systems. There was also no statistical difference noted between smartphone application and tablet application for the same point-of-care tool.

Ease of Use Measure

Ease of use was scored for all point-of-care tools. Each point-of-care tool was reviewed on both a smartphone and tablet of the different operating systems, as we only could get one score from each type of electronic device (n=4). The mean is the average total score for each device. The reviewer would then give a score of 0 or 1 to the different components of ease of use. The components are described in the Methods section.

As Table 2 depicts, DynaMedPlus had the highest score of ease of use (6.00), whereas Essential Evidence Plus was lowest (4.00), $P<.01$. The remaining four tools had medium scores with no significant difference (all $P>.99$).

Quality Measure

Quality measures were scored for all point-of-care tools separately on each device, as we only could get one score from each type of electronic device (n=4). The mean is the average total score for each device. All point-of-care tools scored the same mean, with the exception of Medscape, which entirely lacked one of the components of the quality measure.

The reviewers rated same scores of quality (4.00) for DynaMed, DynaMedPlus, Epocrates, Essential Evidence Plus, and UpToDate (Table 3). Medscape had lower score of quality than others ($P<.01$).

Table 1. Breadth of coverage among six point-of-care tools.

Tool	n[a]	Mean	SD	LS-mean[b]	Multiple comparison	
Medscape	68	6.88	0.32	277.65	A	
DynaMedPlus	68	6.47	1.66	260.35	A	
Essential Evidence Plus	68	6.41	1.66	249.18	A	B
UpToDate	68	6.51	0.72	220.28		B
Epocrates	68	6.12	0.32	132.35		
DynaMed	68	5.53	1.26	87.19		

[a]Total number of diagnoses reviewed on all devices (17 diagnoses times 4 devices).

[b]Least square mean of rank score.

Table 2. Ease of use among six point-of-care tools.

Tool	n[a]	Mean	SD	LS-mean[b]	Multiple comparison
DynaMedPlus	4	6.00	0.00	22.50	
DynaMed	4	5.00	0.00	13.00	A
Epocrates	4	5.00	0.00	13.00	A
UpToDate	4	5.00	0.00	13.00	A
Medscape	4	4.75	0.50	10.50	A
Essential Evidence Plus	4	4.00	0.00	3.00	

[a]Total number of devices.

[b]Least square mean of rank score.

Table 3. Quality among six point-of-care tools.

Tool	n[a]	Mean	SD	LS-mean[b]	Multiple comparison
DynaMed	4	4.00	0.00	14.50	A
DynaMedPlus	4	4.00	0.00	14.50	A
Epocrates	4	4.00	0.00	14.50	A
Essential Evidence Plus	4	4.00	0.00	14.50	A
UpToDate	4	4.00	0.00	14.50	A
Medscape	4	3.00	0.00	2.50	

[a]Total number of devices.

[b]Least square mean of rank score.

Discussion

Principal Findings

While studies exist that analyzed online point-of-care tools, this is the first study assessing access to the point-of-care tools on smartphones and tablet devices. As previous studies have shown, many online point-of-care tools are lacking in the breadth of coverage of medical topics, interface ease of use, and the quality of the contained contents [8,10-18]. The results from this study have been consistent with these findings, suggesting that medical personnel should not rely entirely on any single point-of-care tool.

The breadth of coverage analysis was designed to review the presented information within the point-of-care tool topic. Overall, there was no significant difference between the point-of-care tools for the criteria used to evaluate the breadth of coverage on common internal medicine topics because most point-of-care tools scored within the same range, with the exception of DynaMed. Clinicians usually refer to the point-of-care tools when seeking information on a clinical question and, while DynaMed provides excellent evidence-based information, it does not help with reasoning and arriving at an answer for the clinical question. The publishers may have realized this, addressing the issue with new product DynaMedPlus in 2015. This new product scored higher within the assessment of breadth of coverage, providing methods for reasoning through the care of a patient more than DynaMed.

The evaluation of ease of use showed the greatest variation in results among these tools. These criteria address the needs of the clinician's experience, seeing how well these tools would be adapted into their workflow. All point-of-care tools are using appropriate methods like scrolling and using appropriate cues for linked items for easier mobile usage of the tools. The criteria of searching within a summary topic, as provided by UpToDate and DynaMedPlus, would be extremely useful to help clinicians easily locate what they are looking for during patient care. This feature would have been extremely helpful with Epocrates because it was more challenging to find an answer for each of the criteria defined in the breadth of coverage; there were many tabs to navigate to find the needed information.

Another component reviewed was the need for Internet connectivity to access the point-of-care tool content. This is an important criterion to consider as there are some clinicians with limited access to Wi-Fi or mobile data in their practice making it challenging to access care information at the time of need. Essential Evidence Plus, Epocrates, and UpToDate require a mobile data or Wi-Fi connection to access content, while Medscape, DynaMed, and DynaMedPlus download the summary topics to the device, needing data connection only for updates to the content. UpToDate offers an option for downloading its content on mobile devices with individual subscriptions for an additional charge.

All point-of-care tools also have the feature of creating a personal account; this allows for saving the user ID and password. All of the evaluated point-of-care tools have remote access through personal accounts and proxy institutional licenses, which is very convenient and beneficial.

XSL•FO

RenderX

The majority of the point-of-care tools were available using a native app, composed of pieces of software completely written in the native language of the mobile device platform such as Objective C for iOS and Java and C++ for Android [26]. The one exception in this study was Essential Evidence Plus, which uses a Web app modified for easy access on a mobile Web browser. The choice of development is dependent on a range of considerations when developing content for mobile devices [27]. The native apps integrate device-specific features (eg, GPS, camera) and provide off-line functioning but can be costly and time consuming to develop. The Web app allows for platform independence with use on any device by use of the Web programing language of HTML5 but can be limited in its features and speed. Overall, mobile point-of-care tool developers platform selection will be determined by optimizing the best user experience in the mobile arena.

All tools except Medscape cover all quality measure components. The missing component in Medscape was that it does not provide a date stamping of individual topic content at the time of the conducted research. While Medscape may have current content, the lack of date labeling of the content review or update can lead to questions of the information's currency in a clinical application.

The researchers' initial hypothesis was to not only test the point-of-care tools but to evaluate if there was any difference between the different devices. This study showed there was no difference based on the defined criteria between the mobile devices or operating systems. The use of a particular mobile device is highly subjective, based on user preferences of size, operating system, screen quality, cost, etc.

Cost

Cost of the point-of-care tools will also influence the selection of the tools. The selected tools studied have differing price ranges for individual licenses depending on the training status of the purchaser. Among the six point-of-care tools, only Medscape is available at no charge.

As of December 2015, the following prices for the remaining five point-of-care tools were available via the product websites. DynaMed or DynaMedPlus costs US $99.95 per year for students, US $149.95 per year for residents, and US $395 per year for physicians. DynaMedPlus is freely available to members of the American College of Physicians at the time of this study. Essential Evidence Plus costs around US $85 per year, not distinguishing between the training status of the individual subscriber. Epocrates provides its drug content package for free; to have access to disease information and clinical practice guidelines, it costs US $174.99 per year for physicians and residents. UpToDate costs US $199 per year for students and residents and US $499 per year for physician. For use of the UpToDate mobile app, there is an additional charge of US $49 per year for mobile access of its content, and this access is restricted to two devices. With above-mentioned costs, all tools give full access to their contents.

Access to pricing information for institutional licensing was not available for inclusion in this study due to nondisclosure agreements.

Limitations

Although the researchers attempted to build on evaluation tools used in the past, this study still has several limitations. The researchers were subject to the rapidly changing and inconsistent design environment of point-of-care tools. The creators/editors are constantly offering updates, so summary content and features may have been added after the dates of review.

Only six point-of-care tools were evaluated in the study; the researchers were limited in what was available via institutional licenses or individual subscription in the United States and what was accessible fitting the point-of-care tool selection criteria. This eliminated several of the previously evaluated point-of-care tools from this study, including BMJ Best Practices and PEPID.

During the interrater reliability measurement, there was a measured difference between the researchers due to a disagreement on the criteria defined for the "date of stamping for application platform." This issue was reflected in the evaluation of the point-of-care tool ACP SmartMedicine, which was discontinued prior to the time of data collection for the full assessment. The definition was reviewed and a consensus was reached before the data collection for the full assessment.

ICD-9 codes selection focused on the specialty of medicine; it did not represent the broad spectrum of care thus limiting the assessment of comprehensiveness within the inpatient/outpatient settings. The breadth of coverage evaluation was impacted by the knowledge of the investigators to translate the given code to a synonym or related term to indicate the concept was covered within the point-of-care tool. The scoring system used was binary for all three areas of evaluation. Either the point-of-care tool contained the sought criteria or not; there was no gradient built within the assessment tool.

The investigators chose not to rank the point-of-care tools, as there is no scientific way to give appropriate weight to each of the components. Ranking would allow for personal bias to impact the weighting scale by considering one aspect of the evaluation criteria over another due to subjectivity.

Conclusions

In evaluating the breadth of coverage, quality of information, and ease of use of six mobile point-of-care tools, the investigators were able to determine there is no significant difference between the point-of-care tools with regards to coverage on common topics used by clinicians within the discipline of medicine. The selection of a mobile point-of-care tool will likely depend on individual preference based on ease of use and cost of the application. For institutions subscribing to point-of-care tools via institutional licensing, it is important to gain the individual users' perspective on selection of mobile point-of-care tools, because it enables choice of adoption by the main users of the product.

Acknowledgments

Conflicts of Interest

None declared.

Multimedia Appendix 1

ICD-9 codes used to study breadth of coverage of point-of-care tools.

[PDF File (Adobe PDF File), 36KB - mhealth_v4i4e117_app1.pdf]

Multimedia Appendix 2

Breadth of coverage criteria and definitions.

[PDF File (Adobe PDF File), 23KB - mhealth_v4i4e117_app2.pdf]

Multimedia Appendix 3

[PDF File (Adobe PDF File), 21KB - mhealth_v4i4e117_app3.pdf]

Multimedia Appendix 4

Quality measures and definitions.

[PDF File (Adobe PDF File), 23KB - mhealth_v4i4e117_app4.pdf]

Multimedia Appendix 5

Ease of use factors and definitions.

[PDF File (Adobe PDF File), 23KB - mhealth_v4i4e117_app5.pdf]

References

1. Smith A, Page D. U.S. smartphone use in 2015. Washington, DC: Pew Internet and American Life Project; 2015. URL: http://www.pewinternet.org/files/2015/03/PI_Smartphones_0401151.pdf [WebCite Cache ID 6hHvxd14m]
2. comScore. 2015. The 2015 U.S. Mobile App Report URL: http://www.comscore.com/Insights/Presentations-and-Whitepapers/2015/The-2015-US-Mobile-App-Report [accessed 2016-05-05] [WebCite Cache ID 6hHvxd15h]
3. Dolan B. mobihealthnews. In-Depth: Mobile Adoption Among US Physicians--2014 URL: http://mobihealthnews.com/32232/in-depth-mobile-adoption-among-us-physicians [accessed 2016-07-23] [WebCite Cache ID 6jCozRdbK]
4. 3rd Annual HIMSS Analytics Mobile Survey. Healthcare Information and Management Systems Society. 2014. URL: http://www.himssanalytics.org/research/essentials-brief-mobile-devices-study [accessed 2016-05-05] [WebCite Cache ID 6hHz1sIGO]
5. Franko OI, Tirrell TF. Smartphone app use among medical providers in ACGME training programs. J Med Syst 2012 Oct;36(5):3135-3139. [doi: 10.1007/s10916-011-9798-7] [Medline: 22052129]
6. Mickan S, Tilson JK, Atherton H, Roberts NW, Heneghan C. Evidence of effectiveness of health care professionals using handheld computers: a scoping review of systematic reviews. J Med Internet Res 2013;15(10):e212 [FREE Full text] [doi: 10.2196/jmir.2530] [Medline: 24165786]
7. HLWIKI International. 2015. Point-of-Care Tools in Medicine URL: http://hlwiki.slais.ubc.ca/index.php/Point-of-care_tools_in_medicine [accessed 2016-05-05] [WebCite Cache ID 6hHvxd156]
8. Prorok JC, Iserman EC, Wilczynski NL, Haynes RB. The quality, breadth, and timeliness of content updating vary substantially for 10 online medical texts: an analytic survey. J Clin Epidemiol 2012 Dec;65(12):1289-1295. [doi: 10.1016/j.jclinepi.2012.05.003] [Medline: 22974495]
9. mHealth App Developer Economics 2014. research2guidance. 2014. URL: http://mhealtheconomics.com/ [accessed 2016-05-05] [WebCite Cache ID 6hHvxd15q]

XSL•FO
RenderX

10. Ahmadi S, Faghankhani M, Javanbakht A, Akbarshahi M, Mirghorbani M, Safarnejad B, et al. A comparison of answer retrieval through four evidence-based textbooks (ACP PIER, Essential Evidence Plus, First Consult, and UpToDate): a randomized controlled trial. Med Teach 2011;33(9):724-730. [doi: 10.3109/0142159X.2010.531155] [Medline: 21854150]

11. Banzi R, Liberati A, Moschetti I, Tagliabue L, Moja L. A review of online evidence-based practice point-of-care information summary providers. J Med Internet Res 2010;12(3):e26 [FREE Full text] [doi: 10.2196/jmir.1288] [Medline: 20610379]

12. Banzi R, Cinquini M, Liberati A, Moschetti I, Pecoraro V, Tagliabue L, et al. Speed of updating online evidence based point of care summaries: prospective cohort analysis. BMJ 2011 Sep 23;343(sep22 2):d5856-d5856 [FREE Full text] [doi: 10.1136/bmj.d5856] [Medline: 21948588]

13. Campbell R, Ash J. Comparing bedside information tools: a user-centered, task-oriented approach. AMIA Annu Symp Proc 2005:101-105 [FREE Full text] [Medline: 16779010]

14. Campbell R, Ash J. An evaluation of five bedside information products using a user-centered, task-oriented approach. J Med Libr Assoc 2006 Oct;94(4):435-441 [FREE Full text] [Medline: 17082836]

15. Clauson KA, Marsh WA, Polen HH, Seamon MJ, Ortiz BI. Clinical decision support tools: analysis of online drug information databases. BMC Med Inform Decis Mak 2007 Mar 08;7(7):7 [FREE Full text] [doi: 10.1186/1472-6947-7-7] [Medline: 17346336]

16. Jeffery R, Navarro T, Lokker C, Haynes RB, Wilczynski NL, Farjou G. How current are leading evidence-based medical textbooks? An analytic survey of four online textbooks. J Med Internet Res 2012 Dec 10;14(6):e175 [FREE Full text] [doi: 10.2196/jmir.2105] [Medline: 23220465]

17. Ketchum AM, Saleh AA, Jeong K. Type of evidence behind point-of-care clinical information products: a bibliometric analysis. J Med Internet Res 2011 Feb 18;13(1):e21. [doi: 10.2196/jmir.1539] [Medline: 21335319]

18. Shurtz S, Foster MJ. Developing and using a rubric for evaluating evidence-based medicine point-of-care tools. J Med Libr Assoc 2011 Jul;99(3):247-254 [FREE Full text] [doi: 10.3163/1536-5050.99.3.012] [Medline: 21753917]

19. Bierbrier R, Lo V, Wu RC. Evaluation of the accuracy of smartphone medical calculation apps. J Med Internet Res 2014 Feb 03;16(2):e32 [FREE Full text] [doi: 10.2196/jmir.3062] [Medline: 24491911]

20. Goldbach H, Chang AY, Kyer A, Ketshogileng D, Taylor L, Chandra A, et al. Evaluation of generic medical information accessed via mobile phones at the point of care in resource-limited settings. J Am Med Inform Assoc 2014;21(1):37-42 [FREE Full text] [doi: 10.1136/amiajnl-2012-001276] [Medline: 23535665]

21. Hauser SE, Demner-Fushman D, Jacobs JL, Humphrey SM, Ford G, Thoma GR. Using wireless handheld computers to seek information at the point of care: an evaluation by clinicians. J Am Med Inform Assoc 2007;14(6):807-815 [FREE Full text] [doi: 10.1197/jamia.M2424] [Medline: 17712085]

22. 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]

23. Su WC. A preliminary survey of knowledge discovery on smartphone applications (apps): Principles, techniques and research directions for e-health. Presented at: International Conference on Complex Medical Engineering; June 26-29, 2014; Taipei, Taiwan.

24. mHIMSS App Usability Work Group. Healthcare Information and Management Systems Society. 2012. Selecting a Mobile Appvaluating the Usability of Medical Applications URL: http://www.himss.org/selecting-mobile-app-evaluating-usability-medical-applications-0 [WebCite Cache ID 6kzQ8eXCo]

25. Nielson Norman Group. 2011. Optimizing a Screen for Mobile Use URL: https://www.nngroup.com/articles/optimizing-a-screen-for-mobile-use/ [accessed 2016-05-05] [WebCite Cache ID 6hHvxd15P]

26. Banga C, Weinhold J. Finding the right design flow. In: Essential Mobile Interaction Design: Perfecting Interface Design in Mobile Apps. Upper Saddle River, NJ: Addison-Wesley; 2014:81-102.

27. Budiu R. 2013. Mobile: Native Apps, Web Apps, and Hybrid Apps. Nielson Normal Group. URL: https://www.nngroup.com/articles/mobile-native-apps [accessed 2016-05-05] [WebCite Cache ID 6hHvxd15X]

Abbreviations

HIMSS: Healthcare Information and Management Systems Society
ICD-9: International Classification of Diseases, Ninth Revision
mHealth: mobile health

XSL•FO
RenderX

XSL•FO
RenderX

Original Paper

Patient-Facing Mobile Apps to Treat High-Need, High-Cost Populations: A Scoping Review

Karandeep Singh[1], MD, MMSc; Kaitlin Drouin[2], MS, MA; Lisa P Newmark[3], BA; Malina Filkins[4], BA; Elizabeth Silvers[3], BA; Paul A Bain[5], MS, PhD; Donna M Zulman[6,7], MD, MS; Jae-Ho Lee[8], MD, PhD; Ronen Rozenblum[5,9], MPH, PhD; Erika Pabo[5,9], MD, MBA; Adam Landman[5,10], MD, MS, MIS, MHS; Elissa V Klinger[9], SM; David W Bates[5,9,11], MD, MSc

[1]Departments of Learning Health Sciences and Internal Medicine, University of Michigan Medical School, Ann Arbor, MI, United States

[2]Department of Pediatric Newborn Medicine, Brigham and Women's Hospital, Boston, MA, United States

[3]Information Systems, Partners HealthCare System, Wellesley, MA, United States

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

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

[6]Division of General Medical Disciplines, Stanford University School of Medicine, Stanford, CA, United States

[7]Center for Innovation to Implementation, VA Palo Alto Health Care System, Menlo Park, CA, United States

[8]Department of Emergency Medicine, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Republic of Korea

[9]Division of General Internal Medicine, Brigham and Women's Hospital, Boston, MA, United States

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

[11]Department of Health Policy and Management, Harvard TH Chan School of Public Health, Boston, MA, United States

Corresponding Author:
Karandeep Singh, MD, MMSc
Departments of Learning Health Sciences and Internal Medicine
University of Michigan Medical School
1161H NIB
300 N Ingalls St
Ann Arbor, MI, 48109-5403
United States
Phone: 1 734 936 1649
Fax: 1 734 647 3914
Email: kdpsingh@umich.edu

Abstract

Background: Self-management is essential to caring for high-need, high-cost (HNHC) populations. Advances in mobile phone technology coupled with increased availability and adoption of health-focused mobile apps have made self-management more achievable, but the extent and quality of the literature supporting their use is not well defined.

Objective: The purpose of this review was to assess the breadth, quality, bias, and types of outcomes measured in the literature supporting the use of apps targeting HNHC populations.

Methods: Data sources included articles in PubMed and MEDLINE (National Center for Biotechnology Information), EMBASE (Elsevier), the Cochrane Central Register of Controlled Trials (EBSCO), Web of Science (Thomson Reuters), and the NTIS (National Technical Information Service) Bibliographic Database (EBSCO) published since 2008. We selected studies involving use of patient-facing iOS or Android mobile health apps. Extraction was performed by 1 reviewer; 40 randomly selected articles were evaluated by 2 reviewers to assess agreement.

Results: Our final analysis included 175 studies. The populations most commonly targeted by apps included patients with obesity, physical handicaps, diabetes, older age, and dementia. Only 30.3% (53/175) of the apps studied in the reviewed literature were identifiable and available to the public through app stores. Many of the studies were cross-sectional analyses (42.9%, 75/175), small (median number of participants=31, interquartile range 11.0-207.2, maximum 11,690), or performed by an app's developers (61.1%, 107/175). Of the 175 studies, only 36 (20.6%, 36/175) studies evaluated a clinical outcome.

XSL•FO

RenderX

Conclusions: Most apps described in the literature could not be located on the iOS or Android app stores, and existing research does not robustly evaluate the potential of mobile apps. Whereas apps may be useful in patients with chronic conditions, data do not support this yet. Although we had 2-3 reviewers to screen and assess abstract eligibility, only 1 reviewer abstracted the data. This is one limitation of our study. With respect to the 40 articles (22.9%, 40/175) that were assigned to 2 reviewers (of which 3 articles were excluded), inter-rater agreement was significant on the majority of items (17 of 30) but fair-to-moderate on others.

Introduction

Caring for high-need, high-cost (HNHC) populations represents a complex problem because these individuals often suffer from multiple chronic conditions, functional limitations, behavioral health problems, socioeconomic challenges, and inadequate coordination of care [1,2]. Nearly half of all US adults suffer from a chronic illness and this group accounts for a large share of health care costs [3]. Advances in mobile phone technology coupled with increased availability and adoption of mobile health apps have changed the landscape of self-management [4]. Data increasingly support the role of patient-facing health information technology tools in improving patient-centered care outcomes, health services efficiency, and health outcomes [5-7]. Community health centers and clinics that care for vulnerable populations overwhelmingly perceive mobile health technologies as an ideal tool to engage their patient populations in chronic disease management [8].

Although more than 165,000 mobile health apps are available on the iTunes (iOS) and Google Play (Android) app stores in the Unites States [9] and billions of dollars are being invested in digital health [10], it is not clear how many of these apps focus on patients with chronic conditions and how well the scientific evidence supports their effectiveness. Prior reviews of the literature evaluating the use of patient-facing health apps have been limited by a narrow scope. Reviews have focused on a single medical condition [11-13], on a single aspect of a broad group of apps (such as identifying target populations, behavioral functionalities, privacy policies, and expert involvement) [14-17], or have included only clinical trial–based evidence [9,18], which represents a minority of the ongoing research. A recent systematic review of apps targeting diabetes mellitus, cardiovascular disease, and lung disease found only 3 studies in which a chronic disease management app was used as an intervention and a clinical outcome was measured [19]. Another review focused on how apps can be leveraged by nonprofessional caregivers to care for patients [20].

Although clinical trial evidence supporting the use of apps is generally lacking, this finding may be explained by several factors. First, health apps are fairly new as a medium for engaging patients in comparison with other digital media; therefore, research supporting their use may be ongoing but not yet published. If that is the case, evidence may be found in "gray literature" such as conference proceedings that has not yet made its way to peer-reviewed journals. Second, app developers may be participating in and using research findings to market their apps, which may favor obtaining lower-quality evidence because

it is less costly and potentially biased toward a favorable result. Third, it is possible that high-quality evidence exists but that prior reviews failed to uncover it because they focused too narrowly on a small set of disease areas. Given this set of limitations, a "scoping review" may better describe the extent and quality of the literature as well as evidence gaps in comparison with a traditional systematic review [21].

To address the need for a comprehensive assessment of health app evidence, we performed a scoping review in order to (1) assess the breadth of app coverage across HNHC populations, (2) characterize the quality of the published literature (including full-length journal articles and work presented at scientific conferences), (3) evaluate the possibility of biases due to conflicts of interest, and (4) evaluate the types of outcomes measured.

Methods

Data Sources and Searches

Studies that evaluated health-related apps for mobile devices were identified by searching PubMed and MEDLINE (National Center for Biotechnology Information), EMBASE (Elsevier), the Cochrane Central Register of Controlled Trials (EBSCO), Web of Science (including the Conference Proceedings Citation Indexes; Thomson Reuters), and the NTIS (National Technical Information Service) Bibliographic Database (EBSCO). The search was conducted between June 20, 2014, and July 14, 2014. The complete search strategy including search terms is available in Multimedia Appendix 1. Our search was designed to identify studies examining applications or software programs running on mobile devices such as mobile phones or tablets that are designed to address the health-related needs of specific HNHC populations. Populations included in the search were older adults (age ≥65 years); individuals with chronic conditions including coronary artery disease, congestive heart failure, hypertension, stroke, chronic obstructive pulmonary disease, cancer, diabetes mellitus, obesity, arthritis, chronic kidney disease, cirrhosis, organ transplantation, or chronic pain; the psychologically or mentally vulnerable who have been diagnosed with depression, bipolar disease, posttraumatic stress disorder, attention-deficit hyperactivity disorder, autism, substance-related disorders, dementia, cognitive impairment, developmental delays, or mental impairment; individuals with medication management issues (multiple medications); individuals with physical handicaps or disabilities, including the blind and deaf; and the socially vulnerable including those with low literacy or numeracy, limited English proficiency, minority status (Native

XSL•FO

RenderX

American, Hispanic, African American), low income or homelessness, or infection with human immunodeficiency virus.

Appropriate controlled vocabulary terms were included when available (Medical Subject Headings and Emtree). The retrieval set was limited to articles published in 2008 or later; this start date was selected to coincide with when the iOS and the Android app stores were established. No language restrictions were applied, although non-English abstracts were excluded during title and abstract review (Figure 1). Articles not pertaining to native iOS and Android apps were excluded during the full manuscript review.

Figure 1. Article selection process.
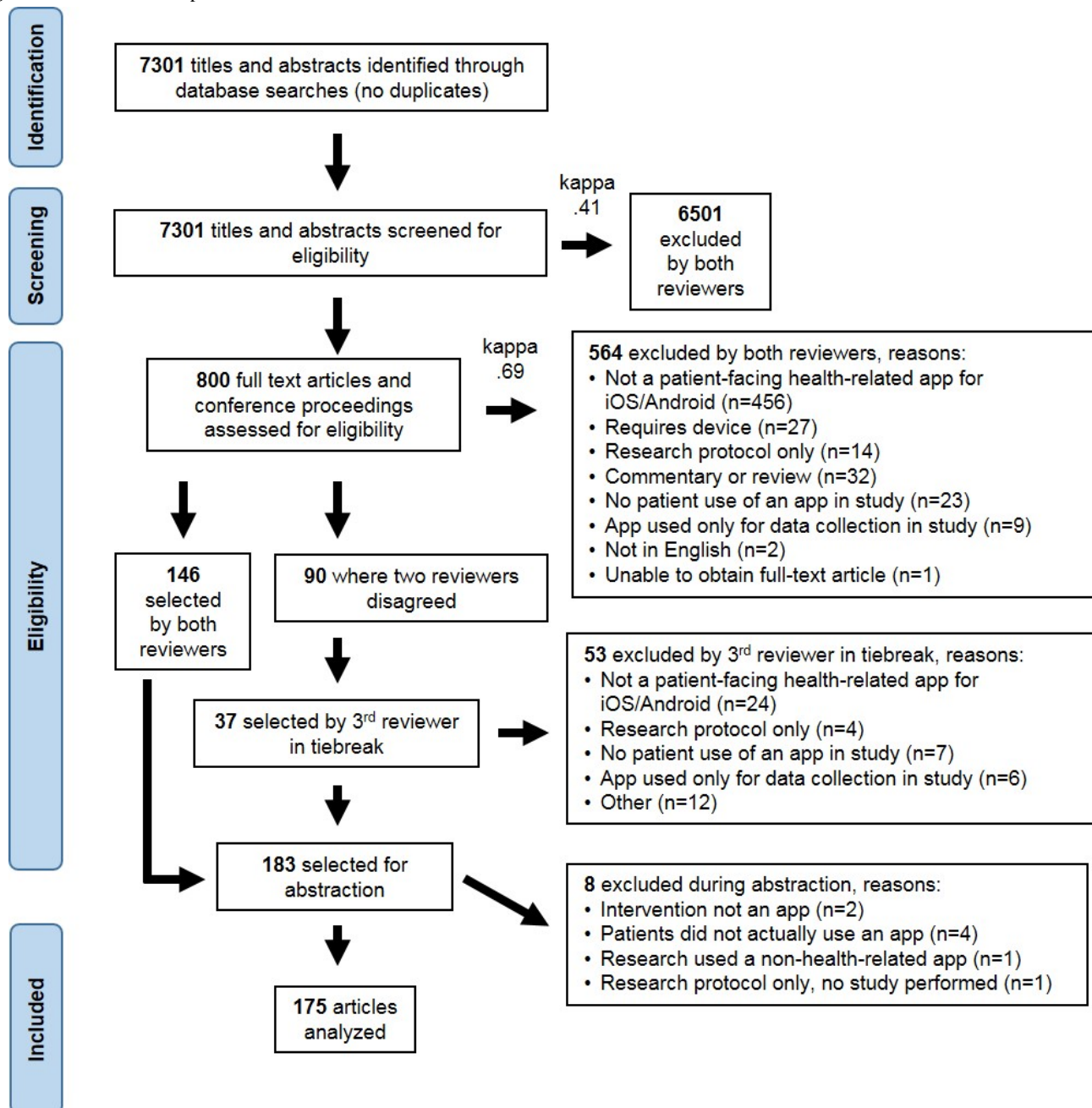
Study Selection

All titles and abstracts were individually examined by 2 reviewers (KS and KD, or KS and LPN). Abstracts were included if they described original research written in the English language involving use of an iOS- or Android-based health-related patient-facing mobile app by study subjects. Patient-facing apps are those intended for use primarily by patients or their caregivers. We selected articles that described either iOS or Android apps because the 2 operating systems serve different demographics, with lower-income individuals, blacks, and younger adults preferring Android devices [22]. Articles describing apps focused on supportive technologies (eg, hearing or vision aids), communication technologies (eg, apps used to help autistic children communicate in school settings), or apps requiring a medical device (eg, an app to interact with artificial pancreas) were excluded. Study design was not a basis for exclusion. Full-length articles were obtained for all abstracts identified for inclusion by *either* reviewer. Certain included abstracts could not be linked to full-length

manuscripts because they were associated with conference proceedings, including oral presentations or poster sessions; these abstracts were included despite the absence of a full-length manuscript as we wanted to capture gray literature in our review.

The full-length manuscripts and conference abstracts were evaluated by 2 reviewers (MF and ES) to confirm that they met the inclusion criteria. Articles identified for inclusion by *both* reviewers were selected for abstraction. Articles where the 2 reviewers disagreed were evaluated by a third reviewer (KD) to break ties.

Data Extraction and Quality Assessment

An abstraction survey tool was created to capture information about both the mobile app as described in the publication and the study itself, including the characteristics of the studied apps, quality of evidence, presence of conflict of interest, and types of outcomes evaluated. During a pilot phase, 8 study investigators each abstracted 3 articles using the tool (24 articles in total); changes were made based on feedback until there was consensus regarding the face validity of the tool.

Abstraction of the selected articles was then performed by 1 reviewer (MF or ES). A total of 40 randomly selected articles were evaluated by both reviewers to assess the level of agreement (Table 1).

Details regarding the abstracted items are presented in Multimedia Appendix 2. App engagement was assessed using a previously described framework [23]. We evaluated the following areas for each article:

General

We captured information about the app studied, including its target population, platform, availability on the app store, and functionalities to support patient engagement.

Quality of the Evidence

We ascertained factors that influence quality and generalizability of the article, including study design, enrollment, follow-up, role of the app in the context of the intervention, and inclusion of relevant patient populations.

Declaration of Conflicts

We determined whether any members of the research team were developers of the app in question (or in a formal role supporting app development such as the advisory board) or whether the app developer directly funded the research. While a conflict of interest does not invalidate the results of a given study, literature written or funded by a company responsible for the product being researched is known to be systematically biased [24].

Outcomes Evaluated

We evaluated the outcomes considered by each study and assessed their direction (ie, positive, neutral, or negative).

Clinical outcomes were those directly related to patient care (eg, decreased hemoglobin A_{1c}). Safety or adverse event outcomes were those relevant to unintended negative consequences of an app. Usability outcomes were those describing an app's ease of use—in some usability studies, multiple rounds of testing are performed, in which case the direction of the outcome was classified based on the final round of testing. Usage describes the amount of time users engaged with the app—this was not reported in a standard fashion and therefore we based the direction on the authors' expectations, considering "sufficient usage" if observed usage matched expectations. Process outcomes refer to measures pertaining to actions taken in response to the app (eg, undergoing testing for hemoglobin A_{1c})—because the result of the action is not considered (eg, decreased hemoglobin A_{1c}), this is not a clinical outcome. A validation outcome was considered present when an app focused on measurement (eg, an app for assessing hepatic encephalopathy) was compared with a non–app-based clinical measure. We evaluated whether the app-based measure performed differently from a non–app-based clinical measure (eg, asterixis). If the article also used a gold standard test (eg, neuropsychiatric testing), we ascertained whether the app-based measure was better or worse than the non–app-based clinical measure. The user satisfaction outcome referred simply to whether users were satisfied with an app.

Data Synthesis and Analysis

Data from the reviewers were imported into R version 3.2.2 (R Foundation for Statistical Computing). Descriptive statistics were calculated and accompanied by a narrative summary.

Results

Article Selection and Abstraction

We identified 7301 titles and abstracts, of which 800 were identified for inclusion by either reviewer (Figure 1). Two reviewers evaluated the full-length manuscripts and 146 articles were selected by both reviewers for abstraction. Of the 90 articles identified for inclusion by only 1 reviewer, a third reviewer selected 37 for abstraction, resulting in a total of 183 articles being selected. During the abstraction process, 8 articles were identified as not meeting the inclusion criteria. After examination by a second reviewer, consensus was achieved on all 8 articles to exclude from the analysis. Thus, in total, 175 articles were abstracted (Multimedia Appendix 3).

Of the 40 articles (22.9%, 40/175) randomly selected for evaluation by both reviewers, 3 were excluded after further examination. As a result, 37 articles were abstracted by both reviewers, and the level of agreement was generally good, with some exceptions such as patient engagement and whether caregivers were included as subjects (Table 1).

Table 1. Level of agreement on items on the abstraction form.

Question[a]	Kappa
General	
Who is the primary population(s) that would benefit from the app studied?	.79[b]
Which platform(s) is used by the app(s) mentioned in the study?	1.0[b]
Is the app(s) studied currently available on the iTunes or Google Play app store?	.85[b]
Based on the app's description in the article, how does it engage patients?	.26[b]
Did the app link to a medical device in the study (eg, glucometer)?	1.0
Did the app link to a consumer wearable device in the study?	1.0
Quality of evidence	
What is the study design?	.62[b]
How many total subjects are enrolled in this study (including controls for controlled trials)?	.72
Is the app studied a standalone intervention (or are there multiple interventions concurrent with app use)?	.59
What was the average length of follow-up reported (in months)?	.54
Was this associated with a conference proceeding (abstract, poster, presentation, etc)?	.93
Does the study have a clinicaltrials.gov registration number?	1.0
Was at least one of the above vulnerable populations included as subjects in the study?	.49
Does the study include children as subjects (people under 18 years old)?	.85
Does the study include people aged 65 or older as subjects?	.79
Were caregivers for at least one of the above vulnerable populations included as subjects in the study?	.29
Conflict of interest	
Did the research team or their employer contribute to the design or development of the app?	.52
What is the source of external funding for this study?	.35[b]
Outcomes evaluated	
Was a clinical outcome considered in this study?	.43
If yes, in what direction was the clinical outcome with use of the app?	.36
Was a safety or adverse event outcome (caused by the use of the app) considered in the study?	0[c]
Was a usability outcome considered in the study?	.53
Was a usage outcome considered in the study?	.71
If yes, in what direction was the usage outcome with use of the app?	.61
Was a process measure considered in this study?	.41
If yes, in what direction was the process measure with use of the app?	.39
Was a validation outcome considered in this study?	.80
If yes, in what direction was the validation outcome with use of the app?	.69
Was user satisfaction considered in this study?	.84
If yes, in what direction was the satisfaction outcome with use of the app?	.71

[a]See Multimedia Appendix 2 for additional information regarding the questions.

[b]Items where reviewers could select multiple options. Only perfect agreement was considered agreement in the kappa calculation.

[c]There was only 1 article evaluated by 2 reviewers in which 1 reviewer marked safety or adverse event outcome as being present.

Identification of Overlapping Research

Of the 175 selected articles, we found 15 sets of articles that assessed the same app, in some instances using different study designs, numbers of participants, or end points. For the purposes of our analysis, we considered each article as separate. We found 2 articles that evaluated the ActiveLifestyle app [25,26], AsthmaCare [27,28], EncephalApp [29,30], iMigraine [31,32], iStepLog [33,34], the Mayo Clinic app [35,36], Multiple Sclerosis Performance Test [37,38], My Meal Mate [39,40],

Ready~Steady [41,42], a cognitive stimulation app for alcoholics [43,44], USMART [45,46], ClinTouch [47,48], and a food addiction intervention [49,50]. We found 3 articles that evaluated a mobile application in the Women with Epilepsy: Pregnancy Outcomes and Deliveries (WEPOD) study [51-53] and 3 that evaluated the SaGAS 20/10 app [54,55].

Characteristics of the Mobile Apps

Of the 27 vulnerable populations targeted by the literature search, the groups most commonly targeted by apps included individuals with obesity, physical handicaps, diabetes, older age, and dementia or mild cognitive impairment (Table 2).

Table 2. Primary population that would benefit from the app studied.

Population[a]	Number of articles (N=175), n (%)
Obesity	24 (13.7)
Physical handicap or disability (including blindness or deafness)	19 (10.9)
Diabetes mellitus	15 (8.6)
Older adults	15 (8.6)
Dementia or mild cognitive impairment	14 (8.0)
Cancer	11 (6.3)
Autism spectrum disorder	10 (5.7)
Alcohol or drug abuse	7 (4.0)
Chronic pain	7 (4.0)
Depression	7 (4.0)
Coronary artery disease	6 (3.4)
Schizophrenia or psychosis	5 (2.9)
Arthritis	4 (2.3)
Stroke	4 (2.3)
Cirrhosis	3 (1.7)
Congestive heart failure	3 (1.7)
Hypertension	3 (1.7)
Posttraumatic stress disorder	3 (1.7)
Developmentally delayed or mentally impaired	2 (1.1)
HIV[b] or AIDS	2 (1.1)
Attention-deficit hyperactivity disorder	1 (0.6)
Bipolar disorder	1 (0.6)
Chronic kidney disease	1 (0.6)
Low income or poor	1 (0.6)
Low literacy or low numeracy	1 (0.6)
Posttransplant	1 (0.6)
Smoking	1 (0.6)
None of the above	38 (21.7)

[a]These are not mutually exclusive categories. Articles may evaluate multiple apps and individual apps may target multiple populations.

[b]HIV: human immunodeficiency virus.

Of the 175 selected articles, 60.6% (106/175) involved iOS apps, 32.0% (56/175) involved Android apps, and 7.4% (13/175) involved both. Reviewers evaluated the availability of these apps on both the iTunes (iOS) and Google Play (Android) app stores. Reviewers were unable to search for the app being studied in 40.0% (70/175) of the articles because the name of the app was not mentioned. Apps from an additional 29.7% (52/175) articles were searched and unable to be found on either app store. Among the articles where an app was found, 66% (35/53) were found on the iOS app store, 6% (3/53) on the Android app store, and 28% (15/53) on both. The ways in which apps engaged patients were assessed based on the functionality described in the articles (Table 3).

Table 3. How health apps engage patients.

Type of engagement[a]	Number of articles (N=175) n (%)
Records information	132 (75.4)
Provides guidance	64 (36.6)
Displays a patient's health information	55 (31.4)
Reminds or alerts patients	45 (25.7)
Provides educational information	36 (20.6)
Enables data sharing with clinician	36 (20.6)
Enables data sharing with caregiver	15 (8.6)
Engages through social media	14 (8.0)
Not enough information to determine	10 (5.7)
None of the above	6 (3.4)

[a]These are not mutually exclusive categories. Apps may engage patients in multiple ways.

The most common functionalities were recording information, providing guidance, and displaying health information, and the least common were engaging with social media and enabling communication with family members. A total of 6 (3.4%, 6/175) articles described apps with the ability to connect with a medical device and 5 (2.9%, 5/175) described apps able to connect with a consumer wearable device.

Quality of Evidence

The method of dissemination involved full-text publications for 136 articles and conference proceedings (eg, oral or poster presentation) for 39 articles. Cross-sectional studies accounted for 42.9% (75/175) of studies (Table 4). Methodologies with lower bias were represented as follows: randomized controlled trials (10.3%, 18/175), nonrandomized controlled trials (2.9%, 5/175), and prospective cohort studies (21.7%, 38/175). The median number of participants in the studies was 31 (interquartile range, IQR, 11.0-207.2, maximum 11,690). The median length of follow-up for non–cross-sectional studies—weighted for the number of participants when articles involved multiple substudies—was 1.4 months (IQR 0.6-3, maximum 42.6).

Table 4. Study designs used in abstracted articles.

Study design[a]	Number of articles (N=175) n (%)
Cross-sectional study	75 (42.9)
Prospective cohort study	38 (21.7)
Qualitative research	34 (19.4)
Before-after study	22 (12.6)
Randomized controlled trial	18 (10.3)
Nonrandomized controlled trial	5 (2.9)
Case report or case series	3 (1.7)
Randomized trial with no control	1 (0.6)
Interrupted time series	1 (0.6)
Not enough information to determine	1 (0.6)

[a]These are not mutually exclusive categories. Articles may use multiple study designs or may describe multiple substudies.

Among 26 studies with a control arm, the app was the sole intervention in 21 (81%, 21/26) articles and just one part of a multipart intervention in 5 (19%, 5/26) articles. In the remaining 149 articles, all study subjects were exposed to the app. Only 7.4% (13/175) studies were registered on ClinicalTrials.gov. Of the 137 articles for which a specific population was identified that may benefit from the app (eg, patients with heart disease), 78.1% (107/137) included members of that particular population in the study; 13.1% (18/137) of studies focused on screening or prevention and therefore participants were healthy individuals; the remaining 8.8% (12/137) of studies did not include participants from the relevant population.

Of the 175 studies, 38 (21.7%, 38/175) studies included children, 53 (30.3%, 53/175) included adults aged 65 years or older, and 17 (9.7%, 17/175) included caregivers of HNHC patients.

Declaration of Conflicts

The authors of the 175 identified studies or their employer directly contributed to the design or development of the app in 107 (61.1%, 107/175) articles; among these, however, 5 (2.9%, 5/175) did not state this explicitly in the body of the paper. The authors of the identified studies were not involved in app development in 28 (16.0%, 28/175) articles, and in 40 (22.9%, 40/175) we were unable to confirm the presence or absence of involvement. Of the 175 studies, 61 (34.9%, 61/175) studies were funded by a government agency, 41 (23.4%, 41/175) by a nonprofit organization, 12 (6.9%, 12/175) by a for-profit company, and 6 (3.4%, 6/175) by a medical professional society; 5 (2.9%, 5/175) studies reported they had no external funding. No statement about the funding source was present in 82 (46.9%, 82/175) articles.

Types of Outcomes Evaluated

Among the 175 articles, 87 (49.7%, 87/175) articles evaluated user satisfaction, finding users to be generally satisfied in 74 (85%, 74/87), generally unsatisfied in 2 (2%, 2/87), and neutral in the remaining 11 (13%, 11/87) articles. A total of 74 (42.3%, 74/175) articles evaluated usability—often multiple cycles of testing were described, and the first cycle typically had worse performance than later cycles after modifications were made. A total of 61 (34.9%, 61/175) articles looked at usage, finding "sufficient" usage in 53 (87%, 53/61) and lower-than-expected use in 8 (13%, 8/61). Of the 175 articles, 56 (32.0%, 56/175) articles validated the measurement ability of apps in comparison with a clinical measure, finding the app to perform better than the clinical measure in 6 (11%, 6/56), worse than the clinical measure in 17 (30%, 17/56), and no different from the clinical measure in 34 (61%, 34/56) studies. A total of 40 (22.9%, 40/175) studies assessed a process measure (eg, increased administration of smoking cessation counseling), as opposed to a clinical outcome (eg, decreased rate of lung cancer). Of these, there were 35 (88%, 35/40) studies with improvement, 1 (2%, 1/40) with worsening, and 4 (10%, 4/40) with no change in the process outcome. A total of 36 (20.6%, 36/175) articles evaluated clinical outcomes, with 26 (72%, 26/36) demonstrating improvement in clinical outcomes and 10 (28%, 10/36) with no change. Only 9 (5.1%, 9/175) articles considered a safety or adverse event outcome caused by use of the app.

Discussion

Principal Findings

While there is optimism that mobile health apps may support the health of HNHC populations, existing research does not robustly evaluate this potential. Our review of the evidence supporting patient-facing mobile health apps identified a number of gaps in the current body of research. A few HNHC groups (older adults and people with obesity, physical handicaps, diabetes, and dementia) are more commonly studied, and we found less than 10 studies published for 20 of the 27 HNHC groups included in our review. The majority of apps studied were unavailable to consumers, the study designs were primarily cross-sectional, non–cross-sectional studies had a fairly short length of follow-up, and study sizes were small. In most cases, developers were often the ones evaluating the apps, sample sizes

were small, funding sources were ambiguous, and clinical outcomes were evaluated in a minority of studies. Even among high-quality studies, drawing an inference about the usefulness of an app was frequently limited by intervention arms in which the app was a small piece of a much larger intervention.

Some of the methodological problems we identified such as small sample sizes and short length of follow-up could be addressed if apps incorporated the consent process and data collection into the apps' functionality. Many of the studies used a traditional "in-person" consent process in order to enroll study subjects. While this may conform to the standards of traditional clinical research, using this approach may limit the number of subjects who can be enrolled and the length of follow-up. New methodological approaches that enable large-scale app outcomes research are needed [56]. Controlled trials where the consent process and data collection occur entirely in the context of a publicly available app may enable such work. The barrier to entry for integrating research into apps has been lowered by frameworks such as Apple ResearchKit, which was used to enroll 11,000 participants for a cardiovascular study in 24 hours [57].

Recommendations

On the basis of our findings, we make the following recommendations for researchers undertaking the study of mobile apps for the purposes of dissemination:

First, the researchers should consider evaluating apps in understudied HNHC groups to address the current imbalance in the body of research between HNHC groups. Second, reports should include the name of the app or intervention, so that literature about the app can be linked to it definitively; every effort should be made to include a bundle ID, permanent app store weblink, or other unique identifier to facilitate identification of the app. Third, researchers conducting interventional studies should consider the inclusion of both a control arm and an app-only intervention arm to make clearer the link between the app and the outcome. Fourth, studies should clearly state the nature of the relationship between the study contributors and the app developers; if the researchers are also the app developers, researchers should consider validating their work at an additional site supervised by a nondeveloper. Fifth, studies should clearly state the funding source or note if no external funding was used. Finally, researchers should report negative results.

In addition, funders will need to support additional evaluations of apps and should target evaluations that target clinically important outcomes and are large enough to deliver meaningful results. With newer enrollment approaches, it may be possible to enable much larger clinical trials, which may be feasible at low expense because much of the data usually collected may be extracted from existing electronic health records.

We used a robust multistage scoping review process involving 2 reviewers in most steps. We included gray literature in our analysis through a search of conference proceedings and did not limit our analysis to only high-quality evidence.

Limitations

Although 2 reviewers participated in the process of screening abstracts and full-length manuscripts for eligibility, data abstraction was carried out by only 1 reviewer for most studies. While the majority (17 of 30) of abstracted items had good agreement among the reviewers, interrater agreement was moderate (kappa .41-.6) for 7 items and fair (kappa .21-.40) for 6 questions, which limits the reliability of conclusions drawn from them. To put the kappa values into context, 2 reviewers agreed on ascertainment of a clinical outcome (kappa .43) on 31 out of 37 articles. In particular, the questions with the lowest interrater agreement involved determining whether safety or adverse event outcomes were considered, whether caregivers were included the study, how the app mentioned in the study engaged patients, and the funding source. The low agreement levels for these questions may be attributed to the heterogeneity in the detail to which studies reported information about the apps and study conducted, which is partly due to the inclusion of conference abstracts in our analysis. The agreement was moderate when reviewers abstracted the types of outcomes measured in the studies. We attribute this partly to the breadth of populations that we considered because what constitutes a clinical outcome differs significantly between chronic conditions. Additionally, differentiating clinical outcomes from process outcomes carries some subjectivity and may introduce disagreements. We did revise our abstraction form based on a review of 24 articles during a pilot phase, but additional cycles of revision may have further improved interrater agreement.

We did not evaluate articles that were not in English, which limits our generalizability toward apps targeting non-English speakers. Finally, we conducted our literature search in 2014, which does not capture recent trends.

Recent Trends in the Literature

Recent studies of patient-facing apps have provided supporting evidence for the role of apps in several areas. In a 2016 randomized controlled crossover study of a mobile app focused on supporting drug intake and vital sign documentation, researchers found that patients who used the iPad app showed greater adherence for both medication intake and blood pressure measurement than a paper-based control group [58]. Another randomized controlled study published in 2016 found that overweight and obese adults who used a social support app lost on average 3 kg more than patients using a self-monitoring control app over the course of the study [59]. Evidence from other recent trials has demonstrated the ability of apps to reduce consumption of sugar-sweetened beverages in women and nutrient-poor foods in men, increase activity level and reduce fatigue following stroke, and improve respiratory parameters with a reduction in corticosteroid usage among individuals with uncontrolled asthma [60-62].

Conclusions

In the future, providers may routinely prescribe apps for their HNHC patients, and health care systems may invest in them. However, given the limited availability of high-quality evidence for most of the HNHC groups included in our review, we would not expect systematic reviews or meta-analyses focused on these groups individually to yield enough evidence to assess the effectiveness of disease-specific apps. Additionally, apps are being lost in translation from research to the app stores, resulting in a lack of commercial impact of existing research. Despite these limitations, the body of evidence overwhelmingly reports early results that favor the use of mobile health apps.

Conflicts of Interest

EP serves as a mentor at Rock Health and previously served as Chief Medical Officer at Twine Health, a mobile health company. AL currently serves as an advisor to the Hacking Medicine Institute, a nonprofit organization that evaluates digital health apps, where he also serves as senior editor on the RANKED Health project.

DWB is a coinventor on Patent No. 6029138 held by Brigham and Women's Hospital on the use of decision support software for medical management, licensed to the Medicalis Corporation. He holds a minority equity position in the privately held company Medicalis, which develops Web-based decision support for radiology test ordering. He serves on the board for SEA Medical Systems, which makes intravenous pump technology. He is on the clinical advisory board for Zynx, Inc, which develops evidence-based algorithms. He consults for EarlySense, which makes patient safety monitoring systems. He receives equity and cash compensation from QPID, Inc, a company focused on intelligence systems for electronic health records. He receives cash compensation from CDI (Negev), Ltd, which is a not-for-profit incubator for health information technology start-ups. He receives equity from Enelgy, which makes software to support evidence-based clinical decisions. He receives equity from Ethosmart, which makes mobile apps to help patients with chronic diseases. He receives equity from Intensix, which makes software to support clinical decision making in intensive care. He receives equity from MDClone, which takes clinical data and produces deidentified versions of the data. DWB's financial interests have been reviewed by Brigham and Women's Hospital and Partners HealthCare in accordance with their institutional policies.

The views presented here are those of the authors and not necessarily those of The Commonwealth Fund or its directors, officers, or staff. The Commonwealth Fund was not involved in any of the following: design or conduct of the study; collection, management,

analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Multimedia Appendix 1

Search Terms.

[PDF File (Adobe PDF File), 37KB - mhealth_v4i4e136_app1.pdf]

Multimedia Appendix 2

Abstraction Form.

[PDF File (Adobe PDF File), 56KB - mhealth_v4i4e136_app2.pdf]

Multimedia Appendix 3

Articles included in review (n=175) and complete results of abstraction.

[XLSX File (Microsoft Excel File), 60KB - mhealth_v4i4e136_app3.xlsx]

References

1. Hong CS, Siegel AL, Ferris TG. Caring for high-need, high-cost patients: what makes for a successful care management program? Issue Brief (Commonw Fund) 2014 Aug;19:1-19. [Medline: 25115035]
2. McCarthy D, Ryan J, Klein S. Models of care for high-need, high-cost patients: an evidence synthesis. Issue Brief (Commonw Fund) 2015 Oct;31:1-19. [Medline: 26591906]
3. 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]
4. Pew Research Center. 2014. Cell Phone and Smartphone Ownership Demographics URL: http://www.pewinternet.org/data-trend/mobile/cell-phone-and-smartphone-ownership-demographics [accessed 2016-12-07] [WebCite Cache ID 6maR4Vf0h]
5. Greene J, Hibbard JH, Sacks R, Overton V, Parrotta CD. When patient activation levels change, health outcomes and costs change, too. Health Aff (Millwood) 2015 Mar;34(3):431-437. [doi: 10.1377/hlthaff.2014.0452] [Medline: 25732493]
6. Grando MA, Rozenblum R, Bates DW, editors. Information Technology for Patient Empowerment in Healthcare. 1st ed. Berlin: Walter de Gruyter Inc; 2015.
7. Hibbard JH, Greene J, Overton V. Patients with lower activation associated with higher costs; delivery systems should know their patients' 'scores'. Health Aff (Millwood) 2013 Feb;32(2):216-222 [FREE Full text] [doi: 10.1377/hlthaff.2012.1064] [Medline: 23381513]
8. Broderick A, Haque F. Mobile health and patient engagement in the safety net: a survey of community health centers and clinics. Issue Brief (Commonw Fund) 2015 May;9:1-9. [Medline: 26040018]
9. IMSHealth. 2015. Patient Adoption of mHealth URL: http://www.imshealth.com/files/web/IMSH%20Institute/Reports/Patient%20Adoption%20of%20mHealth/IIHI_Patient_Adoption_of_mHealth.pdf [accessed 2016-12-07] [WebCite Cache ID 6maRVuzDh]
10. Mercurynews. Apple, Google, VCs invest in health technology URL: http://www.mercurynews.com/2014/08/13/apple-google-vcs-invest-in-health-technology/ [accessed 2016-12-07] [WebCite Cache ID 6maRbIZCz]
11. Martínez-Pérez B, de la Torre-Díez I, López-Coronado M, Herreros-González J. Mobile apps in cardiology: review. JMIR Mhealth Uhealth 2013 Jul;1(2):e15 [FREE Full text] [doi: 10.2196/mhealth.2737] [Medline: 25098320]
12. 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;13(3):e65 [FREE Full text] [doi: 10.2196/jmir.1874] [Medline: 21979293]
13. Wesley KM, Fizur PJ. A review of mobile applications to help adolescent and young adult cancer patients. Adolesc Health Med Ther 2015 Aug;6:141-148 [FREE Full text] [doi: 10.2147/AHMT.S69209] [Medline: 26316835]
14. 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]
15. Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015 Feb;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]
16. Sunyaev A, Dehling T, Taylor PL, Mandl KD. Availability and quality of mobile health app privacy policies. J Am Med Inform Assoc 2015 Apr;22(e1):e28-e33 [FREE Full text] [doi: 10.1136/amiajnl-2013-002605] [Medline: 25147247]

17. Subhi Y, Bube SH, Rolskov Bojsen S, Skou Thomsen AS, Konge L. Expert involvement and adherence to medical evidence in medical mobile phone apps: a systematic review. JMIR Mhealth Uhealth 2015 Jul;3(3):e79 [FREE Full text] [doi: 10.2196/mhealth.4169] [Medline: 26215371]

18. IMSHealth. Patient Apps for Improved Healthcare: From Novelty to Mainstream URL: http://www.imshealth.com/en/thought-leadership/quintilesims-institute/reports/patient-apps-for-improved-healthcare [accessed 2016-12-07] [WebCite Cache ID 6maRqxy4S]

19. 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;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]

20. Dyer DE, Kansagara D, McInnes DK, Freeman M, Woods S. Mobile applications and Internet-based approaches for supporting non-professional caregivers: a systematic review. Department of Veterans Affairs (US) 2012 Nov. [Medline: 23285508]

21. Armstrong R, Hall BJ, Doyle J, Waters E. Cochrane Update. 'Scoping the scope' of a cochrane review. J Public Health (Oxf) 2011 Mar;33(1):147-150 [FREE Full text] [doi: 10.1093/pubmed/fdr015] [Medline: 21345890]

22. Pew Research Center. Smartphone Ownership 2013 URL: http://www.pewinternet.org/2013/06/05/smartphone-ownership-2013/ [accessed 2016-12-07] [WebCite Cache ID 6maS8HFss]

23. Singh K, Drouin K, Newmark LP, Rozenblum R, Lee J, Landman A, et al. Developing a framework for evaluating the patient engagement, quality, and safety of mobile health applications. Issue Brief (Commonw Fund) 2016 Feb;5:1-11. [Medline: 26934758]

24. Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. Br Med J 2003 May 31;326(7400):1167-1170 [FREE Full text] [doi: 10.1136/bmj.326.7400.1167] [Medline: 12775614]

25. Silveira P, van het Reve E, Daniel F, Casati F, de Bruin ED. Motivating and assisting physical exercise in independently living older adults: a pilot study. Int J Med Inform 2013 May;82(5):325-334. [doi: 10.1016/j.ijmedinf.2012.11.015] [Medline: 23273420]

26. 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]

27. Farooqui N, Phillips G, Stukus DR. Feasibility and acceptability of a novel asthma self-management smartphone application for children and adolescents. J Allergy Clin Immunol 2014 Feb;133(2):AB127. [doi: 10.1016/j.jaci.2013.12.473]

28. Stukus DR, Phillips G, Farooqui N. Improved education and self-management in children and adolescents with asthma using a personalized smartphone application. J Allergy Clin Immunol 2014 Feb;133(2):AB156. [doi: 10.1016/j.jaci.2013.12.568]

29. Bajaj JS, Thacker LR, Heuman DM, Fuchs M, Sterling RK, Sanyal AJ, et al. The Stroop smartphone application is a short and valid method to screen for minimal hepatic encephalopathy. Hepatology 2013 Sep;58(3):1122-1132 [FREE Full text] [doi: 10.1002/hep.26309] [Medline: 23389962]

30. Bajaj JS, Heuman DM, Sterling RK, Sanyal AJ, Siddiqui M, Matherly S, et al. Validation of encephalApp, smartphone-based stroop test, for the diagnosis of covert hepatic encephalopathy. Clin Gastroenterol Hepatol 2015 Oct;13(10):1828-1835.e1 [FREE Full text] [doi: 10.1016/j.cgh.2014.05.011] [Medline: 24846278]

31. Nelson TD, Aylward BS, Peugh J, Kroner J, Sullivan S, Hershey AD, et al. Variability in adolescent migraine symptoms and correlates of individual symptom variability. 2013 Presented at: 2013 International Headache Congress; June 27-30, 2013; Boston, MA p. 186.

32. Sullivan S, Kroner J, Aylward B, Allen JR, Kabbouche M, O'Brien H, et al. A mobile application to track pain and functioning in adolescents: feasibility and initial results. 2015 Presented at: 54th American Headache Society Annual Meeting; June 21-24, 2012; Los Angeles, CA p. 897.

33. Kirwan M, Duncan MJ, Vandelanotte C, Mummery WK. Design, development, and formative evaluation of a smartphone application for recording and monitoring physical activity levels: the 10,000 Steps "iStepLog". Health Educ Behav 2013 Apr;40(2):140-151. [doi: 10.1177/1090198112449460] [Medline: 22984196]

34. Kirwan M, Duncan MJ, Vandelanotte C, Mummery WK. 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]

35. Cook DJ, Manning DM, Holland DE, Prinsen SK, Rudzik SD, Roger VL, et al. Patient engagement and reported outcomes in surgical recovery: effectiveness of an e-health platform. J Am Coll Surg 2013 Oct;217(4):648-655. [doi: 10.1016/j.jamcollsurg.2013.05.003] [Medline: 23891066]

36. Cook DJ, Moradkhani A, Douglas KSV, Prinsen SK, Fischer EN, Schroeder DR. Patient education self-management during surgical recovery: combining mobile (iPad) and a content management system. Telemed J E Health 2014 Apr;20(4):312-317 [FREE Full text] [doi: 10.1089/tmj.2013.0219] [Medline: 24443928]

37. Miller D, Lee J, Bethoux F, Rao S, Alberts J, Schindler D, et al. Multiple sclerosis performance test: patients express high satisfaction with the application. Neurology 2014 Apr 8;82(10 Supplement P3.140).

38. Rudick RA, Miller D, Bethoux F, Rao SM, Lee J, Stough D, et al. The multiple sclerosis performance test (MSPT): an iPad-based disability assessment tool. J Vis Exp 2014 Jun 30(88):e51318 [FREE Full text] [doi: 10.3791/51318] [Medline: 25046650]

39. 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 Apr;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]

40. Carter MC, Burley VJ, Nykjaer C, Cade JE. 'My Meal Mate' (MMM): validation of the diet measures captured on a smartphone application to facilitate weight loss. Br J Nutr 2013 Feb 14;109(3):539-546. [doi: 10.1017/S0007114512001353] [Medline: 22717334]

41. Vankipuram M, McMahon S, Fleury J. ReadySteady: app for accelerometer-based activity monitoring and wellness-motivation feedback system for older adults. 2012 Presented at: AMIA Annual Symposium; November 3, 2012 - November 7, 2012; Chicago, IL.

42. McMahon S, Vankipuram M, Hekler EB, Fleury J. Design and evaluation of theory-informed technology to augment a wellness motivation intervention. Transl Behav Med 2014 Mar;4(1):95-107 [FREE Full text] [doi: 10.1007/s13142-013-0221-4] [Medline: 24653780]

43. Gamito P, Oliveira J, Lopes P, Morais D, Brito R, Saraiva T, et al. Assessment of frontal brain functions in alcoholics following a health mobile cognitive stimulation approach. Stud Health Technol Inform 2013;191:110-114. [Medline: 23792854]

44. Gamito P, Oliveira J, Lopes P, Brito R, Morais D, Silva D, et al. Executive functioning in alcoholics following an mHealth cognitive stimulation program: randomized controlled trial. J Med Internet Res 2014 Apr 17;16(4):e102 [FREE Full text] [doi: 10.2196/jmir.2923] [Medline: 24742381]

45. Han JW, Oh K, Kim T, Chi Y, Kim K. Spaced retrieval training in mild cognitive impairment and early alzheimer's disease. Int Psychogeriatr 2013;25:S12-S13.

46. Han JW, Oh K, Yoo S, Kim E, Ahn K, Son Y, et al. Development of the ubiquitous spaced retrieval-based memory advancement and rehabilitation training program. Psychiatry Investig 2014 Jan;11(1):52-58 [FREE Full text] [doi: 10.4306/pi.2014.11.1.52] [Medline: 24605124]

47. Palmier-Claus JE, Ainsworth J, Machin M, Barrowclough C, Dunn G, Barkus E, et al. The feasibility and validity of ambulatory self-report of psychotic symptoms using a smartphone software application. BMC Psychiatry 2012 Oct 17;12:172 [FREE Full text] [doi: 10.1186/1471-244X-12-172] [Medline: 23075387]

48. Palmier-Claus JE, Rogers A, Ainsworth J, Machin M, Barrowclough C, Laverty L, et al. Integrating mobile-phone based assessment for psychosis into people's everyday lives and clinical care: a qualitative study. BMC Psychiatry 2013 Jan 23;13:34 [FREE Full text] [doi: 10.1186/1471-244X-13-34] [Medline: 23343329]

49. Pretlow R. Substance dependence (problem food) approach to childhood obesity, implemented as a smartphone APP: a pilot study. Obes Facts 2012;5:195.

50. Pretlow R. Addiction model intervention for obesity, implemented as a smartphone App: A pilot study. J Behav Addict 2013;2:29.

51. Ernst L, Llewellyn N, French J, Pennell PB, Bartfield E, Lau C, et al. Medication adherence in WWE via use of IPOD application tracking device. Epilepsia 2013;54:131.

52. Llewellyn NG, Harden CL, French J, Pennell PB, Bartfeld E, Davis AR, et al. Maintenance of subject adherence to daily diary entry facilitated by use of a mobile application in the WEPOD study. Epilepsy Curr 2013;13:69-70.

53. Pennell P, French J, Harden CL. Evaluation of a mobile application tool in the wepod study. Epilepsy Curr 2012;12(1).

54. Zeber AC, Vaney C, Vaney S. Displaying the motor components of the MSFC as an interval score on an iPhone App with SaGAS 20/10. Clin Neurophysiol 2012;123(10):e113-e114.

55. Vaney C, Forkel N, Hilfiker R. Validity and responsiveness of SaGAS 20/10, a revised and simplified version of SaGAS (Short and Graphic Ability Score) in the rehabilitation of patients with neurological diseases. Mult Scler 2013;19(11):121-122.

56. Powell AC, Landman AB, Bates DW. In search of a few good apps. J Am Med Assoc 2014 May 14;311(18):1851-1852. [doi: 10.1001/jama.2014.2564] [Medline: 24664278]

57. Rosa C, Campbell AN, Miele GM, Brunner M, Winstanley EL. Using e-technologies in clinical trials. Contemp Clin Trials 2015 Nov;45(Pt A):41-54. [doi: 10.1016/j.cct.2015.07.007] [Medline: 26176884]

58. Mertens A, Brandl C, Miron-Shatz T, Schlick C, Neumann T, Kribben A, et al. A mobile application improves therapy-adherence rates in elderly patients undergoing rehabilitation: a crossover design study comparing documentation via iPad with paper-based control. Medicine (Baltimore) 2016 Sep;95(36):e4446 [FREE Full text] [doi: 10.1097/MD.0000000000004446] [Medline: 27603339]

59. 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]

60. 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]

61. Paul L, Wyke S, Brewster S, Sattar N, Gill JM, Alexander G, et al. Increasing physical activity in stroke survivors using STARFISH, an interactive mobile phone application: a pilot study. Top Stroke Rehabil 2016 Jun;23(3):170-177. [doi: 10.1080/10749357.2015.1122266] [Medline: 27077973]

62. Cook KA, Modena BD, Simon RA. Improvement in asthma control using a minimally burdensome and proactive smartphone application. J Allergy Clin Immunol Pract 2016 Jul;4(4):730-737.e1. [doi: 10.1016/j.jaip.2016.03.005] [Medline: 27107690]

Abbreviations

HNHC: high-need, high-cost
IQR: interquartile range

Original Paper

Benefits of Mobile Phone Technology for Personal Environmental Monitoring

David Donaire-Gonzalez[1,2,3,4*], PhD; Antònia Valentín[1,2,3*], MSc; Audrey de Nazelle[5*], PhD; Albert Ambros[1,2,3*], MSc; Glòria Carrasco-Turigas[1,2,3], MSc; Edmund Seto[6], PhD; Michael Jerrett[7,8], PhD; Mark J Nieuwenhuijsen[1,2,3], PhD

[1]ISGlobal, Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain

[2]Pompeu Fabra University (UPF), Barcelona, Spain

[3]Ciber on Epidemiology and Public Health (CIBERESP), Barcelona, Spain

[4]Physical Activity and Sports Sciences Department, Fundació Blanquerna, Ramon Llull University, Barcelona, Spain

[5]Center for Environmental Policy, Imperial College London, London, United Kingdom

[6]Department of Environmental and Occupational Health Services, University of Washington, Seattle, WA, United States

[7]Environmental Health Sciences, School of Public Health, University of California, Berkeley, CA, United States

[8]Department of Environmental Health, Fielding School of Public Health, University of California, Los Angeles, CA, United States

[*]these authors contributed equally

Corresponding Author:
David Donaire-Gonzalez, PhD
ISGlobal, Centre for Research in Environmental Epidemiology (CREAL)
Doctor Aiguader, 88
Barcelona, 08003
Spain
Phone: 34 93 214 73 17
Fax: 34 93 214 73 01
Email: david.donaire@isglobal.org

Abstract

Background: Tracking individuals in environmental epidemiological studies using novel mobile phone technologies can provide valuable information on geolocation and physical activity, which will improve our understanding of environmental exposures.

Objective: The objective of this study was to assess the performance of one of the least expensive mobile phones on the market to track people's travel-activity pattern.

Methods: Adults living and working in Barcelona (72/162 bicycle commuters) carried simultaneously a mobile phone and a Global Positioning System (GPS) tracker and filled in a travel-activity diary (TAD) for 1 week (N=162). The CalFit app for mobile phones was used to log participants' geographical location and physical activity. The geographical location data were assigned to different microenvironments (home, work or school, in transit, others) with a newly developed spatiotemporal map-matching algorithm. The tracking performance of the mobile phones was compared with that of the GPS trackers using chi-square test and Kruskal-Wallis rank sum test. The minute agreement across all microenvironments between the TAD and the algorithm was compared using the Gwet agreement coefficient (AC1).

Results: The mobile phone acquired locations for 905 (29.2%) more trips reported in travel diaries than the GPS tracker ($P<.001$) and had a median accuracy of 25 m. Subjects spent on average 57.9%, 19.9%, 9.0%, and 13.2% of time at home, work, in transit, and other places, respectively, according to the TAD and 57.5%, 18.8%, 11.6%, and 12.1%, respectively, according to the map-matching algorithm. The overall minute agreement between both methods was high (AC1 .811, 95% CI .810-.812).

Conclusions: The use of mobile phones running the CalFit app provides better information on which microenvironments people spend their time in than previous approaches based only on GPS trackers. The improvements of mobile phone technology in microenvironment determination are because the mobile phones are faster at identifying first locations and capable of getting location in challenging environments thanks to the combination of assisted-GPS technology and network positioning systems. Moreover, collecting location information from mobile phones, which are already carried by individuals, allows monitoring more people with a cheaper and less burdensome method than deploying GPS trackers.

XSL•FO
RenderX

Introduction

Environmental exposures are crucial determinants of people's health [1]. Despite their importance for health, historically, the assessment of exposure to environmental health risks has been mainly based on exposures derived from the occupational or residential microenvironments [2]. Now, however, it is well known that this approach is inaccurate to represent most of the environmental exposures, given the large spatiotemporal variability of people's activities and microenvironments visited in a day [3-5]. As a result, researchers have started to perform small quasi-experimental studies seeking approaches that move the exposure assessment science from the microenvironment level to the individual level [6-9]. The finding of a reliable and accurate remote tracking tool will provide researchers with the opportunity to determine and understand the causal and temporal relationship of natural and urban environments with health-related behaviors and exposures as well as physical and mental health conditions [10]. However, the previously used tracking tools, such as travel diaries, questionnaires, and Global Positioning System (GPS) technology, along with postprocessing methods have prevented the instauration of this new paradigm of exposure science in epidemiological studies because they are limited and suffer from important weaknesses [11,12].

Mobile phone technology may help to overcome the previous limitations because of its widespread use around the world and the combination of assisted GPS technology and network positioning systems [13-16]. The assisted GPS technology makes use of remote GPS location servers to reduce both power consumption and the time to first fix position [14]. Network positioning systems get geolocation using Wi-Fi signals or, in their absence, from cellular network signals to complement location information from the assisted GPS technology when there is limited satellite visibility. However, until now, mobile phone technology has been assessed only regarding its tracking performance and accuracy mainly in experimental or quasi-experimental studies (ie, scripted studies in a controlled environment and small population sizes) [12,14,17-22]. Moreover, mobile phone technology might not be accurate enough for street-level tracking because of the limitations in GPS antenna, digital interface, GPS chipset, or mobile platform [23].

In this context, the aim of this study was to assess the performance of mobile phone technology in tracking people's travel-activity pattern in a dense city while they perform their daily life activities.

Methods

Study Design and Sample

This is a concurrent validation study comparing the tracking and travel-activity determination of a mobile phone versus a GPS tracker and a travel-activity diary (TAD), respectively. The study is nested in the Transportation, Air Pollution and Physical Activities (TAPAS) Travel Survey study [24]. In brief, the TAPAS Travel Survey study aimed to understand the barriers and benefits of bicycle commuting among those who commute by a motorized mode within Barcelona city [25-28].

For this study, a convenience sample of 178 participants from TAPAS Travel Survey study was used. The TAPAS sample was composed of 815 healthy participants recruited following stratified sampling according to commute mode (bicycle vs motorized commuters) in 4 randomized spatiotemporal sampling points, for each of the 10 districts of Barcelona [25]. The TAPAS participants were aged 18-65 years, and they lived and worked or studied in Barcelona city (area: 102 km^2; population density: 15,687 persons/km^2). Of these 178 eligible participants, 8 were excluded because of incomplete or imprecise travel diaries, 6 because they did not wear either the mobile phone or the GPS tracker for at least 10 hours on any of the days with TAD information, and 2 because of technical problems with the GPS tracker, leaving 162 participants for analysis.

The study protocol was approved by the Clinical Research Ethical Committee of the Parc de Salut Mar (CEIC-Parc de Salut Mar), and written informed consent was obtained from all participants.

Instruments and Variables

Participants were instructed to wear a belt with a Samsung Galaxy Y S5360 mobile phone (Samsung Electronics Co Ltd, Suwon, South Korea) and a GlobalSat BT-335 GPS tracker (GlobalSat WorldCom Corp, Taipei, Taiwan) during waking hours for 7 consecutive days and to fill in a TAD for all of their trips throughout the day.

The Samsung Galaxy Y S5360 mobile phone was selected because it has a built-in accelerometer and GPS sensor, it was available in many countries, and it was among the cheapest mobile phones on the market when the study began. It uses Android 2.3.6 and operates with a Broadcom BCM21553 chipset and a BCM4751 GPS module. The Broadcom BCM4751 is a single-chip GPS receiver with 12 channels all-in-view tracking receiver [29]. Its position fix update rate is 1 second. Its accuracy is within 4.8 m for 95% of its measures. Its average reacquisition, warm start, and cold start are done in 1, 30, and 30 seconds, respectively. The mobile phones were equipped with SIM (subscriber identity module) cards with an Internet data allowance of 500 MB/month, an SD card, and an app called CalFit. The CalFit app was selected because it was free, open source, specifically developed for research purpose, and guaranteed the confidentiality of the information collected.

CalFit is a software for Android mobile phones developed by the University of California, Berkeley [30-33], which can be downloaded from the website edmunseto.com [34]. The software

XSL•FO

RenderX

uses mobile telecommunications technology (network), assisted GPS, and accelerometer sensors built in the mobile phones to record the time-resolved location and physical activity of participants [6]. The network uses Wi-Fi or cellular network signals to get the geolocation even when there is limited sky visibility. The assisted GPS technology makes use of remote GPS location servers to reduce power consumption and time to first fix position [14]. Each geographical coordinate recorded is differentiated in either the network or assisted GPS according to its origin, and its precision is estimated in meters (ie, accuracy). The accelerometer information is recorded at 10 Hz, in meters per second squared, and is transformed into physical activity intensity in metabolic equivalents (METs) per minute using validated equations [35]. From the physical activity intensity, it is produced the measure of moderate-to-vigorous physical activity duration, which has been shown interchangeable with the ActiGraph accelerometer [35]. For this study, CalFit was set to provide the geographical coordinates and physical activity intensity every 10 seconds.

Once data collection was completed, each geographical coordinate provided by the mobile phone was assigned to 1 of the 4 predefined microenvironments (home, work or school, in transit, and others) using a newly developed spatiotemporal map-matching algorithm. This map-matching algorithm was developed for this study because of the absence of available algorithms for postprocessing the clouds of geographical coordinates generated when participants are at a place. The chosen cutoff points are based on the extensive revision of the mobile phones' location data. In brief, the algorithm computes the azimuth between sequential coordinates and calculates the circular variance within groups of 30 coordinates in less than 100 m in linear distance. When the circular variance is greater than 0.7, the group of coordinates is identified as a potential place. Then, all coordinates within 30 minutes and 150 m are considered to belong to this spatiotemporal place. Finally, these spatiotemporal places are assigned to a specific microenvironment when distance between the group of coordinates and the geocoded microenvironment is less than 150 m. The rest of the groups of coordinates that do not belong to previous microenvironments are classified as other microenvironments and their central coordinates are calculated.

The GlobalSat BT-335 GPS tracker was selected because of its good performance in the study by Wu and colleagues [12]. This tracker uses the GPS chipset SiRFstarIII and 20 channels all-in-view tracking receiver [36]. Its position fix update rate is variable. Its accuracy is within 10 m for 95% of its measures. Its average reacquisition, warm start, and cold start are done in 0.1, 38, and 42 seconds, respectively. The tracker was calibrated before each deployment following user manual instructions and configured to provide data on date, time, geographical coordinates, speed, bearing, and altitude every 10 seconds. Finally, the TAD used for this study was similar to most travel logs used in transportation studies. It was previously pilot-tested within a convenience sample of 36 participants [6]. It includes questions on start and end time at minute resolution, travel modes, purpose, and destination address for each trip and monitors incidences for each day.

At the end of the study week participants returned the TAD, which was checked, day by day and trip by trip, ensuring that all trips and destinations and their durations and addresses were congruent, helping the participant to correct any illogical situation found. The main travel mode of all multimodal trips of the TAD (n=177 trips) was defined as the most motorized travel mode according to the following ranking: car> motorcycle> bus> metro> bicycle> walk. The geographical coordinates of both mobile phone–based CalFit and the GPS tracker that did not belong to the European continent or with a speed of ≥200 km/h were flagged. Finally, owing to schedule incompatibilities, not all participants were sampled for a week (3 had less than 7 days and 25 more than 7 days). As a result, the total number of monitored days was 1173. Among these, 187 (16%) days were excluded because of the following reasons: (1) the sensors were worn less than 10 hours during waking hours according to the wearing time estimates derived from CalFit physical activity measurements [37]; (2) there was incomplete TAD information; or (3) the participants reported issues with the sensors.

For the analysis, 2 datasets were generated. The first dataset was a trip-level spatiotemporal dataset, with the geographical coordinates of both mobile phone and GPS tracker at 10-second resolution for the episodes identified as trips by the TAD to compare their tracking performance and accuracy. Tracking performance of the trips reported in the TAD was measured by 2 dimensions, identifiability and traceability. Identifiability of TAD trips was defined as having ≥30% of trip duration with geolocation information because it was understood as the minimum cutoff point to distinguish between a real displacement and a measurement error. Traceability of TAD trips was quantified for each identifiable trip by the percentage of the trip duration with geolocation information. On the other hand, the tracking accuracy was quantified by the distance between the geographical coordinates of mobile phone and GPS tracker throughout TAD trips. This distance was calculated between concomitant locations (locations with a difference in time of <10 seconds between both monitors) and corrected for time difference and traveling speed. The second dataset was a microenvironment-level dataset, with information on whether participants were at home, work or school, in transit, or other locations at 1-minute resolution to assess the agreement and variability between map-matching algorithm and TAD.

Finally, other measurements to contextualize participants' characteristics and built environment around the home included sociodemographic characteristics (eg, age, sex, civil and working status), health status (the question "In general, would you say your health is: Excellent, Very Good, Good, Fair, or Poor" from the SF-36 Health Survey [38]), smoking habits, body mass index, main commute mode, and objectively measured social, physical, and built environment variables of a participant's residence or neighborhood (eg, deprivation index, population density, distance to work, altitude, slope, and walkability index). Details of these procedures have been previously published [28].

Statistical Analysis

To assess the validity of the tracking performance of the mobile phone, the identifiability and average traceability of the mobile phone for all trips and for each travel mode were compared with that of the GPS tracker using chi-square test and Kruskal-Wallis rank sum test, respectively. The validity of mobile phone tracking accuracy was assessed by the distance between the concomitant geographical coordinates of the mobile phone and GPS tracker. The tracking accuracy of each mobile phone location was overlapped on a Catalonia street map and a district map of Barcelona city to inspect the spatial coverage and distribution.

On the other hand, the validity of our map-matching algorithm to determine the time in each microenvironment (home, work or school, in transit, and others) was estimated by building a misclassification matrix versus the TAD. From this matrix, the sensitivity (recall), positive predictive value (precision), specificity, negative predictive value, F-score, and Gwet agreement coefficient (AC1) statistics were computed. F-score is the harmonic mean of recall and precision. The AC1 is similar to the multicategory kappa statistic but circumvents the known weakness of kappa [39]. In the AC1 calculation [AC1 = (p0 – pe)/(1 – pe)], "p0" is the concordance observed and "pe" is the concordance expected under the null hypothesis (no relationship).

Finally, two sensitivity analyses were performed. The first one was a comparison between the used geolocation accuracy (based on distance to GPS tracker) and the usual geolocation accuracy (based on distance to nearest street), using only a subset of mobile phone geolocations. The subset includes the geographical locations between the latitudes 41.59 and 41.62 and longitudes 2.605 and 2.645, which belong to the village of Sant Pol de Mar and its surroundings. In the second sensitivity analysis, we assessed the effect of participants' characteristics on the performance of our travel-activity algorithm and the need for specific calibration. The characteristics of participants studied

were as follows: (1) main travel mode for commuting; (2) weekdays versus weekend days; (3) median distance from home to work; and (4) working versus studying status. The effect of the characteristics on the performance of the algorithm was assessed by comparing the agreement between abovementioned characteristics using Kruskal-Wallis rank sum test.

All analyses were conducted during 2014-2015, using R 3.1.3 (The R Foundation for Statistical Computing), Python 2.7 (Python Software Foundation), NumPy $\geq$ 1.6.1 (Travis Oliphant), Pandas $\geq$ 0.12 (Wes McKinney), SQLite $\geq$ 3.7.13 (D. Richard Hipp), and SpatiaLite $\geq$ 4.0.0-RC1 (Alessandro Furieri).

Results

The 162 participants were on average 33 years old, 50% were female, 20% were single, 40% had at least 1 child, 77% were currently employed, and 50% were bicycle commuters (Table 1). Participants resided in densely populated neighborhoods (mean 30,319 persons/km^2) and had a relatively short commute distance (mean 3.4 km; Table 1). The total number of valid monitoring days was 986 days (out of 1173 possible days), which represented 3098 trips (mean 19 trips per participant). The average trip frequency was 3 per day, with each trip having an average duration of 28 minutes (Table 2). 85.99% (2633/3098) of these trips were made within Barcelona city (data not shown).

The mobile phone obtained locations for 905 (29%) more TAD trips than the GPS tracker ($P<.001$; Table 2). Their overall traceability, though, was comparable and the median distance between mobile phone and GPS tracker concomitant coordinates (or accuracy) was 24 m overall, 22 m using satellite signal, and 97 m using network signals (from the 1,294,805 compared geographical coordinates). Moreover, the stratified analysis of traceability of trips according to travel mode showed that the mobile phone had better traceability than the GPS tracker, with the exception of car trips.

XSL•FO

RenderX

Table 1. Description of participants' sociodemographic and home characteristics according to the main commute mode.

Characteristics	All Participants (N=162)	Bicycle (n=72)	Car, motorcycle, or bus (n=47)	Underground (n=43)
Sociodemographic[a], n (%)				
Sex, female	83 (51.2)	34 (47.2)	26 (55.3)	23 (53.5)
Age in years, median (25th-75th)	33 (26-41)	34 (29-41)	34 (28-42)	27 (21-39)
Civil status: single	29 (19.3)	17 (25.4)	5 (11.6)	7 (17.1)
Has at least 1 child: yes	59 (39.3)	22 (32.8)	22 (51.2)	15 (37.5)
Education level: more than secondary	100 (66.9)	54 (80.6)	25 (58.1)	21 (53.7)
Working status: yes	115 (76.8)	59 (88.1)	31 (72.1)	25 (63.4)
Nationality: Spanish	133 (88.7)	59 (88.1)	40 (93.0)	34 (85.4)
Smoking status: current smoker	42 (28.0)	21 (31.3)	15 (34.9)	6 (15.0)
Body mass index, ≥25	37 (24.5)	14 (20.9)	15 (34.9)	8 (19.5)
High stress level: yes[b]	33 (22.1)	10 (15.2)	12 (27.9)	11 (27.5)
Health status: very good or excellent	73 (48.3)	35 (52.2)	20 (46.5)	18 (43.9)
Built environment at home level, mean (SD)				
Deprivation index, z score	−0.2 (0.7)	−0.3 (0.7)	−0.1 (0.7)	0.0 (0.7)
Population density[a], persons/km^2	30295 (12212)	32002 (11372)	29575 (12753)	28175 (12859)
Distance to work, kilometers	3.4 (1.8)	2.8 (1.4)	3.3 (1.8)	4.6 (2.0)
Slope, %	4.0 (5.3)	3.4 (3.7)	4.5 (6.7)	4.4 (5.8)
Altitude, meters	41 (42.7)	37 (28.8)	44 (54.2)	44 (48.2)
Walkability index[a]	0.4 (2.1)	0.6 (2.1)	0.4 (2.1)	0.0 (2.0)

[a]Variables Age, Civil status, Has at least 1 child, Education level, Working status, Nationality, Smoking status, Body mass index and Health status have 12 missing values, High stress level has 14 missing, and Population density and Walkability index have 1 missing.

[b]High stress levels: having a score of ≥4 in each question of the short form of the Perceived Stress Scale [40].

Table 2. Comparison of Global Positioning System and mobile phone tracking performance and description of mobile phone tracking accuracy.

Measures	Travel mode from travel-activity diary							
	All	Motorcycle	Walk	Metro	Bicycle	Car	Bus	Others
Tracking performance								
Travel diary								
No. of trips, n	3098	358	706	581	839	409	199	6
Duration, minutes, mean (SD)	28 (12)	21 (9)	27 (13)	35 (11)	26 (11)	29 (13)	32 (11)	32 (11)
Identifiability[a]								
GPS[b] logger, n (%)	1803 (58.2)	280 (78.2)	409 (57.9)	161 (27.7)	574 (68.4)	257 (62.8)	121 (60.8)	1 (16.7)
Mobile phone, n (%)	2708 (87.4)	311 (86.9)	623 (88.2)	511 (88.0)	766 (91.3)	320 (78.2)	172 (86.4)	5 (83.3)
P value	<.001	<.001	<.001	<.001	<.001	<.001	<.001	.08
Traceability[c]								
GPS logger, median (25th-75th)	74 (55-88)	80 (65-94)	74 (55-88)	44 (36-53)	76 (61-90)	79 (61-89)	74 (54-87)	86 (86-86)
Mobile phone, median (25th-75th)	76 (58-90)	77 (60-92)	79 (60-91)	53 (47-62)	85 (71-95)	64 (46-83)	65 (51-84)	87 (87-87)
P value	.009	.60	.009	<.001	<.001	<.001	.05	.32
Tracking accuracy[d]								
Overall, m, median (25th-75th)	24 (10-51)	22 (11-47)	21 (9-46)	22 (11-43)	23 (10-49)	34 (13-76)	25 (12-51)	12 (4-29)
Satellite, m, median (25th-75th)	22 (10-47)	21 (10-43)	20 (8-43)	21 (10-40)	22 (9-45)	29 (11-61)	23 (11-45)	12 (4-28)
Network, m, median (25th-75th)	97 (26-574)	66 (22-290)	42 (19-183)	104 (29-609)	69 (23-223)	464 (80-2311)	54 (22-372)	177 (104-242)

[a]Identifiability of travel-activity diary trips was defined as having ≥30% of trip duration with location information.

[b]GPS: Global Positioning System.

[c]Traceability of travel-activity diary trips was quantified among the identifiable trips by the percentage of the trip duration with location information.

[d]Tracking accuracy was quantified by the distance between the geographical coordinates of the mobile phone and the GPS tracker throughout travel-activity diary trips. This distance was calculated between concomitant geographical coordinates (geolocations with a difference in time of <10 seconds between both monitors) and corrected for time difference and traveling speed. Overall includes satellite and network locations, while satellite and network refer to the specific accuracy for each signal.

Figure 1 shows the spatial coverage of all sampled trips and how the distance between mobile phone and GPS tracker is greater for the intercity trips. Moreover, the detailed map of Barcelona districts shows that the accuracy of the mobile phone while traveling is almost equal across districts with the exception of Nou Barris district. Figure 2 shows that the distances to the street were less, median 8.7 m (25th-75th, 4.9-17.6 m), compared with the distances between concomitant coordinates, median 46.3 m (25th-75th, 32.7-62.5 m).

The comparison of the overall time in each microenvironment between map-matching algorithm and TAD showed that there is overall a good agreement on time spent in microenvironments,

with only 0.1% (work) to 1.2% (other) difference estimated in each type of microenvironment.

The confusion matrix (Table 3) between map-matching algorithm and TAD showed an overall accuracy of 83%. Moreover, the map-matching algorithm in comparison with TAD was able to properly identify the minutes spent at home (recall 94% and precision 93%), at work (recall 85% and precision 90%), in transit (recall 61% and precision 55%), and at other places (recall 61% and precision 64%). Furthermore, sensitivity analyses of the microenvironment agreement between map-matching algorithm and TAD did not show significant differences in any of the subpopulations (data not shown).

Table 3. Travel-activity confusion matrix between the travel-activity diary and our map-matching algorithm and its interrelationship statistics (cluster defined as 150 m; Gwet agreement coefficient AC1=81%).

Map-matching algorithm	Travel-activity diary				Sens[a] %	Spec[b] %	PPV[c] %	NPV[d] %	ACC[e] %	F score %
	Home	Work	Others	Trip						
Home	758495	5235	29575	26580	94	90	93	91	92	93
Work	10320	257200	10110	9140	85	97	90	96	95	87
Others	20575	23000	104315	15485	61	95	64	95	91	62
Trip	20755	15880	27445	79160	61	95	55	96	92	58

[a]Sens: sensitivity.

[b]Spec: specificity.

[c]PPV: positive predictive value.

[d]NPV: negative predictive value.

[e]ACC: Accuracy

Figure 1. Catalonia street map and Barcelona district map with the spatial distribution of the mobile phone tracking accuracy among the 986 person-days monitored. Gray points represent those locations without concomitant locations from Global Positioning System (GPS) tracker to estimate accuracy. In the district map of Barcelona (inset), the median geolocation accuracy of the mobile phone is shown in the 10 districts of Barcelona city.
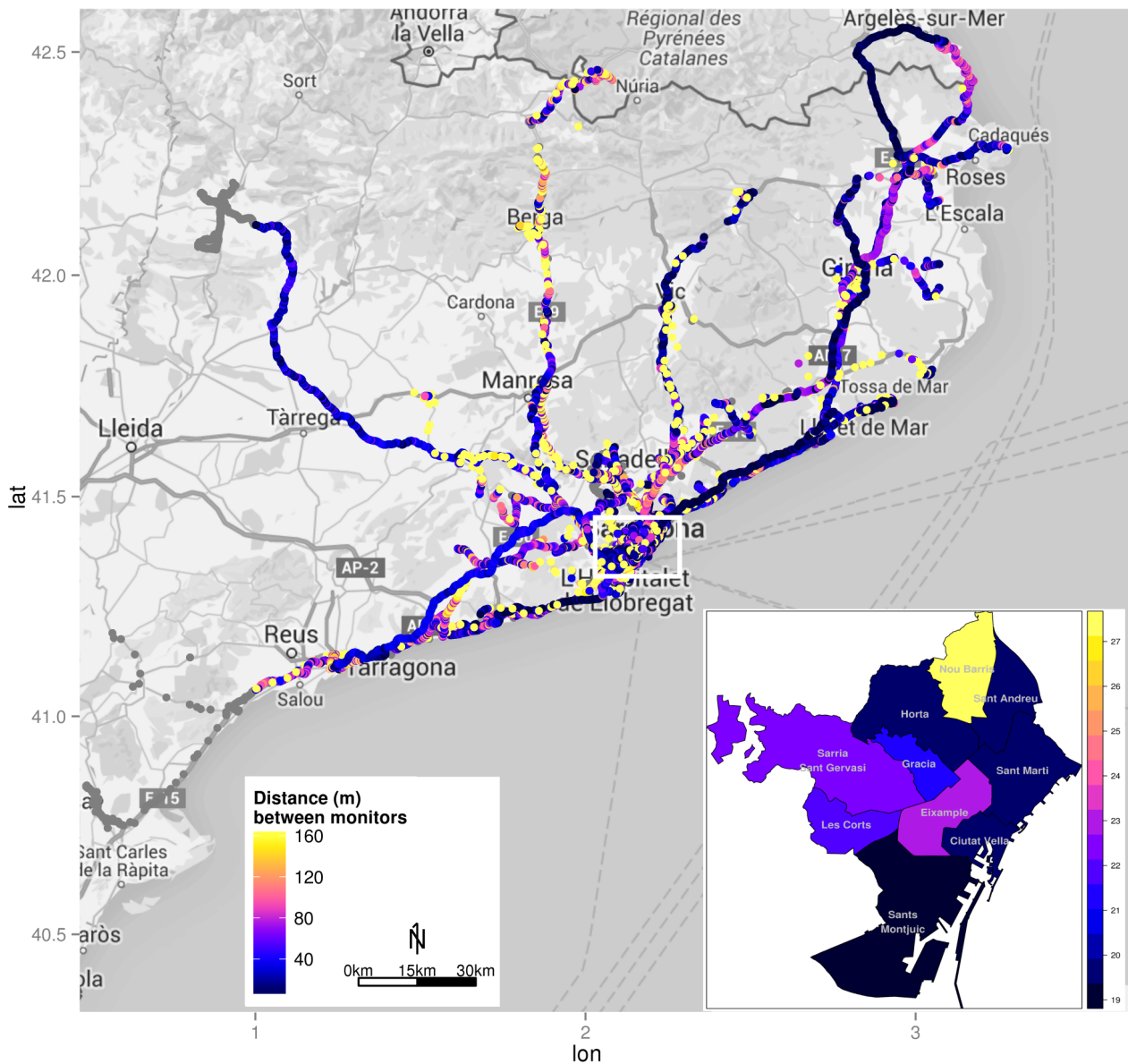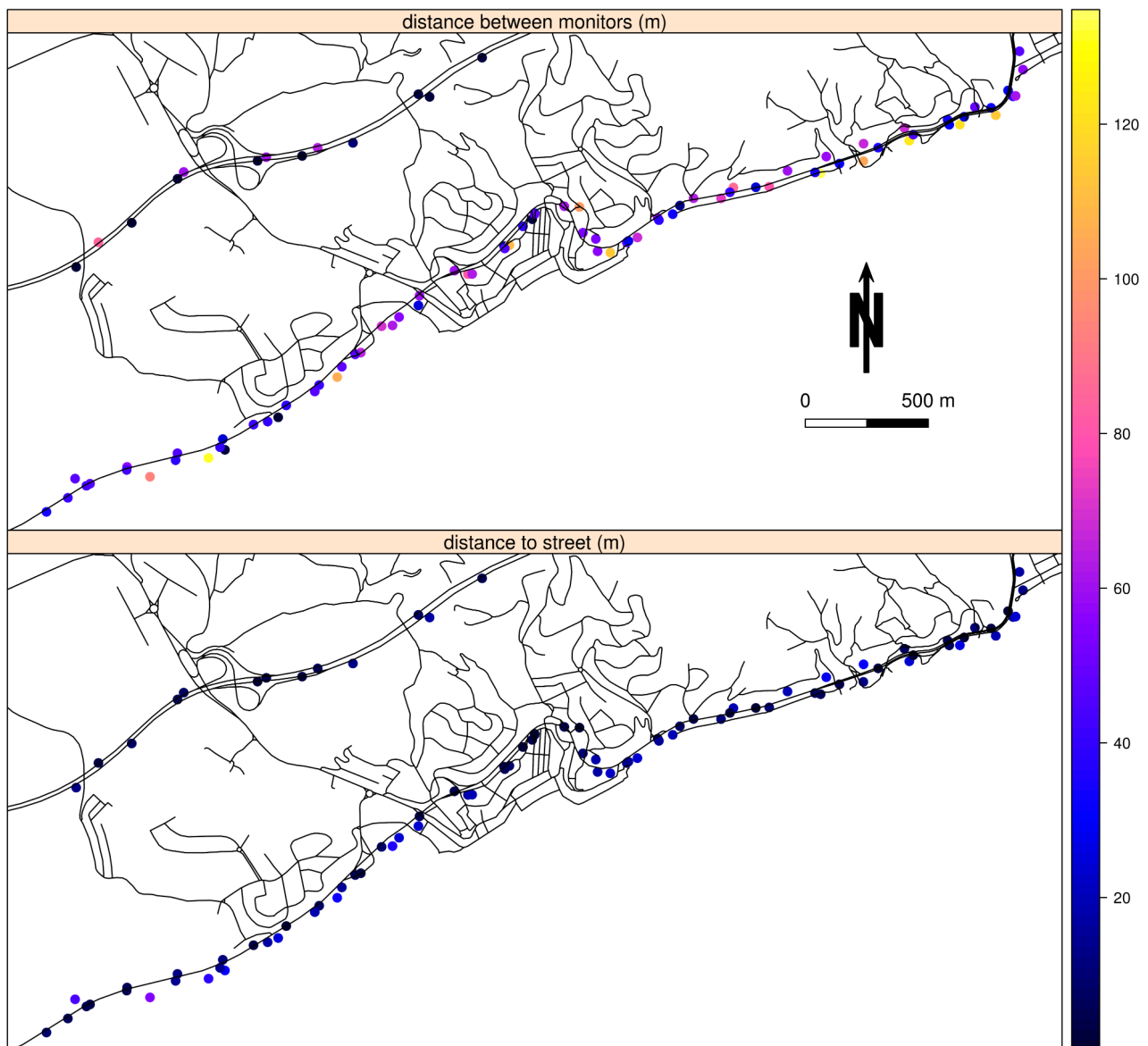
Figure 2. Comparison of distances between smartphone and Global Positioning System (GPS) tracker concomitant locations and to nearest street, while travelling, through the highway C-32 and N-II, at the surroundings of the village of Sant Pol de Mar.

Discussion

Principal Findings

The main findings of this study are that (1) the mobile phone obtained locations for 905 (29%) more trips than a commercial GPS tracker; (2) mobile phone had enough geolocation accuracy to locate the participants at the street level; and (3) the developed map-matching algorithm was able to determine people's travel-activity pattern with an overall accuracy of 83% and in-transit time with a recall of 61% and precision of 55%.

Comparison With Previous Studies

Tracking Performance and Accuracy of Mobile Phone–Based CalFit

To our knowledge, this is the first study describing and comparing tracking performance and accuracy between a mobile phone and a GPS tracker in free-living conditions. Previous studies were mainly focused on evaluating the geolocation accuracy of mobile phones and were conducted in more car-dependent environments and through experimental designs [12,14,17,18,41]. These previous experimental designs tested the accuracy of the location sensor mainly in favorable environments, such as long unimodal trips, constant speed trips, big and wide streets (eg, Interstate 4 or 5, US Highway 301, or downtown Los Angeles), and used accessories to facilitate signal acquisition (eg, cars with roof carrier to hold mobile phone) [12,18].

In this deployment, the mobile phone–based CalFit obtained locations for 905 (29%) more trips than the GPS tracker. According to previous literature, this could be a consequence of the faster time to first fix position and the use of network positioning systems [12]. On the other hand, the traceability of trips with the mobile phone–based CalFit was almost double the traceability found by Michael and colleagues [17] in their scripted study with the Motorola i760 mobile phone (59% vs 35%). Moreover, in contrast to our results, Michael and colleagues [17] also found that traceability was higher when

participants traveled by car, whereas in our study traveling by car was the least traceable travel mode. This difference could be explained because Michael and colleagues had more car measurements in suburban areas, where there are lower barriers for satellite coverage if compared with urban or forest areas. However, we cannot disentangle to what extent these differences could be a result of the placement of the device [42] (in this study, mobile phones were worn on the lower abdomen and the GPS tracker on the left hip), signal problem (lack of Wi-Fi signal during intercity trips), or a real mobile phone technical weakness (such as limitations on GPS antenna, digital interface, or GPS chipset built in the mobile phone).

The geolocation accuracy of mobile phones using only satellite signal in previous dynamic experimental studies was between 2 m and 8 m [12,18], which is more precise than the 22 m found in this study. However, previous studies have assessed only the horizontal error, and because it is known that horizontal error could be different from vertical error depending on satellite distribution [43], differences encountered should be interpreted with caution. Moreover, our results are consistent with Duncan and colleagues' [41] finding regarding the accuracy of GPS trackers in mixed use environments (mean 21 m). On the other hand, the accuracy reached by the mobile phones' network signal in this study, median 97 m (25th-75th, 26-574 m), is comparable to that found by Zandbergen (74 m) [14]. This could be due to the high density of wireless network access points in Barcelona city (>600 access points only from municipal network) [44]. Furthermore, it is worth noting that the sensitivity analysis on the spatial accuracy (Figure 2) has shown that differences between sensors are more a matter of time alignment or data acquisition rather than spatial error on the coordinates.

Travel-Activity Determination Performance

This is the first study to monitor a large sample of adults during a full week while they are performing real-life activities using mobile phone technology. Previous studies mainly focused on commercial GPS trackers and experimental or quasi-experimental designs (Multimedia Appendix 1). Moreover, the built environment conditions of our study area add an extra challenge, when compared with previous study areas (Multimedia Appendix 1), because Barcelona is a smaller city for daily commuting (102 km^2), has higher population density (15,686 persons/km^2), and lower prevalence of private motorized transport use (15%). In addition, previous studies did not provide information about travel behavior of participants (frequency, duration, travel mode, or number of other places visited per day). Finally, the definitions of "in-transit" and "other" microenvironments are not consistent across previous studies. Briefly, "in-transit" microenvironments have been defined in two different approaches: a comprehensive approach, which includes all travel modes (car, bus, metro, motorcycle, bicycle, foot, or skate), and a more restrictive approach, which only includes in-vehicle and/or walking modes (the category "others" then includes the rest of the travel modes).

The performance of the map-matching algorithm to determine the time spent at home or work has been shown to be very sensitive and precise, which is consistent with previous research [42,45-47]. Another relevant point in this study is the

confirmation that a combined use of the mobile phone–based CalFit and the map-matching algorithm provides a better performance to identify the in-transit microenvironments than previous approaches using only GPS trackers [45,46]. However, our in-transit results, which include all travel modes, are still poor in comparison with those focused on the restrictive definition of in-transit microenvironment (ie, mainly in-vehicle mode) [42,47]. On the other hand, and in agreement with previous literature [41], it is still a challenge to distinguish between places that are very close to each other and to detect very short trips (eg, <10 minutes). Both are very important challenges that need to be addressed carefully by researchers during the study design process because they depend on the urban design of each city and activity pattern of the population.

Strength and Limitations

The use of the GlobalSat BT-335 as the GPS tracker, which was found by Wu and colleagues [12] to be among the faster devices in terms of time to first fix and with better geolocation accuracy, reinforces the internal validity of our results. Moreover, the extensive and heterogeneous sample of real-life trips, composed by commutes within the very dense city of Barcelona and leisure trips out of the city, shows the robustness of the external validity of our findings. Therefore, and because this study used one of the least expensive mobile phones on the market, in one of the most dense and complex built environments of Europe, one would expect to find similar or even better results with higher range of mobile phones or in less dense cities, which would confirm mobile phones as the reference tool for personal exposure research. Despite the encouraging findings of this study, caution is required until future multicenter studies engaged in different cities replicate these findings with different populations and other mobile phones and settings (ie, other urban design environments or environments with less dense Wi-Fi access points).

The interpretation of the results on tracking performance of TAD trips by mobile phone–based CalFit calls for prudence because this tracking definition is based on the percentage of the trip duration with location information, which does not take into account geolocation accuracy. On the other hand, the present assessment of geolocation accuracy is based on the comparison against a GPS tracker, and it is well known that GPS trackers are affected by environmental factors (ie, visibility and geometry of satellites) [43,48]. So, the distance between mobile phone–based CalFit and GPS tracker may not reflect the lack of geolocation accuracy of mobile phone when the geometry and visibility of satellites were challenging [14,49], but this approach was the most feasible because the 3098 unscripted trips assessed. Furthermore, because of the comparison between the distances to GPS tracker and the distances to nearest street, we know that the lack of accuracy was more a matter of data acquisition than spatial error on the coordinates. On the other hand, the use of the TAD, as a reference value for participants' travel-activity pattern, could have penalized the recall and precision of the map-matching algorithm because of the recall and response biases. Finally, another limitation of this study was not having collected the information regarding the number of satellites in view, the number of satellites used, and the horizontal dilution of precision

(HDOP) from the satellite signal in order to improve the recall in the detection of the transitions between microenvironments.

Applicability and Future Developments

The mobile phone–based CalFit, together with our map-matching algorithm, provides a clean tracking of people's activities, which provides researchers with the opportunity to determine and understand the causal and temporal relationship of natural and urban environments with health-related behaviors and exposures as well as physical and mental health conditions. Moreover, this study is the basis for future studies aiming to assess if this map-matching algorithm of mobile phone geolocation shows the same feasibility and precision in other built environments.

Finally, future improvements in personal monitoring must include making the apps downloadable from the Internet and transferring the measurements through the Internet directly to a cloud server, which we believe will minimize efforts during the deployment and the burden on participants and will increase participants' compliance. Furthermore, future developments should also add automatic algorithms for travel mode recognition and outdoor time determination, probably using additional recorded information from location provider (ie, number of satellites in view, number of satellites used during location determination, and HDOP) and from other mobile phone built-in sensors (ie, barometer and light and sound sensors).

Therefore, the use of mobile phones running the CalFit app provides better information on which microenvironments people spend their time in than previous approaches based only on GPS trackers. The improvements of mobile phone technology in microenvironment determination are because the mobile phones are faster at identifying first locations and capable of getting location in challenging environments thanks to the combination of assisted-GPS technology and network positioning systems. Moreover, collecting location information from mobile phones, which are already carried by individuals, allows monitoring more people with a cheaper and less burdensome method.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Comparison of sample, monitoring duration, setting, travel behavior, and time-activity microenvironments definition across studies focused on time-activity pattern.

[PDF File (Adobe PDF File), 38KB - mhealth_v4i4e126_app1.pdf]

References

1. Forouzanfar MH, Alexander L, Anderson HR, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks in 188 countries, 1990–2013: a systematic analysis for the Global Burden of Disease Study 2013. The Lancet 2015 Nov;386(10010):2287-2323. [doi: 10.1016/S0140-6736(15)00128-2] [Medline: 26364544]
2. Jerrett M, Arain A, Kanaroglou P, Beckerman B, Potoglou D, Sahsuvaroglu T, et al. A review and evaluation of intraurban air pollution exposure models. J Expo Anal Environ Epidemiol 2005 Mar;15(2):185-204. [doi: 10.1038/sj.jea.7500388] [Medline: 15292906]
3. Avery CL, Mills KT, Williams R, McGraw KA, Poole C, Smith RL, et al. Estimating error in using ambient PM2.5 concentrations as proxies for personal exposures: a review. Epidemiology 2010 Mar;21(2):215-223. [doi: 10.1097/EDE.0b013e3181cb41f7] [Medline: 20087191]
4. Meng QY, Svendsgaard D, Kotchmar D, Pinto JP. Associations between personal exposures and ambient concentrations of nitrogen dioxide: A quantitative research synthesis. Atmospheric Environment 2012 Sep;57:322-329. [doi: 10.1016/j.atmosenv.2012.04.035]
5. Nieuwenhuijsen MJ, Donaire-Gonzalez D, Rivas I, de Castro M, Cirach M, Hoek G, et al. Variability in and agreement between modeled and personal continuously measured black carbon levels using novel smartphone and sensor technologies. Environ Sci Technol 2015 Mar 3;49(5):2977-2982. [doi: 10.1021/es505362x] [Medline: 25621420]

6. de Nazelle A, Seto E, Donaire-Gonzalez D, Mendez M, Matamala J, Nieuwenhuijsen MJ, et al. Improving estimates of air pollution exposure through ubiquitous sensing technologies. Environ Pollut 2013 May;176:92-99 [FREE Full text] [doi: 10.1016/j.envpol.2012.12.032] [Medline: 23416743]

7. Dons E, Int Panis L, Van Poppel M, Theunis J, Wets G. Personal exposure to Black Carbon in transport microenvironments. Atmospheric Environment 2012 Aug;55:392-398. [doi: 10.1016/j.atmosenv.2012.03.020]

8. Dons E, Int Panis L, Van Poppel M, Theunis J, Willems H, Torfs R, et al. Impact of time-activity patterns on personal exposure to black carbon. Atmospheric Environment 2011 Jul;45(21):3594-3602. [doi: 10.1016/j.atmosenv.2011.03.064]

9. Buonanno G, Stabile L, Morawska L, Russi A. Children exposure assessment to ultrafine particles and black carbon: The role of transport and cooking activities. Atmospheric Environment 2013 Nov;79:53-58. [doi: 10.1016/j.atmosenv.2013.06.041]

10. Vrijheid M. The exposome: a new paradigm to study the impact of environment on health. Thorax 2014 Sep;69(9):876-878. [doi: 10.1136/thoraxjnl-2013-204949] [Medline: 24906490]

11. Krenn PJ, Titze S, Oja P, Jones A, Ogilvie D. Use of global positioning systems to study physical activity and the environment: a systematic review. Am J Prev Med 2011 Nov;41(5):508-515 [FREE Full text] [doi: 10.1016/j.amepre.2011.06.046] [Medline: 22011423]

12. Wu J, Jiang C, Liu Z, Houston D, Jaimes G, McConnell R. Performances of different global positioning system devices for time-location tracking in air pollution epidemiological studies. Environ Health Insights 2010;4:93-108 [FREE Full text] [doi: 10.4137/EHI.S6246] [Medline: 21151593]

13. Pratt M, Sarmiento OL, Montes F, Ogilvie D, Marcus BH, Perez LG, Lancet Physical Activity Series Working Group. The implications of megatrends in information and communication technology and transportation for changes in global physical activity. Lancet 2012 Jul 21;380(9838):282-293. [doi: 10.1016/S0140-6736(12)60736-3] [Medline: 22818940]

14. Zandbergen PA. Accuracy of iPhone Locations: A Comparison of Assisted GPS, WiFiCellular Positioning. Transactions in GIS 2009;13:5-25. [doi: 10.1111/j.1467-9671.2009.01152.x]

15. Liu HY, Skjetne E, Kobernus M. Mobile phone tracking: in support of modelling traffic-related air pollution contribution to individual exposure and its implications for public health impact assessment. Environ Health 2013;12:93 [FREE Full text] [doi: 10.1186/1476-069X-12-93] [Medline: 24188173]

16. Su JG, Jerrett M, Meng YY, Pickett M, Ritz B. Integrating smart-phone based momentary location tracking with fixed site air quality monitoring for personal exposure assessment. Sci Total Environ 2015 Feb 15;506-507:518-526. [doi: 10.1016/j.scitotenv.2014.11.022] [Medline: 25437768]

17. Michael Y, McGregor EM, Allen J, Fickas S. Observing Outdoor Activity Using Global Positioning System-Enabled Cell Phones. In: Helal S, Mitra S, Wong J, Chang CK, Monhtari M, editors. Smart Homes and Health Telematics. Berlin: Springer Berlin Heidelberg; 2008:177-184.

18. Zandbergen PA, Barbeau SJ. Positional Accuracy of Assisted GPS Data from High-Sensitivity GPS-enabled Mobile Phones. J. Navigation 2011 Jun 7;64(03):381-399. [doi: 10.1017/S0373463311000051]

19. Garnett R, Stewart R. Comparison of GPS units and mobile Apple GPS capabilities in an urban landscape. Cartography and Geographic Information Science 2014 Oct 28;42(1):1-8. [doi: 10.1080/15230406.2014.974074]

20. Mavoa S, Oliver M, Witten K, Badland HM. Linking GPS and travel diary data using sequence alignment in a study of children's independent mobility. Int J Health Geogr 2011;10:64 [FREE Full text] [doi: 10.1186/1476-072X-10-64] [Medline: 22142322]

21. Kang B, Moudon AV, Hurvitz PM, Reichley L, Saelens BE. Walking objectively measured: classifying accelerometer data with GPS and travel diaries. Med Sci Sports Exerc 2013 Jul;45(7):1419-1428 [FREE Full text] [doi: 10.1249/MSS.0b013e318285f202] [Medline: 23439414]

22. Hurvitz PM, Moudon AV. Home versus nonhome neighborhood: quantifying differences in exposure to the built environment. Am J Prev Med 2012 Apr;42(4):411-417 [FREE Full text] [doi: 10.1016/j.amepre.2011.11.015] [Medline: 22424255]

23. Hess B, Farahani A, Tschirschnitz F, von RF. Evaluation of Fine-granular GPS Tracking on Smartphones. In: Proceedings of the First ACM SIGSPATIAL International Workshop on Mobile Geographic Information Systems. New York, NY, USA: ACM; 2012:33-40.

24. TAPAS. Transportation, Air pollution and Physical ActivitieS: An integrated health risk assessment programme of climate change and urban policies URL: http://www.tapas-program.org/ [accessed 2016-10-31] [WebCite Cache ID 6leHYuvhY]

25. Cole-Hunter T, Donaire-Gonzalez D, Curto A, Ambros A, Valentin A, Garcia-Aymerich J, et al. Objective correlates and determinants of bicycle commuting propensity in an urban environment. Transportation Research Part D: Transport and Environment 2015 Oct;40:132-143. [doi: 10.1016/j.trd.2015.07.004]

26. Curto A, de Nazelle A, Donaire-Gonzalez D, Cole-Hunter T, Garcia-Aymerich J, Martínez D, et al. Private and public modes of bicycle commuting: a perspective on attitude and perception. Eur J Public Health 2016 Aug;26(4):717-723. [doi: 10.1093/eurpub/ckv235] [Medline: 26748099]

27. Braun LM, Rodriguez DA, Cole-Hunter T, Ambros A, Donaire-Gonzalez D, Jerrett M, et al. Short-term planning and policy interventions to promote cycling in urban centers: Findings from a commute mode choice analysis in Barcelona, Spain. Transportation Research Part A: Policy and Practice 2016 Jul;89:164-183. [doi: 10.1016/j.tra.2016.05.007]

XSL•FO

RenderX

28. Donaire-Gonzalez D, de Nazelle A, Cole-Hunter T, Curto A, Rodriguez DA, Mendez MA, et al. The Added Benefit of Bicycle Commuting on the Regular Amount of Physical Activity Performed. Am J Prev Med 2015 Dec;49(6):842-849. [doi: 10.1016/j.amepre.2015.03.036] [Medline: 26228005]

29. GPSworld. GPS World Receiver Survey 2013 URL: http://gpsworld.com/resources/gps-world-receiver-survey/ [accessed 2016-07-19] [WebCite Cache ID 6j7S5AwVi]

30. Gravina R, Alessandro A, Salmeri A, Buondonno L, Raveendranathan N, Loseu V, et al. Enabling Multiple BSN Applications Using the SPINE Framework. Wearable and Implantable Body Sensor Networks, International Workshop on 2010:228-233. [doi: 10.1109/BSN.2010.34]

31. Kuryloski P, Giani A, Giannantonio R, Gilani K, Gravina R, Seppa V, et al. DexterNet: An Open Platform for Heterogeneous Body Sensor Networksits Applications. In: Proceedings - 2009 6th International Workshop on Wearable and Implantable Body Sensor Networks, BSN 2009. Berkeley: IEEE Computer Society; 2009:92-97.

32. Seto E, Martin E, Yang A, Yan P, Gravina R, Lin I, et al. Opportunistic strategies for lightweight signal processing for body sensor networks. In: Proceedings of the 3rd International Conference on PErvasive Technologies Related to Assistive Environments. New York, NY, USA: ACM; 2010:1-6.

33. Yan P, Lin I, Roy M, Seto E, Wang C, Bajcsy R. WAVE and CalFit - Towards social interaction in mobile body sensor networks. 2010 Presented at: 2010 The 5th Annual ICST Wireless Internet Conference; 2010; Singapore.

34. Edmunseto. CalFit for the European PHENOTYPE Study URL: http://www.edmundseto.com/website/researchblog/?P=781 [accessed 2016-10-31] [WebCite Cache ID 6leItrlg1]

35. Donaire-Gonzalez D, de Nazelle A, Seto E, Mendez M, Nieuwenhuijsen MJ, Jerrett M. Comparison of physical activity measures using mobile phone-based CalFit and Actigraph. J Med Internet Res 2013;15(6):e111 [FREE Full text] [doi: 10.2196/jmir.2470] [Medline: 23896156]

36. Globalsat. GlobalSat Wired by Wireless URL: http://www.globalsat.co.uk/product_pages/product_bt335.htm [accessed 2016-07-19] [WebCite Cache ID 6j7DGuzIj]

37. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sports Exerc 2011 Feb;43(2):357-364 [FREE Full text] [doi: 10.1249/MSS.0b013e3181ed61a3] [Medline: 20581716]

38. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992 Jun;30(6):473-483. [Medline: 1593914]

39. Wongpakaran N, Wongpakaran T, Wedding D, Gwet KL. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. BMC Med Res Methodol 2013;13:61 [FREE Full text] [doi: 10.1186/1471-2288-13-61] [Medline: 23627889]

40. Psy.cmu. 1988. Perceived stress in a probability sample of the United States URL: http://www.psy.cmu.edu/~scohen/Cohen,%20S.%20&%20Williamson,%20G.%20(1988).pdf

41. Duncan S, Stewart TI, Oliver M, Mavoa S, MacRae D, Badland HM, et al. Portable global positioning system receivers: static validity and environmental conditions. Am J Prev Med 2013 Feb;44(2):e19-e29. [doi: 10.1016/j.amepre.2012.10.013] [Medline: 23332343]

42. Breen MS, Long TC, Schultz BD, Crooks J, Breen M, Langstaff JE, et al. GPS-based microenvironment tracker (MicroTrac) model to estimate time-location of individuals for air pollution exposure assessments: model evaluation in central North Carolina. J Expo Sci Environ Epidemiol 2014 Jul;24(4):412-420 [FREE Full text] [doi: 10.1038/jes.2014.13] [Medline: 24619294]

43. Schipperijn J, Kerr J, Duncan S, Madsen T, Klinker CD, Troelsen J. Dynamic Accuracy of GPS Receivers for Use in Health Research: A Novel Method to Assess GPS Accuracy in Real-World Settings. Front Public Health 2014;2:21 [FREE Full text] [doi: 10.3389/fpubh.2014.00021] [Medline: 24653984]

44. BCN. Barcelona Wifi | El web de la ciutat de Barcelona URL: http://www.bcn.cat/barcelonawifi/es/ [accessed 2016-03-15] [WebCite Cache ID 6g1WE2eri]

45. Elgethun K, Yost MG, Fitzpatrick CT, Nyerges TL, Fenske RA. Comparison of global positioning system (GPS) tracking and parent-report diaries to characterize children's time-location patterns. J Expo Sci Environ Epidemiol 2007 Mar;17(2):196-206. [doi: 10.1038/sj.jes.7500496] [Medline: 16773123]

46. Nethery E, Mallach G, Rainham D, Goldberg MS, Wheeler AJ. Using Global Positioning Systems (GPS) and temperature data to generate time-activity classifications for estimating personal exposure in air monitoring studies: an automated method. Environ Health 2014;13(1):33 [FREE Full text] [doi: 10.1186/1476-069X-13-33] [Medline: 24885722]

47. Wu J, Jiang C, Houston D, Baker D, Delfino R. Automated time activity classification based on global positioning system (GPS) tracking data. Environ Health 2011;10:101 [FREE Full text] [doi: 10.1186/1476-069X-10-101] [Medline: 22082316]

48. Beekhuizen J, Kromhout H, Huss A, Vermeulen R. Performance of GPS-devices for environmental exposure assessment. J Expo Sci Environ Epidemiol 2013;23(5):498-505. [doi: 10.1038/jes.2012.81] [Medline: 22829049]

49. AlSharabati T, Chen Y. Statistical Approach to Mitigating 3G Interference to GPS in 3G Handset. IJCNS 2010;03(09):730-736. [doi: 10.4236/ijcns.2010.39097]

XSL•FO
RenderX

Abbreviations

GPS: Global Positioning System
HDOP: horizontal dilution of precision
TAD: travel-activity diary
TAPAS: Transportation, Air Pollution and Physical Activities

XSL·FO
RenderX

Short Paper

Using Knowledge Translation to Craft "Sticky" Social Media Health Messages That Provoke Interest, Raise Awareness, Impart Knowledge, and Inspire Change

Sanchia Shibasaki[1*], PhD; Karen Gardner[2*], PhD; Beverly Sibthorpe[3*], PhD

[1]ThinkThrough Consultancy Services, Holland Park, Australia

[2]Centre for Primary Health Care and Equity, University of New South Wales, Sydney, Australia

[3]Consultancy, Port Macquarie, Australia

[*]all authors contributed equally

Corresponding Author:
Sanchia Shibasaki, PhD
ThinkThrough Consultancy Services
P O Box 7083
Holland Park, 4121
Australia
Phone: 61 0447040224
Fax: 61 0447040224
Email: sanchia.shibasaki@gmail.com

Abstract

Background: In Australia, there is growing use of technology supported knowledge translation (KT) strategies such as social media and mobile apps in health promotion and in Indigenous health. However, little is known about how individuals use technologies and the evidence base for the impact of these health interventions on health behavior change is meager.

Objective: The objective of our study was to examine how Facebook is used to promote health messages to Indigenous people and discuss how KT can support planning and implementing health messages to ensure chosen strategies are fit for the purpose and achieve impact.

Methods: A desktop audit of health promotion campaigns on smoking prevention and cessation for Australian Indigenous people using Facebook was conducted.

Results: Our audit identified 13 out of 21 eligible campaigns that used Facebook. Facebook pages with the highest number of likes (more than 5000) were linked to a website and to other social media applications and demonstrated stickiness characteristics by posting frequently (triggers and unexpected), recruiting sporting or public personalities to promote campaigns (social currency and public), recruiting Indigenous people from the local region (stories and emotion), and sharing stories and experiences based on real-life events (credible and practical value).

Conclusions: KT planning may support campaigns to identify and select KT strategies that are best suited and well-aligned to the campaign's goals, messages, and target audiences. KT planning can also help mitigate unforeseen and expected risks, reduce unwarranted costs and expenses, achieve goals, and limit the peer pressure of using strategies that may not be fit for purpose. One of the main challenges in using KT systems and processes involves coming to an adequate conceptualization of the KT process itself.

Introduction

In Australia, the use of technology supported knowledge translation (KT) strategies like social media, mobile software apps, patient-mediated tools, and clinical decision support systems in health promotion and in Aboriginal and Torres Strait Islander (hereafter respectfully referred to as Indigenous) health is growing [1,2]. This corresponds with a growing use of social

media among Australians in general, and in particular among Indigenous people, whose use of Facebook is 20% higher than the national average [3].

Despite this growing use, little is known about how individuals use technologies and evidence of the benefit and impact of these social media applications on health behavior change is meager. In their review of use in Indigenous populations, Brusse and colleagues found that the benefit and impact of social media applications was tentative and scattered, suggesting that producers of health promotion projects needed to obtain a thorough understanding about who engages with these strategies, why they engage, and how they engage [2]. The authors recommended further research in KT and implementation to better understand how to translate principles of commercial success in social media and mobile software into effective health promotion interventions and how to better integrate these methods into health research [2].

KT in health is defined as a "dynamic and iterative process that includes synthesis, dissemination, exchange, and ethically sound application of knowledge to improve health, provide more effective health services and products, and strengthen the health care system" [4]. Using KT strategies to support the design and deployment of health technologies is likely to increase their effectiveness and facilitate more efficient use of resources. However, Goering and colleagues suggest that whereas researchers and others are being encouraged to incorporate KT activities and strategies into their research applications, many are unclear about precisely what this means or how it should be assessed [5].

Three applications of KT, the Barwick KT planning Template [6] and the simple, unexpected, concrete, emotional, stories (SUCCESS) and social currency, triggers, emotion, public, practical value, stories (STEPPS) frameworks [7,8] are processes that have been developed to identify and shape KT strategies so they are fit for purpose for a particular context and a defined audience and for achieving a set of goals and impact. The Barwick Template incorporates a set of guiding questions and evidence-based checklists and refers to other KT models and frameworks that are associated with planning for implementation (eg, Consolidated Framework for Implementation Research, Knowledge to Action, Reach Effectiveness Adoption Implementation Maintenance RE-AIM) and for impact (eg, stickiness frameworks such as SUCCESS and STEPPS, and planning for evaluation eg, measurement indicators) [6]. The SUCCESS and STEPPS frameworks support planning for impact [7,8]. A health message or strategy should have impact such that it "catches on" or is understandable, memorable, and effective in changing thought or behavior. These characteristics are known as stickiness factors (Table 1) [7,8].

In this study, we examined how Facebook is used to promote health messages to Indigenous people on tobacco smoking prevention and cessation. We discuss how KT planning can support teams to plan, develop, and implement health messages to ensure chosen strategies are fit for purpose and designed to achieve impact. Although this study focussed on health promotion campaigns about tobacco use for Indigenous people, the outcomes are generalizable to other types of health campaigns and topics.

Methods

Data Collection

In 2015, a desktop audit was undertaken of the Indigenous HealthInfoNet health promotion resource database for tobacco campaigns published from 2005 to 2015. Campaigns were selected for inclusion in the study if they addressed smoking cessation and prevention, were nonpaper-based with an associated social media presence, and were audiovisual. Social media presence was confirmed by a search using Google Chrome (Google, Mountain View, CA, USA), Facebook (Facebook, Inc, Menlo Park, CA, USA), YouTube (YouTube, LLC, San Bruno, CA, USA), and Twitter (Twitter, Inc, San Francisco, CA, USA). Campaigns were excluded if they could not be retrieved, were duplicates, or also targeted a mainstream audience. A subset using Facebook were subjected to further analysis.

Data Extraction and Analysis

Data about the campaign, the year and state in which it was developed, the producer, type of media strategy, number of likes, links to other websites, and average posts per month were entered into a Microsoft Excel database.

Campaigns were ranked from the highest number of likes to the least number of likes. A "like" indicates an appreciation, enjoyment, or support of the content posted on the Facebook page. The total number and mean number of posts per month were calculated for each Facebook page and pages were categorized into 2 groups: pages with more than 5000 likes and pages with less than 1800 likes (range is 35 to 11,000 likes). The number and type linkages between Facebook and Web-based applications such as websites and other social media applications (eg, Twitter and YouTube) were also analyzed.

The KT Planning Template and stickiness frameworks, SUCCESS and STEPPS, were then used to identify factors that may have contributed to the differences in overall likes, and to analyze key characteristics of interventions to assess the extent to which strategies are "fit for purpose" and to explain their uptake or impact.

Ethics Approval

Ethics approval was not sought as the desktop audit collected data that were publicly available and freely accessible from public profiles on Facebook.

Results

We identified 113 tobacco-related campaigns. Of these, 30 were selected based on our inclusion criteria. Of these, 6 were identified as duplicates and 3 were excluded because they were for Indigenous as well as mainstream audiences resulting in 21 campaigns for further examination (Table 2).

A range of social media applications were used in the 21 campaigns. The most popular were websites, YouTube,

Facebook, and Twitter (Table 3). The most popular format to promote messages was videos.

Facebook pages with the highest number of likes (more than 5000) were linked to a website and to other social media applications. Linking social media applications and websites allows owners to consistently promote campaigns across all social media applications to ensure a broad range of audiences is captured and to "trigger" a reminder for audiences that use one or more applications.

Based on the average number of posts per month and posts' content, it appears pages with the highest number of likes (Deadly Choices, 11,000; Drug Info, 6829; Rockhole, 6304; Indigenous lung cancer ads, 5326) also shared similar stickiness characteristics such as social currency, triggers, emotion, public, stories, simple, credible, and practical value (Table 5).

Table 1. List of stickiness factors.

Characteristic	Short description
Making ideas stick framework	
Simple	Finding and delivering the core of message in a way that is so profound that a person could spend a lifetime learning to follow it.
Unexpected	Engaging people's curiosity over long periods of time by systematically opening gaps in their knowledge and filling those gaps. Involves attracting a person's attention (surprise) and holding that attention (interest).
Concrete	Helping people understand and remember messages through the use of concrete images such as the use of proverbs.
Credible	Ensuring messages carry their own credentials through the use of external (eg, an expert or authority figure) and internal credibility (eg, use of evidence and statistics).
Emotional	Messages that make people feel something by using the power of association, appealing to self-interest, and identify.
Stories	Stories can tell people how to act or how they can inspire (ie, give people the energy to act).
Contagious framework	
Social currency	People like to make a good impression, so products and ideas that make people look good are more likely to be shared.
Triggers	Triggers and cues lead people to talk, choose, and use. Social currency gets people talking. Triggers keep people talking. Top of the mind means tip of the tongue.
Emotion	Similar to making ideas stick framework—see above. Activating the right type of emotions is the key to transmission. When we care, we share.
Public	People are said to imitate one another. So if people can't see what others are doing, they can't imitate them. Making products and ideas popular means making them more publicly observable. If something is built to show, it's built to grow.
Practical value	Practical value is about helping. Information that contributes to something being useful in terms of saving money, making people happier, or saving time is news you can use.
Stories	Similar to making ideas stick framework—see above. A narrative that people will want to share.

Table 2. Campaigns selected for further examination.

No.	Title	Year	Producer	State[a]	Format
1	Give up the smokes	2015	Bega Garnbirringu	WA	Video
2	My QuitBuddy	2015	Quit Now	National	Mobile app
3	Quit for you - quit for two	2014	Quit Now	National	Mobile app
4	Indigenous mothers talk	2014	Rural Health Channel	QLD	Video
5	Puyu paki (Don't smoke - give it up)	2014	Puntukurnu Aboriginal Medical Service	WA	Video
6	Breathe clearly, live healthy, quit smoking	2014	Mawarnkarra Health Service	WA	Video
7	Skinnyfish music health promotion videos	2013	Skinnyfish music	NT	Video
8	Indigenous lung cancer ads	2013	The Lung Foundation	National	Video
9	Rockhole	2013	Indigenous Hip Hop Projects	National	Video
10	Tomorrow's dream advertisement	2013	Aboriginal Health Council of Western Australia	WA	Video
11	Deadly Choices - smoking television commercial #1	2013	Deadly Choices	QLD	Video
12	Stickin it up the smokes: Ellie Lovegrove and Daniel Summer	2013	Lovegrove E. and Summer D.	SA	Video
13	Tobacco addiction story – English	2013	No Smokes	NT	Video
14	Smoking and pregnancy	2013	No Smokes	NT	Video
15	Quit Now online calculator	2013	Quit Now	National	Online calculator
16	Smoke-free homes and cars	2013	Aboriginal Tobacco Control Project	NSW	Video
17	Blow away the smokes: A guide to quitting cigarettes for Aboriginal and Torres Strait Islander people	2012	Baker F. and Gould G.	NSW	Video
18	Stay strong and healthy: Pregnancy resources for Aboriginal women	2012	NSW Ministry of Health	NSW	Mix
19	No durri for this Murri	2012	North Coast Aboriginal Corporation for Community Health	QLD	Video
20	VACCHO[b] World No Tobacco Day	2011	Gallagher, J. Victorian Aboriginal Community Controlled Health Organisation	VIC	Video
21	DrugInfo	2011	Australian Drug Foundation	National	Website

[a]NSW: New South Wales; NT: Northern Territory; QLD: Queensland; SA: South Australia; VIC: Victoria; WA: Western Australia.
[b]VACCHO: Victorian Aboriginal Community Controlled Health Organisation.

Table 3. Social media applications used by 21 smoking cessation and prevention campaigns for Indigenous people.

Social media application	Number of tobacco prevention and cessation campaigns
Website	20
YouTube	18
Facebook	13
Twitter	11
Mobile app	2
Other (SoundCloud, Pinterest, Vimeo)	5

Table 4. Ranking of Facebook tobacco prevention and cessation campaigns.

No.	Title of health promotion program	Year	Producer	State[a]	Type	Likes	Linking	Average number of posts per month
1	Deadly Choices (smoking television commercial #1) [9]	2013	Deadly Choices	QLD	Video	11,000	Facebook to embedded videos and website / Website to Facebook and Twitter	13
2	DrugInfo [10]	2011	Australian Drug Foundation	National	Website	6829	Facebook to Website / Website to Twitter / Twitter to Website	13
3	Rockhole [11]	2013	Indigenous Hip Hop Projects	National	Video	6304	Facebook to embedded videos / Website to videos	15
4	Indigenous lung cancer ads [12]	2013	The Lung Foundation	National	Video	5326	Website to Facebook and videos	12
5	Tobacco addiction story - English [13]	2013	No Smokes	NT	Video	1741	Facebook to Twitter and videos (Youtube and Vimeo) / Twitter to Website and videos	8
6	Smoking and pregnancy [13]	2013	No Smokes	NT	Video	1741	Facebook to Twitter and videos (Youtube and Vimeo) / Twitter to Website and videos	8
7	VACCHO World No Tobacco Day [14]	2011	Victorian Aboriginal Community Controlled Health Organisation	VIC	Video	989	Facebook to Website and videos (Youtube) / Twitter to Website / Website to Facebook, Twitter, Youtube, and Soundcloud (audio)	5
8	Stay strong and healthy: Pregnancy resources for Aboriginal women [15]	2012	NSW Ministry of Health	NSW	Mix	831	Facebook to Website	4
9	Smoke-free homes and cars (Facebook- I quit because)[16]	2013	Aboriginal Tobacco Control Project	NSW	Video	466	Facebook to Website and embedded videos / Website to videos	12
10	Blow away the smokes: A guide to quitting cigarettes for Aboriginal and Torres Strait Islander people [17]	2012	Baker, F. and Gould, G.	NSW	Video	55	Website to Facebook and videos (vimeo)	0
11	Give up the smokes [18]	2015	Bega Garnbirringu	WA	Video	52	Website to Youtube	0.1
12	Tomorrow's dream advert [19]	2013	Aboriginal Health Council of Western Australia	WA	Video	42	Facebook to Website / Website to Facebook, Twitter, Google Plus, PinInterest, Soundcloud	0.1
13	Breathe clearly, live healthy, quit smoking [20]	2014	Mawarnkarra Health Service	WA	Video	35	Twitter to Facebook	0

[a]NSW: New South Wales; NT: Northern Territory; QLD: Queensland; SA: South Australia; VIC: Victoria; WA: Western Australia.

Table 5. Stickiness factors associated with Facebook pages with likes more than 1000.

Stickiness factors	Stickiness factors from Facebook pages with high number of likes more than 5000
Social currency	Demonstrating the adoption of current trends or the rollout of new initiatives such as tackling smoking
	Promoting and sharing photos and video posts that showed people with a known sporting identity at a health promotion event where participants received a t-shirt or some other incentive
Triggers	Routine and regular posts of photos, videos, and information at different times of the day about upcoming campaigns
	Linking of social media applications to allow sharing of posts between platforms, thereby providing regular reminders
Emotion	Posting stories about personal life journeys and experiences such as quit smoking stories
	Use of music and dance to promote health promotion messages
Public	All profiles were public profiles - free to access
	Using tags to share posts with other Facebook profiles
	Linking of social media applications to allow sharing of posts between platforms such as YouTube and Facebook
Stories	Posting videos of ex-smokers sharing their stories about quitting smoking or about the death of a loved one due to lung cancer
Credible	Stories and videos were from an individual's personal life journey and experience
	Advertisements promoted statistics about smoking and lung cancer
Practical value	Posting of videos about how to stay fit and healthy
	Posting of educational messages

For example, the Deadly Choices Facebook page demonstrated the following characteristics:

- Simple: Core message promoted consistently through written posts, photos, and videos.
- Unexpected: Frequently promoted competitions, free giveaways, and meet and greets.
- Triggers: The page frequently posted messages at different times of the day. It used different forms of media such as text, photos, and videos. The page was also linked to other social media platforms to allow sharing of posts to Twitter and YouTube.
- Social Currency and Public: The campaign appeared to have recruited known sporting and public personalities to promote campaigns. The page also posted photos and videos of people with sporting and public personalities at various health promotion campaigns.
- Stories, Emotion, and Credibility: The page posted or linked to videos of Indigenous people telling their stories about their quit smoking journey. There were also good news stories about the benefits of a healthy lifestyle through diet and exercise.
- Practical Value: Provided information about how to stay fit and healthy or how to cease smoking and get fit and healthy.

Discussion

Principal Findings

Using social media like Facebook has certain appeal: the potential to reach large numbers of people with ease of set up and at a minimum cost. There appears to be a perception that few resources, if any, are required to make health messages "sticky."

However, this type of thinking is deceiving. The findings from this study show that the use of social media applications like Facebook do not guarantee that a campaign will have the desired impact and reach, such as through high numbers of likes, shares, and comments. Making health messages sticky through social media such as Facebook requires, at a minimum, content to be sourced and translated into a format that incorporates stickiness characteristics, routine posts that maintain page currency, routine monitoring and evaluating to assess impact and effectiveness, and a skilled and experienced workforce. Workforce expenses are the "hidden costs" of social media applications.

In addition to social media, there are several other strategies to translate knowledge. For example, strategies like knowledge brokers, champions, media campaigns, and pop-up stalls may be more suitable for campaigns that wish to provoke interest and raise awareness in groups that do not use social media or have limited access to the Internet. These strategies may have been more suitable for health promotion campaigns that received less than 500 Facebook likes. However, strategies like financial incentives, new policies, patient education sessions, and communities of practice could be used in combination with social media applications to impart knowledge and inspire change. For example, the Deadly Choices page promotes Indigenous designed t-shirts that are given to individuals who complete the annual health assessment. Once strategies are identified, the next step makes them sticky.

KT planning can be used as a tool to support individuals and teams to craft the delivery of health messages so that they are best suited and well-aligned to the campaigns' goals, messages, and target audiences. KT planning has clear potential to help mitigate unforeseen and expected risks, reduce unwarranted costs and expenses, achieve goals, and limit the peer pressure of using strategies that may not be fit for purpose.

Limitations

Due to privacy requirements, the desktop review did not have access to Facebook metrics, such as Page Insights, to measure the reach and uptake of posts. Page insights provide information

about the number of people to whom a post has reached; who have clicked on a post; liked, commented, or shared a post; or viewed a video. This level of analysis would be important in any evaluation of specific campaigns.

Future Implications

The uptake and use of frameworks and practices like KT planning will take time and will undoubtedly face challenges and barriers. The main challenges are conceptualizing KT and then applying it effectively to the local context [5]. Until this and other challenges (eg, limited organisational KT capacity, limited to access to KT workforce) are addressed, we will continue to use systems and processes that are familiar or easy to use but may be ineffective or have variable uptake and impact.

Acknowledgments

Conflicts of Interest

None declared.

References

1. Sweet MA. Social media: new links for Indigenous health. Med J Aust 2013 Jul 8;199(1):18. [Medline: 23829246]
2. Brusse C, Gardner K, McAullay D, Dowden M. Social media and mobile apps for health promotion in Australian Indigenous populations: scoping review. J Med Internet Res 2014;16(12):e280 [FREE Full text] [doi: 10.2196/jmir.3614] [Medline: 25498835]
3. Callinan T. Remote Australians Rely on Facebook to Stay in Touch. NITV; 2014 Aug 27 URL: http://www.sbs.com.au/nitv/nitv-news/article/2014/08/26/remote-indigenous-australians-rely-facebook-stay-touch [accessed 2016-09-28] [WebCite Cache ID 6kqT13Gv1]
4. CIHR-IRSC. Knowledge Translation - Definition About Us URL: http://www.cihr-irsc.gc.ca/e/29418.html [accessed 2016-09-27] [WebCite Cache ID 6kqTOlfex]
5. Goering P, Ross S, Jacobson N, Butterill D. Developing a guide to support the knowledge translation component of the grant application process. Evidence & Policy 2010;6(1):91-102.
6. Barwick M. 2008. Knowledge Translation Training and Tools. Toronto, ON: Melanie Barwick; 2008. URL: http://www.melaniebarwick.com/KTTemplateFillable_dl.php [WebCite Cache ID 6kqTuEyAY]
7. Heath C, Heath D. Made to Stick: Why Some Ideas Survive and Others Die. New York, NY: Random House; 2007.
8. Berger J. Contagious: How to Build Word of Mouth in the Digital Age. London: Simon & Schuster UK Ltd; 2013.
9. Facebook. Deadly Choices Smoking Television Commercial. Facebook. 2013. URL: https://www.facebook.com/deadlychoices/ [accessed 2016-09-28] [WebCite Cache ID 6kqUO26Nm]
10. Australian Drug Foundation. DrugInfo. 2011. URL: https://www.facebook.com/AustralianDrugFoundation [accessed 2016-09-28] [WebCite Cache ID 6kqUgJWeg]
11. Indigenous Hip Hip Projects. Facebook. 2013. URL: https://www.facebook.com/Indigenoushiphopprojects/ [accessed 2016-09-28] [WebCite Cache ID 6kqUWQZeY]
12. Facebook. Lung Foundation Australia. 2013. URL: https://www.facebook.com/lungfoundation/ [accessed 2016-09-28] [WebCite Cache ID 6kqUnRB8M]
13. Facebook. No Smokes. 2013. URL: https://www.facebook.com/NoSmokes.com.au/info/?entry_point=page_nav_about_item&tab=page_info [accessed 2016-09-01]
14. Facebook. Victorian Aboriginal Community Controlled Health Organisation. 2011. URL: https://www.facebook.com/VACCHO/?fref=nf [accessed 2016-09-28] [WebCite Cache ID 6kqUwoXpy]
15. New South Wales Ministry of Health. Stay strong and healthy: pregnancy resources for Aboriginal women. Facebook. 2012. URL: https://www.facebook.com/StayStrongAndHealthy/info/?entry_point=page_nav_about_item&tab=page_info [accessed 2016-09-01]
16. Aboriginal Tobacco Control Project. Facebook. 2013. I Quit Because URL: https://www.facebook.com/IQuitBecause/info/?entry_point=page_nav_about_item&tab=page_info [accessed 2016-09-01]
17. No Smokes North Coast. Facebook. Blow Away the Smokes - Smoke Free Aboriginal Country. 2012. URL: https://www.facebook.com/Blow-Away-The-Smokes-Smoke-Free-Aboriginal-Country-271034862949684/info/?entry_point=page_nav_about_item&tab=page_info [accessed 2016-09-01]
18. Bega Garnbirringu. Facebook. Give Up the Smokes. 2015. URL: https://www.facebook.com/pages/Bega-Garnbirringu-Health-Services/150729521633534 [accessed 2016-09-27] [WebCite Cache ID 6kqVE2Ye3]
19. Aboriginal Health Council of Western Australia. Facebook. 2013. URL: https://www.facebook.com/AHCWA/?fref=ts [accessed 2016-09-27] [WebCite Cache ID 6kqVIj9jG]

XSL•FO

RenderX

20.   Facebook. Mawarnkarra Health Service. 2014. URL: https://www.facebook.com/Mawarnkarra/?fref=ts [accessed 2016-09-28]
       [WebCite Cache ID 6kqVN1cGf]

Abbreviations

KT: knowledge translation
STEPPS: social currency, triggers, emotion, public, practical value, stories
SUCCESS: simple, unexpected, concrete, emotional, stories
VACCHO: Victorian Aboriginal Community Controlled Health Organisation

Original Paper

Sexual Preferences and Presentation on Geosocial Networking Apps by Indian Men Who Have Sex With Men in Maharashtra

Jayson Rhoton[1*], MA; J Michael Wilkerson[1*], MPH, PhD; Shruta Mengle[2*], MSc; Pallav Patankar[2], MBA; BR Simon Rosser[3*], LP, MPH, PhD; Maria L Ekstrand[4], PhD

[1]Department of Health Promotion and Behavioral Sciences, The University of Texas Health Sceince Center Houston, Houston, TX, United States

[2]The Humsafar Trust, Mumbai, Maharashtra, India

[3]School of Public Health, University of Minnesota, Minneapolis, MN, United States

[4]Center for AIDS Prevention Studies, University of California San Francisco, San Francisco, CA, United States

[*]these authors contributed equally

Corresponding Author:
Jayson Rhoton, MA
Department of Health Promotion and Behavioral Sciences
The University of Texas Health Sceince Center Houston
7000 Fannin Street
2610 I
Houston, TX, 77030
United States
Phone: 1 713 500 9757
Fax: 1 713 500 9750
Email: Jayson.Rhoton@uth.tmc.edu

Abstract

Background: The affordability of smartphones and improved mobile networks globally has increased the popularity of geosocial networking (GSN) apps (eg, Grindr, Scruff, Planetromeo) as a method for men who have sex with men (MSM) to seek causal sex partners and engage with the queer community. As mobile penetration continues to grow in India, it is important to understand how self-presentation on GSN app is relevant because it offers insight into a population that has not been largely studied. There is very little information about how Indian MSM discuss their sexual preferences and condom preferences and disclose their human immunodeficiency virus (HIV) status with potential sex partners on Web-based platforms.

Objective: The objective of this study was to describe how self-presentation by Indian MSM on GSN apps contributes to sexual preferences, HIV or sexually transmitted infection (STI) disclosure, and if the presentation differs due to proximity to the Greater Mumbai or Thane region.

Methods: Between September 2013 and May 2014, participants were recruited through banner advertisements on gay websites, social media advertisements and posts, and distribution of print materials at outreach events hosted by lesbian, gay, bisexual, transgender (LGBT) and HIV service organizations in Maharashtra, India. Eligible participants self-identified as being MSM or hijra (transgender) women, living in Maharashtra, aged above 18 years, having regular Internet access, and having at least one male sex partner in the previous 90 days.

Results: Indian MSM living inside and outside the Greater Mumbai or Thane region reported an average of 6.7 (SD 11.8) male sex partners in the last 3 months; on average HIV status of the sex partners was disclosed to 2.9 (SD 8.9). The most commonly used websites and GSN apps by MSM living inside Greater Mumbai or Thane region were Planetromeo, Grindr, and Gaydar. Results demonstrated that MSM used smartphones to access GSN apps and stated a preference for both condomless and protected anal sex but did not disclose their HIV status. This low level of HIV disclosure potentially increases risk of HIV or STI transmission; therefore, trends in use should be monitored.

Conclusions: Our data helps to fill the gap in understanding how Indian MSM use technology to find casual sex partners, disclose their sexual preference, and their HIV status on Web-based platforms. As mobile penetration in India continues to grow and smartphone use increases, the use of GSN sex-seeking apps by MSM should also increase, potentially increasing the risk of HIV or STI transmission within the app's closed sexual networks.

XSL·FO
RenderX

Introduction

The affordability of smartphones and improved mobile networks globally has likely contributed to the increased popularity of geosocial networking (GSN) apps (eg, Grindr, Scruff, and Planetromeo) as a method for men who have sex with men (MSM) to seek causal sex partners and engage with the queer community [1-3]. The growth in popularity of GSN apps has been attributed to the global positioning systems and algorithms that match men based on shared interests, sexual attraction, muscularity, age, and proximity [4,5]. In Western MSM populations, GSN apps have shown to be more popular than other online peer-to-peer social networking sites (eg, Facebook) and traditional methods (eg, cursing area and gay bars) to find causal sex partners [4-6]. Several studies have found that MSM report logging in to GSN apps at least three times a day, with an average of 12 minutes being spent per log-in [3,7-9].

Till date, research on MSM's use of GSN apps has focused on urban and rural Western MSM populations, examining sexual risk behaviors, partner preferences, condom use, and human immunodeficiency virus (HIV) or sexually transmitted infection (STI) disclosure [3]. Western MSM who use GSN apps to find causal sex partners who tend to be younger and have more sexual encounters, which potentially increases their risk of HIV or STI infection [5,7,10]. While the evidence is unclear if GSN apps actually contribute to increase in sexual risk behavior [4,7], several studies have reported that MSM who used GSN apps had higher lifetime male anal sexual partners and reported greater condomless anal-receptive sex, condomless oral-sex, and rimming, than MSM who used offline venues [11-14]. Additionally, MSM using GSN apps report to have on average more sexual partners, greater encounters with partners known to have HIV and STI diagnoses [4,11], than MSM meeting sex partners offline.

The use of GSN apps as a tool for seeking sexual partners raises the questions of how urban and rural MSM present themselves on Web-based platforms and how they disclose their sexual and condom preferences and HIV or STI status [3,15]. Suler and others exploration of self-presentation on GSN and social networking apps for MSM has found that MSM Web profiles offer dissociative anonymity, which offers MSM the ability to adjust their nonvirtual identities based on their sexual preference (eg, topping, bottoming, versatile, condomless sex and protected sex), which have been found to vary across multiple GSN apps [3]. These findings suggest that MSM are creating multiple profiles on GSN apps with different biographical information, sexual position preferences, condom use preferences, and HIV or STI disclosure.

To our knowledge, there is no published data on how Indian MSM use GSN apps to seek sex partners. It is particularly important to understand this gap in the literature as the Telecom Regulatory Authority of India reports that India has 979 million wireless phone users [1]. As Internet and mobile penetration increases in India [2], it is anticipated that the number of MSM using GSN apps to identify male sex partners will increase, similar to Western MSM [11].

As a first step to understanding to what extent the use of GSN apps by Indian MSM mirror the behavioral risks contributing to HIV or STI transmission among Western MSM, descriptive data is needed about which apps Indian MSM use and how they construct an online identity when using these apps. Presentation of self and disclosure of HIV or STI status is important in this context because virtual communities, such as the sexual communities created by GSN apps, provide a space where individuals can explore the boundaries of their sexual attraction and potentially identify persons willing to engage in casual sex. This has been supported in the literature when looking at unidentified gay men, who create multiple profiles on GSN apps to have same-sex encounters [15,16]. Additionally, it is important to understand if there is a difference in sexual preference and HIV disclosure on GSN apps between men living inside or outside the Greater Mumbai or Thane region. We believe that this is an important distinction to make, as men living inside the Greater Mumbai or Thane region are more likely to be exposed to more HIV prevention messaging than men living outside the region. Therefore, the aim of the study was to describe how Indian MSM self-presentation on GSN apps contributes to sexual preferences, HIV or STI disclosure, and if the presentation differs due to proximity to the Greater Mumbai or Thane region.

Methods

Recruitment

The recruitment procedure is described in detail elsewhere [17]. Broadly, between September 2013 and May 2014, participants were recruited through banner advertisements on gay websites, social media advertisements and posts, and distribution of print materials at outreach events hosted by LGBT and HIV services organizations in Maharashtra, India. Most participants were recruited from the Greater Mumbai or Thane region. This region has a population of approximately 20 million and is the headquarters for many of India's largest corporations, the National Stock Exchange of India, and Bollywood. It attracts migrants from other areas of India seeking employment opportunities and an urban lifestyle. Recruitment materials directed participants to the study website where they could complete the eligibility screen. Eligible participants self-identified as being MSM or hijra or transgender woman, living in Maharashtra, above 18 years, having regular Internet access, and having at least one male sex partner in the previous 90 days. Once eligibility was determined, participants were provided with an overview of the study. After completing an online consent procedure [10], an online cross-sectional survey was administered, which took approximately 30 minutes to complete. Participants were able to complete the eligibility

screener, consent process, and survey in Hindi, Marathi, or English. Participants were compensated 300 rupees (approximately 4 USD) for completion of the survey. Of the 6049 individuals who clicked on the survey link from an online advertisement, 745 completed consent, 617 initiated the survey, 477 completed the survey, and 449 were eligible. The institutional review boards of the authors' institutions approved study procedures.

Measures

Use of Geosocial Networking Apps and Self-Presentation

Participants were asked about which websites or GSN apps they used to meet sex partners, the number of profiles on each (eg, Facebook, Desicrossdressers, Planetromeo, Scruff, Grindr, Manjam, and Gaydar), and if they disclosed their preferred sexual position and HIV status on their profile; participants were provided a write-in option. To identify all possible websites or GSN apps, participants were provided with a list that was generated by the authors who work for a nongovernmental organization serving MSM in the Greater Mumbai or Thane region; a write-in option was provided, so that websites and apps not included in the provided list could be captured. To minimize data-entry errors by participants, an algorithm was employed that reminded participants how many profiles they identified; as participants' wrotein responses, the algorithm subtracted the responses from the total number of profiles identified.

Proximity to the Greater Mumbai or Thane Region

Participants were asked in which district of Maharashtra they lived in. Persons who indicated living in Greater Mumbai or Thane were placed in one category and all other participants were placed in another.

Demographic Characteristics and Self-Reported Questionnaire

Participants were asked to identify their gender, age, educational degrees obtained, employment status, sexual orientation,

openness (outness) about their same-sex attraction, HIV status, and number of female, hijra (transgender) women and male sex partners they had in the last 3 months. Participants self-reported how they stated a preference for condoms and sexual position of their GSN apps.

Statistical Analysis

Bivariate analyses were used to identify which websites and GSN apps Indian MSM access and the frequency of disclosure of their sexual preferences and HIV status. Chi-square tests (χ^2) of independence were used to examine the relation between MSM living within or outside of the Greater Mumbai or Thane region, demographic characteristics, Internet-enabled device, and social media platforms. Chi-square tests of independence were also used to examine the differences between disclosure of sexual preferences and HIV status on websites and GSN apps. Fisher's exact tests were used to identify differences when counts were small.

Results

A total of 449 MSM completed the survey; 96% (433/449) identified as cisgender males (congruency between gender assigned at birth and current gender identity; this term is used to refer to nontransgender persons). Most participants had completed college and were employed. Over three-quarters of participants self-identified as being gay or bisexual, but only 19% (86/445) reported being open about their same-sex attraction (out) to most or everyone in their life. Participants inside and outside the Greater Mumbai or Thane region reported an average of 6.7 male sex partners (SD 11.8) in the last 3 months; on average HIV status was disclosed to 2.9 (SD 8.9) of the sex partners. Participant's demographics are presented in Table 1 and Table 2.

Table 1. Demographic characteristics of participants (mean, SD; N=449).

Demographics	Mean (SD)
Age in years	29.46 (8.20)
Male sexual partners in the previous 90 days	6.77 (11.80)
Hours using the Internet	
Work-related	22.05 (22.60)
Searching for sex	7.99 (11.60)
Meeting potential sex partners	2.36 (5.10)
Looking at pornography	3.74 (6.70)
Other activities not related to work, sex, or education	10.27 (13.90)
HIV[a] status on GSN[b] app	2.98 (8.96)

[a]HIV: human immunodeficiency virus.

[b]GSN: geosocial networking.

Table 2. Demographic characteristics of participants (% values; N=449).

Demographics	n (%)
Live in Greater Mumbai or Thane region	
Yes	335 (74.6)
No	114 (25.3)
Earn Rs 25,001 or more	
Yes	233 (55.7)
No	185 (44.2)
Completed college	
Yes	57 (12)
No	384 (87.0)
Employment status	
Not employed	47 (10)
Student-not employed	63 (14)
Student-employed full or part time	33 (7)
Employed part-time	31 (7)
Employed full-time	267 (60.5)
Outness or openness	
Out to none	110 (24.7)
Out to few to half	249 (55.9)
Out to most to all	86 (19)
HIV status	
Positive	10 (2)
Negative	336 (75.6)
Status unknown	98 (22)
Sexual orientation[a]	
Gay or homosexual	2 (0)
Bisexual	260 (63.7)
Straight or heterosexual	102 (25.0)
Double decker or versatile	8 (2)
Kothi	16 (3)
Panthi	5 (1)
Hijra or transgender	3 (0)
Queer	4 (1)
Other	8 (1)

[a]In addition to identifying with sexual orientations commonly used in the West, some Indian men who have sex with men identify with sexual orientations specific to the Indian context and frequently associated with a preferred sexual position, including double decker or versatile (someone who typically engages in receptive and insertive sex), kothi (someone who typically engages in receptive sex), and panthi (someone who typically engages in insertive sex).

XSL•FO
RenderX

Table 3. Number of sex-seeking profiles used to meet men in the past 12 months by place of residence, September 9, 2013 through June 30, 2014 (N=449).

Sex-seeking platforms	Within the Greater Mumbai or Thane region (N=332)			Outside the Greater Mumbai or Thane region (N=114)			χ^2_2	*P* value
Variable	No profiles (n)	1 profile (n)	2 or more profiles (n)	No profiles (n)	1 profile (n)	2 or more profiles (n)		
Planetromeo	68	196	71	12	67	35	7.8	.02
Scruf	293	38	4	105	6	3	4.5	.1
Grindr	201	118	16	87	23	4	9.9	<.001
Desicrossdressers	322	10	3	110	1	3	3.4	.18
Manjam	263	62	10	92	16	6	2.2	.32
Gaydar	287	45	3	102	8	4	6.8	.03
Facebook	166	130	39	58	39	17	1.2	.54

Indian MSM in our sample had GSN apps running in the background for much of the day, reporting an average of 8 hours per day looking for potential sex partners. The most commonly used Internet-enabled devices for sex-seeking were smartphones (292/442, 66%), personal desktop (154/443, 65%), and laptop computers (272/441, 62%). The most commonly used websites and GSN apps by MSM living inside Greater Mumbai or Thane region were Planetromeo, Grindr, and Gaydar. Additionally, MSM living in the Greater Mumbai or Thane region had created more profiles on GSN apps and websites than MSM living outside the Greater Mumbai or Thane region. Despite the difference in websites and GSN apps between groups, a majority of participants had multiple profiles on websites and GSN app (see Table 3).

MSM living within Greater Mumbai or Thane region stated a greater preference for protected insertive and receptive anal sex (SD 6.35, *P*<.001 and SD 5.34, *P*=.02, respectively) on GSN apps and websites compared with MSM living outside Mumbai or Thane region. MSM living in Greater Mumbai or Thane region also showed significant difference in preference for more condomless insertive (SD 5.61, *P*=.01), condomless receptive (SD 4.38, *P*=.04) and condomless versatile anal sex (SD 4.88, *P*=.03) than MSM living outside Greater Mumbai or Thane region. There were no differences between the MSM within and outside of Greater Mumbai or Thane region in stated preference for versatile anal sex with condoms (SD 2.72, *P*=.09) and disclosing their HIV status on websites or GSN apps (SD 0.86, *P*=.65).

Discussion

Principal Findings

The purpose of this study was to describe how Indian MSM present themselves on GSN apps, how their presentation contributes to sexual preferences, and if the presentation differs due to proximity to the Mumbai or Thane region. Indian MSM are using smartphone (eg, BlackBerry Bold Touch) as their primary means to access sex-seeking GSN apps to search for potential sex partners. Most men in the study had multiple active accounts on sex-seeking GSN apps and allowed these apps to run much of the day. These findings are consistent with Western

MSM use of GSN apps to connect with potential sex partners [12].

That Indian MSM have multiple profiles potentially allowing them to present themselves as someone who prefers to be insertive, receptive, or versatile and someone who does or does not endorse condom use during sex complicates how we understand the presentation of self on GSN apps. That MSM in the Greater Mumbai or Thane region endorsed a preference for both using condoms for anal sex and engaging in condomless anal sex further complicates how we understand the way in which Indian MSM present themselves in an online space. Regarding the use of multiple profiles, Ross and colleagues [3,18] and other researchers [5] suggest that MSM, who are often stigmatized due to a lack of acceptance of their same-sex attraction, use online sex-seek spaces to experiment with sexual identities. In India where homosexual behavior is stigmatized and same-sex behavior is illegal [19], the use of multiple profiles with different sexual preferences could allow men the freedom to virtually experiment and fantasize about engaging in various sexual behaviors, which could aid in sexual identity development and connectedness to other MSM. Distinguishing between profiles they use to meet sex partners offline versus meeting partners to engage in chatting or virtual sex—which would allow for the enactment of sexual behaviors without placing oneself at risk of HIV or STI transmission—was beyond the scope of this study. Few MSM in our sample disclosed their HIV or STI status on their profiles. This result could be explained in the context of the stigmatized culture that Indian MSM encounter or that a majority of the sample did not know their HIV or STI status. We believe it is important to consider the low levels of HIV disclosure on GSN apps as an important behavior to understand, as we know very little about the online Indian MSM community. Western literature suggests that online spaces are typically anonymous and offer an opportunity to create a virtual identity that may not match the nonvirtual identity. Therefore, it is important to explore how virtual identities impact HIV disclosure for Indian MSM. If in future research we find that the use of some profiles contributes to risk while others are protective, there would be an opportunity to educate MSM about how to use the profiles created on GSN apps to fulfill their sexual desires while minimizing risk.

Men living inside the Greater Mumbai or Thane region were more likely to endorse using condoms than MSM living outside the MSA. The difference between groups in stated condom preference could be the result of HIV prevention services and outreach being concentrated within the physical cruising sites of Greater Mumbai or Thane region, suggesting a possible overlap between populations on physical and virtual cruising platforms. HIV prevention services should now also be targeted toward encouraging Internet-using MSM to disclose their condom use preferences on GSN apps to facilitate conversations between potential sex partners about condom use before meeting for a sexual encounter.

Limitations

The results of this study should be considered within the limitations of this understudied population. Due to the rapid changes in technology and rapid creation of mobile apps, the GSN apps reported in this paper may increase or decrease in use overtime. On average, the participants were highly educated, employed, and self-identified as either bisexual or gay, but were not out or open about their sexual identity to most people in their lives. Because the study was cross-sectional, we could not determine if disclosing a preference for condomless anal sex on GSN apps leads to an increased risk for HIV or STIs among Indian MSM or vice versa. There is robust evidence supporting the connection between condomless anal sex and increased risk for HIV transmission [20]. However, there is no evidence that we are aware of, that confirms a connection between stating a preference for condomless anal sex and risk for HIV

transmission within online venues among Indian MSM populations. Furthermore, stated preferences for condom use were based on self-reports, which might explain the discrepancy between high reporting for both protected and condomless anal sex. The discrepancy in self-reported condom preference on GSN apps could be explained by participants having multiple profiles to endorse their current condom-use desires. Therefore, more research is needed to understand how Indian MSM report condom preferences on GSN apps and its connection to actual condom-use.

Conclusions

The purpose of this study was to describe how Indian MSM present themselves on GSN apps, how their presentation contributes to sexual preferences, and if the presentation differs due to proximity to the Greater Mumbai or Thane region. Our data begins to fill the gap in understanding these aspects. As mobile penetration in India continues to grow and smartphone use increases, the use of GSN sex-seeking apps by MSM should also increase, potentially increasing the risk of HIV or STI transmission within the app's closed sexual networks. Considering that HIV interventions are solely focused on physical cruising sites in India, our findings highlight the need for Indian HIV Interventions to look beyond current scopes and expand the ambit of HIV interventions to virtual platforms as well. Interested researchers, including mobile interventions, should continue to monitor how MSM use mobile technology to meet male sex partners.

Conflicts of Interest

None declared.

References

1. Telecom Regulatory Authority of India. Highlights of telecom subsriptions data as on 28th February, 2015. URL: http://www.trai.gov.in/WriteReadData/WhatsNew/Documents/PR-TSD-Feb-10042015.pdf [WebCite Cache ID queryurlhttp3A2F2Fwwwtraigovin2Fampdate20160203]

2. International Data Corporation. India continues as one of the fastest growing smart phone markets in Asia Pacific in 1Q 2014, Says IDC. 2016. URL: http://www.idc.com/research/container_error.jsp [WebCite Cache ID 6f21bSNOW]

3. Miller B. "Dude, where's your face?" Self-presentation, self-description, and partner preferences on a social networking application for men who have sex with men: a content analysis. Sexuality Cult 2015 Apr 10;19(4):637-658. [doi: 10.1007/s12119-015-9283-4]

4. Rice E, Holloway I, Winetrobe H, Rhoades H, Barman-Adhikari A, Gibbs J, et al. Sex risk among young men who have sex with men who use Grindr, a smartphone geo- social networking application. J AIDS Clinic Res 2012:S4-005. [doi: 10.4172/2155-6113]

5. Beymer MR, Weiss RE, Bolan RK, Rudy ET, Bourque LB, Rodriguez JP, et al. Sex on demand: geosocial networking phone apps and risk of sexually transmitted infections among a cross-sectional sample of men who have sex with men in Los Angeles County. Sex Transm Infect 2014 Nov;90(7):567-572 [FREE Full text] [doi: 10.1136/sextrans-2013-051494] [Medline: 24926041]

XSL•FO

RenderX

6. Queermeup. Queer me up: optimum gay culture. 2014. URL: http://queermeup.com/lgbt-community-2/ grindr-turns-five%E2%80%8F/?utm_source=rss&utm_medium=rss&utm_campaign=grindr-turns-five%25e2%2580%258f

7. Phillips G, Magnus M, Kuo I, Rawls A, Peterson J, Jia Y, et al. Use of geosocial networking (GSN) mobile phone applications to find men for sex by men who have sex with men (MSM) in Washington, DC. AIDS Behav 2014 Sep;18(9):1630-1637. [doi: 10.1007/s10461-014-0760-9] [Medline: 24682866]

8. Liau A, Millett G, Marks G. Meta-analytic examination of online sex-seeking and sexual risk behavior among men who have sex with men. Sex Transm Dis 2006 Sep;33(9):576-584. [doi: 10.1097/01.olq.0000204710.35332.c5] [Medline: 16540884]

9. Young SD, Szekeres G, Coates T. The relationship between online social networking and sexual risk behaviors among men who have sex with men (MSM). PLoS One 2013;8(5):e62271 [FREE Full text] [doi: 10.1371/journal.pone.0062271] [Medline: 23658716]

10. Rosser BR, Oakes JM, Horvath KJ, Konstan JA, Danilenko GP, Peterson JL. HIV sexual risk behavior by men who use the Internet to seek sex with men: results of the Men's INTernet Sex Study-II (MINTS-II). AIDS Behav 2009 Jun;13(3):488-498 [FREE Full text] [doi: 10.1007/s10461-009-9524-3] [Medline: 19205866]

11. Bolding G, Davis M, Hart G, Sherr L, Elford J. Gay men who look for sex on the Internet: is there more HIV/STI risk with online partners? AIDS 2005 Jun 10;19(9):961-968. [Medline: 15905678]

12. Goedel WC, Duncan DT. Geosocial-networking app usage patterns of gay, bisexual, and other men who have sex with men: survey among users of Grindr, a mobile dating app. JMIR Public Health Surveill 2015 May 08;1(1):e4. [doi: 10.2196/publichealth.4353]

13. Chiasson MA, Hirshfield S, Remien RH, Humberstone M, Wong T, Wolitski RJ. A comparison of on-line and off-line sexual risk in men who have sex with men: an event-based on-line survey. J Acquir Immune Defic Syndr 2007 Feb 1;44(2):235-243. [doi: 10.1097/QAI.0b013e31802e298c] [Medline: 17179769]

14. Winetrobe H, Rice E, Bauermeister J, Petering R, Holloway IW. Associations of unprotected anal intercourse with Grindr-met partners among Grindr-using young men who have sex with men in Los Angeles. AIDS Care 2014;26(10):1303-1308. [doi: 10.1080/09540121.2014.911811] [Medline: 24754563]

15. Downing MJ, Schrimshaw EW. Self-presentation, desired partner characteristics, and sexual behavior preferences in online personal advertisements of men seeking non-gay-identified men. Psychol Sex Orientat Gend Divers 2014 Mar 14;1(1):30-39 [FREE Full text] [doi: 10.1037/sgd0000022] [Medline: 25750927]

16. Wilkerson JM, Noor SW, Galos DL, Rosser BR. Correlates of a single-item indicator versus a multi-item scale of outness about same-sex attraction. Arch Sex Behav 2016 Jul;45(5):1269-1277. [doi: 10.1007/s10508-015-0605-2] [Medline: 26292840]

17. Wilkerson J, Patankar P, Rawat SM, Simon Rosser BR, Shukla KM, Rhoton J, et al. Recruitment strategies of Indian men who have sex with men in the State of Maharashtra into an online survey. Int J Sex Health 2016;28(3):221-227. [doi: 10.1080/19317611.2016.1193079] [Medline: 27668029]

18. Ross MW, Månsson SA, Daneback K, Cooper A, Tikkanen R. Biases in internet sexual health samples: comparison of an internet sexuality survey and a national sexual health survey in Sweden. Soc Sci Med 2005 Jul;61(1):245-252. [doi: 10.1016/j.socscimed.2005.01.019] [Medline: 15847976]

19. Lawyers Collective. LGBT Section 377. 2015. URL: http://www.lawyerscollective.org/vulnerable-communities/lgbt/ section-377.html [WebCite Cache ID 6iyV2HZlx]

20. Crepaz N, Marks G, Liau A, Mullins MM, Aupont LW, Marshall KJ, HIV/AIDS Prevention Research Synthesis (PRS) Team. Prevalence of unprotected anal intercourse among HIV-diagnosed MSM in the United States: a meta-analysis. AIDS 2009 Aug 24;23(13):1617-1629. [doi: 10.1097/QAD.0b013e32832effae] [Medline: 19584704]

Abbreviations

GSN: geosocial networking
HIV: human immunodeficiency virus
LGBT: lesbian, gay, bisexual, transgender
MSM: men who have sex with men
STI: sexually transmitted infection

XSL•FO
RenderX

Original Paper

Sleep Quality Prediction From Wearable Data Using Deep Learning

Aarti Sathyanarayana[1], MSc; Shafiq Joty[1], PhD; Luis Fernandez-Luque[1], PhD; Ferda Ofli[1], PhD; Jaideep Srivastava[1], PhD; Ahmed Elmagarmid[1], PhD; Teresa Arora[2], PhD; Shahrad Taheri[2], MBBS, PhD

[1]Qatar Computing Research Institute, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar
[2]Department of Medicine, Weill Cornell Medical College in Qatar, Qatar Foundation, Doha, Qatar

Corresponding Author:
Luis Fernandez-Luque, PhD
Qatar Computing Research Institute
Hamad Bin Khalifa University, Qatar Foundation
HBKU Research Complex
Doha, 5825
Qatar
Phone: 974 50173040
Fax: 974 44540630
Email: lluque@qf.org.qa

Related Article:

This is a corrected version. See correction statement: http://mhealth.jmir.org/2016/4/e130/

Abstract

Background: The importance of sleep is paramount to health. Insufficient sleep can reduce physical, emotional, and mental well-being and can lead to a multitude of health complications among people with chronic conditions. Physical activity and sleep are highly interrelated health behaviors. Our physical activity during the day (ie, awake time) influences our quality of sleep, and vice versa. The current popularity of wearables for tracking physical activity and sleep, including actigraphy devices, can foster the development of new advanced data analytics. This can help to develop new electronic health (eHealth) applications and provide more insights into sleep science.

Objective: The objective of this study was to evaluate the feasibility of predicting sleep quality (ie, poor or adequate sleep efficiency) given the physical activity wearable data during awake time. In this study, we focused on predicting good or poor sleep efficiency as an indicator of sleep quality.

Methods: Actigraphy sensors are wearable medical devices used to study sleep and physical activity patterns. The dataset used in our experiments contained the complete actigraphy data from a subset of 92 adolescents over 1 full week. Physical activity data during awake time was used to create predictive models for sleep quality, in particular, poor or good sleep efficiency. The physical activity data from sleep time was used for the evaluation. We compared the predictive performance of traditional logistic regression with more advanced deep learning methods: multilayer perceptron (MLP), convolutional neural network (CNN), simple Elman-type recurrent neural network (RNN), long short-term memory (LSTM-RNN), and a time-batched version of LSTM-RNN (TB-LSTM).

Results: Deep learning models were able to predict the quality of sleep (ie, poor or good sleep efficiency) based on wearable data from awake periods. More specifically, the deep learning methods performed better than traditional logistic regression. "CNN had the highest specificity and sensitivity, and an overall area under the receiver operating characteristic (ROC) curve (AUC) of 0.9449, which was 46% better as compared with traditional logistic regression (0.6463).

Conclusions: Deep learning methods can predict the quality of sleep based on actigraphy data from awake periods. These predictive models can be an important tool for sleep research and to improve eHealth solutions for sleep.

(JMIR Mhealth Uhealth 2016;4(4):e125)   doi:10.2196/mhealth.6562

XSL•FO
RenderX

Introduction

Background

The importance of sleep is paramount to health and performance. Insufficient sleep can impede physical, emotional, and mental well-being [1,2] and can lead to a multitude of health complications such as insulin resistance [3-5], high blood pressure [6], cardiovascular disease [7,8], a compromised immune or metabolic system [9,10], mood disorders (such as depression or anxiety) [11,12], and decreased cognitive function for memory and judgment [13-15].

There are many indicators of sleep quality, an important one being sleep efficiency. Sleep efficiency is a metric that takes into consideration the time spent asleep, the time it takes to fall asleep, and the time asleep but with disturbance (ie, awakenings). Poor sleep efficiency can lead to sleep deprivation, which is found to be a major health risk with links to diseases such as diabetes and obesity [16]. Also, sleep behavior has been found to impact adolescent health [17-19]. Recent systematic reviews have shown the relevance of physical activity to sleep, including sleep efficiency [20-22]. Although the relationship between physical activity and sleep is not yet fully understood, it is thought to be a strong and complex correlation contributing to multiple lifestyle diseases such as type 2 diabetes mellitus and obesity [20-22]. However, the underlying mechanism between physical activity and sleep is not yet fully understood.

Role of Wearables in Sleep Health and eHealth

There are many research tools to study the link between physical activity and sleep, including standardized questionnaires, actigraphy sensors, and polysomnography (PSG). All of these methods have different clinical indications (eg, diagnosis of different sleep disorders) [23]. For example, PSG is considered the "gold standard" for sleep medicine, as it involves the use of multiple sensors during sleep such as electroencephalogram, motion sensors, breathing sensors, SpO_2(oxygen saturation), and so forth. These sensors monitor and observe a patient overnight [24] and can be used for diagnosing different sleep disorders. The elaborate nature of PSG means that it is generally limited to one overnight observation. Even the portable solutions, which permit PSG assessment in the patient's home, are complex to perform and not without limitations [25,26].

To better understand the impact of daily physical activity on sleep behavior, new tools are needed. Sleep researchers in the early 1990s developed a technique called *actigraphy* to study sleep interactions using wearable devices [27]. Although actigraphy traditionally uses wearable devices to evaluate the sleep period of a patient, it can also be used to observe physical activity. Actigraphy has become a widely used tool, as it has been found to be much more reliable than subjective or self-reported sleep diaries and behavior logs [28]. Patients wear the device for a period of time as they continue their daily routines. The technique has been especially influential for large

cohort studies where PSG is not feasible [29]. Moreover, actigraphy allows for the continuous longitudinal monitoring of a patient. This is particularly impactful for the study of diseases such as chronic obstructive pulmonary disease where sleep disturbances can be a predictor of exacerbation of the disease [30]. Current approaches to the analysis of actigraphy data involve sleep experts performing a number of steps manually. This is a bottleneck, and hence there may be troves of actigraphy data left unanalyzed.

Furthermore, there are hundreds of consumer-grade physical activity and sleep tracking devices (eg, Fitbit) collecting motion data similarly to actigraphy devices. These consumer devices are being used by millions of people collecting huge amounts of data. Some devices even allow the collection of data for very long periods (eg, several months), as they require only an occasional need for battery recharge (ie, Garmin VivoFit). Recent studies have found that these consumer-grade devices can sometimes have similar precision to clinical-grade actigraphy sensors [31]. There are successful examples of the integration of physical activity wearable data into eHealth tailored applications [32], including smart-watch health applications [33] that can collect physical activity and sleep data directly from the watch.

Objective

As previously noted, physical activity and sleep are interrelated. In this study, we tested the feasibility of predicting poor or good sleep efficiency based on physical activity data from the awake periods from a wearable device (ie, actigraphy). We took a detour from classical methods and proposed deep learning approaches to modeling the relationship between sleep and physical activity. The importance of this research is two-fold.

First, since our approach can be used in cases where sleep sensor data is not available, our models can be used in the early detection of potential low sleep efficiency. This is a common problem with consumer-grade wearable devices, as users might not wear them during the night (battery recharging, sensors embedded in smart jewelry, and so forth).

Second, our study was focused on advanced deep learning methods. Traditional prediction models applied to activity raw accelerometer data (eg, logistic regression) suffer from at least 2 key limitations: (1) They are not robust enough to learn useful patterns from noisy raw accelerometer output. As a result, existing methods for classification and analysis of physical activity rely on extracting higher-level features that can be fed into prediction models [34]. This process often requires domain expertise and can be time consuming. (2) Traditional methods do not exploit task labels for feature construction, and thus can be limited in their ability to learn task-specific features. Deep learning has the advantage that it is robust to raw noisy data, and can learn, automatically, higher level abstract features by passing raw input signals through nonlinear hidden layers while also optimizing on the target prediction tasks. We leveraged this characteristic by building models using a range of deep

learning methods on raw accelerometer data. This reduced the need for data preprocessing and feature space construction and simplified the overall workflow for clinical practice and sleep researchers.

Methods

Overview

The methodology of this study involved several steps: (1) the collection of wearable data with regards to physical activity and sleep patterns, (2) data processing and representation, (3) data modeling, and (4) performance evaluation. The following subsections explain each of these steps.

Data Collection

The deidentified data used in this study were collected by Weill Cornell Medical College-Qatar for a research study called Qatar's Ultimate Education for Sleep in Teenagers. The aim of the study was to determine different sleep patterns in adolescents residing in Qatar and how these related to body weight status. The institutional review board (IRB) approval was initially obtained by the joint IRB for Hamad Medical Corporation and Weill Cornell Medical College. The selected cohort was chosen from adolescents attending 1 of the 2 high schools and registered in grades 7-11. The dataset used contained complete actigraphy data from a subset of 92 adolescents over 1 week. There were 322 total sleep instances: 102 from boys and 220 from girls.

After agreeing to participate in the study, the adolescent participants were provided with an ActiGraph GT3X+ device, placed on their nondominant wrist, and were instructed to wear it at all times for 7 consecutive days and nights. The ActiGraph GT3X+, shown in Figure 1, is a clinical-grade wearable device that samples a user's activity at 30-100 Hz (we used 30 Hz in this study). The effectiveness of this device has been successfully validated against clinical polysomnography [35]. The participants were instructed not to remove the water-resistant device at any time. We used the manufacturer's software (ActiLife version 6; available from ActiGraph, LLC) to export the data. Although the device is triaxial, our methods used the vertical axis only.

The subjective self-reported sleep diaries were collected, but they were not considered for the study, since actigraphy data provides a more objective measurement of physical activity and sleep patterns. Sadeh's and American Academy of Sleep Medicine sleep definitions were used in this interpretation [27,36,37]. For calculating the sleep efficiency score, we considered each individual sleeping period as sleep.

Figure 1. ActiGraph Gt3X+.

Data Processing and Representation

Sleep Quality Definitions

In a person's activity time series, that is, the continuous data collected from a wearable device, there are moments when an individual is awake and when they are asleep. The latter is referred to as the sleep period. The boundary from awake time to the sleep period is called the sleep onset time, and the boundary from sleep period to awake time is referred to as the sleep awakening time. The period of time between the self-reported time to bed and the sleep onset time is called the latency.

To measure sleep quality we determined sleep efficiency (see Figure 2), which is the ratio of total minutes asleep to total minutes in bed. Those achieving a sleep efficiency score of ≥85% are thought to be good-quality sleepers, and those with a score of <85% are thought to have poor-quality sleep [38].

Total minutes in bed represents the amount of time that an individual spends asleep as well as the amount of time the individual takes to fall asleep, that is, latency. Total sleep time represents the amount of time that an individual spends asleep, less the amount of time the person awakens. This is calculated by subtracting the wake after sleep onset (WASO) from the

duration of the sleep period. WASO is the sum of all moments of wakefulness lasting longer than 5 minutes (see Figure 3).

Figure 2. Sleep efficiency equation as defined by sleep specialists.

$$Sleep\ Efficiency = \frac{TotalSleepTime}{TotalMinutesinBed}$$

$$= \frac{||SleepPeriod|| - WASO}{||SleepPeriod|| + Latency}$$

Figure 3. The adapted wake after sleep onset calculation.

$$WASO = \sum_{\substack{Sleep\\Onset\\Time}}^{\substack{Sleep\\Awakening\\Time}} \begin{cases} ||WakefulPeriod||, & ||WakefulPeriod|| > 5 \\ 0, & otherwise \end{cases}$$

Data Processing

The data collected from the actigraphy device contained triaxial accelerometer movement. Previous research has identified that for devices worn on the wrist, the vertical axis from accelerometer data is the most indicative of physical activity [39]. Consequently, we switched to 1-dimensional input for a simpler and less noisy modeling (ie, each row in our dataset contained a 1-dimensional time series of vertical axis accelerometer data).

The raw accelerometer data was aggregated into minute-by-minute epochs using a script written in R version 3.3.1, a statistical Open Source computing software developed as part of the R Project. In this dataset, the time series contains activity data during both the *awake time* and the *sleep time* (as shown in Figure 4). We used the awake time to form the prediction (ie, as the model input) and the sleep time to determine the ground truth of sleep quality. Sleep quality (good or poor) was determined from the sleep time actigraphy data, and each time series was labeled as such.

As mentioned earlier, sleep onset time and sleep awakening time are metrics that form the boundary of the sleep period [36]. We interpreted and expanded these values for accelerometer data according to Sadeh's actigraphy definitions [37].

Sleep onset time is traditionally defined as the first minute of 15 continuous minutes of sleep after a self-reported bedtime,

and the *sleep awakening time* is the last minute of 15 continuous minutes of sleep that is followed by 30 minutes of movement [37]. To automate this interpretation directly from the accelerometer output, we developed the concept of candidate rows. *Candidate rows* denote moments (or designated epochs) in time with a lack of triaxial movement and require a subsequent pass to determine whether the individual is asleep or awake. Each row is iterated upon and run through a state machine, as illustrated in Figure 5. Since there is no self-reported bedtime, we inferred it as the beginning of sedentary behavior immediately preceding and adjacent to the start of the sleep period. The duration of this sedentary time is the *latency*.

The handling of nonwearing time is very important and should be minimized wherever possible. The device was water-resistant and did not require recharging during the time period the participants were instructed to wear it. If a participant removed the device, the triaxial accelerometer would record zero values for the time it was not worn. Natural human behavior makes micro-movements that are sensed by the accelerometer even during sleep, and so periods with a continuous lack of movement indicate device removal. In other words, during the time the device was not worn, our algorithm would identify candidate rows for the entire period, denote them as sleep time, and score the sleep as having a perfect sleep efficiency of 1. Alternatively, the ActiLife software includes an algorithm to remove nonwear periods from physical activity calculations [40].

Figure 4. Example of sleep definitions on accelerometer data of an actigraphy device.
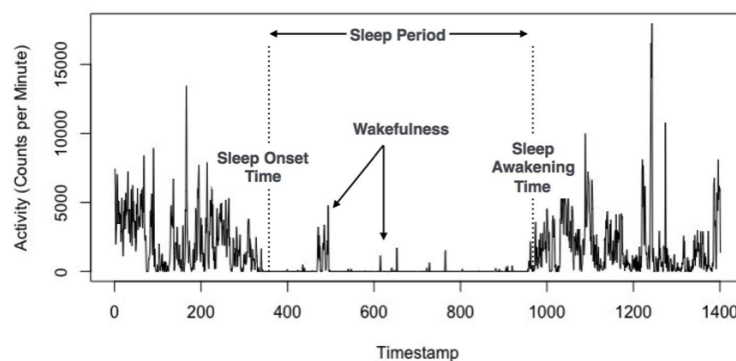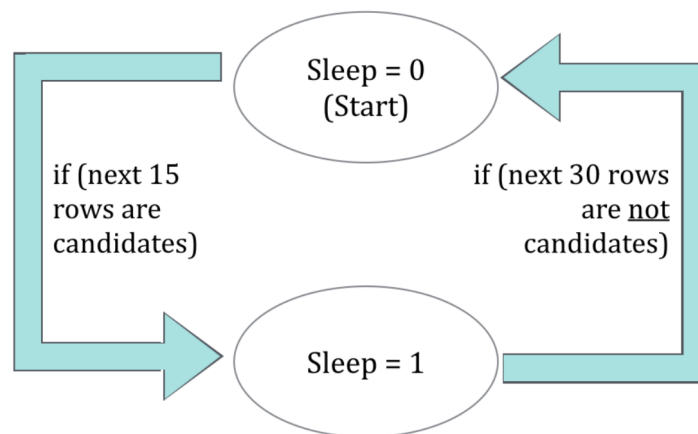
Figure 5. State machine diagram explaining the designation of sleep or awake time periods.

Models for Sleep Quality

As explained in more detail below, in this study we explored the use of deep learning methods to predict sleep quality based on actigraphy data. We compared the results of our deep learning models to those of logistic regression, a standard statistical method. In this section, we explain how each model was built. The models used in our study were as follows:

- Logistic regression, a nondeep learning model
- Multi-layer perceptrons (MLPs), a deep learning model
- Convolutional neural network (CNN), a deep learning model
- Recurrent neural networks (RNN), a deep learning model
- Long short-term memory (LSTM) RNN, a deep learning model
- Time-batched long short-term memory (TB-LSTM) RNN, a deep learning model

For logistic regression, we fed the input signals directly into the output layer for prediction. In contrast, for the deep learning models, we passed the inputs through one or more hidden layers before they were fed to the output layer for prediction. Each hidden unit in the hidden layers used a nonlinear activation function. In our study, we experimented with different hidden layer units using rectified linear unit as our activation function.

Data Partitioning

To train our models without over-fitting and test their performance afterward, we created a random partitioning of the dataset. Each time series was assigned to a partition randomly while maintaining an even class distribution of the target variable, sleep quality. The data were split with a 70%-15%-15% ratio for training, testing, and validation sets, respectively.

Input and Output of the Models

The input of the models were time series vectors, $X=(x_1, \cdots, x_T)$, representing the physical activity of a person's awake time. Each vector corresponded to a continuous period of awake time, and so for each individual, there might be multiple such vectors over the 7 days. Each x_T represented the value of the vertical axis at time t.

The output of the model was a binary classification decision between *good* and *poor* sleep quality based on sleep efficiency (%). These classifications corresponded to the sleep efficiency definitions as described earlier. In addition to the binary decision, the model also gave its confidence (a score between 0.0 and 1.0) in that decision.

Training the Models

To be able to predict, we first trained the models on the training dataset. We used an online training algorithm RMSprop [41], which relied on a number of preset parameters:

- Mini-batch size: how many training instances to consider at one time.
- Learning rate: the rate at which parameters are updated.
- Max epoch: maximum number of iterations over the training set.
- Dropout ratio: ratio of hidden units to turn off in each mini-batch training.

The training algorithm minimizes the cross-entropy between the predicted distribution and the actual (gold) target labels. To avoid over-fitting, we used early stopping based on the model's performance on the validation set. In particular, we evaluated the model after every epoch on the validation set and stopped when its accuracy went down. To reduce the cross-entropy between the predicted distributions and the target distributions, RMSprop was used setting the maximum number of epochs to 50 as suggested by the authors [41].

Logistic Regression (a Nondeep Learning Model)

As a baseline, we used logistic regression (LR) to predict the sleep quality. LR is a generalized linear classification model that does not have any hidden layers. For the LR, the raw input signals X are directly fed to the output layer for prediction without any nonlinear hidden layer transformations. The optimal setting for logistic regression (LR) was with a mini-batch size of 5 and a dropout ratio of 0.5.

Multilayer Perceptrons (a Deep Learning Model)

MLPs, also known as feed-forward neural networks, are the simplest models in the deep learning family. They have one or more hidden layers. In fact, MLP without any hidden layers is equivalent to logistic regression. In MLP, all the units of a

hidden layer are fully connected to the units in the previous layer. The best parameter configuration for MLP was with a mini-batch size of 20, a dropout ratio of 0.1, and a hidden layer size of 15.

Convolutional Neural Network (a Deep Learning Model)

CNNs are a more complex type of deep learning method that includes repetitive filters or kernels applied to local time slots, thereby composing a high level of abstract features. This operation is called *convolution*. After convolution, a *max-pooling* operation is performed to select the most significant abstract features. This design of CNNs yields fewer parameters than its fully connected counterpart (MLP), and therefore generalizes well for target prediction tasks. For its best configuration, we used 25 hidden nodes, filter length of 5 and pooling length of 4, 5 mini-batch size, and 0.0 dropout ratio.

Recurrent Neural Networks (a Deep Learning Models)

RNNs compose abstract features by processing activity measures in an awake time sequentially, at each time step combining the current input with the previous hidden state. RNNs create internal states by remembering the previous hidden layer, which allows them to exhibit dynamic temporal behavior. These features make RNNs a good deep learning method for temporal series. RNNs performed best with a mini-batch size of 5, a dropout ratio of 0.1 and a hidden layer size of 75. To avoid over-fitting, we used a technique based on dropout of hidden units and early stopping based on the loss on the development set [42].

Long Short-Term Memory (a Deep Learning Model)

A subtype of RNN, *LSTM* uses specifically designed memory blocks as units in the recurrent layer to capture longer-range dependencies. The optimal configuration values for LSTM were a mini batch size of 5, dropout ratio of 0.5, and hidden layer size of 100.

Time-Batched Long Short-Term Memory (a Deep Learning Model)

To further improve our implementation of LSTM, we constructed batches of time steps by merging accelerometer measures over time steps. We referred to this version of the model as *TB-LSTM*. The configuration values for TB-LSTM were mini-batch size of 5, dropout ratio of 0.5, and hidden layer size of 100.

Performance Evaluation

For the evaluation of the performance of the different models, we reported on several well-known metrics such as accuracy, precision, recall, F1-score, and area under the receiver operating characteristic (ROC) curve (AUC). These metrics are commonly used in data mining and clinical decision support systems.

Accuracy

It is computed as the proportion of correct predictions, both positive and negative (sum of true positives and true negatives divided by the number of all instances in the dataset).

Precision

It is the fraction of the number of true positive predictions to the number of all positive predictions (ie, true positives divided by the sum of true positives and false positives). In our case, precision described what percentage of the time the model predicted "good-quality sleep" correctly. Note that precision is also known as positive predictive value.

Specificity

It is the fraction of the number of true negative predictions to the actual number of negative instances in the dataset (ie, true negatives divided by the sum of true negatives and false positives). In our case, specificity referred to the percentage of the correctly predicted "poor-quality sleep" to the total number of "poor-quality sleep" instances in the dataset. Note that specificity is also known as true negative rate.

Recall or Sensitivity

It is the fraction of the number of true positive predictions to the actual number of positive instances in the dataset (ie, true positives divided by the sum of true positives and false negatives). In our case, recall referred to the percentage of the correctly predicted "good-quality sleep" to the total number of "good-quality sleep" instances in the dataset. Note that recall is also known as true positive rate or sensitivity.

F1-Score

There is usually an inverse relationship between precision and recall. That is, it is possible to increase the precision at the cost of decreasing the recall, or vice versa. Therefore, it is more useful to combine them into a single measure such as F1 score, which computes the harmonic mean of precision and recall.

Area Under the ROC Curve

It represents the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative instance. Hence, AUC defines an effective and combined measure of sensitivity and specificity (which are often inversely related, just like precision and recall) for assessing inherent validity of a classifier.

Results

Comparison Between Deep Learning and Logistic Regression

As shown in Table 1 and Figure 6, the performance of the logistic regression in the metrics previously explained performed worse than the models based on deep learning. "Only the simple RNN performed worse than logistic regression in both F1-score (harmonic mean of precision and recall) and accuracy.

Table 1. Results on raw accelerometer data.

Data models	AUC[a]	F1-Score	Precision	Recall	Accuracy
LR[b]	0.6463	0.8193	0.7083	0.9714	0.7321
MLP[c]	0.9449	0.9118	0.9394	0.8857	0.8929
CNN[d]	0.9456	0.9444	0.9189	0.9714	0.9286
RNN[e]	0.7143	0.7711	0.6667	0.9143	0.6607
LSTM[f]-RNN	0.8531	0.8500	0.7556	0.9714	0.7857
TB-LSTM[g]	0.9714	0.9211	0.8537	1.000	0.8929

[a]AUC: Area under receiver operating characteristic (ROC) curve.

[b]LR: logistic regression.

[c]MLP: Multilayer perceptrons.

[d]CNN: convolutional neural networks.

[e]RNN: recurrent neural networks.

[f]LSTM: long short-term memory.

[g]TB-LSTM: time-batched LSTM.

As shown in Table 1, the AUC of the logistic regression model was low. The AUC value for LR was 0.6463, which was close to 0.5 (equivalent to a random prediction). This showed the limitation of classical models in analyzing raw accelerometry. In contrast, all the AUC values for the deep learning models were better with a range from 0.7143 to 0.9714, TB-LSTM being the best performer and RNN, the worst. Time-batched LSTM, CNN, and MLP performed the best with AUC scores showing an improvement over LR by 50%, 46%, and 46%, respectively.

Figure 6. Receiver operating characteristic (ROC) curves for each model's prediction of sleep efficiency.
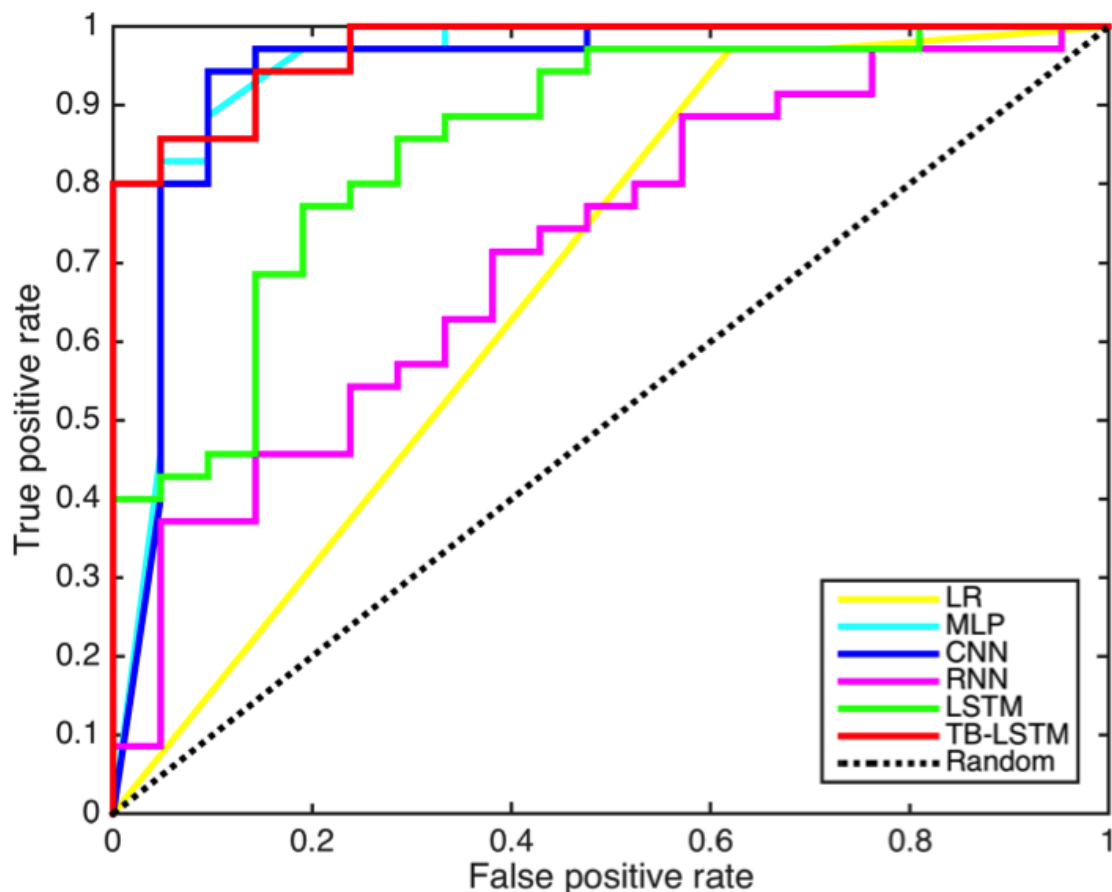
Comparison Between Deep Learning Models

Upon comparing the deep learning neural network models, we noticed that CNN yielded slight improvement over MLP in AUC (0.07% absolute), but more in F1 (3.57%) and accuracy (4.00%). These improvements could be attributed to the time-invariant convolution-pooling operations of the CNN model to pick key local patterns, which generalized well for small training data. The F1-score improved by 12% for time-batched LSTM, by 15% for CNN, and by 11% for MLP. The accuracy improved by 22% for time-batched LSTM, by 27% for CNN, and by 22% for MLP.

A comparison of the RNN models revealed that LSTM outperformed simple RNN by a wide margin; 19%, 10%, and 19% in AUC, F1, and accuracy, respectively. These gains over simple RNNs could be attributed to the specially designed gates of LSTMs that could capture long-range dependencies between physical activities in the sequences.

However, this is not surprising. Both simple and LSTM RNNs operate on sequences, where each time step comprises only one activity value. This often results in very long sequences. As mentioned earlier, in this setting RNNs cannot compose higher-level features effectively because of low-dimensional input at each time step, and also suffer from vanishing gradient problems due to lengthy sequences.

Our solution to surmount this problem was to use a time-batched input to LSTM. When we compared the results of our time-batched LSTM (TB-LSTM) with those of MLP and CNN, we found that TB-LSTM outperformed both MLP and CNN in AUC by 3%; in fact, it had the highest AUC score. It achieved better F1 score than MLP (1%), but worse than CNN (–2%). When we observed their precision and recall values, we found

that TB-LSTM had a very high recall but lower precision, which meant that it tended to predict more goods than gold standard. For the same reason, its accuracy was also lower than that of CNN.

Discussion

Principal Findings

In our study, we focused on the prediction of poor versus good sleep efficiency. That is a simple, but important, problem, as sleep efficiency has been found to be a crucial sleep parameter with important health consequences [38,43,44]. Furthermore, we did not quantify in our prediction the overall sleep efficiency but simply the differentiation between two classes (poor versus good sleep efficiency). This classification is consequently not an indicator of sleep patterns, but the prediction of a sleep quality parameter that might indicate a potential sleep problem.

As in prediction or diagnostic problems, our results need to be discussed in terms of sensitivity and specificity (see Table 2). The deep learning methods of CNN and TB-LSTM were the best performers overall. Their sensitivity (0.97 and 1, respectively) showed that these models were able to detect nearly all the cases of "good-quality sleep," meaning that in a tool for screening potential sleep problems these models will be able to detect easily people with normal sleep quality. Often high sensitivity comes at the price of low specificity (ie, failing to identify negative cases, or true negative error). This was the case of logistic regression, which had a high sensitivity but a specificity of 0.3, meaning that in such models many "poor sleeps" would have been wrongly classified as good sleep. This is very important, since misidentifying poor sleep cases can lead to underdiagnosis of problematic sleep.

Table 2. Sensitivity and specificity results.

Data models	Sensitivity or recall	Specificity
LR[a]	0.9714	0.333
MLP[b]	0.8857	0.9048
CNN[c]	0.9714	0.8571
RNN[d]	0.9143	0.2381
LSTM[e]	0.9714	0.4762
TB-LSTM[f]	1.000	0.7143

[a]LR: logistic regression.

[b]MLP: multilayer perceptrons.

[c]CNN: convolutional neural networks.

[d]RNN: recurrent neural networks.

[e]LSTM: long short-term memory.

[f]TB-LSTM: time-batched LSTM.

The sensitivity (also known as recall) and specificity of each of the models are reported in Table 2. The high sensitivity values of each of the models indicate that deep learning has a strong capability of correctly identifying individuals with good sleep patterns from their preceding awake activity. The specificity is high for TB-LSTM, MLP, and CNN, indicating that these

models were also able to successfully distinguish those with poor sleep patterns.

Relevance of Findings

Previous algorithms, such as by Sadeh et al [27,37], do not focus on the prediction of sleep quality based on physical activity

during awake time, but rather on prediction of sleep quality based on accelerometer data during the sleep periods. The objective of those studies was the validation of the actigraphy data compared with PSG. Consequently, the objectives and findings of those studies cannot be compared with our study.

There are several data analytics areas relevant to our work. To our knowledge, our research was the first one looking into the use of deep learning for the study of actigraphy data related to physical activity and sleep. Previous research on the use of deep learning for sleep science has been focused on PSG data [45,46]. In other application areas, deep learning has been used for human activity recognition [47,48] which is a similar technical problem. In a previous study, we combined human recognition of actigraphy data with other machine learning algorithms, but not deep learning [49].

Impact in Sleep Science

Sleep insufficiency is highly prevalent in contemporary society, and has been shown to influence energy balance by altering metabolic hormone regulation. Consequently, health researchers are exploring the impact of sleep and physical activity on many health conditions. A major bottleneck for this research is that current approaches for studying actigraphy data require intensive manual work by human experts. Furthermore, huge datasets of actigraphy data are emerging from health research, including the study of sleep disorder patients, healthy populations, and epidemiological studies. Furthermore, millions of consumers are buying wearables that incorporate activity sensors. This burst of human activity data is a great opportunity for health research, but to achieve this paradigm shift, it is necessary to develop new algorithms and tools to analyze this type of data.

As explained in the results, our findings supported the feasibility of using physical activity data to predict the quality of sleep in terms of sleep efficiency. These findings were by no means aiming to substitute well-studied algorithms and methods for studying sleep and physical activity data, such as the methods described by Sadeh [37]. Improved algorithms, such as the ones we presented in this study, for actigraphy analysis can lead to a paradigm shift in the study of lifestyle behaviors such as sleep and physical activity, just as electrocardiography became crucial for cardiology and clinical research.

Our research showed that deep learning performed better than classical methods in terms of learning useful patterns from raw accelerometer data for the task of sleep quality prediction. Since deep learning models compute abstract features from raw input signals while optimizing on the actual sleep quality, this process yields a more robust solution. Furthermore, the good results of deep learning showed that raw accelerometer data had more "signal" regarding sleep quality, which traditional models such as logistic regression are not able to capture right now. More research needs to be done to understand why deep learning performs better, which eventually can help in identifying new factors influencing the quality of sleep.

Impact in eHealth

Our study provided an early example of how advanced deep learning methods could be used to infer new insights from raw actigraphy data. Our focus on predicting and forecasting can help design new eHealth applications where predictions are made to personalize coaching for patients or to facilitate decision making of professionals.

There is an increased interest in sleep in the health domain. This is consequently being reflected in an increase use of eHealth for sleep [50-54] and also in the use of social media for sharing sleep logs [53]. The expansion of eHealth into sleep is not limited to sleep disorders, but also to improve sleep for people living with chronic conditions such as cancer [53]. These developments are closely related to the concept of Quantified Self for health [55,56]. Furthermore, we can assume that predicting sleep quality based on physical activity data acquired by accelerometer data (both from actigraphy or activity trackers) can be used to provide personalized feedback, such as momentary ecological interventions based on mobile technology [57].

In our study, we attempted to predict a parameter regarding the quality of sleep solely relying on the physical activity during the awake time. To our knowledge, this was not done earlier. The advantage of this approach is that eventually the same approach can be used to predict sleep quality with data from smart watches and other wearable devices that are not necessarily used during sleep. Therefore, our models can eventually be used within eHealth applications that do not require wearing a sensor during sleep. This is of special interest for the development of smart watch health apps [33], as they might require frequent battery charging.

Our research has many limitations as explained in the next subsection. However, this is the first study, to our knowledge, that focused on the prediction of sleep quality from physical activity accelerometer data during awake periods. Our methodology and results can be used as the baseline for further studies looking into predicting sleep quality from mobile and wearable devices. This is a source of major concern, since many sleep apps in the making predict with unclear methodology and performance [50-52]. Although more studies are highlighting the increasing reliability of consumer sleep wearables [31], we do not know how they calculate or predict sleep quality parameters. To maximize the potential of the wearable mass adoption for sleep health, we need research on not only the reliability of consumer-grade devices but also their data processing and modeling techniques.

Limitations

There were some limitations in our study regarding the generalization of our results. Sleep behavior can be affected by cultural aspects and also change with age. Our study sample drew sleep data from adolescents, aged 10-17 years, living in Qatar. Future research will need to evaluate whether applications of deep learning for sleep research using actigraphy will yield similar results in different populations (eg, adults and people with chronic conditions).

In our study, the prediction was simplified to "good" and "poor" sleep quality with regards to sleep efficiency. This may be an oversimplification of complex sleep problems. To provide more precise predictions (eg, quantitative value of sleep quality), these techniques will need to be validated. A prediction, such

as the one presented in this study, might be useful for the detection of people with unhealthy sleep patterns, but not to identify the causes of poor sleep efficiency.

There is also a limitation in the interpretation of deep learning. Deep learning models are "black boxes" and do not provide explanation of their sleep efficiency prediction. Other techniques such as logistic regression can provide insights on which features contribute to the prediction. However, this study showed that the performance of such models was much lower than that of deep learning. New techniques in deep learning are being researched to facilitate the interpretation of such models.

In our study, the models used the data from the individual's awake time to predict of sleep quality. The prediction was made at the last moment before sleep using the full awake time activity. If these models were to be used to provide personalized feedback to individuals with sleep problems, they will need to be tested with fragments of the awake time, giving an individual time to alter their behavior. Since our data was fragmented into sleep periods and awake times specific to an individual, the models would be able to handle varying durations of awake time.

Conclusions

Our study showed the feasibility of deep learning in predicting sleep efficiency using wearable data from awake periods. This is of paramount importance because deep learning eliminated the need for data preprocessing and simplified the overall workflow in sleep data research. The feasibility of our approach can lead to new applications in sleep science and also to the development of more complex eHealth sleep applications for both professionals and patients. These models can also be integrated in the broader context of quantified self [55,56].

Conflicts of Interest

None declared.

Authors' Contributions

AS was involved in all the research and manuscript preparation. SJ and FO advised AS on the technical aspects of the study and participated in manuscript preparation. LFL, AE, and JS advised AS on the study design and participated in manuscript preparation. LFL is the Principal Investigator of the eHealth project 360QS [58] at QCRI. ST and TA participated as health researchers in the study design, data collection, and the institutional review board application, and reviewed and approved the manuscript. ST is the Principal Investigator of the clinical study.

References

1. Strine TW, Chapman DP. Associations of frequent sleep insufficiency with health-related quality of life and health behaviors. Sleep Med 2005 Jan;6(1):23-27. [doi: 10.1016/j.sleep.2004.06.003] [Medline: 15680291]
2. Colten HR, Altevogt BR, Institute of Medicine (US) Committee on Sleep Medicine and Research. Sleep Disorders and Sleep Deprivation: An Unmet Public Health Problem. Washington, DC: National Academies Press; 2006.
3. Knutson KL, Ryden AM, Mander BA, Van CE. Role of sleep duration and quality in the risk and severity of type 2 diabetes mellitus. Arch Intern Med 2006 Sep 18;166(16):1768-1774. [doi: 10.1001/archinte.166.16.1768] [Medline: 16983057]
4. Gottlieb DJ, Punjabi NM, Newman AB, Resnick HE, Redline S, Baldwin CM, et al. Association of sleep time with diabetes mellitus and impaired glucose tolerance. Arch Intern Med 2005 Apr 25;165(8):863-867. [doi: 10.1001/archinte.165.8.863] [Medline: 15851636]
5. Nilsson PM, Rööst M, Engström G, Hedblad B, Berglund G. Incidence of diabetes in middle-aged men is related to sleep disturbances. Diabetes Care 2004 Oct;27(10):2464-2469. [Medline: 15451917]
6. Palagini L, Bruno RM, Gemignani A, Baglioni C, Ghiadoni L, Riemann D. Sleep loss and hypertension: a systematic review. Curr Pharm Des 2013;19(13):2409-2419. [Medline: 23173590]
7. Kasasbeh E, Chi DS, Krishnaswamy G. Inflammatory aspects of sleep apnea and their cardiovascular consequences. South Med J 2006 Jan;99(1):58-67; quiz 68-9. [doi: 10.1097/01.smj.0000197705.99639.50] [Medline: 16466124]
8. Meier-Ewert HK, Ridker PM, Rifai N, Regan MM, Price NJ, Dinges DF, et al. Effect of sleep loss on C-reactive protein, an inflammatory marker of cardiovascular risk. J Am Coll Cardiol 2004 Feb 18;43(4):678-683 [FREE Full text] [doi: 10.1016/j.jacc.2003.07.050] [Medline: 14975482]
9. Cohen S, Doyle WJ, Alper CM, Janicki-Deverts D, Turner RB. Sleep habits and susceptibility to the common cold. Arch Intern Med 2009 Jan 12;169(1):62-67 [FREE Full text] [doi: 10.1001/archinternmed.2008.505] [Medline: 19139325]
10. Opp MR, Toth LA. Neural-immune interactions in the regulation of sleep. Front Biosci 2003 May 01;8:d768-d779. [Medline: 12700057]

XSL•FO

RenderX

11.   Peterman JS, Carper MM, Kendall PC. Anxiety disorders and comorbid sleep problems in school-aged youth: review and future research directions. Child Psychiatry Hum Dev 2015 Jun;46(3):376-392. [doi: 10.1007/s10578-014-0478-y] [Medline: 24962165]

12.   Murphy MJ, Peterson MJ. Sleep disturbances in depression. Sleep Med Clin 2015 Mar;10(1):17-23. [doi: 10.1016/j.jsmc.2014.11.009] [Medline: 26055669]

13.   Williamson AM, Feyer AM. Moderate sleep deprivation produces impairments in cognitive and motor performance equivalent to legally prescribed levels of alcohol intoxication. Occup Environ Med 2000 Oct;57(10):649-655 [FREE Full text] [Medline: 10984335]

14.   Van Dongen HP, Maislin G, Mullington JM, Dinges DF. The cumulative cost of additional wakefulness: dose-response effects on neurobehavioral functions and sleep physiology from chronic sleep restriction and total sleep deprivation. Sleep 2003 Mar 15;26(2):117-126. [Medline: 12683469]

15.   Ellenbogen JM, Payne JD, Stickgold R. The role of sleep in declarative memory consolidation: passive, permissive, active or none? Curr Opin Neurobiol 2006 Dec;16(6):716-722. [doi: 10.1016/j.conb.2006.10.006] [Medline: 17085038]

16.   Taheri S, Lin L, Austin D, Young T, Mignot E. Short sleep duration is associated with reduced leptin, elevated ghrelin, and increased body mass index. PLoS Med 2004 Dec;1(3):e62 [FREE Full text] [doi: 10.1371/journal.pmed.0010062] [Medline: 15602591]

17.   Arora T, Taheri S. Associations among late chronotype, body mass index and dietary behaviors in young adolescents. Int J Obes (Lond) 2015 Jan;39(1):39-44. [doi: 10.1038/ijo.2014.157] [Medline: 25135376]

18.   Arora T, Hosseini-Araghi M, Bishop J, Yao GL, Thomas GN, Taheri S. The complexity of obesity in U.K. adolescents: relationships with quantity and type of technology, sleep duration and quality, academic performance and aspiration. Pediatr Obes 2013 Oct;8(5):358-366. [doi: 10.1111/j.2047-6310.2012.00119.x] [Medline: 23239604]

19.   Paruthi S, Brooks LJ, D'Ambrosio C, Hall WA, Kotagal S, Lloyd RM, et al. Recommended amount of sleep for pediatric populations: a consensus statement of the American Academy of Sleep Medicine. J Clin Sleep Med 2016 Jun 15;12(6):785-786. [doi: 10.5664/jcsm.5866] [Medline: 27250809]

20.   Kredlow MA, Capozzoli MC, Hearon BA, Calkins AW, Otto MW. The effects of physical activity on sleep: a meta-analytic review. J Behav Med 2015 Jun;38(3):427-449. [doi: 10.1007/s10865-015-9617-6] [Medline: 25596964]

21.   Driver HS, Taylor SR. Exercise and sleep. Sleep Med Rev 2000 Aug;4(4):387-402. [doi: 10.1053/smrv.2000.0110] [Medline: 12531177]

22.   Chennaoui M, Arnal PJ, Sauvet F, Léger D. Sleep and exercise: a reciprocal issue? Sleep Med Rev 2015 Apr;20:59-72. [doi: 10.1016/j.smrv.2014.06.008] [Medline: 25127157]

23.   Morgenthaler T, Alessi C, Friedman L, Owens J, Kapur V, Boehlecke B, Standards of Practice Committee, American Academy of Sleep Medicine. Practice parameters for the use of actigraphy in the assessment of sleep and sleep disorders: an update for 2007. Sleep 2007 Apr;30(4):519-529. [Medline: 17520797]

24.   Hirshkowitz M. The history of polysomnography: tool of scientific discovery. In: Chokroverty S, Billiard M, editors. Sleep Medicine: A Comprehensive Guide to Its Development, Clinical Milestones, and Advances in Treatment. New York, USA: Springer; 2015:91-100.

25.   Bruyneel M, Van den Broecke S, Libert W, Ninane V. Real-time attended home-polysomnography with telematic data transmission. Int J Med Inform 2013 Aug;82(8):696-701. [doi: 10.1016/j.ijmedinf.2013.02.008] [Medline: 23529100]

26.   Masa JF, Corral J, Pereira R, Duran-Cantolla J, Cabello M, Hernández-Blasco L, et al. Therapeutic decision-making for sleep apnea and hypopnea syndrome using home respiratory polygraphy: a large multicentric study. Am J Respir Crit Care Med 2011 Oct 15;184(8):964-971. [doi: 10.1164/rccm.201103-0428OC] [Medline: 21737584]

27.   Sadeh A. The role and validity of actigraphy in Sleep Medicine: an update. Sleep Med Rev 2011 Aug;15(4):259-267. [doi: 10.1016/j.smrv.2010.10.001] [Medline: 21237680]

28.   Arora T, Broglia E, Pushpakumar D, Lodhi T, Taheri S. An investigation into the strength of the association and agreement levels between subjective and objective sleep duration in adolescents. PLoS One 2013;8(8):e72406 [FREE Full text] [doi: 10.1371/journal.pone.0072406] [Medline: 23951321]

29.   Lee I, Shiroma EJ. Using accelerometers to measure physical activity in large-scale epidemiological studies: issues and challenges. Br J Sports Med 2014 Feb;48(3):197-201 [FREE Full text] [doi: 10.1136/bjsports-2013-093154] [Medline: 24297837]

30.   Ehsan M, Khan R, Wakefield D, Qureshi A, Murray L, Zuwallack R, et al. A longitudinal study evaluating the effect of exacerbations on physical activity in patients with chronic obstructive pulmonary disease. Ann Am Thorac Soc 2013 Dec;10(6):559-564. [doi: 10.1513/AnnalsATS.201304-100OC] [Medline: 24053416]

31.   Montgomery-Downs HE, Insana SP, Bond JA. Movement toward a novel activity monitoring device. Sleep Breath 2012 Sep;16(3):913-917. [doi: 10.1007/s11325-011-0585-y] [Medline: 21971963]

32.   Poirier J, Bennett WL, Jerome GJ, Shah NG, Lazo M, Yeh H, et al. Effectiveness of an activity tracker- and Internet-based adaptive walking program for adults: a randomized controlled trial. J Med Internet Res 2016;18(2):e34 [FREE Full text] [doi: 10.2196/jmir.5295] [Medline: 26860434]

33.   Lu T, Fu C, Ma MH, Fang C, Turner AM. Healthcare applications of smart watches: a systematic review. Appl Clin Inform 2016 Sep 14;7(3):850-869. [doi: 10.4338/ACI-2016-03-R-0042] [Medline: 27623763]

XSL•FO
RenderX

34. Wu W, Dasgupta S, Ramirez EE, Peterson C, Norman GJ. Classification accuracies of physical activities using smartphone motion sensors. J Med Internet Res 2012;14(5):e130 [FREE Full text] [doi: 10.2196/jmir.2208] [Medline: 23041431]

35. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. Med Sci Sports Exerc 1998 May;30(5):777-781. [Medline: 9588623]

36. Grigg-Damberger MM. The AASM Scoring Manual four years later. J Clin Sleep Med 2012 Jun 15;8(3):323-332 [FREE Full text] [doi: 10.5664/jcsm.1928] [Medline: 22701392]

37. Sadeh A, Raviv A, Gruber R. Sleep patterns and sleep disruptions in school-age children. Dev Psychol 2000 May;36(3):291-301. [Medline: 10830974]

38. Reed DL, Sacco WP. Measuring sleep efficiency: what should the denominator be? J Clin Sleep Med 2016 Feb;12(2):263-266 [FREE Full text] [doi: 10.5664/jcsm.5498] [Medline: 26194727]

39. Kim Y, Lee J, Peters BP, Gaesser GA, Welk GJ. Examination of different accelerometer cut-points for assessing sedentary behaviors in children. PLoS One 2014;9(4):e90630 [FREE Full text] [doi: 10.1371/journal.pone.0090630] [Medline: 24699259]

40. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. Med Sci Sports Exerc 2011 Feb;43(2):357-364 [FREE Full text] [doi: 10.1249/MSS.0b013e3181ed61a3] [Medline: 20581716]

41. Tieleman T, Hinton G. COURSERA: Neural Networks for Machine Learning. YouTube. 2012. URL: https://www.youtube.com/watch?v=LGA-gRkLEsI [accessed 2016-10-27] [WebCite Cache ID 6lZHTyQZN]

42. Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. J Mach Learn Res. 2014 Jan. URL: https://www.cs.toronto.edu/~hinton/absps/JMLRdropout.pdf [accessed 2016-10-28] [WebCite Cache ID 6lacTnQRx]

43. Roeser K, Eichholz R, Schwerdtle B, Schlarb AA, Kübler A. Relationship of sleep quality and health-related quality of life in adolescents according to self- and proxy ratings: a questionnaire survey. Front Psychiatry 2012;3:76 [FREE Full text] [doi: 10.3389/fpsyt.2012.00076] [Medline: 22969731]

44. Gruber R, Somerville G, Enros P, Paquin S, Kestler M, Gillies-Poitras E. Sleep efficiency (but not sleep duration) of healthy school-age children is associated with grades in math and languages. Sleep Med 2014 Dec;15(12):1517-1525. [doi: 10.1016/j.sleep.2014.08.009] [Medline: 25441747]

45. Shmiel O, Shmiel T, Dagan Y, Teicher M. Data mining techniques for detection of sleep arousals. J Neurosci Methods 2009 May 15;179(2):331-337. [doi: 10.1016/j.jneumeth.2009.01.035] [Medline: 19428545]

46. Längkvist M, Karlsson L, Loutfi A. Sleep stage classification using unsupervised feature learning. Adv Artif Neural Syst 2012:1-9. [doi: 10.1155/2012/107046]

47. Wang L, Qiao Y, Tang X. Action recognition with trajectory- pooled deep-convolutional descriptors. Presented at: IEEE Conference on Computer Vision and Pattern Recognition (CVPR); June, 2015; Boston, MA, USA p. 4305-4314. [doi: 10.1109/CVPR.2015.7299059]

48. Sun L, Jia K, Yeung Y, Shi H. Human action recognition using factorized spatio-temporal convolutional networks. Presented at: IEEE International Conference on Computer Vision (ICCV). IEEE; December, 2015; Santiago, Chile p. 4597-4605. [doi: 10.1109/ICCV.2015.522]

49. Sathyanarayana, A, Ofli F, Fernandez-Luque L, Srivastava J, Elmagarmid A, Arora T, et al. Robust Automated Human Activity Recognition and its Application to Sleep Research. Presented at: IEEE International Conference on Data Mining (ICDM), Data Mining in Human Activity Analysis Workshop; 2016; Barcelona, Spain URL: https://arxiv.org/abs/1607.04867

50. Behar J, Roebuck A, Domingos JS, Gederi E, Clifford GD. A review of current sleep screening applications for smartphones. Physiol Meas 2013 Jul;34(7):R29-R46. [doi: 10.1088/0967-3334/34/7/R29] [Medline: 23771214]

51. de ZM, Baker FC, Colrain IM. Validation of sleep-tracking technology compared with polysomnography in adolescents. Sleep 2015 Sep 01;38(9):1461-1468 [FREE Full text] [doi: 10.5665/sleep.4990] [Medline: 26158896]

52. Bhat S, Ferraris A, Gupta D, Mozafarian M, DeBari VA, Gushway-Henry N, et al. Is there a clinical role for smartphone sleep apps? Comparison of sleep cycle detection by a smartphone application to polysomnography. J Clin Sleep Med 2015 Jul 15;11(7):709-715 [FREE Full text] [doi: 10.5664/jcsm.4840] [Medline: 25766719]

53. Min YH, Lee JW, Shin Y, Jo M, Sohn G, Lee J, et al. Daily collection of self-reporting sleep disturbance data via a smartphone app in breast cancer patients receiving chemotherapy: a feasibility study. J Med Internet Res 2014;16(5):e135 [FREE Full text] [doi: 10.2196/jmir.3421] [Medline: 24860070]

54. McIver DJ, Hawkins JB, Chunara R, Chatterjee AK, Bhandari A, Fitzgerald TP, et al. Characterizing sleep issues using Twitter. J Med Internet Res 2015 Jun 08;17(6):e140 [FREE Full text] [doi: 10.2196/jmir.4476] [Medline: 26054530]

55. Almalki M, Gray K, Martin-Sanchez F. Activity theory as a theoretical framework for health self-quantification: a systematic review of empirical studies. J Med Internet Res 2016 May 27;18(5):e131 [FREE Full text] [doi: 10.2196/jmir.5000] [Medline: 27234343]

56. Paton C, Hansen M, Fernandez-Luque L, Lau AY. Self-tracking, social media and personal health records for patient empowered self-care: contribution of the IMIA Social Media Working Group. Yearb Med Inform 2012;7:16-24. [Medline: 22890336]

57.   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]
58.   Haddadi H, Ofli F, Mejova Y, Weber I, Srivastava J. 360 Quantified Self. Presented at: IEEE International Conference on
       Healthcare Informatics (ICHI); 2015; Dalla, TX, USA p. 587-592. [doi: 10.1109/ICHI.2015.95]

Abbreviations

AUC: Area under ROC curve
CNN: convolutional neural network
LR: logistic regression
LSTM-RNN: long short-term memory RNN
MLP: multilayer perceptron
PSG: polysomnography
ROC: receiver operating characteristic
RNN: recurrent neural network
SpO2: oxygen saturation
TB-LSTM: time-batched version of LSTM-RNN

[XSL•FO]
RenderX

Original Paper

A Brief Tool to Assess Image-Based Dietary Records and Guide Nutrition Counselling Among Pregnant Women: An Evaluation

Amy M Ashman[1,2,3], B Nutr & Diet; Clare E Collins[1,2], PhD; Leanne J Brown[4], PhD; Kym M Rae[3,4,5,6], PhD; Megan E Rollo[1,2], PhD

[1]School of Health Sciences, Faculty of Health and Medicine, University of Newcastle, Callaghan, Australia

[2]Priority Research Centre in Physical Activity and Nutrition, Faculty of Health and Medicine, University of Newcastle, Callaghan, Australia

[3]Gomeroi gaaynggal Centre, Faculty of Health and Medicine, University of Newcastle, Tamworth, Australia

[4]Department of Rural Health, Faculty of Health and Medicine, University of Newcastle, Tamworth, Australia

[5]Priority Research Centre in Reproduction, Faculty of Health and Medicine, University of Newcastle, Callaghan, Australia

[6]Mothers and Babies Research Centre, Faculty of Health and Medicine, University of Newcastle, New Lambton Heights, Australia

Corresponding Author:
Megan E Rollo, PhD
Priority Research Centre in Physical Activity and Nutrition
Faculty of Health and Medicine
School of Health Sciences Office
University of Newcastle
Room HA12, Hunter Building, University Drive
Callaghan, 2308
Australia
Phone: 61 (02) 4921 5649
Fax: 61 (02) 4921 2084
Email: megan.rollo@newcastle.edu.au

Abstract

Background: Dietitians ideally should provide personally tailored nutrition advice to pregnant women. Provision is hampered by a lack of appropriate tools for nutrition assessment and counselling in practice settings. Smartphone technology, through the use of image-based dietary records, can address limitations of traditional methods of recording dietary intake. Feedback on these records can then be provided by the dietitian via smartphone. Efficacy and validity of these methods requires examination.

Objective: The aims of the Australian Diet Bytes and Baby Bumps study, which used image-based dietary records and a purpose-built brief Selected Nutrient and Diet Quality (SNaQ) tool to provide tailored nutrition advice to pregnant women, were to assess relative validity of the SNaQ tool for analyzing dietary intake compared with nutrient analysis software, to describe the nutritional intake adequacy of pregnant participants, and to assess acceptability of dietary feedback via smartphone.

Methods: Eligible women used a smartphone app to record everything they consumed over 3 nonconsecutive days. Records consisted of an image of the food or drink item placed next to a fiducial marker, with a voice or text description, or both, providing additional detail. We used the SNaQ tool to analyze participants' intake of daily food group servings and selected key micronutrients for pregnancy relative to Australian guideline recommendations. A visual reference guide consisting of images of foods and drinks in standard serving sizes assisted the dietitian with quantification. Feedback on participants' diets was provided via 2 methods: (1) a short video summary sent to participants' smartphones, and (2) a follow-up telephone consultation with a dietitian. Agreement between dietary intake assessment using the SNaQ tool and nutrient analysis software was evaluated using Spearman rank correlation and Cohen kappa.

Results: We enrolled 27 women (median age 28.8 years, 8 Indigenous Australians, 15 primiparas), of whom 25 completed the image-based dietary record. Median intakes of grains, vegetables, fruit, meat, and dairy were below recommendations. Median (interquartile range) intake of energy-dense, nutrient-poor foods was 3.5 (2.4-3.9) servings/day and exceeded recommendations (0-2.5 servings/day). Positive correlations between the SNaQ tool and nutrient analysis software were observed for energy (ρ=.898, P<.001) and all selected micronutrients (iron, calcium, zinc, folate, and iodine, ρ range .510-.955, all P<.05), both with and without vitamin and mineral supplements included in the analysis. Cohen kappa showed moderate to substantial agreement for selected micronutrients when supplements were included (kappa range .488-.803, all $P \leq$.001) and for calcium, iodine, and zinc

when excluded (kappa range .554-.632, all *P*<.001). A total of 17 women reported changing their diet as a result of the personalized nutrition advice.

Conclusions: The SNaQ tool demonstrated acceptable validity for assessing adequacy of key pregnancy nutrient intakes and preliminary evidence of utility to support dietitians in providing women with personalized advice to optimize nutrition during pregnancy.

Introduction

Dietitians can assess individual dietary needs and provide advice to clients to optimize their nutritional status [1]. In order to deliver personalized nutrition interventions, accurate information about what individuals are eating is required. For collection of such information to be feasible, the dietary data need to be collected and interpreted with a minimum burden on both the client and dietitian [2]. Feedback should be tailored to the individual and provided in a manner that is meaningful to the recipient so as to encourage positive dietary changes.

Traditional prospective methods of dietary assessment, including weighed or estimated food records, require the recording of all food and drinks consumed. These methods can capture day-to-day variation in diets and are used commonly in research [3]. However, keeping food records is associated with a high participant burden involved in the weighing or estimating of foods, may trigger changes in usual eating behaviors [4,5], and requires high levels of motivation to complete records accurately [3]. Reliability of written records decreases over time due to respondent fatigue, especially for recording periods of more than 4 days [6]. Keeping food records also requires literacy and numeracy skills and therefore may not be appropriate for all population groups. In clinical practice, retrospective methods of dietary assessment, such as diet histories and 24-hour food recalls, are more likely to be used. However, self-report places the onus on individuals to estimate food quantities consumed, a limitation that contributes to underreporting [7,8].

Manual analysis of food records by dietitians or other trained individuals is often required to translate reported food intakes into nutrients and food groups. This analysis is usually undertaken using food composition tables, often embedded in food analysis software. Food composition tables provide detailed information on nutrient composition of foods and drinks, giving determined values for quantities of energy, macronutrients (carbohydrate, protein, and fat), micronutrients (vitamins and minerals), and other food components, such as fiber [9].

Once dietary intake is analyzed, nutrient intakes can be compared with national recommendations. In Australia, the *Nutrient Reference Values* (NRVs) provide national intake recommendations for macro- and micronutrients [10]. The *Australian Guide to Healthy Eating* (AGTHE) is a visual food selection guide providing a representation of the proportions of food groups recommended for daily consumption [11]. The AGTHE supports the *Australian Dietary Guidelines* recommendation to "enjoy a wide variety of nutritious foods

from these five food groups every day:" grain and cereal foods, vegetables, fruit, meat or alternatives ("meat"), and dairy or alternatives ("dairy") [11]. The AGTHE is used as an educational and counselling tool by Australian dietitians to advise on the recommended number of daily servings from each food group and serving sizes from core nutrient-dense and noncore or discretionary energy-dense, nutrient-poor foods.

Innovative dietary assessment methods can address some of the limitations associated with current methods in order to improve the quality of data collected and ease of analysis. Image-based dietary records are a novel method for food and nutrient intake assessment [2,3], where images of consumed food and drinks capture a dietary record from which a person's intake is determined [12]. A passive or active approach can be taken to capturing food intake. A passive approach involves wearable cameras that capture eating and drinking occasions [13,14]. While no effort from users is needed, privacy issues associated with this technology make passive methods of image capture challenging to implement. Active methods involve recording dietary intake via stand-alone cameras or those imbedded in handheld devices, such as smartphones. Although the active method relies on participants to capture the images, the burden of estimating portion size is placed on the dietitian or skilled person performing the analysis [15]. Smartphone ownership is increasing, with 77% of Australian adults owning a smartphone in 2015 [16]. Smartphone features such as cameras, microphones, and Internet connectivity make them an ideal mode of dietary assessment, education, and counselling. With access to appropriate technologies and training, dietary intake data can be relayed between clients and dietitians in real time, transcending distance, and potentially overcoming barriers relating to literacy or numeracy skills. These assessment methods support the provision of dietary feedback over distance (eg, through telephone or video consultation), broadening the scope of dietetic services [17]. Practical tools can support the use of image-based dietary records for both the collection of information on dietary intake and the analysis and interpretation of food and nutrient intake data. However, their use in clinical settings is limited if these tools are not convenient, and validation is required to support manual analysis of image-based dietary records by dietitians.

Previous methods of image-based dietary assessment have been examined in healthy adult [18-22], adolescent [23,24], and child [25] populations, in overweight and obese adults [26], and in type 2 diabetes [15,27]. To our knowledge, no studies to date have examined the use of image-based dietary records in pregnant women or in Indigenous Australians. Dietary intake

and nutritional status during pregnancy have important implications for fetal development and growth, and the long-term health of both mother and infant [28-30]. In Australia, women of childbearing age are at risk of not meeting targets for recommended dietary intake (RDI) [31,32]. In particular, Indigenous women may experience structural barriers to optimal nutrition, including economic and geographical constraints to accessing food, and gaps in knowledge for choosing and preparing nutritious foods [33]. Novel lower-burden methods for dietary assessment and provision of feedback on nutrition warrant investigation and may be of benefit in these population groups.

The Diet Bytes and Baby Bumps (DBBB) study used image-based dietary records, captured via smartphone, in pregnant Indigenous and non-Indigenous women. The DBBB study sought to assess intake of AGTHE core and energy-dense, nutrient-poor food groups, total energy, and selected micronutrients, and to provide personalized feedback to these women via their smartphones, in combination with consultation with a dietitian.

The aims of this analysis were to evaluate the use of a brief approach to dietary analysis using a purpose-built Selected Nutrient and Diet Quality (SNaQ) tool to (1) assess nutrient intakes of pregnant women in the DBBB study, (2) assess the validity of the SNaQ tool for nutrient assessment relative to analysis using nutrient analysis software, and (3) assess the acceptability of SNaQ to pregnant women for provision of feedback on dietary intake.

The DBBB study was approved by the Aboriginal Health and Medical Research Council Ethics committee (962/13), Hunter New England Human Research Ethics Committee (13/06/19/4.04), and the University of Newcastle Human Research Ethics Committee (H-2013-0185). The study was conducted in two locations in New South Wales (NSW), Australia: Newcastle, the second largest city in NSW, and Tamworth, a regional inland NSW town.

Methods

Participants and Recruitment

We recruited participants via promotional fliers at hospital antenatal and general practitioner clinics and the University of Newcastle, through social media (including parenting sites), and through direct contact with pregnant women at antenatal clinics. In Tamworth, participants were also invited to participate through the Gomeroi gaaynggal Centre [34], an Indigenous research and ArtsHealth center. Participants were eligible if they were ≥18 years old, ≤24 weeks' gestation, lived in Newcastle or Tamworth, had no current medical conditions, owned a smartphone, and were willing to use it to record their dietary intake for 3 days.

Surveys and Study Timeline

The study ran for 12 weeks (Figure 1). Participants collected image-based dietary records in week 1, completed three 24-hour food recalls (in weeks 2, 3, and 4), received feedback on their dietary intake in week 6, and completed the Australian Eating Survey food frequency questionnaire in week 12 [35]. Participants completed 3 online surveys over the course of the study to provide demographic and background data (week 1, in-person study visit), evaluate the image-based dietary assessment method (week 2, in-person study visit), and evaluate the feedback on dietary intake that participants received (week 8, survey link sent via email).

Diet Bytes Method

We modelled the method of capturing dietary intake using image-based records on our previous validated method in adults with type 2 diabetes [15,27]. However, in this study, to record dietary intake, participants used Evernote (Evernote Corporation, Redwood City, CA, USA), a free file-sharing and note-taking app for computers and smartphones. The Evernote app was downloaded onto participants' smartphones during the first appointment. Participants were not expected to have any prior experience using the Evernote app. They were provided with training at the first appointment on how to use the app to record dietary intake and completed a test entry. The app was used to capture each eating occasion through notes or entries into a notebook (the dietary record). For the purpose of this study, the study team set up a shared notebook to allow the entries to be recorded. This notebook could only be viewed by the individual study participant and the research team who had access to the Evernote Diet Bytes account. We adjusted settings for Evernote so that the contents of the notebook were shared only with the research team over a Wi-Fi connection, so as not to use participants' data. Participants also had the option of disabling their home Wi-Fi connection during the collection period (week 1), with the images then transmitted during the second study appointment (week 2) over the research center's Wi-Fi connection. Participants were asked to collect information on all food, drinks, and nutritional supplements, such as prenatal vitamin and minerals, consumed over 3 nonconsecutive days, including 1 weekend day. Each eating occasion consisted of a note taken through the app, including an image of the food or drink items for consumption, with a fiducial marker (reference object) placed next to the items. Participants were also required to annotate a text or voice description, or both, of the image's contents with information relating to cooking methods, brands, and types of foods (Figure 2). Any food or drink not consumed was captured using the same process. Participants were encouraged to label each eating occasion at data entry (eg, "Breakfast day 1"). However, the Evernote app automatically captured the date and time when records are made, which assisted with determining when meals were consumed.

Figure 1. Diet Bytes and Baby Bumps study protocol.

Week 1
- Evernote app downloaded to participant smartphone and instructions on how to keep the image record are given
- Survey on demographic and background information
- Participants keep an image-based dietary record for 3 nonconsecutive days, including 1 weekend day

Week 2
- Image records uploaded
- Survey evaluating the image-based dietary assessment method
- 24-hour recall x1 (in-person)

Week 3
- 24-hour recall x1 (telephone-administered)

Week 4
- 24-hour recall x1 (telephone-administered)

Week 6
- Participants receive feedback on their dietary intake via short video summary (delivered to smartphone) and telephone follow-up with a dietitian

Week 8
- Survey evaluating the feedback participants received (online)

Week 12
- Australian Eating Survey food frequency questionnaire (online)

XSL•FO

RenderX

Figure 2. Example of an image-based dietary record in the Diet Bytes and Baby Bumps study, consisting of image, fiducial marker, and audio description of the food and drink items.
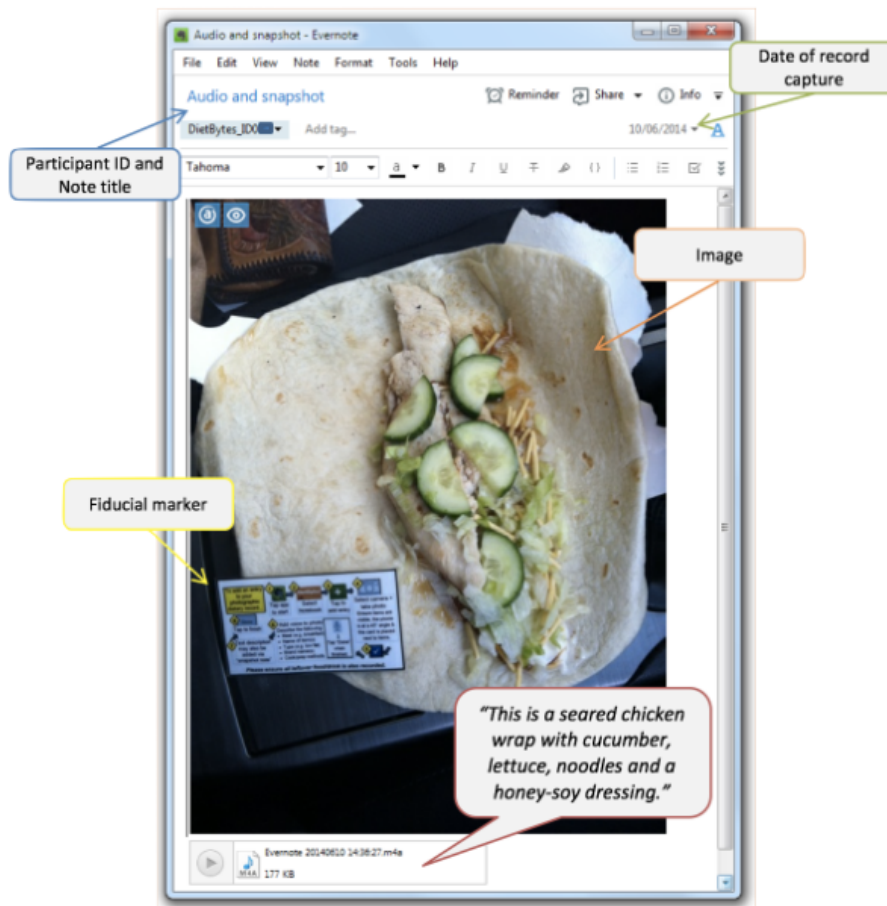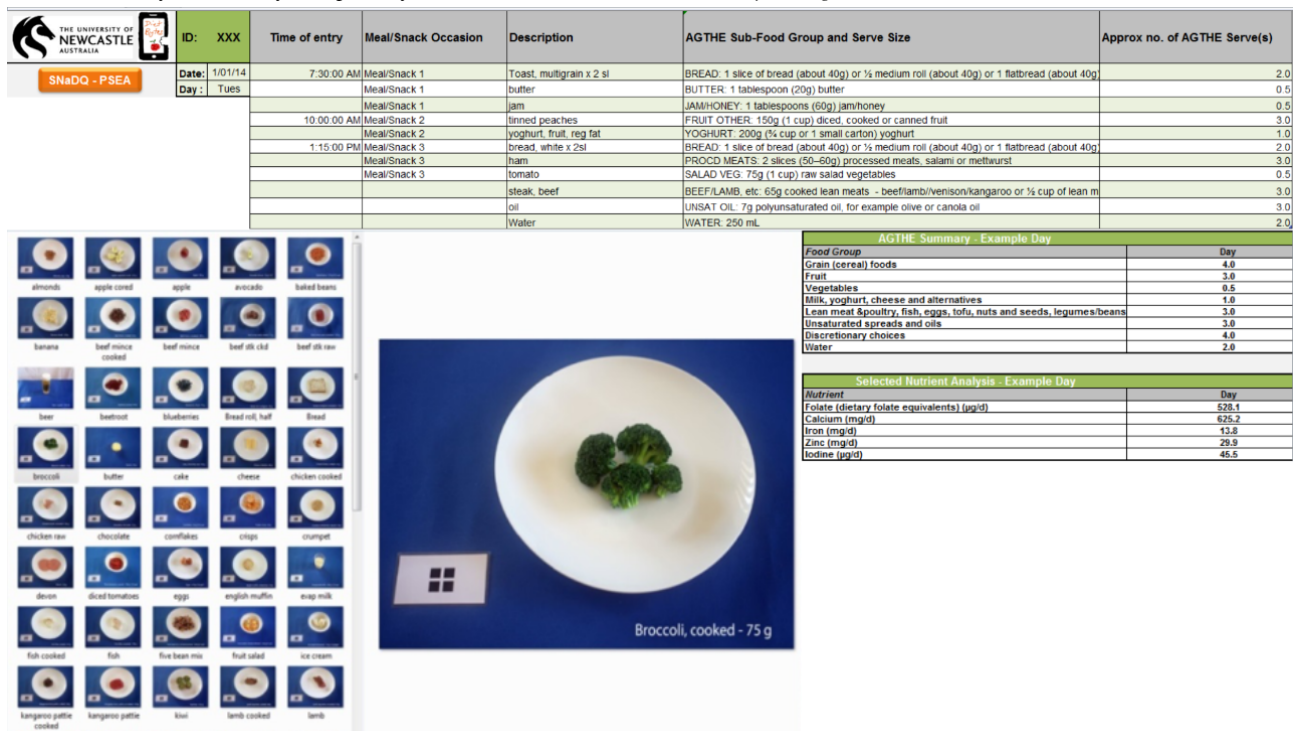

The SNaQ Tool

The SNaQ tool was developed as a brief tool to analyze participants' dietary intake relative to AGTHE daily servings of core and energy-dense, nutrient-poor foods. We estimated key nutrients important during pregnancy (folate, calcium, iron, zinc, and iodine) based on average nutrient composition of the food group servings, using the Australian Food, Supplement & Nutrient Database (AUSNUT) 2007 [36] food composition tables embedded in the SNaQ tool, plus nutrients from micronutrient supplements consumed.

A portion size estimation aid (PSEA) included in the tool assisted with portion size quantification. The PSEA contained 80 photographs of a variety of AGTHE foods and drinks displayed in recommended serving sizes. The dietitian analyzing food portions compared the image from the image-based dietary record with images in the PSEA, in order to quantify portion size of the food and drink items in terms of number of AGTHE servings (see Figure 3). The text or voice description supplementing the image-based record further assisted with

quantification. Mixed dishes and meals were broken down into their composite food groups. The image-based dietary records were first analyzed separately by 2 dietitians, who later conferred to confirm participant dietary intakes.

Feedback was provided to participants in week 6 of the study, via a short (1 minute) video designed to relay a simple, visual summary of food group intake compared with AGTHE recommendations. The video was transmitted to the Diet Bytes notebook, through the Evernote app on participants' smartphones. Participants were sent a text message informing them that their feedback was available to view. The video could be paused and replayed as often as desired. Participants were given a few days to view their feedback and were then contacted later in the week by a dietitian for a telephone consultation. In the telephone conversation, results were discussed in greater detail, including core and energy-dense, nutrient-poor food group results and intakes of selected nutrients, to provide practical tailored examples of foods and serving sizes to optimize the participant's pregnancy dietary intake.

Figure 3. The Selected Nutrient and Diet Quality (SNaQ) analysis tool and portion size estimation aid (PSEA) for analysis of image-based dietary records in the Diet Bytes and Baby Bumps study. AGTHE: *Australian Guide to Healthy Eating*.

Date:	1/01/14	Time of entry	Meal/Snack Occasion	Description	AGTHE Sub-Food Group and Serve Size	Approx no. of AGTHE Serve(s)
Day :	Tues	7:30:00 AM	Meal/Snack 1	Toast, multigrain x 2 sl	BREAD: 1 slice of bread (about 40g) or ½ medium roll (about 40g) or 1 flatbread (about 40g)	2.0
			Meal/Snack 1	butter	BUTTER: 1 tablespoon (20g) butter	0.5
			Meal/Snack 1	jam	JAM/HONEY: 1 tablespoons (60g) jam/honey	0.5
		10:00:00 AM	Meal/Snack 2	tinned peaches	FRUIT OTHER: 150g (1 cup) diced, cooked or canned fruit	3.0
			Meal/Snack 2	yoghurt, fruit, reg fat	YOGHURT: 200g (¾ cup or 1 small carton) yoghurt	1.0
		1:15:00 PM	Meal/Snack 3	bread, white x 2sl	BREAD: 1 slice of bread (about 40g) or ½ medium roll (about 40g) or 1 flatbread (about 40g)	2.0
			Meal/Snack 3	ham	PROCD MEATS: 2 slices (50–60g) processed meats, salami or mettwurst	3.0
			Meal/Snack 3	tomato	SALAD VEG: 75g (1 cup) raw salad vegetables	0.5
				steak, beef	BEEF/LAMB, etc: 65g cooked lean meats - beef/lamb//venison/kangaroo or ⅓ cup of lean m	3.0
				oil	UNSAT OIL: 7g polyunsaturated oil, for example olive or canola oil	3.0
				Water	WATER: 250 mL	2.0

AGTHE Summary - Example Day

Food Group	Day
Grain (cereal) foods	4.0
Fruit	3.0
Vegetables	0.5
Milk, yoghurt, cheese and alternatives	1.0
Lean meat &poultry, fish, eggs, tofu, nuts and seeds, legumes/beans	3.0
Unsaturated spreads and oils	3.0
Discretionary choices	4.0
Water	2.0

Selected Nutrient Analysis - Example Day

Nutrient	Day
Folate (dietary folate equivalents) (µg/d)	528.1
Calcium (mg/d)	625.2
Iron (mg/d)	13.8
Zinc (mg/d)	29.9
Iodine (µg/d)	45.5

Broccoli, cooked - 75 g

Statistical Analysis

We entered image-based dietary records into the nutrient composition software FoodWorks Professional version 7.0.3016 (Xyris Software [Australia] Pty Ltd) using the nutrient composition tables AUSNUT 2007 [36] (with "foods," "brands," and "supplements" selected). The PSEA assisted with the portion size estimation of the images for the FoodWorks entry using the same approach as for the SNaQ analysis, including the use of the image and text description for clarification of quantities, types, and brands of food and cooking methods. Data entered into the SNaQ tool, including information on the estimation of portion size, were not used during the analysis of the image-based records in FoodWorks software. We developed a protocol to standardize the entry of image-based dietary records into FoodWorks, including common assumptions made. For example, if the amount of butter or margarine on a piece of bread was unspecified, we assumed 1 teaspoon per slice and used the "not further specified" option for food types where possible when further details were not provided. Intraclass correlation coefficients for FoodWorks entries of the image-based records between 2 dietitians for a subsample of 10 participant records showed substantial agreement for energy and the selected micronutrients iron, folate, calcium, iodine, and zinc, in the range of .79-.99, all $P<.05$. One dietitian subsequently entered all image records into FoodWorks. We ascertained relative validity of the SNaQ tool in estimating participants' total energy and selected nutrient intakes by comparison with the FoodWorks nutrient assessment of the image-based dietary records, and assessed by the strength of the relationship using Spearman rank correlation coefficients (ρ) and agreement between the methods using Cohen kappa. Analyses were performed using IBM SPSS statistical software version 23.0 (IBM Corporation). We took an inductive approach to analyze short qualitative responses on participants' perceived acceptability of the feedback received [37].

Results

Characteristics of Participants

We enrolled 27 women in the DBBB study, with a median (interquartile range) age 28.8 (27.5-32.5) years, with 1 participant withdrawing due to time constraints. Of the remaining 26 participants, all were born in Australia, 8 (31%) identified as being of Indigenous descent, and all spoke only English at home. At study enrollment, 4 (15%) participants smoked tobacco products. At enrollment, participants ranged from 6 to 24 weeks' gestation, with a mean (SD) of 18 (5) weeks. A total of 4 participants were in their first trimester of pregnancy, and 22 in their second trimester. For 15 (58%) participants it was the first pregnancy; 14 (54%) participants had an undergraduate or postgraduate university degree; and 2 developed health conditions (gestational diabetes and anemia) during the study.

Over half (n=17, 65%) had received nutrition advice from a health professional previously, although only 5 (19%) had received advice from a dietitian. Other sources of nutrition advice came from a general practitioner (n=10, 38%), midwife (n=5, 19%), obstetrician (n=1, 4%), or an antenatal clinic (n=1, 4%). Advice received focused on use of multivitamin supplements (n=12, 46%), managing morning sickness (n=7, 27%), healthy eating throughout pregnancy (n=7, 27%), weight gain during pregnancy (n=5, 19%), healthy eating during breastfeeding (n=5, 19%), or breastfeeding (n=4, 15%). Participants had also accessed pregnancy nutrition information from other sources, including friends (n=11, 42%),

nongovernment websites (n=11, 42%), family (n=10, 38%), government websites (n=9, 35%), smartphone apps (n=7, 27%), and community groups, including mothers' groups (n=2, 8%); 3 (12%) participants had not accessed any of these sources of information. A total of 11 (42%) participants felt they had received enough information about healthy eating for themselves and their baby at the time of enrollment, 13 (50%) were unsure, and 2 (8%) said they had not received enough information.

All participants used their smartphones for sending text messages (short message service, SMS) (n=26, 100%), and the majority for receiving SMS (n=25, 96%), searching or browsing the Internet (n=25, 96%), making voice calls (n=24, 92%), taking photos (n=24, 92%), sending or uploading photos (n=24, 92%), using apps (n=22, 85%), and taking notes (n=20, 77%). Over half (n=16, 62%) used their smartphones for taking videos and 12 (46%) to send or upload these videos. The majority of participants (n=18, 69%) had an Apple iPhone, and 8 (31%) had a Google Android phone. Only 4 (15%) had used their smartphones for making voice recordings.

Food Group Intakes of Pregnant Women

Of the 26 participants, 24 (92%) recorded on all 3 days of the image-based dietary record, 1 participant recorded 2 days, and 1 recorded only 1 day. The participant recording on only 1 day was subsequently excluded from further analyses, and therefore further results are for the 25 participants with dietary records adequate for analysis. We used average food group and micronutrient intakes from participants' multiple-day image records for this analyses.

Table 1 summarizes intakes of core and energy-dense, nutrient-poor foods. Median intakes of core food groups were close to recommendations for fruit and dairy, but did not meet recommendations for grains and cereals, vegetables, or meat, and exceeded recommendations for energy-dense, nutrient-poor foods. All Indigenous participants and approximately half (n=8, 47%) of non-Indigenous participants met recommendations for 0-2.5 daily servings of unsaturated spreads and oils.

Relative Validity of the SNaQ Tool for Nutrient Assessment

Table 2 reports the correlations (Spearman correlation coefficients) and agreement (Cohen kappa) between nutrient values assessed from the SNaQ tool and from nutrient analysis software. Agreement was not substantial between the two methods for total energy (kappa=.031, P=.67). Correlation coefficients for nutrient intakes assessed by the two methods of analyzing the image-based dietary records ranged from ρ=.791 to ρ=.955 (all P<.001) for key micronutrients (iron, folate, calcium, zinc, and iodine) when supplements were included in the analysis (kappa range .488-.803, all P≤.001). With supplement use excluded, correlations ranged from ρ=.510 to ρ=.888 (all P<.05). Agreement between the two analysis methods, ascertained via Cohen kappa, was significant for calcium (kappa=.544, P<.001), iodine (kappa=.632, P<.001), and zinc (kappa=.572, P<.001). Agreement was poor for folate when supplement use was not included (kappa=-.068, P=.52). Both the SNaQ tool and FoodWorks analyses identified that no participant met the estimated average requirement (EAR) for iron of 22 mg when supplement use was not included.

Table 1. Intake of core foods as assessed by the Selected Nutrient and Diet Quality (SNaQ) brief analysis tool from the Diet Bytes and Baby Bumps image-based dietary records (n=25).

Food group	Food group intake in servings/day		AGTHE[b] recommended intake during pregnancy in servings/day	Meeting recommended intake of servings	
	Mean (SD)	Median (IQR[a])		No. of servings	n (%)
All participants combined (n=25)					
Grains and cereals	4.8 (2.0)	4.7 (3.6-6.5)	8.5	≥8.5	1 (4)
Vegetables	2.4 (1.4)	2.2 (1.2-3.5)	5	≥5	1 (4)
Fruit	1.9 (1.6)	1.7 (0.9-2.5)	2	≥2	10 (40)
Lean meat	2.0 (1.0)	1.9 (1.4-2.9)	3.5	≥3.5	2 (8)
Dairy	2.1 (1.3)	1.8 (1.3-2.7)	2.5	≥2.5	10 (40)
Unsaturated spreads and oils	1.9 (1.4)	2.0 (0.5-3.0)	0-2.5	0-2.5	16 (64)
Energy-dense, nutrient-poor foods	3.7 (1.9)	3.5 (2.4-3.9)	0-2.5	0-2.5	7 (28)
Indigenous participants (n=8)					
Grains and cereals	4.7 (2.3)	4.3 (3.4-6.1)	8.5	≥8.5	1 (13)
Vegetables	2.0 (1.4)	1.6 (1.1-3.2)	5	≥5	0 (0)
Fruit	1.4 (1.9)	0.9 (0.0-2.3)	2	≥2	2 (25)
Lean meat	1.6 (0.9)	1.5 (0.8-2.0)	3.5	≥3.5	0 (0)
Dairy	2.5 (1.9)	2.3 (1.0-3.4)	2.5	≥2.5	4 (50)
Unsaturated spreads and oils	0.8 (0.8)	0.7 (0.8-1.7)	0-2.5	0-2.5	8 (100)
Energy-dense, nutrient-poor foods	4.1 (2.9)	3.7 (1.6-7.1)	0-2.5	0-2.5	2 (25)
Non-Indigenous participants (n=17)					
Grains and cereals	4.9 (1.9)	4.9 (3.6-6.9)	8.5	≥8.5	0 (0)
Vegetables	2.6 (1.4)	2.4 (1.7-3.5)	5	≥5	1 (6)
Fruit	2.2 (1.4)	1.8 (1.4-2.7)	2	≥2	8 (47)
Lean meat	2.2 (1.0)	2.0 (1.7-3.1)	3.5	≥3.5	2 (12)
Dairy	1.9 (0.9)	1.7 (1.3-2.7)	2.5	≥2.5	6 (36)
Unsaturated spreads and oils	2.3 (1.4)	2.8 (1.0-3.3)	0-2.5	0-2.5	8 (47)
Energy-dense, nutrient-poor foods	3.5 (1.3)	3.5 (2.4-3.9)	0-2.5	0-2.5	5 (29)

[a]IQR: interquartile range (25th-75th percentiles).

[b]AGTHE: *Australian Guide to Healthy Eating* [11]. Examples of serving sizes of foods: grains and cereals (standard serving 500 kJ), eg, 1 slice of bread, 0.5 cup cooked grain; vegetables (standard serving 75 g, 100-350 kJ), eg, 0.5 cup cooked vegetables, 1 cup raw vegetables, 0.5 medium potato; fruit (standard serving 150 g, 350 kJ), eg, 1 medium piece, 2 small pieces, 125 mL fruit juice (no added sugar, only occasionally); lean meats and alternatives (standard serving 500-600 kJ), eg, 65 g cooked lean red meats, 80 g cooked lean poultry, 100 g cooked fish, 2 large eggs, 1 cup cooked legumes or beans; dairy and alternatives (standard serving 500-600 kJ), eg, 1 cup milk, 2 slices (40 g) hard cheese, 0.75 cup yoghurt, 60 g sardines; unsaturated spreads and oils (standard serving 250 kJ), eg, 10 g unsaturated spread, 7 g unsaturated oil, 10 g nuts; energy-dense, nutrient-poor foods (standard serving 600 kJ), eg, 2 scoops ice cream, 50-60 g processed meats, 1 can soft drink, 12 hot chips, 200 mL wine.

Table 2. Correlation and agreement for energy and selected nutrient intake from mean 3-day image-based dietary records in the Diet Bytes and Baby Bumps study (n=25 participants) analyzed by the Selected Nutrient and Diet Quality (SNaQ) tool and FoodWorks (FW) nutrient analysis software.

Nutrient	Method	Input, median (IQR[a])	ρ (P value)	n (%) <EAR[b]	n (%) ≥EAR to <RDI[c]	n (%) ≥RDI	Cohen kappa (P value)
Intake from food and supplements							
Energy (kJ/day)	SNaQ	8418.33 (7755.83-10,004.17)	.898 (<.001)	N/A[d]			.031[e] (.67)
	FW	7738.89 (6329.94-8995.05)					
Iron (mg/day)	SNaQ	11.30 (8.93-15.08)	.812 (<.001)	21 (84)	0 (0)	4 (16)	.533 (<.001)
	FW	13.54 (10.75-21.47)		19 (76)	3 (12)	3 (12)	
Calcium (mg/day)	SNaQ	877.36 (653.74-1181.60)	.791 (<.001)	12 (48)	4 (16)	9 (36)	.488 (.001)
	FW	831.01 (672.39-1000.89)		13 (52)	6 (24)	6 (24)	
Folate, total DFE[f] (μg/day)	SNaQ	851.90 (225.15-1156.15)	.893 (<.001)	11 (44)	1 (4)	13 (52)	.559 (.001)
	FW	820.20 (393.53-1383.00)		8 (32)	2 (8)	15 (60)	
Iodine (μg/day)	SNaQ	167.00 (93.52-311.28)	.955 (<.001)	11 (44)	4 (16)	10 (40)	.803 (<.001)
	FW	171.42 (92.58-300.20)		12 (48)	3 (12)	10 (40)	
Zinc (mg/day)	SNaQ	13.09 (10.46-19.56)	.905 (<.001)	3 (12)	4 (16)	18 (72)	.741 (<.001)
	FW	14.66 (10.24-21.24)		3 (12)	5 (20)	17 (68)	
Intake from food only, supplements excluded							
Energy (kJ/day)	SNaQ	8418.33 (7755.83-10,004.17)	.898 (.000)	N/A			.031[e] (.67)
	FW	7738.89 (6329.94-8995.05)					
Iron (mg/day)	SNaQ	9.50 (7.70-10.85)	.510 (.009)	25 (100)	0 (0)	0 (0)	Constants (no statistics computed)
	FW	11.78 (8.53, 13.73)		25 (100)	0 (0)	0 (0)	
Calcium (mg/day)	SNaQ	809.90 (653.75-1181.70)	.888 (<.001)	14 (56)	2 (8)	9 (36)	.554 (<.001)
	FW	736.61 (663.19-927.37)		17 (68)	3 (12)	5 (20)	
Folate, total DFE (μg/day)	SNaQ	319.00 (240.25-433.35)	.600 (.002)	21 (84)	4 (16)	0 (0)	-.068 (.52)
	FW	409.79 (259.74-642.22)		16 (64)	2 (8)	7 (28)	
Iodine (μg/day)	SNaQ	99.00 (79.80-139.05)	.850 (<.001)	22 (88)	2 (8)	1 (4)	.632 (<.001)
	FW	104.25 (86.46-130.95)		22 (88)	2 (8)	1 (4)	
Zinc (mg/day)	SNaQ	10.60 (8.40-13.10)	.745 (<.001)	7 (28)	6 (24)	12 (48)	.572 (<.001)
	FW	10.63 (8.89-13.47)		6 (24)	9 (36)	10 (40)	

[a]IQR: interquartile range (25th-75th percentiles).

[b]EAR: estimated average requirement. EAR is a nutrient level estimated to meet the requirements of 50% of the healthy individuals in a life stage or gender group, per day (EARs for nutrients as follows: iron 22 mg, calcium 840 mg, folate 520 μg, iodine 160 μg, zinc 9 mg) [10].

[c]RDI: recommended dietary intake. RDI is the average dietary intake level sufficient to meet nutrient requirements of 97% to 98% of healthy individuals in a life stage or gender group, per day (RDIs for nutrients as follows: iron 27 mg, calcium 1000 mg, folate 600 μg, iodine 220 μg, zinc 11 mg) [10].

Table 3. Participant's perceived acceptability for receiving dietary counselling in the Diet Bytes and Baby Bumps study (n=22[a]) (survey questions with agree-disagree responses).

Questions	Strongly agree n (%)	Agree n (%)	Neutral n (%)	Disagree n (%)	Strongly disagree n (%)
I believe that the combination of the summary of my dietary intake that I received via my mobile/smartphone and the follow-up with the dietitian was helpful.	12 (55)	9 (41)	1 (5)	0 (0)	0 (0)
The summary of my dietary intake that I received via my mobile/smartphone was easy to understand on its own. I did not need to speak to a dietitian to clarify.	2 (9)	5 (23)	6 (27)	8 (36)	1 (5)
The summary of my dietary intake that I received via my mobile/smartphone was difficult to understand.	0 (0)	2 (9)	3 (14)	13 (59)	4 (18)
Neither the summary of the results from the analysis of my photographic dietary record that I received on my mobile/smartphone nor the advice that I received from the dietitian was helpful.	0 (0)	1 (5)	1 (5)	4 (18)	16 (73)

[a]n=22. Two participants did not receive the telephone counselling (1 gave birth before it could be given and 1 did not respond to contact) and 2 participants did not answer this survey.

Table 4. Participant's perceived acceptability for receiving dietary counselling in the Diet Bytes and Baby Bumps study (n=22[a]) (survey questions with yes/no responses).

Questions	Yes, n (%)	No, n (%)
I have changed my diet as a result of the nutrition advice that I received as part of this study.	17 (77)	5 (23)
I have changed the kinds of foods I eat.	16 (73)	6 (27)
I have changed the amount of food I eat.	8 (36)	14 (64)
I have changed the cooking methods I use.	3 (14)	19 (86)
I have changed how I keep track of what I eat and drink.	5 (23)	17 (77)
I have made other changes.	1 (5)	21 (95)

[a]n=22. Two participants did not receive the telephone counselling (1 gave birth before it could be given and 1 did not respond to contact) and 2 participants did not answer this survey.

Acceptability of Receiving Feedback on Dietary Intake

Table 3 and Table 4 summarize participants' perceived acceptability of receiving nutrition feedback. Over three-quarters (n=17, 77%) of the 22 participants who responded to the final survey reported that they had made dietary changes as a result of the personalized nutrition feedback. Changes to the type of foods consumed fell into three categories: (1) food groups or individual foods, including eating more red meat, vegetables, fruit, and individual foods like Milo, yoghurt, cheese, and crackers); (2) nutrients, including consuming more foods higher in iron, calcium, and protein, or continuing or starting to take a prenatal vitamin or mineral supplement; and (3) changes to eating behaviors, including increasing snack occasions. Some participants reported eating greater quantities of foods from the core food groups, while others reported eating smaller amounts of some unspecified foods, and consuming less soft drink and "junk" foods. Some participants reported changes to cooking methods related to meat and vegetables, such as steaming vegetables, and using cooking spray rather than oil or butter to cook meat. When asked if participants had changed how they monitored their dietary intake, 1 participant reported sometimes using an app (not Evernote) to record her intake, although this was a behavior in place prior to the study.

Some participants thought the advice from a dietitian was useful and helped to clarify the feedback provided via the video summary; for example:

> ...the phone consult was very useful to me. Without this the written feedback would have been far less meaningful. I did like the visual graphs to help me understand the information. [27 years old, first baby]

Additionally, another participant commented:

> It was very detailed and thorough and easier to understand what should be done to improve my diet compared to the diet summary received on Evernote. [32 years old, first baby]

Others reported not making changes as a result of the feedback, due to already meeting requirements or not being able to fit all the recommended servings into their daily intake.

Some participants felt that DBBB could be improved by keeping the image-based records for a longer duration and by taking notes, rather than images, for certain foods such as snacks and water. More SMS reminders were requested, as some participants reported forgetting to take images. One commented that having to take an image before eating when you were hungry was inconvenient:

> It's inconvenient to take pictures of food before eating when hungry (which is most of the time), however I think this is a useful way of assessing dietary intake.
> [27 years old, first baby]

The majority of respondents (n=18, 82%), preferred receiving nutrition feedback via the combination of the video summary and follow-up telephone consultation with a dietitian. One indicated that she preferred the video feedback alone, and 3 preferred the consultation with the dietitian alone. Only 1 participant indicated an alternative method for receiving nutrition advice, via a printable email summary.

Discussion

We observed strong positive correlations between the SNaQ tool and the nutrient analysis software for estimates of total energy intake and all selected micronutrients (iron, calcium, zinc, folate, and iodine), both with and without micronutrient supplements included in the analysis. However, SNaQ overestimated energy intake compared with the FoodWorks analysis (8418 kJ vs 7739 kJ) and underestimated intakes of some micronutrients (iron, iodine, and zinc when supplements were included in the analysis; iron, folate, and iodine when supplements were excluded). The relatively minor differences in intakes were not clinically important differences, as evidenced by the comparison of classifications of nutrient intake adequacy (EAR, RDI) using Cohen kappa (in Table 2). Agreement is considered moderate if .41 ≤kappa ≤.6 and substantial if .61≤ kappa ≤.8 [38]. Cohen kappa indicated moderate agreement (kappa range .488-.559, all P ≤.001) between the two methods for assessing adequacy of nutrient intakes of iron, calcium, and folate, and substantial agreement for iodine (kappa=.803, P<.001) and zinc (kappa=.741, P<.001), when supplements were included in the analysis. When supplements were excluded from the analysis, there was moderate agreement for calcium (kappa=.554, P<.001) and zinc (kappa=.572, P<.001), and substantial agreement for iodine (kappa=.632, P<.001). Future estimation of bias could be explored through criterion validity (ie, comparison with objective measures of dietary intake such as nutritional biomarkers). However, this was beyond the scope of our study, which aimed to assess the relative validity of the SNaQ tool compared with nutrient intakes assessed using dietary composition software.

Specifically designed as a brief tool, the SNaQ tool therefore did not include all foods within the food composition database, and as such may have underestimated some micronutrients. When we removed supplements from the analysis, the SNaQ

tool did not show significant agreement with the nutrient software analysis for folate (kappa=-.068, P=.52). While the nutrient software analysis indicated that 9 participants had nutrient intakes meeting or greater than the EAR of 520 µg, SNaQ showed that only 4 participants had intakes that met the EAR. This may be related to the inclusion of Vegemite (a yeast-based spread) in the image-based records of 7 participants (28%) who ate this food on at least 1 record day. A serving (5 g) of Vegemite provides 100 µg folate (19% of the pregnancy EAR) [39], and so pregnant women may be able to meet their requirements without supplements on certain days, if they consume specific folate-rich foods. As the aim of the SNaQ tool was to also provide a food group analysis of participant diets, we did not include some foods that do not fall into food group categories in the SNaQ (eg, gravy, Vegemite, tomato sauce and some other condiments, salt, and fortified foods like Nestlé Milo). This has highlighted that future modifications to SNaQ may be required to better reflect foods commonly consumed by pregnant women.

The majority of pregnant women in the DBBB study did not meet the recommended AGTHE target for daily servings of grain and cereal foods, vegetables, fruit, meat, and dairy. The median daily servings of unsaturated spreads and oils met recommendations, while median intakes of energy-dense, nutrient-poor foods exceeded recommendations, with less than a third of participants consuming within the target 0-2.5 servings/day. When we evaluated food intakes excluding micronutrient supplements, both the SNaQ and nutrient composition software showed that median intakes of selected key micronutrients important in pregnancy were lower than the EAR for iron, calcium, folate, and iodine. When we included vitamin and mineral supplements use, the median intake of iron was still below the EAR.

Intakes of energy-dense, nutrient-poor foods were high, with the majority (n=18, 72%) exceeding the maximum target of 2.5 servings/day. In other cohorts of pregnant Australian women, it has been reported that meeting AGTHE and NRV targets is challenging [40]. Pregnant women have higher requirements for some nutrients, including folate, iron, zinc, and iodine [10]. Australiawide in 2011-2012, 11.7% of nonpregnant women aged 19-30 years did not meet the EAR for iodine, 10.9% for folate, 13.5% for zinc, 37.5% for iron, and 71.3% for calcium [32]. In addition, within the Australian Longitudinal Study on Women's Health (ALSWH) cohort, suboptimal intakes of core foods and nutrients in pregnant women were common [40,41]. Only 1.5% of the 606 pregnant participants achieved the NRVs for key micronutrients, with no pregnant woman meeting AGTHE target intakes for all food groups and median intakes of energy-dense, nutrient-poor choices exceeding recommendations [40]. However, ALSWH highlighted that women who consumed more daily servings of fruit and dairy than the AGTHE targets met pregnancy NRVs, as did consuming more than the 2.5 daily servings of energy-dense, nutrient-poor foods, although this was associated with higher total energy and saturated fat intakes [40]. This indicates that provision of personalized nutrition advice to optimize diet quality and nutrient intake in pregnancy is warranted. While AGTHE serving sizes and recommended numbers of servings

XSL·FO

RenderX

have been revised since this time, including an increase in recommendations from 1.5 to 3.5 servings of meat, and an additional half serving of dairy foods, dietary intakes are still of concern. In the DBBB cohort the median intake from AGTHE food groups did not meet the revised targets for nonpregnant women, with less than 10% meeting targets for daily servings of meat (n=2, 8%), vegetables (n=1, 4%), and grain foods (n=1, 4%), and less than half meeting targets for fruit (n=10, 40%) and dairy (n=10, 40%).

Prior to being in the study, less than half (n=12, 46%) of participants had received information on prenatal nutrient supplements (including folic acid and iron) during their pregnancy, although the results of this study imply that micronutrient supplementation use may help women meet pregnancy EARs, particularly for iron, folate, and iodine. Only approximately a quarter of participants (n=7, 27%) had received advice on healthy eating during pregnancy prior to the study. Nutrition knowledge among pregnant women in Australia is suboptimal, with one cross-sectional study of 400 pregnant women showing that over half (65%) of participants were not familiar with AGTHE recommendations [42,43]. However, high motivation among pregnant women to adopt healthy eating behaviors [43] and increased awareness of nutrition during pregnancy [44] imply that pregnancy may be an opportune time for health professionals to intervene to improve women's nutrition-related knowledge. The results from the DBBB study further suggest that pregnant women could potentially benefit from receiving personalized nutrient intake assessment and provision of information. This was supported by the less than half (n=11, 42%) of participants who reported they felt they had received enough information about healthy eating for themselves and their baby at the time of enrollment, and a high proportion of participants who reported changes to their dietary intake in response to receiving tailored feedback.

The majority (n=17, 77%) of participants who completed the final survey reported that they had made changes to their dietary intake as a result of receiving the personalized feedback, which consisted of the video summary and the telephone consultation with the dietitian. The preferred method of receiving dietary advice was from the video summary and the dietitian consultation combined, with 95% (n=21) of participants agreeing that this combined way of receiving feedback was helpful. Previous research in the area of apps for dietary feedback during pregnancy supports our findings in this study. A recent evaluation of a Dutch online coaching program delivered by a mobile health platform (called Smarter Pregnancy) resulted in improvements in vegetable, fruit, and folic acid intake in pregnant women, although these were not statistically significant, and high compliance with positive feedback from participants was reported [45]. Likewise, results from the pilot study of the Eating4two app, to monitor gestational weight, was viewed favorably by participants as a method to assist in supporting healthy pregnancy dietary behaviors [46]. Dietary advice during pregnancy can come from multiple sources (as reported by DBBB participants), and can be confusing and contradictory [47], and it is therefore promising that this method was perceived as acceptable by participants. Furthermore, the feedback received indicates that the Diet Bytes method is

promising and warrants future testing in randomized controlled studies to establish the efficacy of using a personalized smartphone method for improving pregnancy food and nutrient intakes [48].

Limitations

Limitations of our study include the small sample size of 25 participants completing the study protocol. A review [12] of image-based dietary assessment methods found that validation studies using this method to date have been conducted in sample sizes ranging from 9 [49] to 75 [19] participants. Given the small sample size and the wide variety of foods available in the Australian food supply, more days of recorded dietary intake may have been required to optimize accuracy of estimated food intake. In relation to the quantification of food portions contained in the image-based records, we attempted to reduce the introduction of bias during the analysis. Coding and entry of the records using the SNaQ tool and FoodWorks were performed independently at separate time points, and information on the estimations of portion size made using the SNaQ tool was not available to the dietitian during the FoodWorks analysis. Despite this, it is possible that estimations of portion size made using the SNaQ tool may have influenced the FoodWorks analysis for the first dietitian. However, the subsample (10 participants) analysis of the image-based records by a second dietitian (using FoodWorks only) showed high agreement with the analysis of the first dietitian, suggesting that any impact may have been small. At the time of developing the SNaQ tool, AUSNUT 2007 was the most recent nutrient composition database available, and this was embedded in the SNaQ tool and also used for the FoodWorks analysis. This database does not contain food group equivalents for each of the food items, and we therefore could not establish intakes of AGTHE food groups from FoodWorks. It is therefore a limitation of this study that we were unable to compare estimates of food group intakes between the two methods.

Participants in the DBBB study may not be representative of all Australian women. Those without smartphones were excluded from participating, and therefore these results may not be representative of women who are economically vulnerable or who have other reasons for not owning a smartphone. We did not collect data on prepregnancy weight and weight gain during pregnancy, and therefore we do not know whether study participants were achieving recommendations for appropriate pregnancy weight gain. The median age of DBBB study participants (28.8 years) was slightly lower than the NSW state median age of women giving birth in 2014 (31.2 years) [50] and may be indicative of the number of rural and Indigenous women in this study, who tend to have their children at a younger age [51]. Over half the participants had completed a university degree, compared with 29% of all women Australiawide of working age (15-64 years) in 2015 [52]. However, this study purposively recruited Indigenous participants (n=8, 31%), to ensure their representation in the study was adequate for separate analyses. It should be noted that Australiawide it is estimated that 3% of the population is Aboriginal or Torres Strait Islander [53]. Given that one of the recruitment avenues was through an Indigenous birth cohort to specifically target this population group, the high representation

of Indigenous participants in this study is to be expected and is desirable given that the use of image-based dietary records has not been previously evaluated in this population.

Conclusions

To our knowledge, this study is the first to evaluate the use of image-based dietary records for dietary assessment in pregnant women, including Indigenous Australian women, and demonstrated that the SNaQ tool can adequately assess key nutrient intakes during pregnancy. With training and practice,

the SNaQ tool has the potential to be both time and resource saving as a dietary assessment tool for dietitians, while reducing the burden of recording associated with traditional methods for participants. Importantly this study highlights that using an image-based dietary record in combination with individual phone consultation with a dietitian for the provision of dietary feedback during pregnancy is acceptable. The Diet Bytes method for nutrition assessment and provision of personally tailored feedback may be a useful method for dietitians to assist women in optimizing their food and nutrient intakes during pregnancy.

Conflicts of Interest

None declared.

References

1. Dietitians Association of Australia. What is an accredited practising dietitian. URL: http://daa.asn.au/for-the-public/find-an-apd/what-is-an-accredited-practising-dietitian/ [accessed 2016-08-09] [WebCite Cache ID 6jcWDMGWm]

2. McCabe-Sellers B. Advancing the art and science of dietary assessment through technology. J Am Diet Assoc 2010 Jan;110(1):52-54. [doi: 10.1016/j.jada.2009.10.014] [Medline: 20102827]

3. Thompson FE, Subar AF, Loria CM, Reedy JL, Baranowski T. Need for technological innovation in dietary assessment. J Am Diet Assoc 2010 Jan;110(1):48-51 [FREE Full text] [doi: 10.1016/j.jada.2009.10.008] [Medline: 20102826]

4. Mela DJ, Aaron JI. Honest but invalid: what subjects say about recording their food intake. J Am Diet Assoc 1997 Jul;97(7):791-793. [doi: 10.1016/S0002-8223(97)00195-8] [Medline: 9216559]

5. Rebro SM, Patterson RE, Kristal AR, Cheney CL. The effect of keeping food records on eating patterns. J Am Diet Assoc 1998 Oct;98(10):1163-1165. [doi: 10.1016/S0002-8223(98)00269-7] [Medline: 9787724]

6. Gersovitz M, Madden JP, Smiciklas-Wright H. Validity of the 24-hr. dietary recall and seven-day record for group comparisons. J Am Diet Assoc 1978 Jul;73(1):48-55. [Medline: 659761]

7. Australian Bureau of Statistics. Australian Health Survey: Users' Guide, 2011-13: Under-Reporting in Nutrition Surveys.. 2014. URL: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4363.0.55.001Chapter651512011-13 [accessed 2016-08-09] [WebCite Cache ID 6jcWW3HQR]

8. Moshfegh AJ, Rhodes DG, Baer DJ, Murayi T, Clemens JC, Rumpler WV, et al. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. Am J Clin Nutr 2008 Aug;88(2):324-332 [FREE Full text] [Medline: 18689367]

9. Food Standards Australia New Zealand. Monitoring Nutrients in Our Food Supply. 2016. URL: http://www.foodstandards.gov.au/science/monitoringnutrients/Pages/default.aspx [accessed 2016-07-28] [WebCite Cache ID 6jcWp90wO]

10. National Health and Medical Research Council, New Zealand Ministry of Health. Nutrient Reference Values for Australia and New Zealand Including Recommended Dietary Intakes. Canberra, Australia: Commonwealth of Australia; 2006. URL: https://www.nhmrc.gov.au/_files_NHMRC/publications/attachments/n35.pdf? [accessed 2016-07-28] [WebCite Cache ID 6lKy3cN6k]

11. Australian Government, National Health and Medical Research Council, Department of Health and Ageing. Eat for Health: Australian Dietary Guidelines. Canberra, Australia: National Health and Medical Research Council; 2013. URL: https://www.eatforhealth.gov.au/sites/default/files/files/the_guidelines/n55_australian_dietary_guidelines.pdf [accessed 2016-07-28] [WebCite Cache ID 6lKyGnnvW]

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]

XSL•FO
RenderX

13. 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]

14. 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]

15. 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;7(6):4897-4910 [FREE Full text] [doi: 10.3390/nu7064897] [Medline: 26091234]

16. Poushter J. Smartphone Ownership and Internet Usage Continues to Climb in Emerging Economies. Washington, DC: Pew Research Center; 2016 Feb 22. URL: http://www.pewglobal.org/2016/02/22/smartphone-ownership-and-internet-usage-continues-to-climb-in-emerging-economies/ [accessed 2016-07-04] [WebCite Cache ID 6jcXMKLEk]

17. Rollo ME, Hutchesson MJ, Burrows TL, Krukowski RA, Harvey JR, Hoggle LB, et al. Video consultations and virtual nutrition care for weight management. J Acad Nutr Diet 2015 Aug;115(8):1213-1225. [doi: 10.1016/j.jand.2015.03.016] [Medline: 25986214]

18. Dahl Lassen A, Poulsen S, Ernst L, Kaae Anderson K, Biltoft-Jensen A, Tetens I. Evaluation of a digital method to assess evening meal intake in a free-living adult population. Food Nutr Res 2010;54 [FREE Full text] [doi: 10.3402/fnr.v54i0.5311] [Medline: 21085516]

19. Kikunaga S, Tin T, Ishibashi G, Wang D, Kira S. The application of a handheld personal digital assistant with camera and mobile phone card (Wellnavi) to the general population in a dietary survey. J Nutr Sci Vitaminol (Tokyo) 2007 Apr;53(2):109-116 [FREE Full text] [Medline: 17615997]

20. Martin CK, Han H, Coulon SM, Allen HR, Champagne CM, Anton SD. A novel method to remotely measure food intake of free-living individuals in real time: the remote food photography method. Br J Nutr 2009 Feb;101(3):446-456 [FREE Full text] [doi: 10.1017/S0007114508027438] [Medline: 18616837]

21. Wang D, Kogashiwa M, Ohta S, Kira S. Validity and reliability of a dietary assessment method: the application of a digital camera with a mobile phone card attachment. J Nutr Sci Vitaminol (Tokyo) 2002 Dec;48(6):498-504. [Medline: 12775117]

22. Wang D, Kogashiwa M, Kira S. Development of a new instrument for evaluating individuals' dietary intakes. J Am Diet Assoc 2006 Oct;106(10):1588-1593. [doi: 10.1016/j.jada.2006.07.004] [Medline: 17000191]

23. Daugherty BL, Schap TE, Ettienne-Gittens R, Zhu FM, Bosch M, Delp EJ, et al. Novel technologies for assessing dietary intake: evaluating the usability of a mobile telephone food record among adults and adolescents. J Med Internet Res 2012;14(2):e58 [FREE Full text] [doi: 10.2196/jmir.1967] [Medline: 22504018]

24. Segovia-Siapco G, Sabaté J. Using personal mobile phones to assess dietary intake in free-living adolescents: comparison of face-to-face versus telephone training. JMIR Mhealth Uhealth 2016;4(3):e91 [FREE Full text] [doi: 10.2196/mhealth.5418] [Medline: 27473291]

25. Aflague TF, Boushey CJ, Guerrero R, Ahmad Z, Kerr DA, Delp EJ. Feasibility and use of the mobile food record for capturing eating occasions among children ages 3-10 years in Guam. Nutrients 2015 Jun;7(6):4403-4415 [FREE Full text] [doi: 10.3390/nu7064403] [Medline: 26043037]

26. Martin CK, Correa JB, Han H, Allen HR, Rood JC, Champagne CM, et al. Validity of the Remote Food Photography Method (RFPM) for estimating energy and nutrient intake in near real-time. Obesity (Silver Spring) 2012 Apr;20(4):891-899 [FREE Full text] [doi: 10.1038/oby.2011.344] [Medline: 22134199]

27. Rollo ME, Ash S, Lyons-Wall P, Russell A. Trial of a mobile phone method for recording dietary intake in adults with type 2 diabetes: evaluation and implications for future applications. J Telemed Telecare 2011;17(6):318-323. [doi: 10.1258/jtt.2011.100906] [Medline: 21844173]

28. Gluckman PD, Hanson MA. Developmental plasticity and human disease: research directions. J Intern Med 2007 May;261(5):461-471 [FREE Full text] [doi: 10.1111/j.1365-2796.2007.01802.x] [Medline: 17444885]

29. Godfrey KM, Gluckman PD, Hanson MA. Developmental origins of metabolic disease: life course and intergenerational perspectives. Trends Endocrinol Metab 2010 Apr;21(4):199-205. [doi: 10.1016/j.tem.2009.12.008] [Medline: 20080045]

30. Morrison JL, Regnault T. Optimising maternal diet and fetal adaptations to altered nutrient supply. Nutrients 2016;8(6) [FREE Full text] [doi: 10.3390/nu8060342] [Medline: 27271666]

31. Australian Bureau of Statistics. Australian Health Survey: Biomedical Results for Nutrients, 2011-12: Feature Article: Women of Childbearing Age. 2013. URL: http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/4364.0.55.006Chapter1302011-12 [accessed 2016-08-09] [WebCite Cache ID 6jcXnKowE]

32. Australian Bureau of Statistics. Australian Health Survey: Usual Nutrient Intakes, 2011-12. 2015. URL: http://www.ausstats.abs.gov.au/ausstats/subscriber.nsf/0/31EEAEBF2A09D051CA257DFF000CA865/$File/australian%20health%20survey,%20usual%20nutrient%20intakes%202011-12.pdf [accessed 2016-10-18] [WebCite Cache ID 6lKzE29ti]

33. Browne J, Laurence S, Thorpe S. Acting on Food Insecurity in Urban Aboriginal and Torres Strait Islander Communities: Policy and Practice Interventions to Improve Local Access and Supply of Nutritious Food. Mt Lawley, Western Australia: Australian Indigenous HealthInfoNet; 2009. URL: http://www.healthinfonet.ecu.edu.au/health-risks/nutrition/other-reviews [accessed 2016-08-09] [WebCite Cache ID 6jcZpaccp]

34.   Ashman AM, Collins CE, Weatherall L, Brown LJ, Rollo ME, Clausen D, et al. A cohort of Indigenous Australian women and their children through pregnancy and beyond: the Gomeroi gaaynggal study. J Dev Orig Health Dis 2016 Aug;7(4):357-368. [doi: 10.1017/S204017441600009X] [Medline: 27080434]

35.   Collins CE, Boggess MM, Watson JF, Guest M, Duncanson K, Pezdirc K, et al. Reproducibility and comparative validity of a food frequency questionnaire for Australian adults. Clin Nutr 2014 Oct;33(5):906-914. [doi: 10.1016/j.clnu.2013.09.015] [Medline: 24144913]

36.   Food Standards Australia New Zealand. Australian Food, Supplement & Nutrient Database 2007 for Estimation of Population Nutrient Intakes. Canberra, Australia: Food Standards Australia New Zealand; 2008. URL: https://www.foodstandards.gov.au/science/monitoringnutrients/ausnut/documents/AUSNUT%202007%20-%20Explanatory%20Notes.pdf [accessed 2016-10-18] [WebCite Cache ID 6lL0ZYVyK]

37.   Thomas DR. A general inductive approach for analysing qualitative evaluation data. Am J Eval 2006 Jun 01;27(2):237-246. [doi: 10.1177/1098214005283748]

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

39.   Mondelez Australia Foods. VEGEMITE. 2016. URL: https://www.vegemite.com.au/products/vegemite [accessed 2016-08-09] [WebCite Cache ID 6jcaPFhci]

40.   Blumfield ML, Hure AJ, Macdonald-Wicks L, Patterson AJ, Smith R, Collins CE. Disparities exist between national food group recommendations and the dietary intakes of women. BMC Womens Health 2011;11(37) [FREE Full text] [doi: 10.1186/1472-6874-11-37] [Medline: 21819627]

41.   Hure A, Young A, Smith R, Collins C. Diet and pregnancy status in Australian women. Public Health Nutr 2009 Jun;12(6):853-861. [doi: 10.1017/S1368980008003212] [Medline: 18647425]

42.   Bookari K, Yeatman H, Williamson M. Exploring Australian women's level of nutrition knowledge during pregnancy: a cross-sectional study. Int J Womens Health 2016;8:405-419 [FREE Full text] [doi: 10.2147/IJWH.S110072] [Medline: 27574470]

43.   Bookari K, Yeatman H, Williamson M. Falling short of dietary guidelines: what do Australian pregnant women really know? A cross sectional study. Women Birth 2016 Jun 14. [doi: 10.1016/j.wombi.2016.05.010] [Medline: 27317129]

44.   Szwajcer E, Hiddink GJ, Maas L, Koelen M, van Woerkum C. Nutrition awareness before and throughout different trimesters in pregnancy: a quantitative study among Dutch women. Fam Pract 2012 Apr;29 Suppl 1:i82-i88 [FREE Full text] [doi: 10.1093/fampra/cmr107] [Medline: 22399562]

45.   Van Dijk MR, Huijgen NA, Willemsen SP, Laven JS, Steegers EA, Steegers-Theunissen RP. Impact of an mHealth platform for pregnancy on nutrition and lifestyle of the reproductive population: a survey. JMIR Mhealth Uhealth 2016;4(2). [Medline: 27234926]

46.   Knight-Agarwal C, Davis DL, Williams L, Davey R, Cox R, Clarke A. Development and pilot testing of the Eating4two mobile phone app to monitor gestational weight gain. JMIR Mhealth Uhealth 2015;3(2):e44 [FREE Full text] [doi: 10.2196/mhealth.4071] [Medline: 26048313]

47.   Wennberg AL, Lundqvist A, Högberg U, Sandström H, Hamberg K. Women's experiences of dietary advice and dietary changes during pregnancy. Midwifery 2013 Sep;29(9):1027-1034. [doi: 10.1016/j.midw.2012.09.005] [Medline: 23427852]

48.   O'Brien OA, McCarthy M, Gibney ER, McAuliffe FM. Technology-supported dietary and lifestyle interventions in healthy pregnant women: a systematic review. Eur J Clin Nutr 2014 Jul;68(7):760-766. [doi: 10.1038/ejcn.2014.59] [Medline: 24781682]

49.   Gregory R, Walwyn L, Bloor S, Amin S. A feasibility study of the use of photographic food diaries in the management of obesity. Pract Diabetes 2006;23(2):66-68. [doi: 10.1002/pdi.899]

50.   Australian Bureau of Statistics. Births, Australia, 2014: Inner City Areas Have Lowest Birth Rates. 2015 Oct 29. URL: http://www.abs.gov.au/ausstats%5Cabs@.nsf/0/8668A9A0D4B0156CCA25792F0016186A?Opendocument [accessed 2016-08-09] [WebCite Cache ID 6jcahovpx]

51.   Australian Bureau of Statistics. Births, Australia, 2014: Aboriginal and Torres Strait Islander Births and Fertility. 2015 Oct 29. URL: http://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/3301.0Main%20Features62014?opendocument&tabname=Summary&prodno=3301.0&issue=2014&num=&view= [accessed 2016-08-03] [WebCite Cache ID 6jcaupzh7]

52.   Australian Bureau of Statistics. Education and Work, Australia, May 2015: Media Release: Australians More Educated Than Ever Before. 2015 Nov 27. URL: http://www.abs.gov.au/AUSSTATS/abs@.nsf/Latestproducts/6227.0Media%20Release1May%202015?opendocument&tabname=Summary&prodno=6227.0&issue=May%202015&num=&view= [accessed 2016-08-09] [WebCite Cache ID 6jcbFYjG6]

53.   Australian Bureau of Statistics. Estimates of Aboriginal and Torres Strait Islander Australians, June 2011. 2013 Aug 30. URL: http://www.abs.gov.au/ausstats/abs@.nsf/mf/3238.0.55.001 [accessed 2016-08-09] [WebCite Cache ID 6jcbNqDPj]

Abbreviations

    AGTHE: Australian Guide to Healthy Eating

XSL•FO

RenderX

ALSWH: Australian Longitudinal Study on Women's Health
AUSNUT: Australian Food, Supplement & Nutrient Database
DBBB: Diet Bytes and Baby Bumps
EAR: estimated average requirement
NRV: Nutrient Reference Value
NSW: New South Wales
PSEA: portion size estimation aid
RDI: recommended dietary intake
SMS: short message service
SNaQ: Selected Nutrient and Diet Quality

<u>Original Paper</u>

Text Message-Based Intervention Targeting Alcohol Consumption Among University Students: Findings From a Formative Development Study

Kristin Thomas[1], PhD; Catharina Linderoth[1], BA; Marcus Bendtsen[2], MSc; Preben Bendtsen[1], PhD; Ulrika Müssener[1], PhD

[1]Division of Community Medicine, Department of Medical and Health Sciences, Linköping University, Linköping, Sweden
[2]Database and Information Techniques, Department of Computer and Information Science, Linköping University, Linköping, Sweden

Corresponding Author:
Kristin Thomas, PhD
Division of Community Medicine
Department of Medical and Health Sciences
Linköping University
58183
Linköping,
Sweden
Phone: 46 13282546
Fax: 46 13 145004
Email: kristin.thomas@liu.se

Abstract

Background: Drinking of alcohol among university students is a global phenomenon; heavy episodic drinking is accepted despite several potential negative consequences. There is emerging evidence that short message service (SMS) text messaging interventions are effective to promote behavior change among students. However, it is still unclear how effectiveness can be optimized through intervention design or how user interest and adherence can be maximized.

Objective: The objective of this study was to develop an SMS text message-based intervention targeting alcohol drinking among university students using formative research.

Methods: A formative research design was used including an iterative revision process based on input from end users and experts. Data were collected via seven focus groups with students and a panel evaluation involving students (n=15) and experts (n=5). Student participants were recruited from five universities in Sweden. A semistructured interview guide was used in the focus groups and included questions on alcohol culture, message content, and intervention format. The panel evaluation asked participants to rate to what degree preliminary messages were understandable, usable, and had a good tone on a scale from 1 (very low degree) to 4 (very high degree). Participants could also write their own comments for each message. Qualitative data were analyzed using qualitative descriptive analysis. Quantitative data were analyzed using descriptive statistics. The SMS text messages and the intervention format were revised continuously in parallel with data collection. A behavior change technique (BCT) analysis was conducted on the final version of the program.

Results: Overall, students were positive toward the SMS text message intervention. Messages that were neutral, motivated, clear, and tangible engaged students. Students expressed that they preferred short, concise messages and confirmed that a 6-week intervention was an appropriate duration. However, there was limited consensus regarding SMS text message frequency, personalization of messages, and timing. Overall, messages scored high on understanding (mean 3.86, SD 0.43), usability (mean 3.70, SD 0.61), and tone (mean 3.78, SD 0.53). Participants added comments to 67 of 70 messages, including suggestions for change in wording, order of messages, and feedback on why a message was unclear or needed major revision. Comments also included positive feedback that confirmed the value of the messages. Twenty-three BCTs aimed at addressing self-regulatory skills, for example, were identified in the final program.

Conclusions: The formative research design was valuable and resulted in significant changes to the intervention. All the original SMS text messages were changed and new messages were added. Overall, the findings showed that students were positive toward receiving support through SMS text message and that neutral, motivated, clear, and tangible messages promoted engagement. However, limited consensus was found on the timing, frequency, and tailoring of messages.

XSL•FO
RenderX

KEYWORDS

Introduction

An increasing proportion of the global burden of disease is due to alcohol consumption. In 2010, alcohol caused 5.5% of the total burden of disease and approximately 5 million deaths, which represents a 30% increase from 1990 [1]. Drinking of alcohol among university students is a global phenomenon and heavy drinking is a part of the social norm despite potential negative consequences [2,3]. Approximately half of all young adults in Sweden attend university making the health and well-being of this group an important public health concern [4]. Thus, it is important to develop cost-effective interventions that are able to reach large numbers of students.

There is emerging evidence that short message service (SMS) text messaging is a cost-effective method to support behavior change [5]. A review of controlled trials of SMS text message-based interventions showed that eight of nine studies found support for SMS text message-based interventions on behavior change. Outcomes included weight loss, smoking cessation, and diabetes management [6]. Another review of 14 studies identified evidence supporting SMS text message-based interventions on behavior change in all but one study. Aspects such as tailored content and interactivity were found to be important [7]. Furthermore, a 12-week SMS text message-based intervention targeting heavy drinking among young adults was found to have an effect on the number of heavy drinking days and the number of drinks per drinking day compared with baseline data [8].

Furthermore, SMS text message-based interventions have been shown to be highly accessible in that messages are likely to be read within minutes of being received and that receiving and reading messages require limited time and effort by the user [9-11]. SMS text message-based interventions can enable continuous, real-time, brief support in a real-world setting [9,12,13]. The real-time aspect may even serve a purpose beyond the actual content or meaning of messages [14]. SMS text message-based interventions have been designed based on existing evidence-based interventions and have adopted techniques similar to face-to-face treatments, such as tailored advice or goal setting [9,14]. Thus, SMS text message-based interventions could be cost-effective and applicable to many health behaviors [5,15,16]. However, it is still unclear how effectiveness can be optimized through, for example, message content and structure [17], or how user interest and adherence can be maximized [15]. User compatibility is seldom evaluated; at best, it is performed after delivery of the intervention [15].

Formative research methods, which consider the input from users and experts, have been successfully used previously [18]. The aim is to improve the design and performance of an intervention through input on content and regimen from the target group, experts, and prevalent theory. By engaging users in the development of the intervention, it is believed that compliance and fidelity to the intervention will increase; however, few studies merge the knowledge and interests of users and investigators [15]. Also, more studies describing the process of developing interventions in a transparent way are needed [19].

Furthermore, the use of theory has been limited in the development of SMS text message-based interventions [6,15,20] making it difficult to identify why and how an intervention results in desired outcomes. The association between using theory as a basis for designing interventions and effectiveness is not well understood [20]. It is unclear if it is more effective to use single or multiple theories in intervention development [17]. A multitheory approach without a clear rationale may not always be as effective as an intervention with systematic use of a single theory [20,21]. Taxonomies of behavior change techniques (BCTs) have been developed to aid the process of identifying effective features of an intervention [22,23].

This study reports on the findings from formative development of a text messaging alcohol intervention aimed at university students. The research builds on previous work from the AMADEUS research program [11,24,25].

Methods

Intervention Revision Process

The first version of the intervention, which included 57 messages, was generated by PB inspired by prevalent behavior change theory [11]. A formative research design was used in this study building on this previous work. The formative process included (1) focus groups with students and (2) panel evaluation with students and experts (see Multimedia Appendix 1). The revision process was iterative, including continuous revisions of messages and the intervention format based on data from the focus groups and panel evaluation. Changes included omitting messages, adding messages, and changing wording, content, and order of messages (see Results section). Both CL and KT revised messages and the intervention format and generated a final version. All authors read and approved the final version of the messages and intervention format. Finally, a BCT analysis was conducted to elucidate the content of the messages and their theory base.

Focus Group Discussions

The aim of the focus groups was to explore the perspective of potential users regarding the preliminary content and structure of the messages. In addition, the first two focus groups aimed to investigate student alcohol culture and perspectives of SMS text message interventions in general. Participants were recruited through advertisements and the snowball technique at the five universities included in the study. Purposive sampling was used and included individuals from the target population. The focus group discussions lasted approximately 1.5 hours following a semistructured format and including questions on student alcohol

culture, length, frequency, and timing of messages, as well as the intervention format (see Multimedia Appendix 2). Participants were shown examples of preliminary messages to gain an idea of how a message could look like. All focus group discussions were audio-recorded and transcribed. CL took part in all the focus groups. PB took part in two focus groups. On these occasions, CL and PB alternated between the roles of note taker and group facilitator.

The data from the first two focus groups were compiled by CL and summarized according to aim (ie, students' thoughts on SMS text message support). Qualitative descriptive analysis [26] was used to analyze the data from the next five focus groups. In this analysis, the data were coded using predefined categories based on the aim of the study: length, frequency, and timing of messages and intervention format. Relevant data were identified and summarized for each category. The content of each category was expanded by revisiting the data and comparing data across categories. KT performed the qualitative descriptive data analysis.

Panel Evaluation

The panel evaluation aimed to gain input from students and experts on specific preliminary messages. Three separate questions asked participants to rate, on a scale from 1 to 4 (1=very low degree; 4=very high degree), to what degree messages were understandable, usable, and had good tone. There was also space to write comments for each message. Thirty-three students from the five universities in Sweden that had participated in the focus groups were invited via email to complete the panel evaluation. All students were compensated 500 SEK for taking part. Twenty-one experts were invited to participate and included both researchers and staff from student health care centers. The experts were not compensated for participating.

The qualitative data were analyzed by CL using qualitative descriptive analysis [26]. The data from all participants were summarized for each message. Data on individual messages were then and potential revisions discussed. The quantitative data were analyzed by KT using descriptive statistics. Messages that were rated low or very low by at least three participants on either of the three outcomes were analyzed separately. The qualitative data for these messages were reviewed and revisions made. Each message was also categorized as information- or practice-based. Information-based messages included facts or general tips about behavior change. Behavioral practice-based messages typically asked users to reflect on or practice behavior change. The categorization of messages was performed to investigate if the results differed depending on the type of message. Both CL and KT carried out revisions of messages based on the data from the panel evaluations.

Behavior Change Technique Analysis

The aim of the BCT analysis was to elucidate the content of the messages and their theory base. A taxonomy of BCTs to reduce excessive alcohol consumption was used in the analysis [27]. The taxonomy was developed by Michie and colleagues [22,23] and is based on previous taxonomies for behavior change and stems from behavior change theory. The techniques represent strategies, such as self-monitoring, that have been found to be effective in behavior change interventions. In this study, the BCT analysis was conducted on the final version of the intervention. Multimedia Appendix 3 includes all the SMS text messages as well as their corresponding BCT.

Results

Participants

Focus Group Discussions

Seven focus groups were conducted with students from five universities in Sweden. The participants were between 19 and 25 years of age. Of a total of 43 students, 11 were males and 32 were females. The students were attending a mixture of academic degree courses in arts and sciences.

Panel Evaluation

Sixteen students completed the panel evaluation (aged between 19 and 24 years; 15 females, 1 male). In total, five experts took part; two used the evaluation form and three responded as a group giving general feedback and comments to individual messages. All experts were women with a mean age of 58 (SD 12) years.

Main Findings

Focus Group Discussions

Overall, students expressed positive feedback regarding the duration of the intervention and regarding receiving SMS text message support for reducing drinking. Four message characteristics that were found to be important to engage students emerged from the data: neutrality, motivating content, tangible information, and clarity. These four criteria were used to assess the messages and identify the need for revisions. Neutrality meant that the message was based on facts and, if relevant, highlighted both positive and negative aspects of alcohol:

> If you keep to pure facts then I think that you take it more seriously and you take it in easier. [group 4]

> Exactly...and maybe not in all messages that you should try to do it that all should be a bit positive and negative but it is quite a good aspect in these other messages. [group 1]

The motivating aspect included the wording or tone that aimed to increase self-esteem, self-efficacy, and motivation to make changes. It was deemed important that messages had a coaching approach instead of using scare tactics or creating feelings of anxiety or guilt:

> I like the start-up messages, like, that it's like then it's a bit triggering or coaching like "now let's go" for six weeks. [group 1]

> Yeah you should watch out for using scare tactics 'cause that can make you drop out cause it makes you feel bad cause I make bad choices. [group 4]

A tangible characteristic meant that the messages were easy to read with accessible information, clear and to the point. Another aspect was a preference for contrasting examples (eg, "How

much money do you spend on food versus alcohol, and do you think it is worth it?"):

> *I think that if you compare with food every month and it's maybe half the amount that I spend on food what better food I could get.* [group 2]

> *...it's also good with short and concise like you get a reminder.* [group 5]

Ensuring clarity of messages encompassed providing sufficient information to minimize the risk of misunderstandings and not including multiple themes within one message. The students typically expressed that they preferred short concise messages, and that messages not exceed one page or screen:

> *Yeah or the message is good but I think that you can reword it so that the purpose of it is clearer.* [group 1]

> *...but if there's properly clear words, like the first ones, then you will read it.* [group 3]

Furthermore, the focus group participants perceived 6 weeks of messages to be an appropriate duration for the intervention. There was limited consensus among students regarding the timing and frequency of messages and their personalization. Students expressed that they read messages more thoroughly in the mornings and the evenings. During the daytime, messages could get lost or were given low priority. However, students also expressed that messages were always read, but perhaps not straight away. Moreover, students' preferences for message frequency varied from three to four messages per week to one message every day, at least at the start of the program. On nights out, many students agreed that two messages with one message being sent late at night would be valuable. Some students thought that it would be good to have more personalized messages (eg, "Hi John"), whereas others perceived that personalized messages would be less credible.

Students described a university culture in which alcohol drinking was a social phenomenon and that heavy drinking was normalized. For example, the first weeks of term were described as an induction to heavy drinking rather than an induction to academic studies. Furthermore, students expressed that alcohol consumption was highly emphasized on university campuses through posters, flyers, special offers, and student union activities. Student unions and other student societies were believed to have a role in pushing alcohol consumption based on financial interests (ie, selling alcohol).

Panel Evaluation

Overall, students and experts gave high scores to the majority of the messages. The majority of messages were given a score greater than three out of a top mark of four. Messages were given a mean score of 3.86 (SD 0.43) on understanding, 3.70 (SD 0.61) on usability, and 3.78 (SD 0.53) regarding tone. Messages were rated high irrespective of the type of message. Messages with primarily information content were given a mean score of 3.84 (SD 0.43) on understanding, 3.72 (SD 0.60) on usability, and 3.78 (SD 0.56) on tone. Similarly, practice-based messages were given a mean score of 3.85 (SD 0.44) on understanding, 3.71 (SD 0.60), on usability, and 3.78 (SD 0.54) on the tone of the message. Students and experts scored messages similarly with a mean score greater than three for the majority of messages regarding all three outcome measures: understanding, relevance, and tone. Thirteen of the 70 messages were rated low or very low on either understanding, usability, or tone by at least three participants (experts or participants). These messages were analyzed separately; the qualitative data were consulted and subsequent revisions made. The distribution of low scores was equal for understanding, usability, and tone.

Qualitative feedback was given for 67 of 70 messages and included suggestions for change in wording, change in the order of messages, and feedback on why a message was unclear or needed major revision. For example, a message that included tips on how to prepare for saying no to a drink on the next night out received feedback that it was too long and that it had a confusing message. The message was shortened and the wording made more concise. The data also included positive feedback that confirmed the value of messages. For example, messages that were to the point and included tangible examples (eg, "How much money do you spend on alcohol in a typical week compared with food?").

Behavior Change Technique Analysis

Twenty-three techniques were identified in the final version of the intervention (62 messages). Some techniques were used in more than one message or across two messages. The techniques aimed to motivate students to reduce their alcohol consumption, address self-regulation, increase self-efficacy, and increase students' awareness of social and professional support. Table 1 shows the BCTs for each message and their function presented per week.

Table 1. Behavior change technique and function for each message per week.

Message by week[a]	Behavior change technique	Function
Week 1		
1-2	Provide feedback on performance in relation to behavior goal*	Motivation
3	Provide tips on how to perform the behavior	Motivation
4-5	Provide information on consequences of excessive drinking	Motivation
6	Identify reasons for wanting and not wanting to reduce drinking	Motivation
7	Assess current and past drinking behavior	Motivation
8	Provide tips on how to perform the behavior	Motivation
9	Prompt to practice refusal	Self-efficacy
10	Advise on use of social support	Social influence
11	Provide information on the consequences excessive drinking	Motivation
Week 2		
12	Provide instruction on how to perform the behavior	Self-efficacy
13-14	Provide feedback on performance in relation to behavior goal[b]	Motivation
15	Provide information on the consequences of excessive alcohol consumption	Motivation
16	Identify reasons for (not) wanting to reduce excessive alcohol consumption	Motivation
17	Assess past history of attempts to reduce excessive alcohol consumption	Motivation
18	Provide information on the consequences of excessive alcohol consumption	Motivation
19	Advise on environmental restructuring	Self-regulation
20	Prompt reflection on approval of others	Social influence
21	Provide information on the consequences of excessive alcohol	Motivation
22	Prompt reflection on approval of others	Motivation
Week 3		
23	Prompt reflection on approval of others	Motivation
24-25	Provide feedback on performance in relation to behavior goal[b]	Motivation
26	Provide information on the consequences of excessive alcohol	Motivation
27	Behavior substitution	Motivation
28	Prompt practice	Self-regulation
29	Provide information on the consequences of excessive alcohol	Motivation
30	Identify reasons for (not) wanting to reduce excessive alcohol consumption	Motivation
31	Provide information on the consequences of excessive alcohol	Motivation
32	Facilitate identification of barriers and problem solving	Self-regulation
33	Provide instruction on how to perform the behavior	Self-efficacy
Week 4		
34-35	Provide feedback on performance in relation to behavior goal[b]	Motivation
36-37	Prompt review of goals	Self-regulation
38	Behavior substitution	Self-regulation
39	Provide information on the consequences of excessive alcohol	Motivation
40	Facilitate identification of barriers and problem solving	Self-regulation
41-43	Provide instruction on how to perform the behavior	Self-efficacy
44	Provide information on the consequences of excessive alcohol consumption	Motivation

Message by week[a]	Behavior change technique	Function
Week 5		
45-46	Provide feedback on performance in relation to behavior goal[b]	Motivation
47	Assess current and past drinking behavior	Self-regulation
48	Advise on avoidance of social cues for drinking	Self-regulation
49	Provide information on the consequences of excessive alcohol consumption	Motivation
50	Identify reasons for (not) wanting to reduce excessive alcohol consumption	Motivation
51	Give options for additional and later support	Social influence
52	Prompt self-recording	Self-regulation
53	Identify reasons for (not) wanting to reduce excessive alcohol consumption	Motivation
Week 6		
54	Prompt self-recording	Self-regulation
55-56	Provide feedback on performance in relation to behavior goal[b]	Motivation
57	Emphasize choice	Self-efficacy
58	Environmental restructuring	Self-regulation
59	Prompt practice	Self-efficacy
60	Anticipated regret	Motivation
61-62	Provide feedback on performance in relation to behavior goal[b]	Motivation

[a]Each week started on a Sunday.

[b]Paired messages that were repeated each week.

The Final Program Design

Significant changes were made to the messages in the formative process. The final intervention included a 6-week program with a total of 62 messages (see Multimedia Appendix 3). The initial preliminary intervention consisted of 57 unique messages. Thus, changes included omitting and adding messages, changing the wording and order of messages, and splitting and combining messages. All the original 57 messages were altered at least once.

The final version of the intervention commenced with students receiving an email asking them to set a goal of how much they would like to reduce their drinking. They could then register for the intervention by sending their mobile number to receive the first message. Participants could unsubscribe to the messages at any time by responding the word "STOP" to any of the SMS text messages.

The first 4 weeks of the program had a higher frequency of messages, nine each week, followed by seven messages in weeks 2 to 5 and five messages in week 6. Messages were sent 7 days a week at various times around midday, late afternoon, or early evening. Forty-eight of the 62 messages were unique. Two messages were repeated at the start of each week; students were asked to report the number of drinks they had the previous week via SMS text message. Subsequently, they received a second SMS text message including feedback on their performance in relation to their goal set at the start of the intervention. These "paired" messages were repeated every Sunday. The content of the unique messages were primarily information-based or behavioral practice-based. Information-based messages typically included facts about alcohol and health, the consequences of excessive drinking, or tips about behavior change strategies. Behavioral practice-based messages included asking students to reflect on or practice behavior change; for example, reflect on triggers for excessive drinking or practicing saying no to drinking on a night out.

The SMS text messages will be sent from GSM modems, programmed and operated by one of the authors (MB). The technical platform has been used in previous studies and proven to work without any major interruptions.

Discussion

Principal Findings

We set out to develop an SMS text message intervention targeting excessive drinking among university students using formative research. The formative process resulted in significant changes to the original messages. The findings showed that neutral, motivated, clear, and tangible messages engaged students. However, there was limited consensus among students regarding optimal timing, frequency, and personalization of messages, although the students agreed that a 6-week duration of the program was acceptable. The panel evaluation showed overall positive scores on the understanding, usability, and tone for the majority of the messages. The panel also received comments for most of the messages. The BCT analysis identified 23 techniques used in the final intervention. The final program design resulted in a 6-week intervention including 62 messages sent with varied frequency and timing.

XSL•FO

RenderX

Comparison With Previous Work

In line with previous research, these findings show that text messaging has the potential to be a useful tool for interventions targeting excessive drinking among students [15]. Overall, the findings showed that students were positive toward receiving support through SMS text messages. However, there was limited consensus among students regarding the timing, frequency, and personalization of messages. To create optimal timing schedules for specific target groups could prove to be a challenge for SMS text message interventions. Previous studies have attempted to solve this by allowing users to decide on the content and timing of messages via a website [28]. However, more research is needed on the effectiveness of user-created compared with preset schedules. Furthermore, our findings showed that messages were read, if not straightaway, at least the same day, suggesting that the exact timing of a message may be less relevant. This could be especially true for messages that aim to evoke reflection (that can be done at any time) compared with reminders or prompts, which would be more valuable at certain time points (eg, a reminder to drink water on a night out).

Engaging users is a central aspect of optimizing the effectiveness of SMS text message interventions. A recent study showed that SMS text messages that aimed to reduce hazardous drinking were only perceived to be moderately helpful and interesting among a population of young adults. The study highlighted the challenge of engaging this target group in alcohol prevention [29]. Our findings showed that messages that were neutral, motivated, clear, and tangible were more engaging and likeable. These findings support previous research that messages need to be supportive and have a positive tone. Bock et al [15] found that students appreciated making wise choices and did not want to be told not to drink. Techniques used in face-to-face interventions could be applied in SMS text message interventions [30]. Techniques relating to building rapport and general support can be more difficult to achieve via text messaging due to their interactive components. The findings from this study, in support of previous research, showed that the tone of messages is important to engage users. Other engagement strategies that have been suggested are ease of use, user ability to change the design, and tailored information [15,30]. Furthermore, a study that asked users to choose one of two messages with slightly different wording showed the importance of the wording of messages to engage users. Similar to our findings, users preferred messages that highlighted the positive over the negative, were nonaggressive, and benefit-oriented messages. In addition, aspects such as correct spelling and grammar, directive rather than passive approach, and messages without "textese" (eg, abbreviations) were found to be more preferred [31].

The notion of tailored interventions can be derived from the elaboration likelihood model (ELM). ELM proposes that information that is perceived to be personally relevant and expressed by a reliable sender increases perceived acceptability [32], which has also found empirical support [33]. A review on the main aspects of computer-tailored health interventions showed that tailoring was typically achieved through giving tailored feedback, personalization, and matching content with different groups of users [34]. However, tailoring has also been found to be complex and influenced by, for example, target groups and demographic variables [35]. Similarly, our findings showed limited consensus among students regarding tailoring, for example, personalization of messages. In order to optimize tailoring, intervention developers must have a sound knowledge of the culture, needs, and preferences of the end users. Although formative research designs would be valuable in this process, more research is needed on the opportunities with SMS text message interventions regarding tailoring. That is, what are the technical opportunities and challenges with tailoring and what level of tailoring is optimal, such as individual user-created messages versus group tailored information.

Our BCT analysis identified 23 techniques in the final program. In a Delphie study, international experts reached consensus on four BCTs that were estimated to be effective for mobile interventions targeting alcohol consumption. The techniques that received the highest mean ranking were for example goal setting, self-monitoring, action planning, and feedback in relation to goals [30]. Most of these techniques were used in our intervention and derived from, for example, control theory [36] and goal-setting theory [37]. Furthermore, these BCTs aim to tap into varied aspects of the self-regulatory process whereby the individual continuously monitors and adjusts discrepancies between an ideal or goal and actual states of being [38]. There is empirical support for interventions that address self-regulatory skills. A Cochrane review showed that receiving normative feedback had a small effect on reducing alcohol consumption in student populations [39]. Furthermore, self-monitoring has been found to be associated with improved effectiveness of brief interventions [27]. An earlier study found that reporting alcohol use on a daily basis reduced drinking among heavy drinkers by 20% [40]. There is also empirical support for including "if...then" tasks in interventions (eg, action planning) for a variety of behaviors [41] and student populations [42]. Finally, in a review on effective interventions to reduce risky drinking, goal setting, feedback, and advice were found to be the most effective strategies [43]. Thus, addressing self-regulatory skills in SMS text message interventions is feasible and, considering its empirical support, important in future intervention development.

In addition to self-regulatory skills, our intervention included BCTs that address motivation, such as anticipated regret. Looking at prevalent behavior change theory (eg, social cognition models [44,45] and social cognitive theory [46]), motivation to change is a central tenet. These theories argue that behavior change is a function of modified attitudes and beliefs about a behavior and perceptions about one's ability to make changes. Other BCTs that derive from these theories, and that were included in our intervention, give information on the consequences of drinking or prompt reflection on the reasons to reduce drinking [27].

Strengths and Limitations

A strength of this study could be the formative study design, which entailed revision of all messages based on both user and expert feedback. We believe that the input from the target audience and experts enhanced the intervention. A future randomized controlled trial will investigate the effect of the

XSL•FO

RenderX

intervention on alcohol consumption among students. Another strength of the study is the use of BCT analysis, which elucidated the theory base of the messages, giving readers a better understanding of what the intervention entails and enabling comparison with other interventions. A limitation of the study could be that most of the student participants in both the focus groups and panel evaluation were women. However, men were included in the data and messages were revised to interest both sexes. Another limitation could be that some participants took part in both the focus groups and the panel survey (10 of 21). The fact that individuals participated in focus groups prior completing the panel survey may have introduced bias. However, participating in both focus groups and the panel survey could also have been an advantage because these participants had a thorough understanding of the program. Lastly, the BCT analysis was performed by only one researcher, which may have limited the analysis compared with using several raters.

Implications

The study followed the recommended steps for the development of SMS text message-based interventions [19]. The findings show the importance and value of using a formative process in intervention development. The content of the SMS text messages was significantly revised based on input from end users and experts. The findings showed that neutral, motivated, clear, and tangible messages engaged students. These findings could be used in future studies aiming to develop SMS text message-based interventions. Furthermore, drinking patterns and access to mobile phones among university students in Sweden are similar to those in other Western countries. Therefore, we believe that the findings can be generalized to other university contexts.

Conclusions

The formative research design was valuable and resulted in significant changes in the intervention. All the original SMS text messages were changed and new messages were added. Overall, the findings showed that students were positive toward receiving support through SMS text messages and that neutral, motivated, clear, and tangible messages promoted engagement. More research is needed on the timing, frequency, and tailoring of messages.

Acknowledgments

Authors' Contributions

PB, MB, and UM designed the study. CL undertook preparation of the data collection material and actual data collection. CL and KT analyzed and compiled the focus group data. KT analyzed the quantitative panel evaluation data. CL analyzed the qualitative panel evaluation data. CL and KT performed revisions of the intervention and generated a preliminary final intervention. All authors read and accepted the final intervention. PB and KT wrote the first draft of the manuscript.

Conflicts of Interest

PB and MB own shares in and work in a private company that develops and distributes mobile health interventions.

Multimedia Appendix 1

The formative research design.

[PDF File (Adobe PDF File), 106KB - mhealth_v4i4e119_app1.pdf]

Multimedia Appendix 2

Focus group interview guide.

[PDF File (Adobe PDF File), 180KB - mhealth_v4i4e119_app2.pdf]

Multimedia Appendix 3

List of SMS messages in the final program design.

[PDF File (Adobe PDF File), 488KB - mhealth_v4i4e119_app3.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]

XSL•FO

RenderX

2.   Andersson C, Johnsson K, Berglund M, Ojehagen A. Intervention for hazardous alcohol use and high level of stress in university freshmen: a comparison between an intervention and a control university. Brain Res 2009 Dec 11;1305 Suppl:S61-S71. [doi: 10.1016/j.brainres.2009.08.030] [Medline: 19699184]

3.   Wicki M, Kuntsche E, Gmel G. Drinking at European universities? A review of students' alcohol use. Addict Behav 2010 Nov;35(11):913-924. [doi: 10.1016/j.addbeh.2010.06.015] [Medline: 20624671]

4.   Statistics Sweden. Population census 2014 URL: http://www.scb.se/sv_/Hitta-statistik/Statistik-efter-amne/Befolkning/ [accessed 2016-10-10] [WebCite Cache ID 6l9NMRJAy]

5.   Mason M, Ola B, Zaharakis N, Zhang J. Text messaging interventions for adolescent and young adult substance use: a meta-analysis. Prev Sci 2015 Feb;16(2):181-188. [doi: 10.1007/s11121-014-0498-7] [Medline: 24930386]

6.   Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010;32:56-69 [FREE Full text] [doi: 10.1093/epirev/mxq004] [Medline: 20354039]

7.   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]

8.   Suffoletto B, Callaway C, Kristan J, Kraemer K, Clark DB. Text-message-based drinking assessments and brief interventions for young adults discharged from the emergency department. Alcohol Clin Exp Res 2012 Mar;36(3):552-560. [doi: 10.1111/j.1530-0277.2011.01646.x] [Medline: 22168137]

9.   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]

10.   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]

11.   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 Jun;2(2):e30 [FREE Full text] [doi: 10.2196/mhealth.3233] [Medline: 25098296]

12.   Trull T, Ebner-Priemer U. Ambulatory assessment. Annu Rev Clin Psychol 2013;9:151-176 [FREE Full text] [doi: 10.1146/annurev-clinpsy-050212-185510] [Medline: 23157450]

13.   Beckjord E, Shiffman S. Background for real-time monitoring and intervention related to alcohol use. Alcohol Res 2014;36(1):9-18 [FREE Full text] [Medline: 26258996]

14.   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]

15.   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;3(1):e22 [FREE Full text] [doi: 10.2196/mhealth.3779] [Medline: 25714907]

16.   Mason MJ, Campbell L, Way T, Keyser-Marcus L, Benotsch E, Mennis J, et al. Development and outcomes of a text messaging tobacco cessation intervention with urban adolescents. Subst Abus 2015;36(4):500-506. [doi: 10.1080/08897077.2014.987946] [Medline: 25551337]

17.   Webb TL, Sniehotta FF, Michie S. Using theories of behaviour change to inform interventions for addictive behaviours. Addiction 2010 Nov;105(11):1879-1892. [doi: 10.1111/j.1360-0443.2010.03028.x] [Medline: 20670346]

18.   Hospital M, Wagner E, Morris S, Sawant M, Siqueira L, Soumah M. Developing an SMS intervention for the prevention of underage drinking: results from focus groups. Subst Use Misuse 2016;51(2):155-164. [doi: 10.3109/10826084.2015.1073325]

19.   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;3(4):e107 [FREE Full text] [doi: 10.2196/mhealth.4917] [Medline: 26690917]

20.   Prestwich A, Sniehotta FF, Whittington C, Dombrowski SU, Rogers L, Michie S. Does theory influence the effectiveness of health behavior interventions? Meta-analysis. Health Psychol 2014 May;33(5):465-474. [doi: 10.1037/a0032853] [Medline: 23730717]

21.   Michie S, Prestwich A. Are interventions theory-based? Development of a theory coding scheme. Health Psychol 2010 Jan;29(1):1-8. [doi: 10.1037/a0016939] [Medline: 20063930]

22.   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]

23.   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]

24.   McCambridge J, Bendtsen P, Bendtsen M, Nilsen P. Alcohol email assessment and feedback study dismantling effectiveness for university students (AMADEUS-1): study protocol for a randomized controlled trial. Trials 2012;13:49 [FREE Full text] [doi: 10.1186/1745-6215-13-49] [Medline: 22540638]

25. McCambridge J, Bendtsen M, Karlsson N, White IR, Bendtsen P. Alcohol assessment & feedback by e-mail for university student hazardous and harmful drinkers: study protocol for the AMADEUS-2 randomised controlled trial. BMC Public Health 2013;13:949 [FREE Full text] [doi: 10.1186/1471-2458-13-949] [Medline: 24456668]

26. Sandelowski M. Qualitative analysis: what it is and how to begin. Res Nurs Health 1995;18:371-375.

27. 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]

28. Renner KA, Natalie W, Varsha P, Ross M. Harm reduction text messages delivered during alcohol drinking: feasibility study protocol. JMIR Res Protoc 2012;1(1):e4 [FREE Full text] [doi: 10.2196/resprot.1970] [Medline: 23611773]

29. Kristan J, Suffoletto B. Using online crowdsourcing to understand young adult attitudes toward expert-authored messages aimed at reducing hazardous alcohol consumption and to collect peer-authored messages. Transl Behav Med 2015 Mar;5(1):45-52 [FREE Full text] [doi: 10.1007/s13142-014-0298-4] [Medline: 25729452]

30. 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]

31. Muench F, van Stolk-Cooke K, Morgenstern J, Kuerbis AN, Markle K. Understanding messaging preferences to inform development of mobile goal-directed behavioral interventions. J Med Internet Res 2014;16(2):e14 [FREE Full text] [doi: 10.2196/jmir.2945] [Medline: 24500775]

32. Petty E, Cacioppo J. Communication and Persuasion: Central and Peripheral Routes to Attitude Change. New York: Springer-Verlag; 1986.

33. Stephens PC, Sloboda Z, Grey S, Stephens R, Hammond A, Hawthorne R, et al. Is the receptivity of substance abuse prevention programming affected by students' perceptions of the instructor? Health Educ Behav 2009 Aug;36(4):724-745. [doi: 10.1177/1090198107304388] [Medline: 18809689]

34. Lustria ML, Cortese J, Noar SM, Glueckauf RL. Computer-tailored health interventions delivered over the Web: review and analysis of key components. Patient Educ Couns 2009 Feb;74(2):156-173. [doi: 10.1016/j.pec.2008.08.023] [Medline: 18947966]

35. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007 Jul;133(4):673-693. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]

36. Carver CS, Scheier MF. Control theory: a useful conceptual framework for personality-social, clinical, and health psychology. Psychol Bull 1982 Jul;92(1):111-135. [Medline: 7134324]

37. Locke A, Latham G. A Theory of Goal Setting and Task Performance. Englewood Cliffs, NJ: Prentice Hall; 1990.

38. Baumeister R, Heatherton T, Tice D. Losing Control: How and Why People Fail at Self-Regulation. San Diego: Academic Press; 1994.

39. Foxcroft DR, Moreira MT, Almeida SN, Smith LA. Social norms information for alcohol misuse in university and college students. Cochrane Database Syst Rev 2015;1:CD006748. [doi: 10.1002/14651858.CD006748.pub3] [Medline: 25622306]

40. Helzer J, Badger G, Rose G, Mongeon J, Searles J. Decline in alcohol consumption during two years of daily reporting. J Stud Alcohol 2002 Sep;63(5):551-558. [Medline: 12380851]

41. Gollwitzer P, Sheeran P. Implementation intentions and goal achievement: a meta-analysis of effects and processes. Adv Exp Soc Psychol 2006;38:69-120.

42. Murgraff V, White D, Phillips K. Moderating binge drinking: it is possible to change behaviour if you plan it in advance. Alcohol Alcohol 1996 Nov;31(6):577-582 [FREE Full text] [Medline: 9010548]

43. Whitlock E, Polen M, Green CA, Orleans T, Klein J, US Preventive Services Task Force. Behavioral counseling interventions in primary care to reduce risky/harmful alcohol use by adults: a summary of the evidence for the U.S. Preventive Services Task Force. Ann Intern Med 2004 Apr 6;140(7):557-568. [Medline: 15068985]

44. Ajzen I. From Intentions to Actions: A Theory of Planned Behaviour. Berlin: Springer Berlin Heidelberg; 1985.

45. Janz N, Becker M. The Health Belief Model: a decade later. Health Educ Q 1984;11(1):1-47. [Medline: 6392204]

46. Bandura A. Self-efficacy: the exercise of control. New York: WH Freeman; 1997.

Abbreviations

BCT: behavior change technique
ELM: elaboration likelihood model
SMS: short message service

Original Paper

Economic Evaluation of Text-Messaging and Smartphone-Based Interventions to Improve Medication Adherence in Adolescents with Chronic Health Conditions: A Systematic Review

Sherif M Badawy[1,2], MBBCh, MS, MD; Lisa M Kuhns[3], MPH, PhD

[1]Ann and Robert H Lurie Children's Hospital of Chicago, Department of Pediatrics, Division of Hematology, Oncology, and Stem Cell Transplant, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

[2]Faculty of Medicine, Department of Pediatrics, Division of Hematology and Oncology, Zagazig University, Zagazig, Egypt

[3]Ann and Robert H Lurie Children's Hospital of Chicago, Department of Pediatrics, Division of Adolescent Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, United States

Corresponding Author:
Sherif M Badawy, MBBCh, MS, MD
Ann and Robert H Lurie Children's Hospital of Chicago
Department of Pediatrics, Division of Hematology, Oncology, and Stem Cell Transplant
Northwestern University Feinberg School of Medicine
225 E Chicago Ave, Box 30
Chicago, IL,
United States
Phone: 1 312 227 4836
Fax: 1 312 227 9376
Email: sbadawy@luriechildrens.org

Abstract

Background: The rate of chronic health conditions (CHCs) in children and adolescents has doubled in the past 20 years, with increased health care costs. Technology-based interventions have demonstrated efficacy to improving medication adherence. However, data to support the cost effectiveness of these interventions are lacking.

Objective: The objective of this study is to conduct an economic evaluation of text-messaging and smartphone-based interventions that focus on improving medication adherence in adolescents with CHCs.

Methods: Searches included PubMed MEDLINE, Embase, Cochrane Central Register of Controlled Trials, Cumulative Index to Nursing and Allied Health Literature, PsycINFO, Web of Science, and Inspec. Eligibility criteria included age (12-24 years old), original articles, outcomes for medication adherence, and economic outcomes.

Results: Our search identified 1118 unique articles that were independently screened. A total of 156 articles met inclusion criteria and were then examined independently with full-text review. A total of 15 articles met most criteria but lacked economic outcomes such as cost effectiveness or cost-utility data. No articles met all predefined criteria to be included for final review. Only 4 articles (text messaging [n=3], electronic directly observed therapy [n=1]) described interventions with possible future cost-saving but no formal economic evaluation.

Conclusions: The evidence to support the cost effectiveness of text-messaging and smartphone-based interventions in improving medication adherence in adolescents with CHCs is insufficient. This lack of research highlights the need for comprehensive economic evaluation of such interventions to better understand their role in cost-savings while improving medication adherence and health outcomes. Economic evaluation of technology-based interventions can contribute to more evidence-based assessment of the scalability, sustainability, and benefits of broader investment of such technology tools in adolescents with CHCs.

XSL·FO
RenderX

Introduction

The rate of chronic health conditions (CHCs) in children and adolescents (eg, asthma, diabetes) has doubled in the past 20 years [1,2]. Adolescents (12-24 years old) with CHCs, a special subpopulation of pediatric patients, face the day-to-day challenges of transitioning to adult responsibilities while simultaneously managing their illness-related routine [3-5]. Adolescence is an important time to develop healthy habits and behaviors, and building self-management skills is a critical component of successful transition to adulthood. Engaging adolescents with CHCs in self-management skill building is an invaluable investment with long-term benefits. In particular, medication adherence is a crucial component of self-management, and poor adherence is a common problem in adolescents with CHCs [3]. Across pediatric chronic conditions, this can have negative effects on morbidity, mortality, and quality of life with increased use of health services and annual health care cost [3,6-10].

Taking daily medications is a challenge irrespective of the frequency, formulation, or patient's age. There are possible differences in adherence barriers across chronic conditions, such as disease-specific treatment regimens and monitoring requirements. However, evidence from a recent systematic review suggests that among adolescents with chronic conditions, most perceived barriers are not unique to a specific disease state [11]. Nevertheless, barriers to medication adherence among adolescents may be multifaceted, and there may be common attributes to this phenomenon that could be amenable to interventions across chronic conditions.

Adolescents have adopted communication technology such as cellular phones, the Internet, and social networking at a rapid rate across levels of social position and status [12-14]. Recent developments in information and communications technologies have opened new opportunities to improve health care and to link patients and their providers. A recent report from a Pew Internet research survey found that teens have access to smartphones, tablets, desktop computers, and laptop computers at rates of 73%, 58%, 87%, and 81%, respectively [12,13]. This presents an opportunity to promote self-management and medication adherence among adolescents via these technologies. The use of portable and easily accessible technology-based interventions, particularly text messaging and smartphone apps, to address health-related problems has been shown to be both feasible and acceptable for different health conditions [15-17]. In addition, previous systematic reviews and meta-analyses of pediatric patients with CHCs have shown considerable positive effects of these types of interventions in improving medication adherence, health-related quality of life (HRQOL), and family functioning [15-21]. However, the cost effectiveness of developing and maintaining such technology-based interventions remains poorly understood.

The effect of technology-based interventions on health care costs in adolescents with CHCs may go beyond the direct cost savings associated with medication adherence and related health outcomes. These interventions may facilitate efficient health care operations (eg, fewer missed clinic, screening, or laboratory

monitoring appointments), increased access to high-quality care (eg, timely referrals to other services for consultations), and potential cost savings of labor. Therefore, economic evaluation of technology-based interventions can contribute to a better understanding of the scalability, sustainability, and benefits of broader investment of such technology tools. Economic data may also raise considerations for third-party reimbursement should interventions prove to be effective in improving health outcomes in this population. The objectives of this systematic review are to (1) conduct an economic evaluation (cost effectiveness and cost-utility analyses) of text-messaging and smartphone-based interventions that focus on improving medication adherence in adolescents with CHCs and (2) determine whether the incremental benefit gained from using such interventions is enough to justify the additional cost required to adopt, develop, and maintain the intervention.

Methods

Search Strategy

The authors collaborated with a librarian who developed the search strategies and from July to September 2015 ran searches in the following databases: PubMed MEDLINE, Embase (embase.com), Cochrane Central Register of Controlled Trials (CENTRAL) on the Wiley platform, Cumulative Index to Nursing and Allied Health Literature (CINAHL) (EBSCO), PsycINFO (EBSCO), Web of Science, Center for Review and Dissemination (CRD); and Inspec (EBSCO). Further searches were run in November 2015 using the following sources: ProQuest dissertations, Scopus, ClinicalTrials.gov, World Health Organization clinical trials, Controlled-Trials.org, Institute for Electrical and Electronics Engineers (IEEE) Xplore, and Google Scholar. Search strategies for all databases except MEDLINE were adapted from the PubMed MEDLINE strategy. All databases were searched back to 1995 with no language limits applied. The search strategy looked for all articles on text messaging, phones, mobile apps, and portable software combined with adherence or compliance, and search terms related to child, pediatric, adolescent, and youth. Multimedia Appendix 1 shows the complete search strategies in each database. The authors also attempted to discover additional studies by searching the reference lists of key studies and relevant systematic reviews. We followed the guidelines for Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) in the report of evidence across the studies reviewed herein [22].

Inclusion and Exclusion Criteria

The inclusion criteria were as follows: (1) adolescents (12-24 years old) with a CHC that requires long-term daily or weekly medications for 12 months or longer [23], (2) original research manuscripts, (3) studies that were either randomized controlled trials, quasi-experimental studies, or pilot/feasibility studies (including single arm, pretest/posttest), (4) text-messaging or smartphone-based interventions (app or mobile intervention), and (5) medication adherence as the primary or the secondary outcome. The exclusion criteria included the following: (1) mean or median age of participants younger than 12 years old or older than 24 years old or mean or median age not specified

in the article, (2) adolescent participants not the focus of the study intervention (eg, interventions that target babies born to adolescent mothers with HIV or target parents of adolescent patients with CHCs), (3) interventions focused on disease monitoring or ecological momentary assessment but not designed to improve medication adherence, (4) technology-based interventions other than text messaging and smartphone apps, and (5) lack of economic outcomes such as cost effectiveness or cost-utility data.

Data Extraction

We developed a standardized form for data extraction from the final included articles, adjusted for this particular study. Data items in the extraction form included the following: first author's name; publication year; country; CHC; participant ages; study design; duration of intervention and follow-up; components of technology intervention (text messaging or smartphone apps); adherence measures and rates; disease-related outcomes; theoretical framework; and economic outcomes including cost effectiveness and/or cost-utility data (eg, cost components of each intervention), incremental cost-effectiveness ratios (ICERs), quality adjusted life years (QALYs), and sensitivity analyses. We planned to evaluate the quality of evidence using the GRADE (Grades of Recommendation, Assessment, Development, and Evaluation) approach [24].

Results

Overview

The initial search retrieved 1137 records from the main electronic databases (PubMed, Embase, PsycINFO, CINAHL, CENTRAL, Web of Science, Inspec, CRD, and IEEE Xplore). We identified an additional 286 records from the gray literature and hand search of the bibliography of other systematic reviews. After removal of the duplicates, 1118 original articles remained (Figure 1). The authors independently screened the article titles and abstracts, and of those screened, 156 articles met the inclusion criteria. The authors then independently reviewed the full text of the 156 articles against the exclusion criteria. No articles met all predefined criteria to be included for final review. The reasons for exclusion of full text papers were documented in an adapted PRISMA study flowchart (Figure 1) [22]. It is worth noting that 15 articles met most predetermined inclusion and exclusion criteria but lacked economic outcomes such as cost effectiveness or cost-utility data. In addition, only 4 articles (text messaging [n=3] and mobile directly observed therapy [DOT] [n=1]) described interventions with possible future cost saving but no formal economic evaluation. Therefore, we summarized data from these 4 articles and suggested an economic evaluation approach for future studies.

Figure 1. Flow of studies through the review according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines.
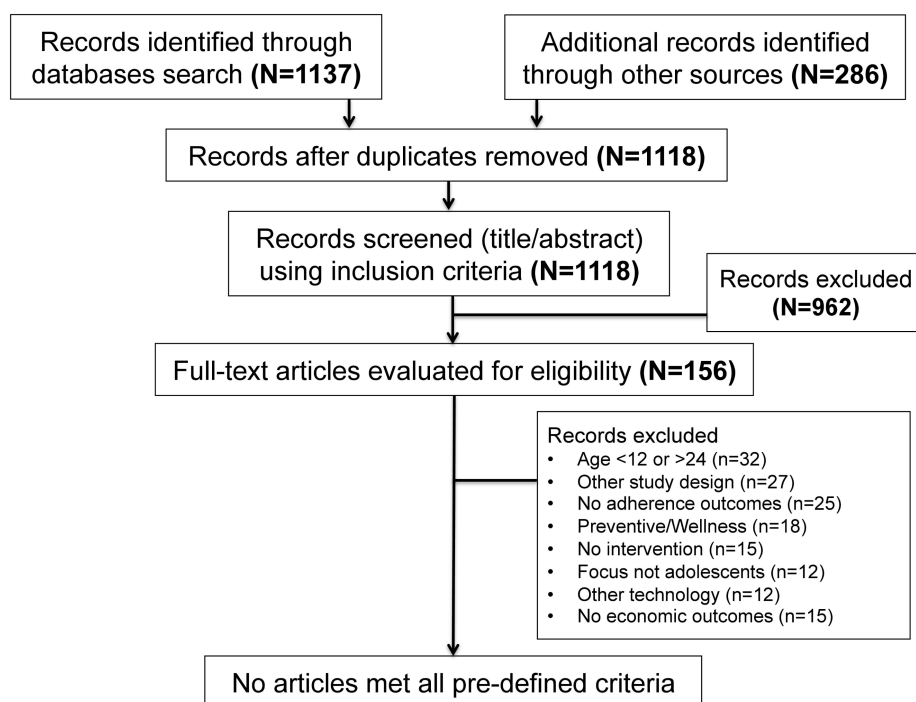
Study Characteristics

Creary et al were able to achieve a significant improvement in hydroxyurea adherence rates (93.3%) among children and adolescents with sickle cell disease by using mobile DOT [25]. The authors suggested that mobile DOT could be a cost-effective intervention as it has the potential for wider application with lower technology cost over time and better facial recognition

capabilities [25]. The authors also projected that mobile DOT could decrease health care use because patients with higher adherence to hydroxyurea would have fewer disease-related complications, hospitalizations, and visits to the emergency department [25]. Another study by Ting et al using text message reminders showed statistically significant improvement in clinic attendance rates among adolescents with childhood-onset systemic lupus erythematosus but no improvement in adherence

to their medication (hydroxychloroquine) [26]. However, the authors suggested that using a text-messaging approach might prove to be cost effective by reducing health care cost while improving clinic adherence and health outcomes overall [26]. Moreover, Miloh et al have shown that using a text-messaging approach not only significantly improved adherence to immune-suppressive medications in pediatric recipients of liver transplant but also reduced rejection episodes, suggesting possible cost-saving effects [27]. The authors further highlighted that the success of an implemented text-messaging approach lies in the characteristics of the intervention itself being personal, discreet, simple, socially acceptable, minimally intrusive, and low cost and requiring minimal time commitment from health care providers [27]. Furthermore, Franklin et al evaluated the efficacy of a text-messaging support system—Sweet Talk—among adolescents with diabetes mellitus type 1 [28]. The authors reported improvement in hemoglobin A_{1c} in patients who received intensive insulin therapy in addition to Sweet Talk and improvement in self-efficacy and adherence by self-report (in comparison to conventional therapy plus Sweet Talk) [28]. The authors highlighted that a text-messaging intervention may help to overcome critical limitations of the current more labor-intensive approach to diabetes education and management, such as cost and time commitment [28]. The authors also projected that text messaging might be a low-cost behavioral support intervention that can address the need for long-term behavior change and be integrated into routine care in the clinic setting with detailed cost-effectiveness evaluations [28].

Suggested Approach for Economic Evaluation

Health economic evaluation of technology-based interventions (eg, text messaging and smartphone apps) helps to highlight the added value of these interventions by addressing two important points: (1) whether the technology-based intervention used to improve medication adherence among adolescents with chronic conditions improves health outcomes relative to other existing interventions and (2) whether the incremental benefit gained from using the technology-based intervention is enough to justify the additional cost required to adopt, develop, and maintain the intervention. Data on ICERs per unit improvement in medication adherence, disease-related outcomes, and QALYs would inform health economic evaluation and aid in the development of a cost-effectiveness model to evaluate whether these improvements are worth the costs required to develop, maintain, and disseminate the intervention. Disease-related outcomes include disease-specific complications, mortality, and HRQOL. In addition, evaluation of health and social care costs are important to consider in relation to HRQOL. Health economic evaluation should include a comprehensive cost analysis of the development, maintenance, and dissemination of technology-based interventions. A cost-effectiveness model using a cost-utility analysis of technology-based interventions for improving patient outcomes could be measured in terms of ICER per unit improvement in disease-related outcomes or QALY gained compared to standard of care. The ICER per extra QALY generated by text-messaging or smartphone app interventions to improve medication adherence over standard of care can be calculated using the following equations:

- $(cost_{text\ messaging} - cost_{standard\ of\ care}) / (QALY_{text\ messaging} - QALY_{standard\ of\ care})$

- $(cost_{smartphone\ app} - cost_{standard\ of\ care}) / (QALY_{smartphone\ app} - QALY_{standard\ of\ care})$

- $(cost_{smartphone\ app} - cost_{text\ messaging}) / (QALY_{smartphone\ app} - QALY_{text\ messaging})$

Discussion

Principal Findings

In our study, we were not able to identify any articles that met all our predefined criteria. Only 4 articles described interventions with possible future cost saving but no formal economic evaluation. Deriving QALY for use in cost-utility analysis in pediatric populations is a challenging task with only a few child-specific preference-based measures available. There is no single preference-based utility measure that has been validated for children of all ages in different health states, including adolescents with chronic conditions [29]. Time period is another important consideration in terms of evaluating short-term and, more commonly, long-term outcomes. Alternatives include using adult instruments, proxy measures, expert opinion, or published catalogs of pediatric utility values for different chronic conditions [29]. Individual preferences for health states can be elicited by either direct or indirect measures. Direct measures include standard gamble and time trade-off. The standard gamble is a technique to measure individuals' preferences under uncertainty and to express the outcome of different therapeutic choices in utility values that can be used in clinical decision analysis and health program evaluation. The time trade-off is another technique to elicit individuals' preferences for health states by letting them imagine living a defined number of years in an imperfect health state then indicate the number of remaining life years in full health at which they are indifferent between the longer period of impaired health and the shorter period of full health. The challenge lies in adolescents' ability to interpret some of those measures. They may lack the cognitive ability to understand the abstract concepts included in standard gambles or time trade-offs, especially the ones related to time spent in different health states and the possibility of death [30]. Chaining has been suggested as a technique to address these issues [31]. For example, the worst possible health state of a disease is used instead of death. Indirect measures include the Health Utilities Index and the European Group Quality of Life 5; both are validated for use in adolescents 12 years and older [32,33]. The problem with these measures is that the predetermined utility weights currently used for these questionnaires are based on adult preferences, which may compromise their use for adolescents [29]. However, an Australian group evaluated another measure, the Assessment of Quality of Life instrument. The authors conducted a recalibration study and were able to derive utility weights specifically for adolescents [34]. The study included adolescents from 4 different countries and used algorithms and multiplicative models to develop age- and country-specific utility weights [34].

XSL•FO

RenderX

Nonadherence to the recommended treatment is a widespread problem in pediatric CHCs. The increasing prevalence of CHCs coupled with management problems in pediatric populations present a barrier to optimal health. Self-management skills in adolescents and young adults are critical to maintain optimal adherence to their chronic medications, especially when they transition from pediatric to adult facilities with more expectations of self-care. Adolescent-centered interventions are needed to optimize their adherence to prescribed medication across CHCs, support the development of self-management skills, and enhance intervention uptake and long-term engagement while transitioning to self-care.

There has been a growing interest in the use of technology to improve medication adherence and self-management skills in the last few years. Similarly, there has been increased interest in the use of portable and easily accessible technology-based approaches to address health-related problems with an overall acceptability and feasibility for different health conditions. Among adults, evidence for the efficacy of text messaging to support medication adherence exists [35,36] and the evidence among adolescents with chronic conditions is emerging [15-17]. While results of these studies are promising, they suggest that additional adherence intervention development is needed and should be tested with more rigorous designs and across a broader range of chronic conditions.

Despite the growing evidence of the efficacy of text-messaging and smartphone app interventions in improving medication adherence in adults with CHCs [35-38], to date there have been no formal economic evaluations of these interventions, and their cost effectiveness remains unclear [38]. However, there is evidence to support the cost effectiveness of different technology-based interventions to promote behavior change among adults, such as smoking cessation [39-41].

The majority of smartphone app initiatives have been pilot studies and the data generated from these studies are limited. In addition to efficacy and effectiveness data, economic evaluation is warranted. The cost to develop and maintain each intervention could be a barrier to the use of these interventions on a broader scale. Additionally, there is variability in patient access to preferred technologies. Formal economic evaluation of various interventions will help health care authorities determine whether the investment required to develop, maintain, and disseminate these interventions is worth the broader benefit, or lack thereof, experienced by patients with chronic conditions. Given the emerging evidence in the field of eHealth, future economic evaluations could consider broader inclusion criteria for different technology-based interventions.

Conclusion

In conclusion, we found no evidence to support the cost effectiveness of technology-based text-messaging and smartphone app interventions. The effect of such technology tools on health care costs in adolescents with CHCs can be beyond medication adherence. Technology-based interventions can facilitate increased operating efficiencies (eg, fewer missed clinic appointments), increased access to high-quality health services (eg, timely referrals to other services for consultations and self-management tools), and potential labor cost savings. Economic evaluation of technology-based interventions can contribute to a better and more evidence-based assessment of the scalability, sustainability, and benefits of broader investment of such technology tools and may raise considerations for third-party reimbursement should interventions prove to be effective in improving health outcomes in this population.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Database search strategies.

[PDF File (Adobe PDF File), 149KB - mhealth_v4i4e121_app1.pdf]

References

1. Van Cleave J, Gortmaker SL, Perrin JM. Dynamics of obesity and chronic health conditions among children and youth. J Am Med Assoc 2010 Feb 17;303(7):623-630. [doi: 10.1001/jama.2010.104] [Medline: 20159870]
2. Berry JG, Hall M, Hall DE, Kuo DZ, Cohen E, Agrawal R, et al. Inpatient growth and resource use in 28 children's hospitals: a longitudinal, multi-institutional study. JAMA Pediatr 2013 Feb;167(2):170-177 [FREE Full text] [doi: 10.1001/jamapediatrics.2013.432] [Medline: 23266509]
3. Rapoff M. Adherence to Pediatric Medical Regimens. New York, NY: Springer Science+Business Media; 2010.
4. Drotar D. Promoting Adherence to Medical Treatment in Chronic Childhood Illness: Concepts, Methods, and Interventions. Mahwah, NJ: Lawrence Erlbaum Associates; 2000.
5. Hommel KA, Davis CM, Baldassano RN. Objective versus subjective assessment of oral medication adherence in pediatric inflammatory bowel disease. Inflamm Bowel Dis 2009 Apr;15(4):589-593 [FREE Full text] [doi: 10.1002/ibd.20798] [Medline: 18985746]
6. McGrady ME, Hommel KA. Medication adherence and health care utilization in pediatric chronic illness: a systematic review. Pediatrics 2013 Oct;132(4):730-740 [FREE Full text] [doi: 10.1542/peds.2013-1451] [Medline: 23999953]

7. DiMatteo MR. Variations in patients' adherence to medical recommendations: a quantitative review of 50 years of research. Med Care 2004 Mar;42(3):200-209. [Medline: 15076819]

8. DiMatteo MR, Giordani PJ, Lepper HS, Croghan TW. Patient adherence and medical treatment outcomes: a meta-analysis. Med Care 2002 Sep;40(9):794-811. [doi: 10.1097/01.MLR.0000024612.61915.2D] [Medline: 12218770]

9. Sokol MC, McGuigan KA, Verbrugge RR, Epstein RS. Impact of medication adherence on hospitalization risk and healthcare cost. Med Care 2005 Jun;43(6):521-530. [Medline: 15908846]

10. Ingerski LM, Modi AC, Hood KK, Pai AL, Zeller M, Piazza-Waggoner C, et al. Health-related quality of life across pediatric chronic conditions. J Pediatr 2010 Apr;156(4):639-644 [FREE Full text] [doi: 10.1016/j.jpeds.2009.11.008] [Medline: 20117793]

11. Hanghøj S, Boisen KA. Self-reported barriers to medication adherence among chronically ill adolescents: a systematic review. J Adolesc Health 2014 Feb;54(2):121-138. [doi: 10.1016/j.jadohealth.2013.08.009] [Medline: 24182940]

12. Lenhart A. Teens, Social Media & Technology Overview 2015. Washington, DC: Pew Internet and American Life Project; 2015. URL: http://www.pewinternet.org/files/2015/04/PI_TeensandTech_Update2015_0409151.pdf [accessed 2016-10-18] [WebCite Cache ID 6jaddHT8x]

13. Smith A. U.S. Smartphone Use in 2015. Washington, DC: Pew Internet and American Life Project; 2015. URL: http://www.pewinternet.org/files/2015/03/PI_Smartphones_0401151.pdf [accessed 2016-10-18] [WebCite Cache ID 6jae5eF6U]

14. Badawy SM, Thompson AA, Liem RI. Technology access and smartphone app preferences for medication adherence in adolescents and young adults with sickle cell disease. Pediatr Blood Cancer 2016 May;63(5):848-852. [doi: 10.1002/pbc.25905] [Medline: 26844685]

15. Mistry N, Keepanasseril A, Wilczynski NL, Nieuwlaat R, Ravall M, Haynes RB, Patient Adherence Review Team. Technology-mediated interventions for enhancing medication adherence. J Am Med Inform Assoc 2015 Apr;22(e1):e177-e193. [doi: 10.1093/jamia/ocu047] [Medline: 25726568]

16. Pai AL, McGrady M. Systematic review and meta-analysis of psychological interventions to promote treatment adherence in children, adolescents, and young adults with chronic illness. J Pediatr Psychol 2014 Sep;39(8):918-931. [doi: 10.1093/jpepsy/jsu038] [Medline: 24952359]

17. McGrady ME, Ryan JL, Gutiérrez-Colina AM, Fredericks EM, Towner EK, Pai AL. The impact of effective paediatric adherence promotion interventions: systematic review and meta-analysis. Child Care Health Dev 2015 Nov;41(6):789-802. [doi: 10.1111/cch.12271] [Medline: 26255643]

18. McGrady ME, Hommel KA. Targeting health behaviors to reduce health care costs in pediatric psychology: descriptive review and recommendations. J Pediatr Psychol 2015 Sep 10. [doi: 10.1093/jpepsy/jsv083] [Medline: 26359311]

19. Graves MM, Roberts MC, Rapoff M, Boyer A. The efficacy of adherence interventions for chronically ill children: a meta-analytic review. J Pediatr Psychol 2010 May;35(4):368-382 [FREE Full text] [doi: 10.1093/jpepsy/jsp072] [Medline: 19710248]

20. Dean AJ, Walters J, Hall A. A systematic review of interventions to enhance medication adherence in children and adolescents with chronic illness. Arch Dis Child 2010 Sep;95(9):717-723. [doi: 10.1136/adc.2009.175125] [Medline: 20522463]

21. Nieuwlaat R, Wilczynski N, Navarro T, Hobson N, Jeffery R, Keepanasseril A, et al. Interventions for enhancing medication adherence. Cochrane Database Syst Rev 2014;CD000011. [doi: 10.1002/14651858.CD000011.pub4] [Medline: 25412402]

22. Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015 Jan;4:1 [FREE Full text] [doi: 10.1186/2046-4053-4-1] [Medline: 25554246]

23. Goodman RA, Posner SF, Huang ES, Parekh AK, Koh HK. Defining and measuring chronic conditions: imperatives for research, policy, program, and practice. Prev Chronic Dis 2013;10:E66 [FREE Full text] [doi: 10.5888/pcd10.120239] [Medline: 23618546]

24. Higgins JPT, Green S, editors. Cochrane Handbook for Systematic Reviews of Interventions, Version 5.1.0. London, England: The Cochrane Collaboration; 2011.

25. Creary SE, Gladwin MT, Byrne M, Hildesheim M, Krishnamurti L. A pilot study of electronic directly observed therapy to improve hydroxyurea adherence in pediatric patients with sickle-cell disease. Pediatr Blood Cancer 2014 Jun;61(6):1068-1073. [doi: 10.1002/pbc.24931] [Medline: 24436121]

26. Ting TV, Kudalkar D, Nelson S, Cortina S, Pendl J, Budhani S, et al. Usefulness of cellular text messaging for improving adherence among adolescents and young adults with systemic lupus erythematosus. J Rheumatol 2012 Jan;39(1):174-179. [doi: 10.3899/jrheum.110771] [Medline: 22089460]

27. Miloh T, Annunziato R, Arnon R, Warshaw J, Parkar S, Suchy FJ, et al. Improved adherence and outcomes for pediatric liver transplant recipients by using text messaging. Pediatrics 2009 Nov;124(5):e844-e850 [FREE Full text] [doi: 10.1542/peds.2009-0415] [Medline: 19822583]

28. Franklin VL, Waller A, Pagliari C, Greene SA. A randomized controlled trial of Sweet Talk, a text-messaging system to support young people with diabetes. Diabet Med 2006 Dec;23(12):1332-1338. [doi: 10.1111/j.1464-5491.2006.01989.x] [Medline: 17116184]

29. Ungar WJ. Challenges in health state valuation in paediatric economic evaluation: are QALYs contraindicated? Pharmacoeconomics 2011 Aug;29(8):641-652. [doi: 10.2165/11591570-000000000-00000] [Medline: 21604821]

30.   Petrou S. Methodological issues raised by preference-based approaches to measuring the health status of children. Health Econ 2003 Aug;12(8):697-702. [doi: 10.1002/hec.775] [Medline: 12898666]

31.   Torrance GW. Utility measurement in healthcare: the things I never got to. Pharmacoeconomics 2006;24(11):1069-1078. [Medline: 17067192]

32.   Horsman J, Furlong W, Feeny D, Torrance G. The Health Utilities Index (HUI): concepts, measurement properties and applications. Health Qual Life Outcomes 2003;1:54 [FREE Full text] [doi: 10.1186/1477-7525-1-54] [Medline: 14613568]

33.   Oppe M, van Reenen M. EQ-5D-3L User Guide, Version 5.1. 2015. URL: http://www.euroqol.org/fileadmin/user_upload/Documenten/PDF/Folders_Flyers/EQ-5D-3L_UserGuide_2015.pdf [accessed 2016-10-18] [WebCite Cache ID 6jaf8Wgjo]

34.   Moodie M, Richardson J, Rankin B, Iezzi A, Sinha K. Predicting time trade-off health state valuations of adolescents in four Pacific countries using the Assessment of Quality-of-Life (AQoL-6D) instrument. Value Health 2010 Dec;13(8):1014-1027 [FREE Full text] [doi: 10.1111/j.1524-4733.2010.00780.x] [Medline: 20825621]

35.   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]

36.   Payne HE, Lister C, West JH, Bernhardt JM. Behavioral functionality of mobile apps in health interventions: a systematic review of the literature. JMIR Mhealth Uhealth 2015;3(1):e20 [FREE Full text] [doi: 10.2196/mhealth.3335] [Medline: 25803705]

37.   Hamine S, Gerth-Guyette E, Faulx D, Green BB, Ginsburg AS. Impact of mHealth chronic disease management on treatment adherence and patient outcomes: a systematic review. J Med Internet Res 2015;17(2):e52 [FREE Full text] [doi: 10.2196/jmir.3951] [Medline: 25803266]

38.   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]

39.   Stanczyk NE, Smit ES, Schulz DN, de VH, Bolman C, Muris JW, et al. An economic evaluation of a video- and text-based computer-tailored intervention for smoking cessation: a cost-effectiveness and cost-utility analysis of a randomized controlled trial. PLoS One 2014;9(10):e110117 [FREE Full text] [doi: 10.1371/journal.pone.0110117] [Medline: 25310007]

40.   Guerriero C, Cairns J, Roberts I, Rodgers A, Whittaker R, Free C. The cost-effectiveness of smoking cessation support delivered by mobile phone text messaging: Txt2stop. Eur J Health Econ 2013 Oct;14(5):789-797 [FREE Full text] [doi: 10.1007/s10198-012-0424-5] [Medline: 22961230]

41.   Chen Y, Madan J, Welton N, Yahaya I, Aveyard P, Bauld L, et al. Effectiveness and cost-effectiveness of computer and other electronic aids for smoking cessation: a systematic review and network meta-analysis. Health Technol Assess 2012;16(38):1-205 [FREE Full text] [doi: 10.3310/hta16380] [Medline: 23046909]

Abbreviations

CENTRAL: Cochrane Central Register of Controlled Trials
CHC: chronic health conditions
CINAHL: Cumulative Index to Nursing and Allied Health Literature
CRD: Center for Review and Dissemination
DOT: directly observed therapy
GRADE: Grades of Recommendation, Assessment, Development, and Evaluation
HRQOL: health-related quality of life
ICER: incremental cost-effectiveness ratio
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
QALY: quality adjusted life years

XSL•FO

RenderX

Original Paper

Design Considerations in Development of a Mobile Health Intervention Program: The TEXT ME and TEXTMEDS Experience

Jay Thakkar[1,2,3], MD; Tony Barry[3,4], BSc; Aravinda Thiagalingam[2,3], PhD; Julie Redfern[1,2], PhD; Alistair L McEwan[4], PhD; Anthony Rodgers[1,2], PhD; Clara K Chow[1,2,3], PhD

[1]The George Institute for Global Health, Camperdown, Australia

[2]Sydney Medical School, The University of Sydney, Sydney, Australia

[3]Westmead Hospital, Sydney, Australia

[4]The University of Sydney, Sydney, Australia

Corresponding Author:
Clara K Chow, PhD
The George Institute for Global Health
Level 10, King George V Building
Missenden Road
Camperdown, 2050
Australia
Phone: 61 2 9993 4500
Fax: 61 2 9993 4501
Email: cchow@georgeinstitute.org.au

Abstract

Background: Mobile health (mHealth) has huge potential to deliver preventative health services. However, there is paucity of literature on theoretical constructs, technical, practical, and regulatory considerations that enable delivery of such services.

Objectives: The objective of this study was to outline the key considerations in the development of a text message-based mHealth program; thus providing broad recommendations and guidance to future researchers designing similar programs.

Methods: We describe the key considerations in designing the intervention with respect to functionality, technical infrastructure, data management, software components, regulatory requirements, and operationalization. We also illustrate some of the potential issues and decision points utilizing our experience of developing text message (short message service, SMS) management systems to support 2 large randomized controlled trials: TEXT messages to improve MEDication adherence & Secondary prevention (TEXTMEDS) and Tobacco, EXercise and dieT MEssages (TEXT ME).

Results: The steps identified in the development process were: (1) background research and development of the text message bank based on scientific evidence and disease-specific guidelines, (2) pilot testing with target audience and incorporating feedback, (3) software-hardware customization to enable delivery of complex personalized programs using prespecified algorithms, and (4) legal and regulatory considerations. Additional considerations in developing text message management systems include: balancing the use of customized versus preexisting software systems, the level of automation versus need for human inputs, monitoring, ensuring data security, interface flexibility, and the ability for upscaling.

Conclusions: A merging of expertise in clinical and behavioral sciences, health and research data management systems, software engineering, and mobile phone regulatory requirements is essential to develop a platform to deliver and manage support programs to hundreds of participants simultaneously as in TEXT ME and TEXTMEDS trials. This research provides broad principles that may assist other researchers in developing mHealth programs.

XSL·FO
RenderX

Introduction

The use of mobile phone technologies in health care has evolved into a new field of medicine known as mobile health (mHealth) [1]. Subscription to mobile phones is ever increasing with an estimated 7.1 billion mobile subscriptions and mobile network population coverage close to 95% [2]. Technology uptake is increasing among people across all socioeconomic classes [3,4], age groups [4], and continents [5]. Texting is a common mode of efficient, cheap, and personalized means of communication [6]. There is growing evidence on the role of text message-based programs for supporting health behavior changes [7-9] and improving adherence to treatment recommendations in the management of chronic diseases [10].

Despite emerging literature on the use of text message-based interventions for health care, there are a few explicit descriptions on the development of text message program content, structure, and message management software. A health researcher naïve to software-hardware complexities may have to rely entirely on an external professional agency; this carries the risk of inability to deliver the product to specifications, within budget or provide the product for long-term use. There are additional conceivable considerations such as legal obligations and privacy and security concerns over telecommunications-based programs.
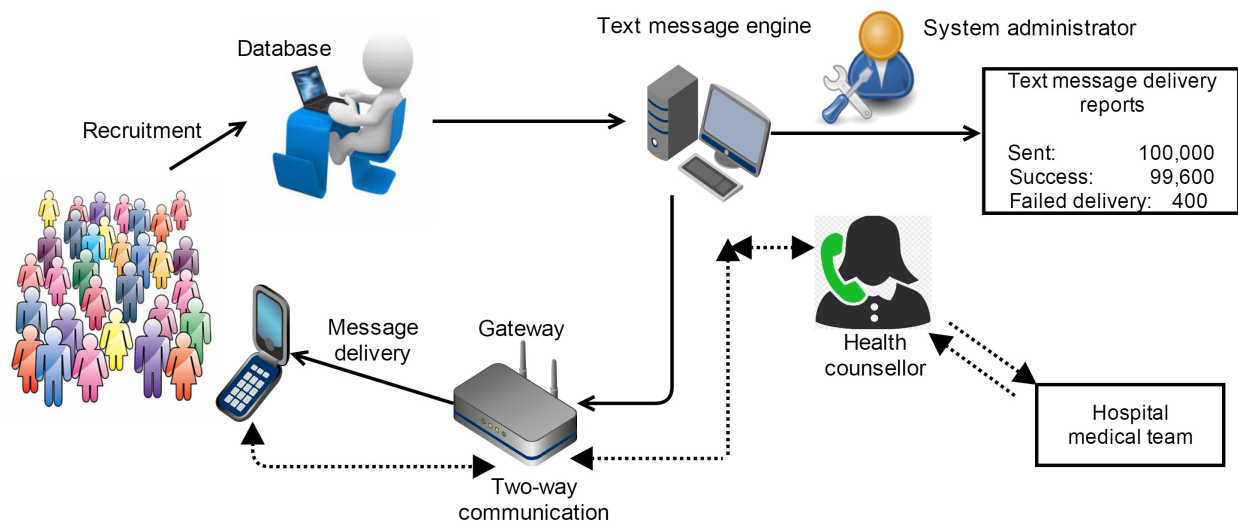
We have developed message management systems to support 2 large randomized controlled trials— Tobacco, EXercise and dieT Messages (TEXT ME; ACTRN 12611000161921) [11,12] and TEXT messages to improve MEDication adherence & Secondary prevention (TEXTMEDS; ACTRN 12613000793718) [13]. Both these trials were designed to evaluate cardiovascular disease secondary prevention support program delivered via mobile phone text messages to patients with coronary heart disease (CHD) (Table 1 and Figure 1). The aim of this paper was to outline the major practical elements that merit consideration when developing text message-based interventions. We do this by leveraging our experiences in developing the computerized message management system adopted in the TEXT ME and TEXTMEDS studies.

Table 1. Overview of TEXT ME and TEXTMEDS trials and intervention programs.

Study characteristics	TEXT ME	TEXTMEDS
Study design	Single blind[a]	Single blind[a]
	Randomized	Randomized
	Single center	Multicenter (Multitime zones)
	6 Month	12 Month
Primary focus	Behavioral change	Medication adherence
Sample size	700	1400
Message content	Lifestyle and general cardiovascular health advice	Medication adherence, lifestyle, and general cardiovascular health advice
Message delivery, time of the day	Randomly send, 10am-4pm	Randomly send, 10am-4pm
Message frequency	1 message per day; 4 random weekdays.	1 message per day; 2-4 random weekdays. Tailored frequency with fewer messages in the mid-trial compared with start and conclusion.
Program structure	Random message order. Each message is complete without reliance on previous content.	Structured delivery of information appropriate to the duration of the participant's inclusion in the study.
Two-way communication	Not encouraged. Replies were monitored for regulatory compliance	Encouraged
Additional support	Not provided	Monthly reminders to participant about availability of health counselor for additional information

[a]Due to the nature of intervention, participants could not be blinded.

Figure 1. Staff, hardware infrastructure, and interactions for the TEXT messages to improve MEDication adherence & Secondary prevention (TEXTMEDS) study.

Methods

We identified the following key stages in the developmental process for the message management system: (1) development of the program content (text message bank), (2) development of the text message engine software and integration with core database (participant data), and (3) text message send and gateway considerations. These are detailed in the following sections. In addition, we have described additional features that merit deliberation; text message identification codes, two-way communication considerations, text message reply monitoring, security-privacy, and legal requirements.

Results

Development of the Program Content (Text Message Bank)

There are a variety of approaches for developing program content; however, a systematic approach with the engagement of end-users is important. The process of development typically involves 3 phases: involving input from a range of experts and consumers, evaluation and refinement, and pilot testing. We have previously described the process of development of message content for the TEXT ME study [14]. During the first phase, a prototype bank was prepared by a multidisciplinary team incorporating various aspects, such as behavior change goals, scientific evidence and facts, and information from national health guidelines. The 160 character limit (including spaces) for a text message required careful wording and clarity of expression to avoid misunderstandings. In the second phase, the prototype bank was examined by practicing clinicians and potential consumers who reviewed each message using a survey that included questions with Likert-type responses about the readability, language appropriateness, and perceived utility of each message. The content was modified based on the survey feedback. The third phase involved pilot testing to ensure the functionality of the software, delivery of the text messages to recipients on different mobile networks, and seeking feedback on real-time experiences with respect to message frequency and timing. Following the pilot testing, further minor modifications were incorporated into the program prior to large-scale implementation of the main study. The final program structure and message content were based on the review of literature, feedback from potential participants, and pilot testing. When developing the text message program, it was essential to have a prespecified framework for how the content would be delivered and what participant information would be utilized to customize the messages. We considered a range of aspects to provide a support program, and enable future upscale of the program with minimal staff support (Table 2).

Table 2. Text message characteristics that merit consideration during the design phase.

Text message characteristic	Features	Features adapted in TEXT ME and TEXTMEDS studies
Message customization	The message content can be generic or individualized	Both studies sent messages with content that was partly generic and partly customized to participant's needs eg, smoking status, diet (vegetarian or non-vegetarian) and types of medications.
	Ability to update custom settings. (particularly relevant for a long-duration study).	TEXT ME: Customization was only at baseline. TEXTMEDS: Allowed flexibility to changes in participant's status such as behaviors or medication eg, if a participant successfully quit smoking and requested to stop smoking-related messages, this could be honored anytime during the 12-month study.
	Simple versus complex customization	TEXT ME: Relatively simple algorithms using minimal baseline data. TEXTMEDS: More complex customization using baseline data, and the ability to modify during the course of the program.
Personalization	This may enhance participant's engagement with the program	Both studies implemented this function.
Delivery timing	Needs to consider intrusiveness of messages delivered during working hours, out off hours, weekends, and public holidays	Both studies sent messages on working days during working hours. Occasional season's greetings message on holidays.
	Random times may prevent habituation.	Both studies implemented random delivery times on random weekdays.
	Specific times—delivery timed with a specific behavior eg, medication intake	We did not implement this aspect, as network latency times cannot be measured with confidence.
Frequency	Number of messages per day or per week	Both studies sent 1 message per day, average 4 messages per week.
	Fixed versus variable frequency	TEXT ME: A consistent schedule of 4 messages per week on random weekdays. TEXTMEDS: The frequency of messages varied from 2-4 per week on random weekdays.
Order of message content	Structured delivery may increase participant interest in program eg, patients hospitalized for myocardial infarction—early messages can focus on recovery and tips like use of prn nitrates; later messages can focus on healthy lifestyle, medium, and long-term goals	TEXT ME: Messages were written to stand on their own and not rely on previous messages. Hence, could be delivered in a random order. TEXTMEDS: Messages were delivered in order, enabling the ability to deliver a structured story. Messages delivered may reference a previous message. Participants received messages according to protocol-driven timing ie, on a given day each participant received a unique message, which may be different from other participants, but was appropriate to their duration on the study.
Unique messaging and repetition	Nonrepetitive messages, repetition of key messages or repetition of key message after rewording	TEXT ME had nonrepetitive messages TEXTMEDS had some repetitive key messages.
Two-way interaction	May increase patient engagement	TEXT ME was a one-way study. TEXTMEDS was two-way study. Participant replies were methodically logged and actioned by the health counselor.
Character set	Non-Latin (Unicode) characters are better avoided; alternatively, they must be tested for correct transmission by network operator and decoding by the recipient's mobile phone	Both the studies had the capacity to support Unicode but did not implement this.
Readability	Avoidance of medical jargon and abbreviations	Both studies considered 5th-8th–grade reading level [15]. Text messages were pilot tested among prospective patients who were specifically asked about readability.
Message length	Preferred length ≤160 characters (including spaces)	In both studies the messages had a character count that ranged from 120-160. Longer messages are often delivered fractionated; consequentially may result in increased message send cost and unintelligible formatting by recipient's handset.
Sender signature	Helps source recognition and distinguish the messages from spam. May be mandatory, refer to country-specific legislation	In both studies, participants were encouraged to save the study mobile number in the "contacts directory" of their mobile handsets. A full signature was included in the first message. Subsequent messages used abbreviated signature due to 160 character limit per text messages.

XSL•FO
RenderX

Text message characteristic	Features	Features adapted in TEXT ME and TEXTMEDS studies
Unsubscription	Clear instructions for unsubscription; keyword - "STOP" or "OPT OUT" or "Unsubscribe"	Both studies included unsubscribe information with the first message. It was not possible to send unsubscribe information to participants with each message due length exceeding 160 characters. Legal advice recommended a way to consent participants so that this was not required.

Considerations in Design of the Text Message Engine Software

A prerequisite to the development of any custom software is writing a "specifications" document to guide the programmers of the software. For example, in the TEXTMEDS study the key requirements identified included:

1. Automated import of data from a secure database into the text message engine software to minimize human errors from repetitive data entry.

2. Data validation of key variables, for example, mobile phone number and participant characteristics such as smoking status, diet, and medication class that guide customization.

3. Flexibility to accommodate changes in participant behavior during the course of the study.

4. The ability to receive text message replies from the participants as a summative digest in the form of a daily email to the health counselor.

5. Maintain chronological log for each participant transactions.

6. Include manual flexibility to send additional broadcast messages to all participants

7. Automated generation of daily encrypted backup files.

A precursor step for the delivery of a personalized and customized program was the necessity to assimilate participants "key characteristics" into the text message management software. This information can be entered manually into the software engine; however, this process remains vulnerable to human errors. To minimize such errors, simple measures can be implemented, for example, disabling copy-paste function and forced double-entry for the mobile phone number and other key variables that determine customization. Alternatively, the software can be configured to automatically import key data variables from a core database. With both the methods, a process of data validation is necessary and should be incorporated in the software.

Text Message Send and Gateway

In order to send messages to the participants, the text message management software generally will have to integrate with a "Gateway." In computer networking, a gateway is a nodal point that allows access to other networks; in this case, allowing Internet connected applications to access the participant's telecom company. In both TEXT ME and TEXTMEDS, we contracted external companies to provide this service. When choosing a gateway for our studies, we considered a local reputable company that offered reasonable pricing, had the capacity to provide a log of delivery and failure reports, could capture participants' replies, and forward them to our team. An important consideration is time taken by the Gateway to deliver a message, that is, latency. This can range from seconds to hours. Latency is primarily dependent on the type of network used and its ability to handle traffic during periods of heightened activity [16]. It is important to clarify this before engaging a prospective company and is especially relevant for a study that intends to deliver time-sensitive text messages, for example, timed with participant's medications.

Text Message Identification Codes

Telecom regulations require that the text messages should be sent with a unique code allocated to the sender. This confirms legitimacy and helps source identification. Some jurisdictions allow short codes (typically a 5-6 digit alpha-numeric), but most jurisdictions allow long codes or standard numeric phone numbers. The choice of the code (Table 3) is primarily dictated by the type of code available in the country, local regulations governing these codes, cost involved, and need for two-way communication. Dedicated short codes take a while for set-up, have substantial set-up fee, and often attract ongoing monthly fees. Premium services as mobile ringtone downloads, television program voting, and charity donations often use short codes. Reply messages sent to short codes by the customer are charged a higher fee and may even automatically subscribe a monthly fee to the recipient's mobile phone account. As a result, many telecom companies have a default policy to block messages originating from the short codes. This can be a major hurdle for program implementation. By comparison, long code set-up is quicker, attracts conventional text message rates, and has the ability to reach all the carriers.

XSL•FO
RenderX

Table 3. Text message sending codes.

Code	Advantages	Disadvantages
Short code	Burst sending (30-40 text messages per second)	Easy to be mistaken as spam by the participant
	Easy to remember	Participants' network provider may have a policy to block messages originating from a short code
		Limited to national borders
Long code	More likely to be acceptable and identifiable as genuine text message	Relatively slower message rate (1 message per second) depending on jurisdiction
	Long-term rental or lease can be cost effective	
	International reception capability	

Two-Way Communication Considerations

Two-way communication may be incorporated into any text message-based study. These bilateral transactions may enhance participant engagement and can be a confirmation that a participant has read the text message. Bidirectional engagement can bolster trust in the patient-provider relationship [17]. Two-way communication, however, requires additional resources: monitoring of the "send-receive" loop, tracking of the conversation (to see which message is linked with which response), personnel to oversee and respond to messages, and the economic cost of replies via text message which must be either borne by the participant or charged back to the study.

Even if a study is designed to be one-way (eg, TEXT ME), it is important to monitor replies for regulatory compliance, for example, for unsubscribe requests. TEXTMEDS, by contrast, allowed and encouraged two-way communication. The TEXTMEDS software had the additional ability to file a chronologic log of all messages sent, the reply from the participant, and a corresponding response (if this was necessary) to the reply. This coherent conversation record is essential to enable tracking of the conversation with participants. Successful implementation of two-way communication requires a dedicated health counselor to monitor all the messages. The TEXMTEDS study counselors were trained to respond in accordance with a standardized manual. The health counselor role is probably best served by an allied health professional (eg, nurse or dietitian) and may need additional support from clinicians for specialist information. We did not implement "artificial intelligence" that is, the ability to automatically update settings and respond to text message responses from participants. This was because we were unsure about the frequency of participant replies and we expected a wide variation in responses; therefore, we considered but did develop this feature. The potential though remains intriguing as algorithmic analysis and computing power is rapidly evolving. This feature could be considered as a supplement to health counselor role in the next generation text message trials.

Text Message "Send-Reply" Loop Monitoring

The breakdown in electronic communication loops is unpredictable and unfortunately not uncommon. Hence, all programs using two-way communication require some method to check the "send-receive" loop. This can be done manually or by using a virtual mobile. We utilized a second text message provider company that maintained a virtual mobile phone on our behalf. Each time a daily send occurred, we also sent a message (which we named the "heartbeat text message") to the virtual mobile phone. It then sends an automated reply back to the primary text message provider, who is expected to send an email to our server registering the reply. If that reply is not received, the text message engine generates an error message to the study team. The "heartbeat text message" results in an additional cost but provides assurance that the loop is operational. Such monitoring is essential for large-scale program implementation. Part of our monitoring processes also involved checking if the messages were delivered uniformly across all major mobile phone service providers.

Text Message Delivery Monitoring

It is important to configure text message management software to record delivery reports (success and failure) and generate an error advisory to the research staff. This is essential for monitoring the program, intervention fidelity analysis as well as early recognition of participants changing their mobile phone number. With some phone contracts, a phone number may be reallocated and messages may, therefore, be delivered to a different recipient. This may compromise privacy and confidentiality. Hence, it is essential to provide explicit instructions (eg, on a participant information sheet) at program initiation to promptly report any changes to their preferred mobile phone number and have a process in place that research staff recognizes this and execute immediate unsubscription till further contact with the participant is reestablished.

Security Considerations

Data security and privacy should be of paramount importance in developing any technology-based programs. Security concerns arise at multiple levels in a text message program. Any application or business, no matter its purpose, is susceptible to hacking [18]. The electronic database, the text message engine software, communication portals between database-engine-gateway, and the text message itself are vulnerable to security and privacy compromise. The key consideration during the designing of any software is the necessity to protect information about the patient's health, contact details, and all communication. The measures adopted in TEXT ME and TEXTMEDS studies are summarized in Table 4. A system backup was generated at regular intervals. With TEXT ME, the local computer maintained file backup of each state change in the data files. The hard drive was backed up to a local external drive after every operation (participant entry,

parse, message-send). A weekly compact disk backup was created and stored offsite. With TEXTMEDS, the files were zipped into an archive and then encrypted using SHA256 hash

function [19]. The encrypted study files were sent via email to the system administrator each day. The restore function was built into TEXTMEDS.

Table 4. Areas at risk of security compromise in a text message study and measures adapted in TEXT ME and TEXTMEDS study.

Components	Measures adapted in TEXT ME and TEXTMEDS study
Text message engine hardware	Access to room by authorized personnel only
	Unit physically locked to bench
Text message engine software	Dedicated computer for sole purpose of running text message engine
	Passwords at screen login and engine software login
	Front facing passwords are hashed and salted
	Computer is configured to prevent reboot from CD[a] or USB[b] drive
Database	Database selected with security levels required for holding potentially identifying patient data.
Communication between portals	Secure HTTP (hypertext transfer protocol) or secure sockets layer (SSL) emails
	Onsite and offsite data backup (encrypted)
Text message	Avoiding identifying information in the body of the text message
	Participants instructed to consider password protection of their mobile phone and disable message preview function

[a]CD: Compact disc.

[b]USB: Universal serial bus.

Consideration of Legal Requirements

Text message communications are subject to 2 important laws—privacy policy and spam acts. It is important to note that approval by an ethics committee does not grant legislative exemption. The messages sent during the study (directly or indirectly) have a potential to disclose individual's health information. Each country may have different legislation on privacy policy and data protection—Health Insurance Portability and Accountability ACT (HIPAA, USA) [20], Personal Information Protection and Electronic Documentation Act (PIPEDA, Canada), and Privacy Act 1988 (Australia) or Data Protection Directive (European Union). There are 2 possible approaches to mitigate this. First, restructure text messages to remove all personal health information; alternatively, retain limited personal information, but conduct a risk analysis to ensure proper security measures are executed [21].

Sending text messages in bulk is recognized as spam in most countries. Anti-spam legislation is governed in the United States by the CAN SPAM Act and Telephone Consumer Protection Act, and in Australia by the Spam Act [22,23]. Many countries have similar legislative requirements and these must be taken into consideration. Noncompliance penalties can be substantial. Texting in certain countries from certain institutions (eg, hospitals, schools, banks) may sometimes be categorized as transactional rather than commercial and awarded a special exemption status [24]. However, appropriate legal consultation where available should be solicited. Three key elements were identified to avoid legal consequences in our jurisdiction and are universally applicable. First, opt-in consent, that is, participants should provide written and explicit consent to receive messages corresponding to the trial duration. Second, each text message communication had to clearly identify the

organization that authorized message send, that is, a unique signature for example, TEXT ME study used "TEXTME" and TEXTMEDS study used "TXTMED-abbreviated hospital name" as a sender signature. Third, opt-out capacity, that is, there had to be a clear method for unsubscription and this must be honored as soon as possible within legally acceptable time frame. It is important to note that as a standard text message length is of 160 characters length, including unsubscription information as well as health information within 160 characters was generally not possible. It was intended that this be highlighted to the participant at program initiation and included as a clause on the consent form as well as participant information sheet; thus, allowing us to avoid the necessity for explicit unsubscribe information with each text message.

Discussion

Principal Findings

This paper describes the key aspects that merit considerations in the design of an mHealth project. We have leveraged our experiences acquired during the development of the TEXT ME and TEXTMEDS text message management systems to provide a framework for other researchers planning similar projects. Automated message management systems can enable scalability, but should be developed and utilized in trial settings to obtain important information on their usability, reliability, and ability to deliver the intervention as intended.

To initiate the build of an mHealth intervention, there are 2 broad choices: (1) vendor-based solution or (2) In-house solution [25]. A vendor-hosted solution is usually a company (generally for profit) that offers Web-based solutions to their clients where the vendor essentially performs all functions other than writing content (eg, in the case of texting intervention—the messages)

and pushing the "send" button. The advantages of this approach may include the relevant experience of the vendor, lack of requirements of up-front investment in software development, and efficient intervention delivery. Disadvantages may include high overall cost primarily determined by the extent of customization. It is important to determine the vendors track-record, service quality, privacy policy and policies on data sharing, and degree of security implemented by the vendor to protect participant's personal information [26]. The alternative approach is an in-house solution. This allows maximum customization possibilities and desired functionality. The disadvantages of this approach include the need for technical expertise, large up-front cost in software development, lag time (software development to program initiation), and ongoing maintenance costs. While assessing the cost and benefits, researchers should also factor in—the duration of the study, size of study population, volume of text message exchange, and if there are plans to run parallel projects using the same software. The text message engine running the TEXT ME and TEXTMEDS studies were in-house designed, primarily because no appropriate vendor system existed (to our knowledge) that offered the level of customization we required. The team running these projects comprised of highly motivated researchers and academics with clinical and technical knowledge, which facilitated smooth delivery while keeping the overheads low. We also desired to deliver multiple similar trials varying by time zone, locations, and message customization [27].

Although this paper details the considerations and methodology relevant to developing a text message-based health intervention, there is a significant overlap with other mHealth interventions such as mobile device apps. The advantage of text message delivery include universal compatibility with all mobile handsets, requires minimal technological skills from participants and is a "push" technology (ie, it is delivered to the participant until they opt out). App-based interventions could be used to deliver a similar program with the potential advantages of lower delivery cost and the ability to deliver more graphical content. One potential advantage of app-based messaging is the ability to transmit highly secure information with end-to-end encryption. An app can also be designed to provide instantaneous automated responses to participant inputs. They also carry higher functionality, support interactivity, and bilateral engagement. The major disadvantages of app-based messaging approach are increased development cost, complexity, and compatibility issues that is, need for mobile phones or equivalent devices and Internet connectivity. This may be a limiting factor when the target audience comprises elderly or socioeconomically disadvantaged population. In addition, app-based interventions are more dependent on the user continuing to choose to use the app, whereas text message would continue to be delivered to the participant until they make an active choice to opt out of the program.

Conclusions

Highly customized multifaceted interventions can be delivered to large patient populations using an automated text message engine. There is a need for further development of customized and structured text message programs that combines patient needs with the potential for large-scale delivery. Researchers running information technology projects may have constitutional obligations and must ensure reasonable steps for secure electronic communication.

References

1. World Health Organization. mHealth - New horizons for health through mobile technologies. 2011. URL: http://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2016-10-29] [WebCite Cache ID 6ldCXD30U]
2. International Telecommunications Union. Measuring the information society report. Geneva; 2015 URL: http://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2015/MISR2015-w5.pdf [WebCite Cache ID 6hYEdysxa]

XSL•FO

RenderX

3. Koivusilta LK, Lintonen TP, Rimpela AH. Orientations in adolescent use of information and communication technology: a digital divide by sociodemographic background, educational career, and health. Scand J Public Health 2007;35(1):95-103. [Medline: 17366093]

4. Atun R, Sittampalan S. The role of mobile phone in increasing accessibility and efficiency in health care. Vodafone. URL: http://www.vodafone.com/content/dam/vodafone/about/public_policy/policy_papers/public_policy_series_4.pdf [accessed 2016-05-16] [WebCite Cache ID 6hYFbEkYC]

5. Goggin G. Cell Phone Culture: Mobile Technology in everyday life. NewYork, NY: Taylor and Francis group; 2006.

6. Shaw RJ, Bosworth HB, Hess JC, Silva SG, Lipkus IM, Davis LL, et al. Development of a theoretically driven mHealth text messaging application for sustaining recent weight loss. JMIR Mhealth Uhealth 2013;1(1):e5 [FREE Full text] [doi: 10.2196/mhealth.2343] [Medline: 25100678]

7. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010;32:56-69 [FREE Full text] [doi: 10.1093/epirev/mxq004] [Medline: 20354039]

8. Kannisto KA, Koivunen MH, Välimäki MA. Use of mobile phone text message reminders in health care services: a narrative literature review. J Med Internet Res 2014;16(10):e222 [FREE Full text] [doi: 10.2196/jmir.3442] [Medline: 25326646]

9. Wei J, Hollin I, Kachnowski S. A review of the use of mobile phone text messaging in clinical and healthy behaviour interventions. J Telemed Telecare 2011;17(1):41-48. [doi: 10.1258/jtt.2010.100322] [Medline: 21097565]

10. Thakkar J, Kurup R, Laba TL, Santo K, Thiagalingam A, Rodgers A, et al. Mobile telephone text messaging for medication adherence in chronic disease: a meta-analysis. JAMA Intern Med 2016 Mar;176(3):340-349. [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]

11. Chow CK, Redfern J, Thiagalingam A, Jan S, Whittaker R, Hackett M, et al. Design and rationale of the tobacco, exercise and diet messages (TEXT ME) trial of a text message-based intervention for ongoing prevention of cardiovascular disease in people with coronary disease: a randomised controlled trial protocol. BMJ Open 2012;2(1):e000606 [FREE Full text] [doi: 10.1136/bmjopen-2011-000606] [Medline: 22267690]

12. Chow CK, Redfern J, Hillis GS, Thakkar J, Santo K, Hackett ML, et al. Effect of lifestyle-focused text messaging on risk factor modification in patients with coronary heart disease: a randomized clinical trial. JAMA 2015;314(12):1255-1263. [doi: 10.1001/jama.2015.10945] [Medline: 26393848]

13. ANZ Clinical Trials Register. Trial review. 2013 Jul 16. URL: https://www.anzctr.org.au/Trial/Registration/TrialReview. aspx?id=364448 [accessed 2016-10-29] [WebCite Cache ID 6ldHVZuvw]

14. Redfern J, Thiagalingam A, Jan S, Whittaker R, Hackett ML, Mooney J, et al. Development of a set of mobile phone text messages designed for prevention of recurrent cardiovascular events. Eur J Prev Cardiol 2014 Apr;21(4):492-499. [doi: 10.1177/2047487312449416] [Medline: 22605787]

15. Safeer RS, Keenan J. Health literacy: the gap between physicians and patients. Am Fam Physician 2005 Aug 1;72(3):463-468 [FREE Full text] [Medline: 16100861]

16. van der Ploeg HP, Chey T, Korda RJ, Banks E, Bauman A. Sitting time and all-cause mortality risk in 222 497 Australian adults. Arch Intern Med 2012 Mar 26;172(6):494-500. [doi: 10.1001/archinternmed.2011.2174] [Medline: 22450936]

17. Wald DS, Butt S, Bestwick JP. One-way versus two-way text messaging on improving medication adherence: meta-analysis of randomized trials. Am J Med 2015 Oct;128(10):1139.e1-1139.e5. [doi: 10.1016/j.amjmed.2015.05.035] [Medline: 26087045]

18. Symentec. ISTR: Internet security threat report, 2013. 2013 Apr. URL: http://www.symantec.com/content/en/us/enterprise/ other_resources/b-istr_main_report_v18_2012_21291018.en-us.pdf [WebCite Cache ID 6hYgiwqcZ]

19. Adriancs, Hornby T. Salted password hashing-doing it right. Codeproject. 2016 Feb 14. URL: http://www.codeproject.com/ Articles/704865/Salted-Password-Hashing-Doing-it-Right [accessed 2016-05-16] [WebCite Cache ID 6hYpBOVBj]

20. Health Insurance Portablity and Accountablity Act of 1996. U.S.Department of Health and Human Services. Human Services URL: http://www.hhs.gov/hipaa/for-professionals/index.html [accessed 2016-05-17] [WebCite Cache ID 6hYkVE23M]

21. Karasz HN, Eiden A, Bogan S. Text messaging to communicate with public health audiences: how the HIPAA Security Rule affects practice. Am J Public Health 2013 Apr;103(4):617-622. [doi: 10.2105/AJPH.2012.300999] [Medline: 23409902]

22. Federal Register of Legislation. Spam Act 2003, Amended. 2014 Jun 24. URL: https://www.legislation.gov.au/Details/ C2014C00214 [accessed 2016-05-16] [WebCite Cache ID 6hYh2CMD4]

23. Federal Communications Commission. Unwanted commercial electronic mail URL: https://www.fcc.gov/general/can-spam [accessed 2016-05-16] [WebCite Cache ID 6hYhtlJVH]

24. Nccptrai. The Telecom commercial communications customer preference regulations, 2010 (6 of 2010). New Delhi; 2010 Dec 01. URL: http://www.nccptrai.gov.in/nccpregistry/regulation1diccndiv.pdf [accessed 2016-10-29] [WebCite Cache ID 6ldGFeLHN]

25. : School of Public Health Karasz H, Bogan S. Investing in text messaging system: a comparison of three solutions. Nwcphp. University of Washington: School of Public Health; 2012. URL: http://www.nwcphp.org/docs/sms-toolkit/overview/ TextingSystemInvestment.pdf [WebCite Cache ID 6ldGQ1Mrv]

26. Center for Research in Implementation Science and Prevention (CRISP). Text messaging in health care: research toolkit. CRISP, University of Colorado School of Medicine; 2013 Oct. URL: http://www.ucdenver.edu/academics/colleges/

medicalschool/programs/crisp/training/toolkits/textingtoolkit/Documents/
Text%20Messaging%20in%20Healthcare%20Research%20Toolkit%202.pdf [WebCite Cache ID 6ldGlq5bo]

27. Thakkar J, Karthikeyan G, Purohit G, Thakkar S, Sharma J, Verma S, et al. Development of macaronic Hindi-English 'Hinglish' text message content for a coronary heart disease secondary prevention programme. Heart Asia 2016;8(2):32-38. [doi: 10.1136/heartasia-2016-010789] [Medline: 27752288]

Abbreviations

CHD: coronary heart disease
HIPAA: Health Insurance Portability and Accountability ACT
HTTP: hypertext transfer protocol
mhealth: mobile health
PIPEDA: Personal Information Protection and Electronic Documentation Act
SSL: secure sockets layer
TEXT ME: Tobacco, EXercise and dieT Messages
TEXTMEDS: TEXT messages to improve MEDication adherence & Secondary prevention

Original Paper

Design and Feasibility of a Text Messaging Intervention to Prevent Indoor Tanning Among Young Adult Women: A Pilot Study

William D Evans[1], PhD; Darren Mays[2], PhD

[1]Department of Prevention and Community Health, Milken Institute School of Public Health, The George Washington University, Washington, DC, United States

[2]Lombardi Comprehensive Cancer Center, Georgetown University Medical Center, Georgetown University, Washington, DC, United States

Corresponding Author:
William D Evans, PhD
Department of Prevention and Community Health
Milken Institute School of Public Health
The George Washington University
950 New Hampshire Avenue
Washington, DC, 20052
United States
Phone: 1 202 994 3632
Fax: 1 202 994 3601
Email: wdevans@gwu.edu

Abstract

Background: Although skin cancer is largely preventable, it affects nearly 1 of 5 US adults. There is a need for research on how to optimally design persuasive public health indoor tanning prevention messages.

Objective: The objective of our study was to examine whether framed messages on indoor tanning behavioral intentions delivered through short message service (SMS) text messaging would produce (1) positive responses to the messages, including message receptivity and emotional response; (2) indoor tanning efficacy beliefs, including response efficacy and self-efficacy; and (3) indoor tanning risk beliefs.

Methods: We conducted a pilot study of indoor tanning prevention messages delivered via mobile phone text messaging in a sample of 21 young adult women who indoor tan. Participants completed baseline measures, were randomly assigned to receive gain-, loss-, or balanced-framed text messages, and completed postexposure outcome measures on indoor tanning cognitions and behaviors. Participants received daily mobile phone indoor tanning prevention text messages for 1 week and completed the same postexposure measures as at baseline.

Results: Over the 1-week period there were trends or significant changes after receipt of the text messages, including increased perceived susceptibility ($P<.001$), response efficacy beliefs ($P<.001$), and message receptivity ($P=.03$). Ordinary least squares stepwise linear regression models showed an effect of text message exposure on self-efficacy to quit indoor tanning ($t_6=-2.475$, $P<.02$). Ordinary least squares linear regression including all measured scales showed a marginal effect of SMS texts on self-efficacy ($t_{20}=1.905$, $P=.08$). Participants endorsed highly favorable views toward the text messaging protocol.

Conclusions: This study supports this use of mobile text messaging as an indoor tanning prevention strategy. Given the nature of skin cancer risk perceptions, the addition of multimedia messaging service is another area of potential innovation for disseminating indoor tanning prevention messages.

Introduction

Although skin cancer is largely preventable by reducing ultraviolet radiation exposure, the incidence has increased [1,2], affecting nearly 1 of 5 US adults and incurring significant costs to society [3-5]. Indoor tanning increases the risks of skin cancer [6,7], causing an estimated 10% of all cases [8]. Among US adults, indoor tanning is most prevalent among white (Hispanic

XSL•FO
RenderX

and non-Hispanic) young adult women aged 18 to 30 years, with up to 30% of this group tanning each year [9-11]. Indoor tanning among young women further increases skin cancer risks [6,7,12,13] and leads to early-onset disease [14].

The 2014 *Surgeon General's Call to Action to Prevent Skin Cancer* called for research on how to optimally design persuasive public health indoor tanning prevention messages targeting young women [5]. In other areas of cancer prevention and control (eg, tobacco control), public health communication messaging is a recommended best practice for preventing and reducing behavioral risk factors for cancer [15,16]. Such messaging approaches are designed for wide reach and impact because they can be delivered through channels such as the Internet, mobile phones, and public health campaigns [15,16]. In settings such as Australia, public health messaging campaigns have helped to curb a growing skin cancer epidemic [17]. However, research conducted to date on skin cancer prevention messaging has focused primarily on sun safety behaviors (eg, sunscreen use) [18,19]. The available evidence on indoor tanning behavior among young women suggests different motives may drive intentional indoor tanning behaviors (eg, improving physical appearance) and override perceived short- and long-term risks [5,20-22] (also DM and WDE, unpublished data, 2016).

The extended parallel process model provides a theoretical perspective for how to design and frame persuasive indoor tanning prevention messages [23]. However, evidence of the effects of persuasive public health messages that attempt to frame the potential benefits (gain) and consequences (loss) of skin cancer preventive behaviors is mixed. Some studies favor gain-framed messages [24] emphasizing the benefits of behaviors such as sun protection, while some favor loss-framed messages conveying potential risks [25]. Other studies show no distinct advantage of either gain- or loss-framed messages [26]. Meta-analyses generally reflect these mixed results [27], with no clear advantage of either message frame emerging for sun protection behaviors. However, prior messaging studies have focused primarily on sun protection behaviors (eg, seeking shade, using sunscreen), and research investigating the effects of persuasive, framed messages for preventing indoor tanning is scarce [28,29].

Research on how to craft effective indoor tanning prevention messaging in such a way that addresses the unique motives of this behavior among young women is needed for national skin cancer prevention efforts. Research is needed not only on the *content* and *framing* of persuasive indoor tanning prevention messaging, but also on the optimal *delivery modality*.

Mobile devices are virtually ubiquitous among US young adults: 85% of US young adults own a smartphone with multimedia and Internet capabilities, and virtually all of these young adults use their device to some extent for short message service (SMS) text messaging [30]. Mobile devices are also a popular medium for delivering behavior change interventions, medication reminders, treatment information, and adherence tools targeted to improve health outcomes [31]. Research in this area indicates mobile text messaging interventions are effective for promoting healthy behavior change, including weight loss [32], nutrition

and physical activity [33,34], smoking cessation [31,35,36], and prenatal behaviors among pregnant women [37,38].

These examples in the literature, and the widespread use of mobile phone text messaging among young adults, provide support for mobile phone text messaging as a medium for delivering persuasive messaging to prevent indoor tanning among young women. However, few previous studies in this area have harnessed the potential of optimally framed messages delivered via text message.

We designed this study to build from previous research and inform indoor tanning prevention efforts targeting young adult women by examining the effects of gain-, loss-, and balanced-framed text messages on indoor tanning behavioral intentions. The overall hypothesis was that delivering indoor tanning prevention messages both on the Web and via mobile phone text messages to young adult women who indoor tan would be feasible and that procedures would be acceptable to study participants. Specifically, this study explored 3 research questions (RQs): (1) Do participants respond positively to the text messages, as measured by message receptivity and emotional response? (2) Is receipt of text messages associated with changes in indoor tanning efficacy beliefs, including response efficacy (ie, perceived benefits of avoiding indoor tanning) and self-efficacy to avoid tanning? (3) Is receipt of text messages associated with changes in indoor tanning risk beliefs, including perceived severity and perceived personal susceptibility to the risks? We examined these questions in a pilot study to evaluate the feasibility of using SMS texts as a delivery modality for indoor tanning prevention messages. Additionally, we collected qualitative feedback from participants on their perceptions of the indoor tanning prevention messages. We report on these data to help interpret our quantitative pilot results.

Methods

Protocol

We conducted a pilot study to test the feasibility and acceptability of delivering persuasive gain-, loss-, and balanced-framed messages on the Web and via mobile phone text messaging in a sample of 21 young adult women who indoor tan (mean age 24.9 years, SD 2.9; 9/21, 43% frequent tanners). The pilot drew from our previous experience in similar studies of persuasive messaging for cancer prevention and control, and other health behavior domains [39,40] (also DM and WDE, unpublished data, 2016). Study inclusion criteria were as follows: (1) white females 18-30 years of age, (2) having had 1 or more indoor tanning exposures in the past 12 months, (3) willing to send and receive text messages via a personal mobile phone, and (4) able to complete all study assessments and procedures in English.

Potential participants were young adult women who had participated previously in an observational study on indoor tanning behavior [41] and met study eligibility criteria. Participants were initially contacted by email with a brief description of study procedures and an invitation to participate, and those who contacted study personnel expressing interest

were rescreened for eligibility for the study by telephone. Eligible participants then received informed consent forms, and enrollment was complete once informed consent forms were signed and returned to study personnel. Nearly all participants screened for the pilot (20/21, 95%) were eligible, consented to participate, and completed study procedures.

Eligible, consenting participants completed baseline measures through a custom website and at the conclusion of the baseline assessment were randomly assigned to receive gain-, loss-, or balanced-framed text messages. Participants were randomly assigned in a 1:1:1 ratio to these conditions using an algorithm embedded within the Web-based survey software (Qualtrics Research Suite, Qualtrics LLC). Participants then completed immediate postexposure measures of behavioral intentions, indoor tanning risk (perceived severity, susceptibility) and efficacy (self-efficacy, response efficacy) beliefs, and message response (emotional response, message receptivity). Then for a 1-week period participants received daily mobile phone indoor tanning prevention text messages. At study enrollment, we

assessed participants' preferred time of message delivery relative to the typical time of day they indoor tan (morning, afternoon, evening) and scheduled message delivery accordingly. Participants responded to an initial message reporting whether they indoor tanned that day. We tailored 2 subsequent messages to their reported tanning behavior each day, as the examples in Table 1 show. The initial message delivered fact-based information on indoor tanning risks and the benefits of avoiding indoor tanning. The second message served as a prompt to encourage behavior change (avoiding indoor tanning) or to maintain indoor tanning avoidance [42]. The tailoring algorithm and delivery schedule was preprogrammed into a commercially available messaging software system (TextIt, Nyaruka Ltd). After receiving messages for 1 week, participants completed a Web-based follow-up with measures similar to the baseline assessment.

All participants provided written informed consent, and study procedures were reviewed and approved by the Georgetown University Institutional Review Board.

Table 1. Examples of initial (SMS1[a]) and follow-up prompt (SMS2) indoor tanning prevention text messages for day 1 (of 7 days), by message framing condition (gain-, loss-, or balanced-framed), based on whether a participant answered "yes" or "no" to whether they had indoor tanned that day.

Text message	Text message condition					
	Gain		Loss		Balanced	
	Yes, tanned	No, did not tan	Yes, tanned	No, did not tan	Yes, tanned	No, did not tan
SMS1, 160-character maximum	Every time you tan, you're damaging the DNA in your skin cells. Keep your skin safe and healthy. Avoid indoor tanning to prevent skin damage.	Every time you tan, you're damaging the DNA in your skin cells. Keep your skin safe and healthy. Avoid indoor tanning to prevent skin damage.	Every time you tan, you're damaging the DNA in your skin cells. Indoor tanning causes skin damage.	Every time you tan, you're damaging the DNA in your skin cells. Indoor tanning causes skin damage.	Every time you tan, you're damaging the DNA in your skin cells. Avoid indoor tanning to prevent skin damage.	Every time you tan, you're damaging the DNA in your skin cells. Avoid indoor tanning to prevent skin damage.
SMS2	Quitting indoor tanning is easy! Quitting indoor tanning now will keep your skin safe and healthy.	Avoiding indoor tanning is easy! Keep avoiding indoor tanning so your skin stays safe and healthy.	Quit indoor tanning now to avoid skin damage.	Avoid indoor tanning so you don't damage your skin.	Quitting indoor tanning is easy! You can quit indoor tanning now to keep your skin safe and healthy.	Avoiding indoor tanning is easy! Keep avoiding indoor tanning so your skin stays safe and healthy.

[a]SMS: short message service.

Preexposure Measures

Prior to any message exposure, we captured demographic characteristics (age, household income, whether participants were current students) and past year indoor tanning behavior using validated items from epidemiological surveys [10]. To characterize the sample, we operationalized frequent indoor tanning based on a binary variable indicating indoor tanning 10 or more times in the past year [10].

Text Messages

We developed a 7-day program of text messages to be delivered to participants after we collected baseline data. Messages were designed based on the lead author's previous studies on cancer prevention messages and SMS for young women [38,39]. We developed message content based on indoor tanning prevention messages we previously tested [28,29], skin cancer prevention text messaging research [42], and research on indoor tanning beliefs, motives, and health risks [5,14,43]. Content was tailored

to participants' reported indoor tanning behavior each day and was framed based on the conditions to which participants were randomly assigned. Table 1 provides examples of the gain, loss, and balanced text messages delivered on day 1 of the pilot. Note that we first sent an initial message to get the participant's attention, followed by a prompt message framed as gain, loss, or balanced.

Postexposure Measures

We administered the following measures immediately after participants were exposed to the Web-based messages at baseline, and at the conclusion of the 1-week text message exposure period to capture participants' responses to the indoor tanning prevention messages.

Emotional Response

Emotional response to the messages was measured with 3 items from prior research assessing whether participants felt frightened, anxious, or nervous while reading the message [44].

Responses were based on a 4-point scale (1=not at all, 4=extremely) and were averaged to create a score, with higher values indicating stronger fear responses (Cronbach alpha=.89 at both time points).

Message Receptivity

Receptivity to the messages was measured using an adapted 7-item scale [28]. Examples of items are "The message was convincing," "The message said something important to me," and "The message gave me a good reason not to tan indoors." Participants responded to the statements on a 7-point scale (1=strongly disagree, 7=strongly agree), and responses were averaged to create a score, with higher values indicating greater message receptivity (Cronbach alpha range .79-.82).

Risk Beliefs

Perceived severity of the risks of skin cancer was measured using 5 items adapted from previous research [45] with a 5-point response scale (1=strongly disagree, 5=strongly agree). The items were averaged to create a score, with higher values indicating greater perceived severity (Cronbach alpha range .70-.73). Perceived susceptibility to skin cancer was measured using 6 items adapted from a previous study [45]. Responses were based on a 5-point scale (1=strongly disagree, 5=strongly agree) and were averaged to create a score, with higher values indicating greater perceived susceptibility (Cronbach alpha range .80-.81).

Efficacy Beliefs

Response efficacy was assessed using a 7-item scale adapted from other cancer prevention risk behavior research to capture the perceived health benefits of avoiding indoor tanning (eg, reducing risks of skin cancer) [28,37]. Responses were based on a 9-point scale (1=no chance, 7=certain to happen) and averaged to create a summery score, with higher values indicating stronger perceived response efficacy (Cronbach alpha range .81-87). Self-efficacy was measured using 2 items assessing how confident participants were and how easy it would be for participants to quit indoor tanning in the next year [28]. Responses were based on a 7-point scale (1=not at all, 7=extremely) and averaged to create a summary score, with higher values indicating greater self-efficacy (Cronbach alpha range .75-.85).

Indoor Tanning Behavioral Intentions

Similar to previous studies [28,29], the primary outcome measured following exposure to the messages at baseline and the 1-week posttest was indoor tanning behavioral intentions. We chose this outcome because behavioral intentions have been demonstrated in previous research to predict future health behavior change [46] and could be assessed as a potential indicator of future behavior change in a pilot study with limited follow-up duration. For this study, we captured behavioral intentions to indoor tan using 2 items assessing intentions to tan even once and intentions to tan regularly in the next year on a 7-point scale (1=definitely will not, 7=definitely will) [28]. Intentions to quit indoor tanning were measured using a single item assessing how much the message made participants want to avoid indoor tanning in the next year on a 7-point response scale (1=not at all, 7=a lot) [28,29]. These items were

moderately correlated (r range .51-.76) and had good internal consistency when we reverse coded the intentions-to-quit item and we considered items as a single behavioral intentions construct (Cronbach alpha range .73-.84). We averaged the items to create a summary variable, where higher values indicated stronger behavioral intentions to indoor tan.

At the 1-week posttest, we also administered items assessing the acceptability of the study procedures and willingness to participate in such studies in the future [47]. Specifically, these measures captured whether participants found completing study procedures to be easy, if they encountered any challenges, and how long they would be willing to participate in such a study in the future.

Data Analysis

All data analyses were performed with IBM SPSS version 19 (IBM Corporation). We examined the variables of interest descriptively to characterize the sample, and used bivariate analyses to test for differences across the experimental conditions. No participant characteristics differed by study condition; therefore, we did not adjust analyses for covariates. To examine changes over the 1-week exposure period in variables of interest, we used paired t tests to compare means at baseline and 1 week posttest. To evaluate the effects of the texts on the dependent measure of intentions to indoor tan, and due to the small sample size, we estimated 2 ordinary least squares (OLS) linear regression models. First, we used a stepwise procedure [48]. In this OLS analysis, the posttest measures of risk and efficacy beliefs and message response were regressed onto the dependent variable of indoor tanning behavioral intentions at posttest. Second, we estimated an OLS model using all posttest measures regressed onto the indoor tanning intentions variable at posttest. As noted above, in all analyses the variable for indoor tanning behavioral intentions was an average of 3 items, with higher values indicating stronger intentions to tan on a 1 to 7 scale.

Results

Descriptive Results

We achieved 100% (21/21) compliance with the daily text messaging protocol, with all participants sending and receiving daily text messages. Table 2 shows a comparison of measures administered after exposure to Web-based messages at baseline and after the text messaging exposure at 1 week. Over this brief period, we observed trends or significant changes in the variables of interest after receipt of the tailored text messages for 1 week, including increased perceived susceptibility (P<.001), response efficacy beliefs (P<.001), and message receptivity (P=.03).

Multivariate Regression Models

Next, we analyzed the data using an OLS linear regression with a stepwise algorithm. One indoor tanning belief scale, self-efficacy to quit indoor tanning, emerged as statistically significant and was positively correlated with text message exposure at follow-up. Table 3 and Table 4 summarize these results.

Table 2. Comparison of measures at baseline and 1 week (N=21).

Measures	Baseline	After 1 week	P value
	Mean (SD)	Mean (SD)	
Perceived susceptibility	3.38 (0.844)	3.84 0 (.746)	<.001
Perceived severity	3.89 (0.736)	4.06 (0.608)	.24
Self-efficacy	5.43 (1.39)	5.45 (1.52)	.64
Response efficacy	4.98 (0.743)	5.65 (0.881)	<.001
Emotional response	2.11 (0.845)	1.98 (0.792)	.31
Message receptivity	5.46 (1.04)	5.80 (0.812)	.03
indoor tanning behavioral intentions	3.30 (1.34)	2.92 (1.39)	.11

Table 3. Ordinary least squares stepwise linear regression of indoor tanning behavioral intentions on beliefs.

Model	B	SE	Beta	t [a]	P value
Constant	5.382	1.031		5.222	<.001
Self-efficacy to quit indoor tanning	−.451	0.182	−.494	−2.475	.02

[a]df=6 for the regression; df=14 for the residual; df=20 for the total model.

Table 4. Excluded variables in stepwise procedure.

Variables	Collinearity statistics				
	Beta	t [a]	P value	Partial correlation	Tolerance
Risk beliefs	−.147	−0.725	.48	−0.168	0.991
Perceived severity	−.158	−0.762	.46	−0.177	0.944
Response efficacy	−.052	−0.232	.82	−0.055	0.828
Emotional response	−.272	−1.343	.20	−0.302	0.93
Message receptivity	−.344	−1.777	.09	−0.386	0.957

[a]df=6 for the regression; df=14 for the residual; df=20 for the total model.

Table 5. Ordinary least squares linear regression with all variables included (no stepwise procedure).

Variables	Unstandardized coefficients			Standardized coefficients	
	B	SE	Beta	t [a]	P value
(Constant)	8.423	2.641		3.1899	.007
Risk beliefs	0.109	0.517	.0599	0.2111	.84
Perceived severity	0.051	0.664	.0222	0.0777	.94
Self-efficacy	−0.432	0.227	−.472	1.905	.08
Response efficacy	−0.018	0.492	−.012	−0.037	.97
Emotional response	−0.316	0.485	−0.181	−0.651	.53
Message receptivity	−0.525	0.458	−0.308	−1.146	.27

[a]df=6 for the regression; df=14 for the residual; df=20 for the total model.

Next we ran the OLS linear regression without the stepwise algorithm, including all indoor tanning belief and message response predictor variables, summarized in Table 5. In this model, self-efficacy was no longer significantly associated with indoor tanning behavioral intentions at P>.05, but this result was near significance at P=.08.

Text Message Acceptability

Finally, we asked participants follow-up questions regarding the text messaging and their participation in the pilot study. In response to these questions, respondents all indicated that it was easy to participate in the pilot (21/21) and that they did not find it challenging or difficult (21/21); 81% of participants (17/21) indicated that they would be willing to participate in a future

XSL•FO

RenderX

text messaging study related to indoor tanning for a period of 4 weeks. These results generally indicate the intervention was well received and feasible for participants, suggesting the opportunity to conduct additional text messaging for indoor tanning prevention studies in the future.

Discussion

Principal Findings

Overall, regarding RQ1, we found that respondents were receptive and responded positively to the text messages. Our descriptive results showed higher receptivity using the validated message receptivity scale at 1-week follow-up, based on our descriptive results. For RQ2, descriptive results showed both that self-efficacy (to avoid indoor tanning) and response efficacy (perceiving that avoiding tanning will be beneficial) increased after text message exposure. Results from the OLS models were mixed, with self-efficacy increasing in the stepwise procedure, but not reaching significance in the full model. Our limited sample size and the pilot nature of the research should be considered in interpreting these findings. Finally, for RQ3, descriptive results showed that perceived severity and perceived personal susceptibility both increased after exposure to the text. However, the OLS models revealed no effects on these risk beliefs.

Regarding our overall goal to assess feasibility of using SMS for indoor tanning prevention, we found that text messaging was an acceptable approach to delivering indoor tanning prevention messages. We achieved total compliance with the intervention in the pilot sample, and participants expressed high levels of satisfaction and ease of use. These results are consistent with previous SMS interventions but are novel results for the subject matter of tanning prevention [34,49]. Self-efficacy appears to be an important belief to promote, as it was associated with higher intentions to avoid future indoor tanning in our sample. These results are novel and suggest that SMS is a promising approach for preventing indoor tanning.

Text messaging is a proven intervention strategy in several domains of public health [49]. While text messaging has long been used for treatment adherence and as a reminder system, there is now solid evidence of texting for behavior change, especially in smoking cessation and antiretroviral therapy promotion [31], and it is growing in popularity in other domains [50]. Texting interventions have been shown to be effective in changing specific targeted behaviors through persuasive messaging [38].

However, there has been virtually no previous work on text messaging for indoor tanning prevention [51]. Given the widespread use of mobile phone text messaging among young adults, using mobile phone text messaging as a medium for delivering persuasive messaging to prevent indoor tanning among young women, the population subgroup where indoor tanning is most prevalent, is especially promising. This study provides initial evidence that indoor tanning prevention text messaging is feasible, is accepted by the at-risk target population of young women, and shows short-term effects of exposure to texting on indoor tanning behavioral intentions and other indoor

tanning belief measures. This suggests that text messaging can potentially be implemented in a longer intervention period in the future.

Future research on indoor tanning messaging should include more measurement and analysis of dosage, and other means of optimizing message delivery. For example, studies have examined point-of-decision prompts to increase exercise and nutrition [52] and use of mobile technologies for health interventions [53], but, to our knowledge, no study has combined both. Text messaging interventions can do much more than simply deliver text reminders—they can deliver, right into the hands of a highly targeted population, the public health messages that in years past would have appeared in mass media [54]. There is potential to extend these areas of future research to indoor tanning messaging.

Future research should consider the nature of skin cancer risk perceptions and explore the addition of multimedia messaging service (MMS) delivery. MMS is another area of potential innovation that may enhance the persuasive appeal of indoor tanning prevention messages. The full capabilities of mobile technology using graphic imagery, video, and other multimedia via MMS to deliver persuasive indoor tanning prevention messages along with text-based content should be explored. MMS also allows for interactive communications, content tailoring, and embedding of photos and images into text messages delivered to mobile devices. These features appear promising for indoor tanning prevention messaging given the importance of skin appearance, especially to priority populations at risk from indoor tanning exposure, such as young adult women.

Study Limitations

Our results should be interpreted in light of two important limitations. First, this was a pilot study with a small convenience sample, and results were evaluated over a short period of 1 week of text messaging exposure. Thus, we did not evaluate the long-term effects of texting exposure and known variables that have affected results of previous mHealth trials, such as dosage, potential "wear-out" effects of long-term message exposure, timing and sequencing, and message content. Second, we did not examine sociodemographic and other subgroup differences due to the small sample size. Some of these variables, such as age and race/ethnicity, have been found to affect results of previous indoor tanning prevention interventions [28,29]. These factors should be part of future studies with larger samples conducted over extended time periods.

Conclusions

Overall, this study provides preliminary evidence for the effectiveness of SMS in promoting indoor tanning prevention beliefs and for increasing risk beliefs concerning the health consequence of tanning. It also shows that an SMS intervention is acceptable and may be a feasible health communication channel for indoor tanning prevention. Based on our results, text messaging appears to be a way in which persuasive messages may be tailored to the at-risk population of young adult women and to an optimal time for delivery [55].

Conflicts of Interest

None declared.

References

1. Jemal A, Saraiya M, Patel P, Cherala SS, Barnholtz-Sloan J, Kim J, et al. Recent trends in cutaneous melanoma incidence and death rates in the United States, 1992-2006. J Am Acad Dermatol 2011 Nov;65(5 Suppl 1):S17-25.e1 [FREE Full text] [doi: 10.1016/j.jaad.2011.04.032] [Medline: 22018063]

2. Xiang F, Lucas R, Hales S, Neale R. Incidence of nonmelanoma skin cancer in relation to ambient UV radiation in white populations, 1978-2012: empirical relationships. JAMA Dermatol 2014 Oct;150(10):1063-1071. [doi: 10.1001/jamadermatol.2014.762] [Medline: 25103031]

3. Rogers HW, Weinstock MA, Harris AR, Hinckley MR, Feldman SR, Fleischer AB, et al. Incidence estimate of nonmelanoma skin cancer in the United States, 2006. Arch Dermatol 2010 Mar;146(3):283-287. [doi: 10.1001/archdermatol.2010.19] [Medline: 20231499]

4. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. CA Cancer J Clin 2014;64(1):9-29 [FREE Full text] [doi: 10.3322/caac.21208] [Medline: 24399786]

5. U.S. Department of Health and Human Services. The surgeon general's call to action to prevent skin cancer. 2014.. Washington, DC: U.S. Department of Health and Human Services, Office of the Surgeon General URL: https://www.surgeongeneral.gov/library/calls/prevent-skin-cancer/call-to-action-prevent-skin-cancer.pdf [accessed 2016-12-20] [WebCite Cache ID 6mtrRIUBv]

6. Boniol M, Autier P, Boyle P, Gandini S. Cutaneous melanoma attributable to sunbed use: systematic review and meta-analysis. BMJ 2012;345:e4757 [FREE Full text] [Medline: 22833605]

7. Wehner MR, Shive ML, Chren M, Han J, Qureshi AA, Linos E. Indoor tanning and non-melanoma skin cancer: systematic review and meta-analysis. BMJ 2012;345:e5909 [FREE Full text] [Medline: 23033409]

8. Wehner MR, Chren M, Nameth D, Choudhry A, Gaskins M, Nead KT, et al. International prevalence of indoor tanning: a systematic review and meta-analysis. JAMA Dermatol 2014 Apr;150(4):390-400 [FREE Full text] [doi: 10.1001/jamadermatol.2013.6896] [Medline: 24477278]

9. Centers for Disease Control and Prevention (CDC). Use of indoor tanning devices by adults--United States, 2010. MMWR Morb Mortal Wkly Rep 2012 May 11;61(18):323-326 [FREE Full text] [Medline: 22572978]

10. Guy GP, Berkowitz Z, Watson M, Holman DM, Richardson LC. Indoor tanning among young non-Hispanic white females. JAMA Intern Med 2013 Nov 11;173(20):1920-1922 [FREE Full text] [doi: 10.1001/jamainternmed.2013.10013] [Medline: 23959651]

11. Buller DB, Cokkinides V, Hall HI, Hartman AM, Saraiya M, Miller E, et al. Prevalence of sunburn, sun protection, and indoor tanning behaviors among Americans: review from national surveys and case studies of 3 states. J Am Acad Dermatol 2011 Nov;65(5 Suppl 1):S114-S123. [doi: 10.1016/j.jaad.2011.05.033] [Medline: 22018060]

12. Lazovich D, Vogel RI, Berwick M, Weinstock MA, Anderson KE, Warshaw EM. Indoor tanning and risk of melanoma: a case-control study in a highly exposed population. Cancer Epidemiol Biomarkers Prev 2010 Jun;19(6):1557-1568 [FREE Full text] [doi: 10.1158/1055-9965.EPI-09-1249] [Medline: 20507845]

13. Zhang M, Qureshi AA, Geller AC, Frazier L, Hunter DJ, Han J. Use of tanning beds and incidence of skin cancer. J Clin Oncol 2012 May 10;30(14):1588-1593 [FREE Full text] [doi: 10.1200/JCO.2011.39.3652] [Medline: 22370316]

14. Ferrucci LM, Cartmel B, Molinaro AM, Leffell DJ, Bale AE, Mayne ST. Indoor tanning and risk of early-onset basal cell carcinoma. J Am Acad Dermatol 2012 Oct;67(4):552-562 [FREE Full text] [doi: 10.1016/j.jaad.2011.11.940] [Medline: 22153793]

15. Centers for Disease Control and Prevention. Best practices for comprehensive tobacco control programs--2014. 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. URL: http://www.cdc.gov/tobacco/stateandcommunity/best_practices/pdfs/2014/comprehensive [accessed 2016-12-20] [WebCite Cache ID 6mu4EvRxz]

16. National Cancer Institute. The role of the media in promoting and reducing tobacco use. Tobacco Control Monograph No 19. NIH Pub. No. 07-6242. Bethesda, MD: U.S. Department of Health and Human Services, National Institutes of Health, National Cancer Institute; 2008. URL: https://cancercontrol.cancer.gov/brp/tcrb/monographs/19/m19_complete.pdf [accessed 2016-12-20] [WebCite Cache ID 6mu4L4gJU]

17. Montague M, Borland R, Sinclair C. Slip! Slop! Slap! and SunSmart, 1980-2000: skin cancer control and 20 years of population-based campaigning. Health Educ Behav 2001 Jun;28(3):290-305. [Medline: 11380050]

18. McWhirter JE, Hoffman-Goetz L. A systematic review of visual image theory, assessment, and use in skin cancer and tanning research. J Health Commun 2014;19(6):738-757. [doi: 10.1080/10810730.2013.837562] [Medline: 24512597]

19. Williams AL, Grogan S, Clark-Carter D, Buckley E. Appearance-based interventions to reduce ultraviolet exposure and/or increase sun protection intentions and behaviours: a systematic review and meta-analyses. Br J Health Psychol 2013 Feb;18(1):182-217. [doi: 10.1111/j.2044-8287.2012.02089.x] [Medline: 22989352]

20.   Banerjee SC, Hay JL, Greene K. College students' cognitive rationalizations for tanning bed use: an exploratory study. Arch Dermatol 2012 Jun;148(6):761-762. [doi: 10.1001/archdermatol.2012.398] [Medline: 22710466]

21.   Blashill AJ, Williams A, Grogan S, Clark-Carter D. Negative appearance evaluation is associated with skin cancer risk behaviors among American men and women. Health Psychol 2015 Jan;34(1):93-96 [FREE Full text] [doi: 10.1037/hea0000100] [Medline: 25133823]

22.   Coups EJ, Phillips LA. A more systematic review of correlates of indoor tanning. J Eur Acad Dermatol Venereol 2011 May;25(5):610-6; author reply 617. [doi: 10.1111/j.1468-3083.2011.03996.x] [Medline: 21349117]

23.   Witte K, Allen M. A meta-analysis of fear appeals: implications for effective public health campaigns. Health Educ Behav 2000 Oct;27(5):591-615. [Medline: 11009129]

24.   Detweiler JB, Bedell BT, Salovey P, Pronin E, Rothman AJ. Message framing and sunscreen use: gain-framed messages motivate beach-goers. Health Psychol 1999 Mar;18(2):189-196. [Medline: 10194055]

25.   Thomas K, Hevey D, Pertl M, Ní Chuinneagain CS, Craig A, Maher L. Appearance matters: the frame and focus of health messages influences beliefs about skin cancer. Br J Health Psychol 2011 May;16(Pt 2):418-429. [doi: 10.1348/135910710X520088] [Medline: 21489067]

26.   van't Riet J, Ruiter RA, Werrij MQ, De Vries H. Self-efficacy moderates message-framing effects: The case of skin-cancer detection. Psychol Health 2010 Mar;25(3):339-349. [doi: 10.1080/08870440802530798] [Medline: 20204928]

27.   Gallagher KM, Updegraff JA. Health message framing effects on attitudes, intentions, and behavior: a meta-analytic review. Ann Behav Med 2012 Feb;43(1):101-116. [doi: 10.1007/s12160-011-9308-7] [Medline: 21993844]

28.   Mays D, Zhao X. The influence of framed messages and self-affirmation on indoor tanning behavioral intentions in 18- to 30-year-old women. Health Psychol 2016 Feb;35(2):123-130. [doi: 10.1037/hea0000253] [Medline: 26192383]

29.   Mays D, Tercyak KP. Framing indoor tanning warning messages to reduce skin cancer risks among young women: implications for research and policy. Am J Public Health 2015 Aug;105(8):e70-e76. [doi: 10.2105/AJPH.2015.302665] [Medline: 26066932]

30.   Smith A. U.S. smartphone use in 2015. Washington, DC: Pew Research Center; 2015. URL: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/ [accessed 2016-12-20] [WebCite Cache ID 6mu4QMT2W]

31.   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]

32.   Patrick K, Raab F, Adams MA, Dillon L, Zabinski M, Rock CL, et al. A text message-based intervention for weight loss: randomized controlled trial. J Med Internet Res 2009;11(1):e1 [FREE Full text] [doi: 10.2196/jmir.1100] [Medline: 19141433]

33.   Napolitano MA, Whiteley JA. Social media obesity treatment for college students. 2016 Dec 15. URL: https://clinicaltrials.gov/ct2/show/NCT02342912?term=R01DK100916&rank=1 [accessed 2016-12-21]

34.   Head KJ, Noar SM, Iannarino NT, Grant Harrington N. 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]

35.   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]

36.   Abroms LC, Johnson PR, Heminger CL, Van Alstyne JM, Leavitt LE, Schindler-Ruwisch JM, et al. Quit4baby: results from a pilot test of a mobile smoking cessation program for pregnant women. JMIR Mhealth Uhealth 2015;3(1):e10 [FREE Full text] [doi: 10.2196/mhealth.3846] [Medline: 25650765]

37.   Evans WD, Wallace BJ, Szekely D, Nielsen P, Murray E, Abroms L, et al. Initial outcomes from a 4-week follow-up study of the Text4baby program in the military women's population: randomized controlled trial. J Med Internet Res 2014;16(5):e131 [FREE Full text] [doi: 10.2196/jmir.3297] [Medline: 24846909]

38.   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;3(1):e12 [FREE Full text] [doi: 10.2196/mhealth.3909] [Medline: 25630361]

39.   Evans WD, Lantz PM, Mead K, Alvarez C, Snider J. Adherence to clinical preventive services guidelines: Population-based online randomized trial. SSM Popul Health 2015 Dec;1:48-55. [doi: 10.1016/j.ssmph.2015.11.003]

40.   Lantz PM, Evans WD, Mead H, Alvarez C, Stewart L. Knowledge of and attitudes toward evidence-based guidelines for and against clinical preventive services: results from a national survey. Milbank Q 2016 Mar;94(1):51-76. [doi: 10.1111/1468-0009.12181] [Medline: 26994709]

41.   Mays D, Murphy SE, Bubly R, Atkins MB, Tercyak KP. Support for indoor tanning policies among young adult women who indoor tan. Transl Behav Med 2016 Dec;6(4):613-621. [doi: 10.1007/s13142-016-0432-6] [Medline: 27496162]

42.   Armstrong AW, Watson AJ, Makredes M, Frangos JE, Kimball AB, Kvedar JC. Text-message reminders to improve sunscreen use: a randomized, controlled trial using electronic monitoring. Arch Dermatol 2009 Nov;145(11):1230-1236. [doi: 10.1001/archdermatol.2009.269] [Medline: 19917951]

43.   Petit A, Lejoyeux M, Reynaud M, Karila L. Excessive indoor tanning as a behavioral addiction: a literature review. Curr Pharm Des 2014;20(25):4070-4075. [Medline: 24001299]

XSL•FO

RenderX

44. Dillard JP, Shen L. On the nature of reactance and its role in persuasive health communication. Commun Monogr 2005 Jun;72(2):144-168. [doi: 10.1080/03637750500111815]

45. Greene K, Brinn LS. Messages influencing college women's tanning bed use: statistical versus narrative evidence format and a self-assessment to increase perceived susceptibility. J Health Commun 2003;8(5):443-461. [doi: 10.1080/713852118] [Medline: 14530147]

46. Wood C, Conner M, Miles E, Sandberg T, Taylor N, Godin G, et al. The impact of asking intention or self-prediction questions on subsequent behavior: a meta-analysis. Pers Soc Psychol Rev 2016 Aug;20(3):245-268 [FREE Full text] [doi: 10.1177/1088868315592334] [Medline: 26162771]

47. Weitzel JA, Bernhardt JM, Usdan S, Mays D, Glanz K. Using wireless handheld computers and tailored text messaging to reduce negative consequences of drinking alcohol. J Stud Alcohol Drugs 2007 Jul;68(4):534-537. [Medline: 17568957]

48. Harrell F. Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York, NY: Springer-Verlag; 2001.

49. Cole-Lewis H, Kershaw T. Text messaging as a tool for behavior change in disease prevention and management. Epidemiol Rev 2010;32:56-69 [FREE Full text] [doi: 10.1093/epirev/mxq004] [Medline: 20354039]

50. Whittaker R, McRobbie H, Bullen C, Borland R, Rodgers A, Gu Y. Mobile phone-based interventions for smoking cessation. Cochrane Database Syst Rev 2012 Nov 14;11:CD006611. [doi: 10.1002/14651858.CD006611.pub3] [Medline: 23152238]

51. Mays D, Kraemer J. FDA regulation of indoor tanning devices and opportunities for skin cancer prevention. JAMA 2015 Jun;313(24):2423-2424. [doi: 10.1001/jama.2015.5975] [Medline: 26018657]

52. Reed JA, Powers A, Greenwood M, Smith W, Underwood R. Using "point of decision" messages to intervene on college students' eating behaviors. Am J Health Promot 2011;25(5):298-300. [doi: 10.4278/ajhp.090511-ARB-162] [Medline: 21534831]

53. Fogg B, Adler R. Texting 4 Health: A Simple, Powerful Way to Change Lives. Stanford, CA: Captology Media; 2009.

54. Snyder LB, Hamilton MA, Mitchell EW, Kiwanuka-Tondo J, Fleming-Milici F, Proctor D. A meta-analysis of the effect of mediated health communication campaigns on behavior change in the United States. J Health Commun 2004;9(Suppl 1):71-96. [doi: 10.1080/10810730490271548] [Medline: 14960405]

55. Freedman MR, Connors R. Point-of-purchase nutrition information influences food-purchasing behaviors of college students: a pilot study. J Am Diet Assoc 2010 Aug;110(8):1222-1226. [doi: 10.1016/j.jada.2010.05.002] [Medline: 20656098]

Abbreviations

MMS: multimedia messaging service
OLS: ordinary least squares
RQ: research question
SMS: short message service

XSL•FO
RenderX

Original Paper

A Review of Persuasive Principles in Mobile Apps for Chronic Arthritis Patients: Opportunities for Improvement

Jonas Geuens[1], MSc; Thijs Willem Swinnen[2,3], PT, MSc; Rene Westhovens[2,3], MD, PhD; Kurt de Vlam[2,3], MD, PhD; Luc Geurts[1], PhD; Vero Vanden Abeele[1], PhD

[1]e-Media Lab, KU Leuven, Leuven, Belgium

[2]UZ Gasthuisberg, Division of Rheumatology, University Hospitals Leuven, Leuven, Belgium

[3]Skeletal Biology and Engineering Research Center, Department of Development and Regeneration, KU Leuven, Leuven, Belgium

Corresponding Author:
Jonas Geuens, MSc
e-Media Lab
KU Leuven
A. Vesaliusstraat 13
Leuven, 3000
Belgium
Phone: 32 472690740
Fax: 32 16313201
Email: jonas.geuens@kuleuven.be

Abstract

Background: Chronic arthritis (CA), an umbrella term for inflammatory rheumatic and other musculoskeletal diseases, is highly prevalent. Effective disease-modifying antirheumatic drugs for CA are available, with the exception of osteoarthritis, but require a long-term commitment of patients to comply with the medication regimen and management program as well as a tight follow-up by the treating physician and health professionals. Additionally, patients are advised to participate in physical exercise programs. Adherence to exercises and physical activity programs is often very low. Patients would benefit from support to increase medication compliance as well as compliance to the physical exercise programs. To address these shortcomings, health apps for CA patients have been created. These mobile apps assist patients in self-management of overall health measures, health prevention, and disease management. By including persuasive principles designed to reinforce, change, or shape attitudes or behaviors, health apps can transform into support tools that motivate and stimulate users to achieve or keep up with target behavior, also called persuasive systems. However, the extent to which health apps for CA patients consciously and successfully employ such persuasive principles remains unknown.

Objective: The objective of this study was to evaluate the number and type of persuasive principles present in current health apps for CA patients.

Methods: A review of apps for arthritis patients was conducted across the three major app stores (Google Play, Apple App Store, and Windows Phone Store). Collected apps were coded according to 37 persuasive principles, based on an altered version of the Persuasive System Design taxonomy of Oinas-Kukkonen and Harjuma and the taxonomy of Behavior Change Techniques of Michie and Abraham. In addition, user ratings, number of installs, and price of the apps were also coded.

Results: We coded 28 apps. On average, 5.8 out of 37 persuasive principles were used in each app. The most used category of persuasive principles was System Credibility with an average of 2.6 principles. Task Support was the second most used, with an average of 2.3 persuasive principles. Next was Dialogue Support with an average of 0.5 principles. Social Support was last with an average of 0.01 persuasive principles only.

Conclusions: Current health apps for CA patients would benefit from adding Social Support techniques (eg, social media, user fora) and extending Dialogue Support techniques (eg, rewards, praise). The addition of automated tracking of health-related parameters (eg, physical activity, step count) could further reduce the effort for CA patients to manage their disease and thus increase Task Support. Finally, apps for health could benefit from a more evidence-based approach, both in developing the app as well as ensuring that content can be verified as scientifically proven, which will result in enhanced System Credibility.

KEYWORDS

persuasive technology; mobile applications; chronic arthritis

Introduction

Chronic arthritis (CA) is an umbrella term for inflammatory rheumatic and musculoskeletal diseases such as rheumatoid arthritis (RA), osteoarthritis (OA), and spondyloarthritis (SpA). CA is highly prevalent. One in five adults in the United States has doctor-diagnosed arthritis [1]. CA is typically associated with processes of inflammation and/or destruction, which cause joint pain, swelling, stiffness and instability, joint destruction, or bony ankylosis resulting in progressive immobility [2]. These disease processes largely contribute to limitations in performing day-to-day activities such as walking, cleaning, and working [3]. Luckily, effective disease-modifying antirheumatic drugs have become available to tackle RA and SpA. However, they require the patient's long-term commitment to tightly follow up on disease parameters by the treating physician and other health professionals, as well as compliance with the medication regimen and other therapy proposals [4,5]. Treatment recommendations for all CA also include physical therapy programs in order to improve aspects of physical fitness such as cardiovascular endurance, muscle strength, posture and movement control, range of motion, and balance [3,6-8]. Ample evidence suggests that physical exercise has strong benefits [9-11]. Unfortunately, between 35% and 75% of patients with CA fail to adhere to the physical exercise recommendations of their therapist [12-14]. Besides arthritis-specific barriers (eg, pain, disability), adherence to therapy is reduced by personal barriers (eg, lack of motivation) and contextual barriers (eg, lack of skilled staff and facilities) [15,16]. Clearly there is a need to support patients in managing their disease and adhering to their physical exercise programs.

Perhaps mobile health (mHealth) apps can be part of a solution for supporting CA patients. Health apps are defined as "mobile applications that assist consumers in self-management of overall wellness, disease prevention and disease management" [17]. These apps are further advancements of the telehealth movement, not only supporting remote interventions but also providing the opportunity to intervene at any time and place [18]. mHealth apps provide clear benefits for treatment, assessment, and self-management of CA [19], for example, the assessment of gait in rheumatic diseases [20] or the logging of pain and physical condition [21]. The advent of health apps is strongly interlinked with the rise of mobile phones, specifically smartphones. Smartphone penetration rate at the end of 2015 was predicted to be 42.6% on a global scale, 59.8% in the United States, and 54.9% in Western Europe [22-24]. Most patients today already own and use smartphones on a daily basis. Many CA researchers have identified clear needs and opportunities for mobile phone apps to tackle educational, lifestyle, and treatment interventions to ease delivery and increase involvement of CA patients [19-21,25].

In addition, there has been a call for a more conscious design of these apps, starting from the CA patients' needs and addressing CA patients' motivations for using health apps [26-28]. In particular, designing for patients who suffer from comorbidities involving chronic pain, depression, and fatigue presents specific challenges [29-31] with respect to sustained motivation. In this study, we want to investigate how to increase the motivation of CA patients to use mHealth apps, and more particularly by including "persuasive principles." Persuasive principles are specific design techniques such as offering praise, providing reminders, imitating social agents, providing "social support," or augmenting "system credibility" [32]. By including persuasive principles, health apps can transform into supportive tools that motivate and stimulate users to achieve or keep up with targeted behavior, also called persuasive systems, defined as "computerized software or information systems designed to reinforce, change or shape attitudes or behaviors or both without using coercion or deception" [33].

It has been argued that the design of current health apps lacks a "conscious" implementation of persuasive principles [27,29]. According to Kelders et al [33] and Tomlinson et al [34], apps in health care are often designed and treated as "black boxes": the technology works but the design itself is not evidence-based nor based on behavior change. This may result in technology that has a low impact on health care practices [35-37].

In this paper, we investigated which persuasive principles are most prevalent in current apps directed at CA management and which principles are lacking. This knowledge may inform and inspire health professionals who are looking for innovative ways to support their interventions with mHealth apps. It is our aspiration that this analysis can help health experts select better quality apps to help patients manage their disease in a better and more effective way.

This insight may equally help health professions and app developers build more effective support tools. Particularly, this knowledge may help how to move beyond the current state and result in mHealth apps that are more motivating for CA patients and hence more effective with respect to, for example, adherence of therapies or the removal of contextual barriers.

Several theories, frameworks, and taxonomies exist to guide designers of persuasive systems. Oinas-Kukkonen and Harjumaa proposed a model of Persuasive Systems Design (PSD) [32]. This model contains 28 persuasive system design techniques divided over four categories: primary task support, dialogue support, social support, and system credibility support. Primary task support aims to persuade the user to complete a task by supporting the user in their execution of the task. This category includes, for example, reducing one complex task into a set of smaller tasks that are easier to complete or tunneling to offer the user a trajectory. Dialogue support provides system feedback to guide the user towards the intended behavior. Examples include providing rewards or praise to the user and providing reminders. Social support strengthens the overall persuasiveness of a software system by leveraging the human nature to interact with others [38,39]. Examples include comparing oneself to the norm of a population, cooperating with other users to achieve a similar goal, and learning behaviors or actions by observing others. Finally, system credibility support includes principles

on how to design a system so that it is more credible and thus more persuasive.

Besides the PSD model by Oinas-Kukkonen and Harjumaa, another established framework is the one offered by Abraham and Michie. These authors proposed a Taxonomy of Behavior Change Techniques (BCT) [40]. Different from the PSD model, the 26 BCT principles are not organized in broader categories, but rather the authors link the principles explicitly to the underlying theoretical models of behavior change such as Operant Conditioning [41], Social Cognitive Theory [42], and Theory of Planned Behavior [43].

Several authors used either the PSD model or BCT to study Web-based health interventions or a combination of both [44-48] to study the effect on motivation or adherence to the intervention. Kelders et al [44] conducted a systematic review of Web-based health interventions, relying on the PSD model, to determine key persuasive factors for adherence to the intervention. They classified which persuasive system design principles were most prevalent and how they impacted adherence to the intervention. They found that primary task support showed the highest mean, while social support showed the lowest mean. However, while primary task support principles were most commonly employed in Web-based interventions, they did not show any predictive value for adherence. In contrast, social support principles showed a significant contribution to better adherence, yet social support principles were least implemented in Web-based interventions. Overall, Kelders et al concluded that the use of persuasive technology elements can explain a significant amount of the variance in adherence.

Lehto et al [45] used the PSD principles as well, to classify Web-based alcohol and smoking interventions. Again, they found primary task support components (eg, reduction of task complexity and self-monitoring) to be widely utilized and reported on widely in the reviewed studies. However, they found a lack of tailoring, which may imply that the interventions are not targeting a specific audience. The conclusion was that more research is needed to increase the understanding of persuasive principles in interventions and their contributions to intervention outcomes.

Overall, these studies found that interventions that incorporated more BCTs also tended to have larger effects as compared to interventions that incorporated fewer BCTs, and that different techniques are beneficial for different types of intended behaviors. In particular, social support principles show significant contribution to adherence, yet they are least implemented.

To date, only a few studies have been conducted on persuasive principles in health apps [49-52]. As mentioned, the ubiquity and pervasiveness of mobile phones enable health interventions to go beyond what can be offered via a Web-based intervention. Additional sensing and networking features bring along opportunities for real-time interactions and monitoring, context awareness (eg, localization), physical activity sensing, etc. It is therefore interesting to investigate whether persuasive principles are also used in health apps on smartphones.

Matthews et al conducted a systematic review of persuasive principles in mobile apps promoting physical activity [52]. They also used the PSD model of Oinas-Kukkonen and Harjumaa and reviewed 20 articles describing the use of mobile apps to promote physical activity. The authors found on average only 4 PSD principles (out of 28) were implemented. The most common PSD principle was self-monitoring (implemented in 14 of 20 analyzed systems), a feature that supports a user's primary task and that was not listed as an important principle in the previous Web-based interventions. In addition, apps frequently implemented a variety of dialogue support and social support principles, as a combination of praise, rewards, reminders, and suggestion to motivate the user to be physically active. They found the most lacking category to be system credibility, in other words, PSD principles that influence the way a user perceives the credibility of the system. Hence, Matthews et al concluded that it is "not clear for users to judge the extent to which the apps are credible or not."

In line with the previous authors, Vollmer et al launched a call for "Apps seeking theories" [50]. Vollmer et al conducted a systematic review of BCTs in cancer survivorship apps. They identified 68 apps for cancer survivorship on both the Android and iOS platform. Interestingly, these authors developed a coding manual specifically for mHealth apps for cancer survivorship, based on both the BCT taxonomy and PSD model. This coding manual contained 17 persuasive principles. Vollmer et al found that their 68 apps on average contained 4 persuasive principles. They highlighted as well as Matthews et al that a more theory-based approach is needed when designing and developing mHealth apps, especially with respect to the persuasive design/behavior techniques that could empower behavior change: "While the current advancements in mobile hardware, sensors and electronic patient records offer opportunities for health apps, a better understanding is needed of how this translates into benefits for patients" [50].

There is a clear need for health apps to provide treatment, assessment, and self-management of chronic arthritis [19-21,25]. However, currently, these apps lack a conscious implementation of persuasive principles [27], [29]. To contribute to the knowledge surrounding persuasive principles in health apps for chronic arthritis, this paper investigates to what extent persuasive principles are found in disease management apps for arthritis and whether these persuasive principles contribute towards a higher rating of apps. It is our hope that this analysis can help health experts select better quality apps to help patients manage their disease in a better and more effective way. This insight may equally help app developers and health professionals to build more effective support tools. CA patients may in turn benefit from newly developed apps that are more motivating and effective.

Methods

Eligibility Criteria for Health Apps

A review of health apps (ie, target intervention) aimed at the management of chronic arthritis (ie, target population) was conducted. This review included all apps available on Google Play (Google, Mountain View, CA), the Apple App Store (Apple

Inc, Cupertino, CA), and the Windows Phone Store (Microsoft, Redmond, WA) between January 15 and June 15, 2015. Criteria for inclusion were that the app (1) targets one or more arthritis-related diseases, (2) aids in the management of arthritis-related disease factors, and (3) targets adult patients. Because Google Play employs a broad-spectrum algorithm to return apps related to the search, a number of apps related to arthritis but not intended for patients with arthritis were excluded. Exclusion criteria were (1) apps intended for medical personnel only, (2) magazines, general fitness or pain trackers, or food guides in app format, and (3) apps that failed to install or open on our test devices. These exclusion criteria were applied during the initial screening. There were no language restrictions.

Search Criteria and Procedures

The used search terms included a range of words related to arthritis as a disease. All three app stores were searched using the same keywords: Arthrit*, Rheuma*, Spondyl*, Osteo*, Bechterew. The search was conducted independently by 2 researchers (JG, VVA). In case one reviewer included an app that was not recorded by the other reviewer, that app was added to the full list to be coded. This was done due to the rapidly changing content of the app stores where new apps are added every day and because both reviewers searched the app stores at different moments.

Because the search algorithm of Google Play continues to load new content related to the search terms, the initial selection stopped when the last 10 app titles were no longer relevant to our search. An initial screening revealed 41 relevant apps across all three app stores. This initial screening was done using keywords described above and by reviewing screenshots and app descriptions. Only apps that met the inclusion criteria were selected, for example, apps that were clearly intended for medical personnel were excluded from the initial selection. After removing duplicates present in one or more of the app stores, 38 apps were left to be screened further. See Figure 1 for the app selection flow.

In the next screening step, each app was installed on a mobile device, either an LG Nexus 4 (LG Electronics, Seoul, South Korea) running Android 5.0.2 (Google) or an iOS (Apple Inc) tablet, running iOS version 8, or a Windows 8 (Microsoft) phone. If multiple platforms were available, the Android platform was chosen first, next iOS, and Windows Phone as a third platform. Each app was then used and tested until a good understanding of the features of the app was established. This was done to determine features not explained in the description of the app. During this screening phase, seven more apps were identified that were intended for therapists or physicians only and not patients. These apps did pass the initial screening (based on their title and screenshots) but were found to meet the exclusion criteria (ie, intended for medical personnel only) after further analysis. Finally, three more apps were removed from the app store by the developers between searching and coding. These were excluded as well. A total of 28 apps remained to be coded.

Figure 1. App selection flow.
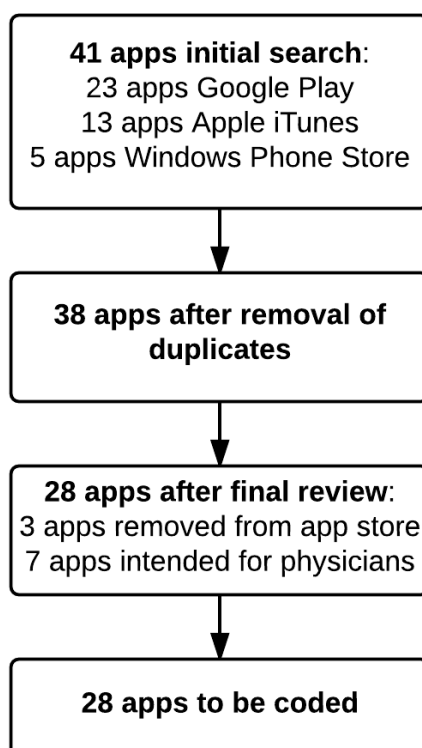
Data Extraction and Analyses

Apart from the behavior change/persuasive design principles (see coding system below), the name, targeted disease, platform(s), price, language, average user rating, number of user ratings, number of installs, and finally the last update were coded for every app.

The average user rating and the number of user ratings were also logged since they can provide a good indication of how the app was valued by its users. However, requiring a minimum of user ratings is good practice in order to limit the influence of biased ratings (eg, user ratings that are influenced by the app owners/developers). Therefore, the average user rating of apps was logged only when it had user ratings coming from at least 5 different users. This is also in line with the Apple App store that also sets the minimum number of user ratings necessary for a rating to appear on five. In summary, 14 apps had sufficient user ratings and 14 apps did not show a number of user ratings above the required threshold of five.

In addition, for the apps originating from the Google Play Store, the range of installs was also logged. The Google Play Store does not provide the exact number of installs but rather provides a general indication according to the following brackets: 1-10, 11-50, 50-100, 100-500, 500-1000, 1000-5000, 5000-10000, 10000-50000, etc. Unfortunately, the number of installs is not available to the public in the Apple App Store or the Windows Phone Store.

Finally, to estimate the effects of the number of implemented persuasive principles on favorability of apps, a correlation analysis was performed between the number of persuasive principles and the number of user ratings as well as the number of installs and the average user rating.

Coding System for Persuasive Design

Apps were coded for persuasive design principles according to an adaptation of the 28 PSD principles by Harjumaa and Oinas-Kukkonen [32] and the Taxonomy of Behavior Change Techniques by Abraham and Michie [40] composed of 26 BCTs.

Principles that were ambiguous (eg, "Liking," being visually attractive) or overlapped with other principles within the same taxonomy (PSD or BCT) (eg, "trustworthiness" and "surface credibility") were merged, hence ending up with 37 principles organized according to the four overarching categories of task support, dialogue support, system credibility, and social support. Prior to coding, both coders discussed every principle and provided an app-related example to minimize misconceptions.

For each app, the specific and total amount of behavior change/persuasive design principles used was coded and calculated by 2 independent coders (JG and VVA) and a third (JH) in case of disagreement. Agreement occurs when 2 coders both mark the presence or absence of a principle. Disagreement occurs when one coder marks the presence of a principle while the other marks an absence. Intercoder reliability was calculated using both Cohen's kappa statistic and as percentages of agreement. Reliability values were found to be between 71.4 and 100% wise agreement and between .65 and 1.0 kappa agreement depending on the principle. The mean value of all kappa scores was .96 (SD .08). All kappa values were significant at $P < .0005$.

The lowest rated principle in terms of intercoder reliability was surface credibility (percentage wise agreement: 96.4, kappa agreement: .65) and micro tailoring (percent wise agreement: 96.4, kappa: .781). The first achieved a low kappa value due to a very high percent wise agreement, high level of occurrence, and chance correction used by the kappa calculation. Overall, all intercoder reliability values are at an acceptable level (ie, >.60 [53,54]) with 26 principles (out of 37) scoring perfect agreement. Intercoder reliability scores for each principles can be found in the Results section.

Results

As mentioned, 28 apps (see Table 1) met the inclusion criteria and were reviewed according to the combination of the behavior change/persuasive design principles.

XSL•FO

RenderX

Table 1. Detailed description of the 28 health apps that met the selection criteria.

Name	Targeted disease[a]	Platform[b]	Installs, n	Rating (out of 5)[c]	Ratings, n	Price, US $	Language	Persuasive principles, n
ArthritisID	A	iOS				0.00	English	13
Arthritis Diary	A	An/WP				4.99	English	10
NHS 24 MSK Help	MSK	iOS				0.00	English	10
RA Patient Companion	RA	iOS				0.00	English	10
Pauseboogie fra Gigtforeningen	A	An/iOS	5000-1000	4.0	39	0.00	Danish	8
Rheuma AKTIV	RA	iOS				0.00	German	8
Andar	RA	An	100-500	3.3	3	0.00	Spanish	7
MyRA	RA	iOS		4.5	62	0.00	English	7
RA Helper	RA	An	500-1000	4.8	5	0.00	English/ Slovenian	7
RAPA – RA betegalkalmazás	RA	An	50-100	4.8	5	0.00	Hungarian	7
RheumaTrack RA	RA/SpA	An/iOS	10,000-50,000	4.2	375	0.00	English	7
Track + React	A	iOS		3.4	72	0.00	English	7
Back to Action	SpA	An/iOS	1000-5000	4.2	51	0.00	English	6
Bewegen met Bechterew	SpA	An	100-500	2.1	8	0.00	Dutch	6
iAnkylosing Spondylitis	SpA	iOS				0.00	English	6
DAS Calculadora	RA	An	10-50			0.00	Spanish	5
RADAI	A	WP				0.99	English	5
Arthritis Symptoms+Treatment	RA	An	500-1000	4.5	31	0.00	English	4
RheumaHelper	RA	An/iOS	5000-10,000	4.3	157	0.00	English	4
RhEumAtic Disease activitY	RA	iOS				0.00	English	4
Arthritis Relief	A	WP		5.0	1	4.99	English	3
DAS28 - RA	RA	An	5000-10,000	4.1	64	0.00	English	3
DAS28 Free	RA	An	1000-5000	3.6	8	0.00	English	3
Juvenile RA	RA	An	10-50	4.5	2	0.00	English	3
Living Well With Arthritis	A	iOS				0.99	English	3
Rheumatoid Arthritis Disease	RA	An	100-500	4.4	11	0.00	English	3
Rheumatoid Arthritis of Knee	RA	An	5000-10,000	4.0	47	0.00	English	3
Arthritis	A	WP				3.99	English	1

[a]A: arthritis, MSK: musculoskeletal diseases, RA: rheumatoid arthritis, SpA: spondyloarthritis.

[b]An: Android, WP: Windows phone.

[c]Rating from 1-5 with 1 being the lowest rating.

XSL•FO
RenderX

Table 2. Persuasive design principles and number of apps (N=28) that used the principle.

Design principles	Apps, n	appa (percentage of agreement)
Task support		
General information[a]	16	.786 (89.3)
Self-monitoring	12	.926 (96.4)
Reduction	9	.836 (92.9)
Logging[a]	9	.92 (96.4)
Instruction[a]	7	.909 (96.4)
Goal setting	3	1.0 (100)
Micro tailoring[a]	3	.781 (96.4)
Macro tailoring[a]	1	1.0 (100)
Simulation	1	1.0 (100)
Contextual cues[a]	1	1.0 (100)
Tunneling	0	1.0 (100)
Tracking[a]	0	1.0 (100)
Rehearsal	0	1.0 (100)
Behavioral contract[a]	0	1.0 (100)
Dialogue support		
Reminders	10	1.0 (100)
Suggestion	2	1.0 (100)
Rewards	1	1.0 (100)
Praise	0	1.0 (100)
Similarity	0	1.0 (100)
Personalization	0	1.0 (100)
Social role	0	1.0 (100)
Prompt self-talk[a]	0	1.0 (100)
System credibility		
Surface credibility	26	.65 (96.4)
Expertise	25	.837 (71.4)
Verifiability	8	1.0 (89.3)
Real-world feel	7	.884 (85.7)
Authority	7	.837 (96.4)
Third-party endorsements	0	1.0 (100)
Social support		
Social interaction[a]	1	1.0 (100)
Social learning	0	1.0 (100)
Social identification[a]	0	1.0 (100)
Social comparison	0	1.0 (100)
Normative influence	0	1.0 (100)
Social facilitation	0	1.0 (100)
Cooperation	0	1.0 (100)
Competition	0	1.0 (100)

XSL•FO

RenderX

Design principles	Apps, n	appa (percentage of agreement)
Recognition	0	1.0 (100)

[a]Principles added from behavior change techniques.

## Prevalence of Behavior Change/Persuasive Design Principles

On average, an app was found to utilize 5.8 (median 6) behavior change/persuasive design principles out of 37. The maximum number of principles used in one app was 10. The least number of principles used was 1 (out of 37) in one app. Of the four behavior change/persuasive design categories, the most used was system credibility with an average of 2.6 (median 2) principles (ie, on average, each app used 2.6/37 principles from the system credibility category), then task support with an average of 2.3 (median 2) principles, next dialogue support with an average of 0.5 (median 0) principles, and finally social support with an average of 0.01 (median 0) (see Figure 2).

Analyzing the principles within each behavior change/persuasive design category more carefully, principles to influence system credibility were limited to offering a level of surface credibility (n=15) (ie, "Do not show advertisements" and "At a first glance seem to offer truthful information") and perceived expertise (ie, "expertise in the information provided") (n=12). Some CA health apps also implemented verifiability (n=8) allowing users to find out more by linking to studies or reports that provide evidence, by providing real-world feel (highlighting the people and organizations behind the app, n=7), or by providing authority figures (medical doctors, n=7). Third-party endorsements were completely lacking.

As for the task support category, the most prevalent principle was to "Provide general information" (n=16). Other common task support principles were self-monitoring (being able to [re]view your own data) (n=12) and reduction (reducing complex behavior into simple tasks) (n=9). All of these apps offered reduction as the calculation of the Disease Activity Score (DAS) and visualized the evolution of the DAS score over time, hence allowing for self-monitoring. Goal setting in the form of setting specific goals and guiding the user towards that goal was less prevalent (n=3). Macrotailoring (tailoring information to specific patient groups) was found in only two apps, and finally both simulation and the use of contextual cues was found in only one app.

Dialogue support was present in the shape of reminders (n=10), mainly reminding patients to exercise or to fill out scores related to CA parameters, for example, number of sore joints, overall pain. One application also provided a reward in the shape of stars that could be collected when inputting information regularly. All other principles were lacking.

Finally, social support principles were less applied, in fact techniques to, for example, influence normative beliefs or to include provide social support were simply lacking in all the apps. Only one app implemented a means for social interaction by providing a link to a community forum where patients could exchange thoughts and feelings. Complete results can be found in Table 2 or Multimedia Appendix 1.

Figure 2. Prevalence of behavior change/persuasive design categories and principles.
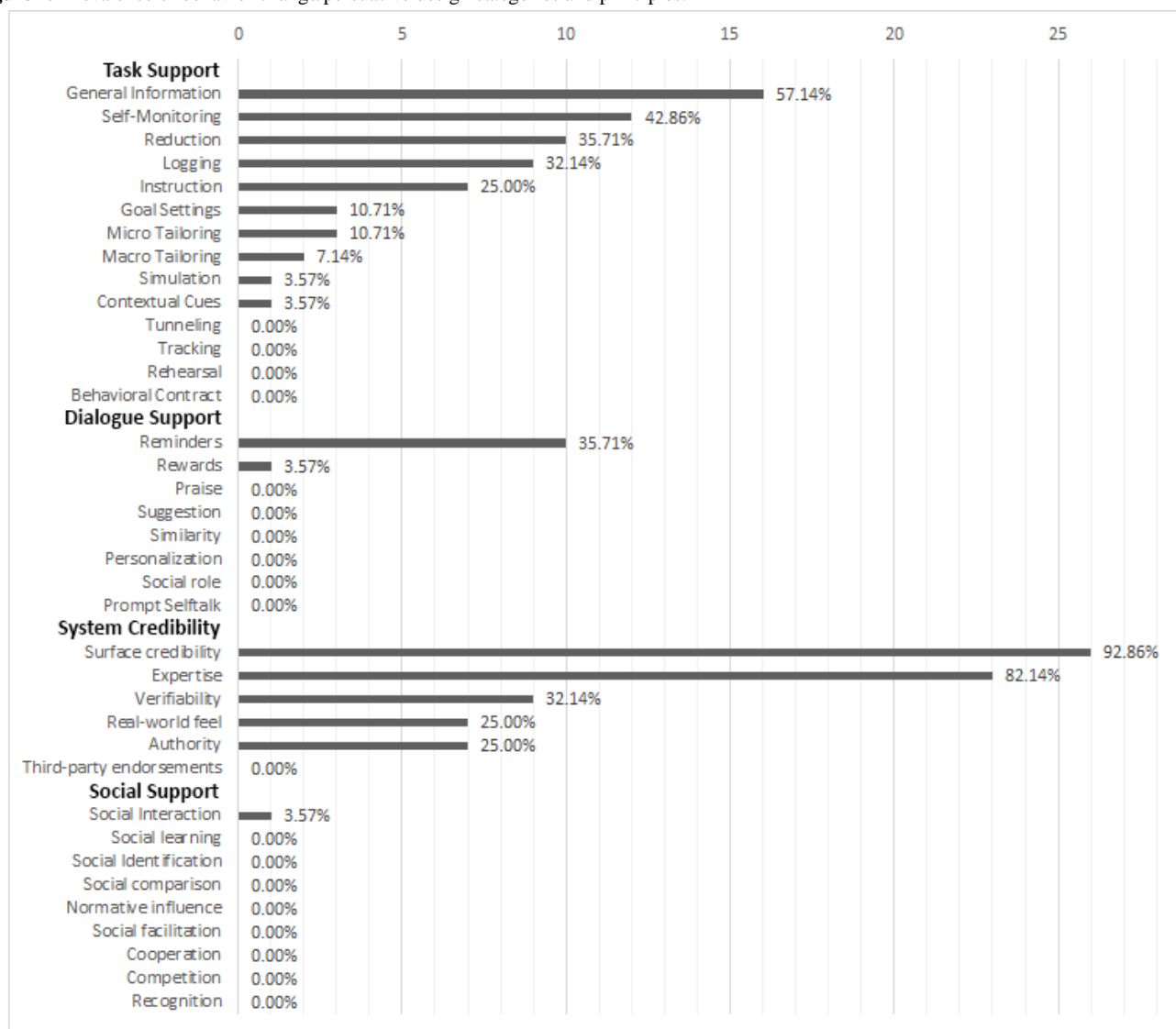
User Ratings, Number of User Ratings, and Number of Installs

The average rating of all eligible apps (14/28, 50%) was 4.06 (SD 0.67) with a median 4.2 (out of 5). As mentioned, this average does not include apps that have fewer than five user ratings to minimize the influence of biased ratings. However, this distribution is negatively skewed (skewness 1.90).

On average (again excluding apps that did not meet the minimum of five user ratings), apps received 52 user ratings (SD 87), median 43, with the maximum number of user ratings recorded being 375. The distribution is heavily positively skewed (skewness 2.80). Only two apps have received +100 user ratings: RheumaTrack and RheumaHelper. RheumaTrack was developed with support from AbbVie, a company focused on pharmaceutical research and development and offers extensive logging of disease-related parameters. RheumaHelper

was developed by an independent developer and offers a wide variety of calculations related to arthritis suitable for both patients and medical personnel.

Finally, the average price was €0.57 (SD €1.45), with a median of €0.00. Of 28 apps, 23 were free. The highest price was €4.99. Figure 3 displays the frequency distribution of apps according to price category.

No significant correlation was found between the number of PSD principles and the number of user ratings and no significant correlation between the number of persuasive principles and average user rating. Figure 4 illustrates this relation between the number of persuasive principles and user rating.

Also, no significant correlation could be found between number of PSD principles and number of installs (including only the apps of Google Play Store).

Figure 3. Number of apps per price category (N=28).

Figure 4. Number of persuasive principles in relation to user rating (n=14).
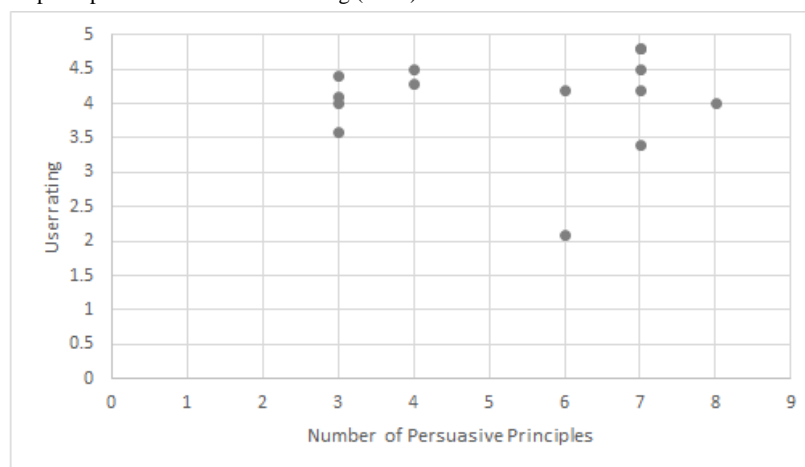
Discussion

Principal Findings

The aim of this review was to study which persuasive principles were used in health apps and specifically in apps for CA patients. In the 28 apps we coded, we found a median of 6 persuasive principles, with most prominent categories being system credibility (median of 2 principles) and task support (median of 2 principles). Surprisingly, social support principles and dialogue support principles were lacking.

High system credibility implies that the apps are perceived as credible and trustworthy. Most of the apps that scored high on system credibility were either developed or funded by health services or pharmaceutical companies. Although the system credibility category is the most prevalent in our study, we argue that scoring only 2.6 (median 2) out of 7 possible system credibility principles per app is still surprisingly low. The first system credibility principle is surface credibility, characterized by the absence of obvious banners/ads. This was found in 26 apps (N=28). In addition, 25 apps equally offered expertise with respect to information about CA they presented. We argue that these principles are "low hanging fruit." All CA apps, as all health apps, should use truthful information that is scientifically proven. Unfortunately, only 8 apps scored on verifiability (ie, the principle that implies a way to verify the content of the app).

Third-party endorsements were completely lacking. Furthermore, information on system credibility was often available only when obtaining and installing the app.

With the open spirit of app stores, anyone can develop an app and publish it without having to provide evidence on its disease management tactics. This results in a substantial number of apps that seem to provide only a copy of static webpage content app format. It is remarkable that several apps were designed several years ago as confirmed by the last update date on their page on the app stores and seemed abandoned without any support from the developers. When browsing in the app store, it is not immediately clear which apps contain scientifically proven content and which do not. In line with the claim of Vollmer et al with "Apps seeking theories" [50], we argue that more apps need to give patients the means to verify the extent to which apps incorporate evidence-based practices.

The second highest prevalent category is task support (an average of 2.3 techniques per app with a median of 2). Task support lowers the effort for users to execute the target behavior by simplifying the task. In this study, apps mainly supported the task by providing more information on disease management or by providing a means to calculate the DAS score. In fact, many apps (11/28, 39%) could actually be classified as a "tracker" apps [55], that is, able to track a patient's disease throughout a timespan of multiple days. Hence, trackers focus

XSL•FO
RenderX

on being able to input disease data (either through manual input or automatic collection of data) and visualizing that data. By creating a place where patients can collect and visualize all their data, obviously "tracker" apps lower the effort for patients to manage their disease. However, the apps reviewed in this study, despite the advanced sensing and hardware opportunities that mobile phones offer, limited tracking to manual input. CA patients still had to manually enter a score about their physical activity level, rather than this being sensed through activity sensors inside the phone. In the authors' opinion, many of the features that fitness and sports apps offer are simply not implemented in the CA apps reviewed in this study. In our humble opinion, much improvement can be made.

Dialogue support principles were implemented in only 10 apps (an average of 0.5 techniques per app with a median of 0). This category contains techniques that focus on feedback from the app to the patient. The most used technique we found in this category was to offer reminders. This is no surprise as patients need to take medication at regular intervals and reminders help them achieve that goal.

We note that dialogue support is a key feature of the more commercially successful fitness and sports apps like Nike+, Runkeeper, or Fitbit. Users can achieve a high score in this category through their use of rewards (commonly implemented as badges, trophies, points [56]), praise (motivational messages), and reminders. These so-called gamification techniques are currently popular in many apps [57-59] but lacking in CA apps, although health apps for CA patients may equally benefit from these techniques to motivate patients. Hence, we argue that CA apps could again benefit from implementing rewards and praise.

The least used category was social support; in fact, principles of this category were completely lacking. Again, this is in contrast with the aforementioned fitness and sports app and contrary to studies stating the importance of social support as mentioned by Kelders et al and others [44,60,61]. These apps include an extensive use of social support techniques to allow users to share their results and current activity on social media. It has been argued by many theories and researchers that, in particular, principles from social psychology may be effective in the long run. For example, relatedness has been linked to intrinsic motivation [62], normative beliefs have shown crucial for planned behavior change [43,63-65], and so on. Although social networks specifically for patients with chronic diseases exist [66], these social support techniques have not been used in any of the apps we reviewed. Again, we argue that particularly in this realm, health apps for CA patients can benefit from employing social support techniques.

To summarize, on the basis of our review, we argue that current health apps for CA patients would benefit from (1) a more evidence-based approach, by using content that can be verified as scientifically proven and endorsed by third parties, which in turn will result in an increase in system credibility, (2) the addition of automated tracking of certain health-related parameters (eg, physical activity, step count) that could further reduce the effort needed to manage a disease and thus increase task support. The current technology incorporated in mobile phones enables us to easily implement this functionality, (3)

the extension of dialogue support techniques (eg, rewards, praise), in other words, the gamification principles also found in fitness and wellness apps, and (4) the addition of social support techniques (eg, social media, user forums).

Limitations and Further Work

Unfortunately, we were not able to link the number of persuasive principles to patient appreciation. We found no correlation with user ratings and no correlation with number of installs. Perhaps there is simply no correlation, but this might also be attributed to our coding scheme. As stated by Lehto et al [45], the current PSD model, while highly valuable, is still evolving. Not much research has been done to validate its workings for assessing health apps, let alone health apps for CA patients in particular. For example, the current taxonomy does not include weights or priorities for the different techniques with respect to the management of CA. Yet, it is highly likely that not all persuasive principles are weighted as equally important by the CA patients. In particular, persuasive principles that lend themselves to long-term engagement may be more appropriate (eg, principles such as general information and goal setting [67,68]) than other principles that are associated more with short-term rewards (eg, reminders) [44,69]. Moreover, comorbidity of depression and chronic pain may suggest the importance of coping principles such as suggestion, social facilitation, and cooperation. Further studies may dissect which persuasive principles are deemed as more important by CA patients.

Moreover, some categories overlap, or remain elusive. Similar to Lehto et al [49], we found it hard to code some of the persuasive principles. Some principles are not hard grained as a feature of the app, but rather rely on an interpretation by the patient. An example of such a subjective principle is "Liking" from the PSD model. This principle is described as "A system that is visually attractive for its users" [32]. While this is a valid principle to increase persuasiveness, it can be argued that researchers are unable to judge whether certain apps are perceived as visually attractive by certain patient groups on the basis of the app alone. Other principles overlap, that is, one principle directly implied the use of the other. For example, an app that lacks surface credibility will not be perceived as trustworthy either. Moreover, the current persuasive principles are not specifically geared towards health or mobile apps. Hence, a taxonomy of persuasive principles geared towards mHealth apps is needed as well as further work on persuasive principles tailored for CA patients.

Conclusion

In this study, we investigated which persuasive principles are most prevalent in current health apps targeted at CA patients. We found a remarkable lack of persuasive techniques to engage patients in digital management of their disease. Although persuasive principles such as social support and dialogue support contribute to the success of most fitness apps, health apps for chronic arthritis patients rarely use them. Several persuasive principles remain unused, leaving opportunities open to support patients in self-managing the disease. In particular, health apps for CA patients would benefit from adding social support techniques (eg, social media, user forums) and extending dialogue support principles (eg, rewards, praise). Tracking of

certain health-related parameters (eg, physical activity, step count) could further reduce the effort needed to manage a disease and thus increase support. This knowledge might inform and inspire developers and health professionals to create apps that move beyond the current state and result in more motivating mHealth apps. Finally, health apps could certainly benefit from a more evidence-based approach, in offering content that can be verified as scientifically proven and endorsed by third parties.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Coding table.

[PDF File (Adobe PDF File), 82KB - mhealth_v4i4e118_app1.pdf]

References

1. Prevalence of Doctor-Diagnosed Arthritis and Arthritis-Attributable Activity Limitation. URL: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6244a1.htm?s_cid=mm6244a1_w [accessed 2016-05-02] [WebCite Cache ID 6hCNYe3g3]
2. Morris W. The American Heritage Dictionary of the English Language. Boston, MA: Houghton Mifflin; 1992.
3. Van den Ende CH, Vlieland TP, Munneke M, Hazes J. Dynamic exercise therapy in rheumatoid arthritis: a systematic review. Br J Rheumatol 1998 Jun;37(6):677-687 [FREE Full text] [Medline: 9667624]
4. Sokol M, McGuigan K, Verbrugge R, Epstein R. Impact of medication adherence on hospitalization risk and healthcare cost. Med Care 2005 Jun;43(6):521-530. [Medline: 15908846]
5. Balkrishnan R. The importance of medication adherence in improving chronic-disease related outcomes: what we know and what we need to further know. Med Care 2005 Jun;43(6):517-520. [Medline: 15908845]
6. Knittle K, DeGucht V, Hurkmans E, Peeters A, Ronday K, Maes S, et al. Targeting motivation and self-regulation to increase physical activity among patients with rheumatoid arthritis: a randomised controlled trial. Clin Rheumatol 2015 Feb;34(2):231-238. [doi: 10.1007/s10067-013-2425-x] [Medline: 24213780]
7. Roddy E, Zhang W, Doherty M, Arden NK, Barlow J, Birrell F, et al. Evidence-based recommendations for the role of exercise in the management of osteoarthritis of the hip or knee--the MOVE consensus. Rheumatology (Oxford) 2005 Jan;44(1):67-73 [FREE Full text] [doi: 10.1093/rheumatology/keh399] [Medline: 15353613]
8. Fernandes L, Hagen KB, Bijlsma JWJ, Andreassen O, Christensen P, Conaghan PG, European League Against Rheumatism (EULAR). EULAR recommendations for the non-pharmacological core management of hip and knee osteoarthritis. Ann Rheum Dis 2013 Jul;72(7):1125-1135 [FREE Full text] [doi: 10.1136/annrheumdis-2012-202745] [Medline: 23595142]
9. Brodin N, Eurenius E, Jensen I, Nisell R, Opava C. Coaching patients with early rheumatoid arthritis to healthy physical activity: a multicenter, randomized, controlled study. Arthritis Rheum 2008 Mar 15;59(3):325-331 [FREE Full text] [doi: 10.1002/art.23327] [Medline: 18311770]
10. Haskell W, Lee I, Pate R, Powell K, Blair S, Franklin B, 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]
11. van den Berg MH, Ronday H, Peeters A, leCessie S, van der Giesen FJ, Breedveld F, et al. Using internet technology to deliver a home-based physical activity intervention for patients with rheumatoid arthritis: A randomized controlled trial. Arthritis Rheum 2006 Dec 15;55(6):935-945 [FREE Full text] [doi: 10.1002/art.22339] [Medline: 17139640]
12. Pradhan E, Baumgarten M, Langenberg P, Handwerger B, Gilpin A, Magyari T, et al. Effect of Mindfulness-Based Stress Reduction in rheumatoid arthritis patients. Arthritis Rheum 2007 Oct 15;57(7):1134-1142 [FREE Full text] [doi: 10.1002/art.23010] [Medline: 17907231]
13. Radojevic V, Nicassio P, Weisman M. Behavioral intervention with and without family support for rheumatoid arthritis. Behavior Therapy 1992;23(1):13-30. [doi: 10.1016/S0005-7894(05)80305-3]
14. Riemsma R, Taal E, Rasker J. Group education for patients with rheumatoid arthritis and their partners. Arthritis Rheum 2003 Aug 15;49(4):556-566 [FREE Full text] [doi: 10.1002/art.11207] [Medline: 12910564]
15. Brittain D, Gyurcsik N, McElroy M, Hillard S. General and arthritis-specific barriers to moderate physical activity in women with arthritis. Womens Health Issues 2011;21(1):57-63. [doi: 10.1016/j.whi.2010.07.010] [Medline: 20833069]
16. Gyurcsik N, Brawley L, Spink K, Brittain D, Fuller D, Chad K. Physical activity in women with arthritis: examining perceived barriers and self-regulatory efficacy to cope. Arthritis Rheum 2009 Aug 15;61(8):1087-1094 [FREE Full text] [doi: 10.1002/art.24697] [Medline: 19644901]
17. IMS Institute for Healthcare Informatics. IMS Institute. 2015 Sep. Patient Adoption of mHealth URL: http://www.imshealth.com/files/web/IMSH%20Institute/Reports/Patient%20Adoption%20of%20mHealth/IIHI_Patient_Adoption_of_mHealth.pdf [accessed 2016-10-06] [WebCite Cache ID 6l3pVbVha]

XSL•FO
RenderX

18. El Khaddar MA, Harroud H, Boulmalf M, Elkoutbi M, Habbani A. Emerging wireless technologies in e-health trends, challenges, and framework design issues. 2012 Presented at: IEEE. International Conference on Multimedia Computing and Systems (ICMCS); May 10-12, 2012; Tangier, Morocco URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6320276 [doi: 10.1109/ICMCS.2012.6320276]

19. Azevedo R, Bernardes M, Fonseca J, Lima A. Smartphone application for rheumatoid arthritis self-management: cross-sectional study revealed the usefulness, willingness to use and patients' needs. Rheumatol Int 2015 Oct;35(10):1675-1685. [doi: 10.1007/s00296-015-3270-9] [Medline: 25903352]

20. Yamada M, Aoyama T, Mori S, Nishiguchi S, Okamoto K, Ito T, et al. Objective assessment of abnormal gait in patients with rheumatoid arthritis using a smartphone. Rheumatol Int 2012 Dec;32(12):3869-3874. [doi: 10.1007/s00296-011-2283-2] [Medline: 22193221]

21. Shinohara A, Ito T, Ura T, Nishiguchi S, Ito H, Yamada M, et al. Development of lifelog sharing system for rheumatoid arthritis patients using smartphone. Conf Proc IEEE Eng Med Biol Soc 2013;2013:7266-7269. [doi: 10.1109/EMBC.2013.6611235] [Medline: 24111422]

22. eMarketer. Smartphone Users Worldwide Will Total 1.75 Billion in 2014. URL: http://www.emarketer.com/Article/Smartphone-Users-Worldwide-Will-Total-175-Billion-2014/1010536 [accessed 2016-05-02] [WebCite Cache ID 6hCNrKs5J]

23. Statista. U.S. smartphone penetration 2010-2017. URL: http://www.statista.com/statistics/201183/forecast-of-smartphone-penetration-in-the-us/ [accessed 2016-05-02] [WebCite Cache ID 6hCNyiday]

24. Statista. Smartphone user penetration Western Europe 2011-2018. URL: http://www.statista.com/statistics/203722/smartphone-penetration-per-capita-in-western-europe-since-2000/ [accessed 2016-05-02] [WebCite Cache ID 6hCO4YsqU]

25. Azevedo AN, de Sousa HML, Monteiro JAF, Lima ARNP. Future perspectives of Smartphone applications for rheumatic diseases self-management. Rheumatol Int 2015 Mar;35(3):419-431. [doi: 10.1007/s00296-014-3117-9] [Medline: 25168866]

26. Singh A, Klapper A, Jia J, Fidalgo A, Tajadura-Jiménez A, Kanakam N. Motivating people with chronic pain to do physical activity: opportunities for technology design. In: CHI 14. 2014 Presented at: ACM: SIGCHI Conference on Human Factors in Computing Systems; April 26-May 1, 2014; Toronto, Canada p. 2803-2812. [doi: 10.1145/2556288.2557268]

27. Swann-Sternberg T, Singh A, Bianchi-Berthouze N, Williams A. User needs for technology supporting physical activity in chronic pain. In: CHI 12. 2012 Presented at: ACM: SIGCHI Conference on Human Factors in Computing Systems; May 5-10, 2012; Austin, TX p. 2441-2446. [doi: 10.1145/2212776.2223816]

28. Revenäs Å, Martin C, Brusewitz M, Keller C, Åsenlöf P. A Mobile Internet Service for Self-Management of Physical Activity in People With Rheumatoid Arthritis: Challenges in Advancing the Co-Design Process During the Requirements Specification Phase. JMIR Res Protoc 2015;4(3):e111 [FREE Full text] [doi: 10.2196/resprot.4824] [Medline: 26381221]

29. Wolfe F, Michaud K. Predicting depression in rheumatoid arthritis: the signal importance of pain extent and fatigue, and comorbidity. Arthritis Rheum 2009 May 15;61(5):667-673 [FREE Full text] [doi: 10.1002/art.24428] [Medline: 19404997]

30. Tsang A, Von Korff M, Lee S, Alonso J, Karam E, Angermeyer M. Common chronic pain conditions in developed and developing countries: gender and age differences and comorbidity with depression-anxiety disorders. J Pain 2008;9(10):891. [doi: 10.1016/j.jpain.2008.05.005]

31. Beckham JC, Keefe FJ, Caldwell DS, Roodman AA. Pain coping strategies in rheumatoid arthritis: Relationships to pain, disability, depression and daily hassles. Behavior Therapy 1991;22(1):113-124. [doi: 10.1016/S0005-7894(05)80249-7]

32. Oinas-Kukkonen H, Harjumaa M. Persuasive systems design: Key issues, process model, and system features. Commun Assoc Inf Syst 2009;24(1):28.

33. Oinas-Kukkonen H, Harjumaa M. Towards Deeper Understanding of Persuasion in Software and Information Systems. In: ACHI '08. 2008 Presented at: IEEE: First International Conference on Advances in Computer-Human Interaction; St. Luce, Martinique; Feb. 10-15, 2008 URL: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4455982 [doi: 10.1109/ACHI.2008.31]

34. Tomlinson M, Rotheram-Borus M, Swartz L, Tsai A. Scaling Up mHealth: Where Is the Evidence? PLoS Med 2013;10(2):-. [doi: 10.1371/journal.pmed.1001382]

35. van Gemert-Pijnen J, Nijland N, van Limburg M, Ossebaard H, Kelders S, Eysenbach G, et al. A holistic framework to improve the uptake and impact of eHealth technologies. J Med Internet Res 2011;13(4):e111 [FREE Full text] [doi: 10.2196/jmir.1672] [Medline: 22155738]

36. Middelweerd A, Mollee J, van der Wal N, Brug J, Te Velde SJ. Apps to promote physical activity among adults: a review and content analysis. Int J Behav Nutr Phys Act 2014;11:97 [FREE Full text] [doi: 10.1186/s12966-014-0097-9] [Medline: 25059981]

37. 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;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]

38. Zajonc RB. Social Facilitation. Science 1965 Jul 16;149(3681):269-274. [Medline: 14300526]

39. Fogg BJ. Persuasive technology. Ubiquity 2002 Dec 01;2002(December):2. [doi: 10.1145/764008.763957]

40. 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]

41. Skinner F. Operant behavior. Am Psych 1963;18(8):503-515. [doi: 10.1037/h0045185]

42. Bandura A. Social Foundations of Thought and Action: A Social Cognitive Theory. Englewood Cliffs, NJ: Prentice-Hall; 1986.

43. Ajzen I. From intentions to actions: A theory of planned behavior. In: Kuhl PDJ, Beckmann DJ, editors. Social Psychology. Berlin: Springer; 1985.

44. Kelders S, Kok R, Ossebaard H, Van Gemert-Pijnen J. 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]

45. Lehto T, Oinas-Kukkonen H. Persuasive features in web-based alcohol and smoking interventions: a systematic review of the literature. J Med Internet Res 2011;13(3):e46 [FREE Full text] [doi: 10.2196/jmir.1559] [Medline: 21795238]

46. Webb T, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. J Med Internet Res 2010;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]

47. 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]

48. Geuens J, Geurts L, Swinnen TW, Huysmans J, Peulen J, Westhovens R, et al. Systematic Review of Persuasive Principles Used in Technological Systems to Support Physical Therapy of Musculoskeletal Diseases. 2015 Jan 1 Presented at: Belgian Congress on Rheumatology; September 23-25, 2015; Genk, Belgium URL: https://lirias.kuleuven.be/handle/123456789/505709

49. Mosa ASM, Yoo I, Sheets L. A systematic review of healthcare applications for smartphones. BMC Med Inform Decis Mak 2012;12:67 [FREE Full text] [doi: 10.1186/1472-6947-12-67] [Medline: 22781312]

50. 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]

51. Donker T, Petrie K, Proudfoot J, Clarke J, Birch M, Christensen H. Smartphones for smarter delivery of mental health programs: a systematic review. J Med Internet Res 2013;15(11):e247 [FREE Full text] [doi: 10.2196/jmir.2791] [Medline: 24240579]

52. 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]

53. Lombard M, Snyder-Duch J, Bracken C. Content analysis in mass communication: assessment and reporting of intercoder reliability. Human Comm Res 2002 Oct;28(4):587-604. [doi: 10.1111/j.1468-2958.2002.tb00826.x]

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

55. myhealthapps. Apps tried and tested by people like you. URL: http://myhealthapps.net/ [accessed 2016-05-02] [WebCite Cache ID 6hCPISwyg]

56. Deterding S. Gamification. interactions 2012 Jul 01;19(4):14. [doi: 10.1145/2212877.2212883]

57. Sailer M, Hense J, Mandl H, Klevers M. Psychological Perspectives on Motivation through Gamification. URL: http://www.fml.mw.tum.de/fml/images/Publikationen/19_2.pdf [accessed 2014-11-25] [WebCite Cache ID 6hE8AWOrq]

58. Zichermann G, Cunningham C. Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps. Sebastopol, CA: O'Reilly Media; 2011.

59. Cafazzo J, Casselman M, Hamming N, Katzman D, Palmert M. Design of an mHealth app for the self-management of adolescent type 1 diabetes: a pilot study. J Med Internet Res 2012;14(3):e70 [FREE Full text] [doi: 10.2196/jmir.2058] [Medline: 22564332]

60. Uchino BN, Cacioppo JT, Kiecolt-Glaser JK. The relationship between social support and physiological processes: a review with emphasis on underlying mechanisms and implications for health. Psychol Bull 1996 May;119(3):488-531. [Medline: 8668748]

61. Kaplan B, Cassel J, Gore S. Social support and health. Med Care 1977 May;15(5 SUPPL):47-58. [Medline: 853781]

62. Deci E, Ryan R. Intrinsic Motivation and Self-Determination in Human Behavior. New York, NY: Plenum; 1985.

63. Fishbein M. A theory of reasoned action: some applications and implications. Nebr Symp Motiv 1980;27:65-116. [Medline: 7242751]

64. Fishbein M, Ajzen I. Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research. Reading, MA: Addison-Wesley Pub. Co; 1975.

65. Cialdini R, Trost M. Social influence: Social norms, conformity and compliance. In: Handbook of Social Psychology, Vols 1-2, 4th ed. New York, NY: McGraw-Hill; 1998:151-192.

66. PatientsLikeMe. Live better, together! URL: https://www.patientslikeme.com/ [accessed 2016-05-02] [WebCite Cache ID 6hCPRStHM]

67. Lorig K, Mazonson P, Holman H. Evidence suggesting that health education for self-management in patients with chronic arthritis has sustained health benefits while reducing health care costs. Arthritis Rheum 1993;36(4):439-446.

XSL•FO

RenderX

68. Fisher J, Fisher W, Amico K, Harman J. An Information-Motivation-Behavioral Skills Model of Adherence to Antiretroviral Therapy. Health Psychol 2006;25(4):462-473. [Medline: 16846321]

69. Fogg B, Hreha J. Behavior wizard: a method for matching target behaviors with solutions. In: PERSUASIVE 10.: Springer; 2010 Presented at: 5th International Conference on Persuasive Technology; June 7-10, 2010; Denmark p. 117-131 URL: http://link.springer.com/chapter/10.1007/978-3-642-13226-1_13 [doi: 10.1007/978-3-642-13226-1_13]

Abbreviations

BCT: behavior change technique
CA: chronic arthritis
DAS: Disease Activity Score
OA: osteoarthritis
PSD: Persuasive Systems Design
RA: rheumatoid arthritis
SpA: spondyloarthritis

Review

Safe Sex Messages Within Dating and Entertainment Smartphone Apps: A Review

Evelyn Tzu-Yen Huang[1,2], MPH, MD; Henrietta Williams[1,3], MBBS, MPH; Jane S Hocking[1], BAppSc, MHlthSc, MPH, PhD; Megan SC Lim[4,5], B BiomedSci (Hons), PhD

[1]Melbourne School of Population and Global Health, University of Melbourne, Melbourne, Australia

[2]Austin Hospital, Austin Health, Heidelberg, Australia

[3]Melbourne Sexual Health Centre, Alfred Health, Carlton, Australia

[4]Centre for Population Health, Burnet Institute, Melbourne, Australia

[5]Department of Epidemiology and Preventive Medicine, Monash University, Melbourne, Australia

Corresponding Author:
Evelyn Tzu-Yen Huang, MPH, MD
Austin Hospital
Austin Health
145 Studley Road
Heidelberg, 3084
Australia
Phone: 61 3 9496 5000
Fax: 61 3 9458 4779
Email: drevelynhuang@gmail.com

Abstract

Background: Smartphone apps provide a new platform for entertainment, information distribution, and health promotion activities, as well as for dating and casual sexual encounters. Previous research has shown high acceptability of sexual health interventions via smartphone apps; however, sexual health promotion apps were infrequently downloaded and underused. Integrating sexual health promotion into established apps might be a more effective method.

Objective: The objective of our study was to critically review popular sex-related apps and dating apps, in order to ascertain whether they contain any sexual health content.

Methods: Part 1: In January 2015, we used the term "sexual" to search for free apps in the Apple iTunes store and Android Google Play store, and categorized the sexual health content of the 137 apps identified. Part 2: We used the term "dating" to search for free geosocial-networking apps in the Apple iTunes and Android Google Play stores. The apps were downloaded to test functionality and to determine whether they included sexual health content.

Results: Part 1: Of the 137 apps identified, 15 (11.0%) had sexual health content and 15 (11.0%) contained messages about sexual assault or violence. The majority of the apps did not contain any sexual health content. Part 2: We reviewed 60 dating apps: 44 (73%) targeting heterosexual users, 9 (15%) targeting men who have sex with men (MSM), 3 (5%) targeting lesbian women, and 4 (7%) for group dating. Only 9 dating apps contained sexual health content, of which 7 targeted MSM.

Conclusions: The majority of sex-related apps and dating apps contained no sexual health content that could educate users about and remind them of their sexual risks. Sexual health practitioners and public health departments will need to work with app developers to promote sexual health within existing popular apps. For those apps that already contain sexual health messages, further study to investigate the effectiveness of the content is needed.

Introduction

In recent years, the number of smartphone users has surged across the world and downloads of smartphone apps have grown significantly [1]. Nielsen's monthly survey found that 71% of American mobile phone users owned a smartphone by mid-2014 [2]. Smartphone apps provide a new platform for information distribution and networking. By 2013, there were over 50 billion app downloads from the Google Play store, and more than 60 billion from the Apple iTunes store [3,4]. This platform creates various opportunities for health promotion activities such as distributing health-related information, offering resources for health care, and providing forums for sharing experiences [1,5,6]. The benefits of using apps for health promotion are many, including low cost to develop and operate, potentially widespread distribution, and convenience for both health care providers and health care seekers [1].

The availability of geosocial-networking smartphone apps—apps that use the global positioning system to locate their subscribers—has created a novel way of networking that is quick, cheap, and convenient [7-10]. Users can easily identify other users by physical proximity. While these geosocial-networking apps can be used for forming friendships and building a community, they are frequently used for dating and to facilitate the process of finding sexual partners [7,8,11,12]. By filtering user profiles, such as age, appearance, and interests, subscribers can select the type of partners they seek [13]. Grindr, a popular dating app that targets men who have sex with men (MSM), had more than 7 million subscribers globally by 2013, and the number is increasing [14]. Previous studies have found that users of these dating apps report more sexual contacts and more casual sexual partners [7,8,10-12,15]. Users also reveal significant increases in casual sex since starting online dating [16]. Rice et al reported that 75% of Grindr subscribers had had sex with people they met through the app, and 15% reported unprotected anal sex with sexual partners from Grindr [8]. The likelihood of young MSM engaging in unprotected anal intercourse was 3 times higher among Grindr users than nonusers [11], and users reported a higher prevalence than nonusers of ever being diagnosed with sexually transmissible infections (STIs) [9].

Sexual health interventions that are integrated with modern technologies have been successful [17-22]. Text messaging has been widely used to promote sexual health, including appointment reminders, partner tracing, and result notification [22]. Studies of Internet-based human immunodeficiency virus (HIV) infection interventions targeting MSM using online questionnaires and tutorial sessions revealed a reduced rate of unprotected anal intercourse and increased condom use [20,21]. In recent years, due to the increasing use of smartphones, apps designed to provide sexual health information and education are readily available on the market [23]. However, these apps are infrequently downloaded, have low user ratings, and are unlikely to reach the target groups [23].

Rather than creating new sexual health apps, leveraging established and popular apps may improve the distribution of health promotion information to a larger number of users [15].

Most important, integrating sexual health information within these apps can be an effective way to reach key populations, such as MSM or people who have casual sexual partners [15,24]. In addition, it is possible for health professionals to harvest global positioning system data from the apps and provide services according to users' physical locations, such as referral to STI testing centers [15]. Several studies have suggested that young adults consider this approach acceptable [10,15,24].

The aim of this study was to systematically and critically review free sex-related apps (including all apps that have sexual content, such as sexual entertainment, sexual health information, and sex enhancement) and popular free dating apps, determine whether they contain any sexual health content, and, if so, what kind of information they provide to educate users about sexual health.

Methods

This review was conducted in 2 parts: a review of sex-related apps and a review of dating apps. Ethical approval was not required, as the research did not involve participants.

Part 1: Review of Sex-Related Apps

Search and Inclusion/Exclusion Criteria

The first part of the study was a content analysis of free sex-related apps. We used the term "sexual" to search the Android Google Play marketplace (Google Inc, Mountain View, CA, USA) and Apple iTunes store (Apple Inc, Cupertino, CA, USA) in January 2015. We conducted the search under the stores' default algorithms, except that we filtered to search for free apps in the Android Google Play marketplace (as the choice was available). The search yielded 250 apps from Android Google Play marketplace and 263 apps from the Apple iTunes store. We then excluded all the paid apps from the Apple iTunes store. Apps were also excluded if they did not have an English-language interface, if they served the function of online dating, or if they were not related to sex (eg, the search found a "find your phone" app).

Data Extraction and Review Methods

We recorded the following information from the individual apps during the review: the app store category (eg, health and fitness, games, education), the app developer, and the user rating (the average of individual user ratings of 1 to 5). Apps were classified according to their primary purposes: "sex aid or sexual exploration" covered apps that provide ideas about sexual positioning or foreplay; "entertainment" included game apps or apps that calculate sexual compatibility based on horoscope; "sex education/information" encompassed apps designed to provide sexual health knowledge; "sexual assault/violence" included apps with the primary purpose of tackling sexual assault or violence or helping the victims of sexual violence; and "other" covered apps that did not fit into the above categories, including period tracking apps and apps for sex offender registries. Chi-square or Fisher exact tests compared the main purposes of the apps and presence of sexual health content between iTunes and Google Play apps. Apps were downloaded and reviewed in February or March 2015 by a

XSL•FO

RenderX

single reviewer (ETH). All the functions of each app were tested and the outcome was categorized as having "sexual health content," "sexual assault/violence information," or "none" (Textbox 1).

Part 2: Review of Dating Apps

Search and Inclusion/Exclusion Criteria

The second part of the study was a review of popular dating apps. We used the term "dating" to search the Apple iTunes store and the Android Google Play marketplace in January 2015. The first 50 free dating apps from each store were included. Apps requiring in-app purchase for basic functions such as receiving messages and online chats were excluded. We included an extra 3 lesbian dating apps that were available in both stores and had the most downloads according to the download numbers available in the Android Google Play marketplace.

Data Extraction and Review Methods

We downloaded and reviewed the apps by creating a user profile and testing the apps' functions during April and May 2015. We used 1 iPhone (Apple Inc) and 1 Android phone (HTC; HTC Corporation, Taoyuan, Taiwan) to test the functions of each app. A female profile was created for each heterosexual app and lesbian app, and a male profile was created for MSM apps. We classified apps as containing no sexual health content and no safe dating tips if we found no relevant information after testing all the functions of the app, and logging in and out on 5 separate days. The following information was extracted from the apps: app store category (social, social networking, and lifestyle), user rating, and the name of the app's developers. Apps were categorized into 4 groups based on their primary target groups as heterosexual, MSM, lesbian, and other (apps for seeking threesomes or group dates). Chi-square or Fisher exact test compared the presence of sexual health content between apps with different target groups.

Textbox 1. Definition of categories of sex-related apps.

- Sexual health content
    - Information about sexually transmissible infections (STIs)
    - STI testing information or resources
    - Information about condom use or assistance locating condoms
    - Information about contraception
- Sexual assault/violence content
    - Identification of signs of sexual assault
    - Prevention of sexual assault
    - Medical and psychological care after sexual assault
    - Information about sexual assault
- None
    - Containing none of the information listed above

Results

Part 1: Sex-Related Apps

Our search yielded 250 apps from the Android Google Play marketplace and 263 apps from the Apple iTunes store. Ultimately, 137 apps were shortlisted for review (Figure 1). Of the 137 apps reviewed, the most common app purpose was sex aid and sexual exploration apps, which included information or advice on sexual positions (such as Kama Sutra apps) and apps that provided tips and ideas for foreplay and other techniques for promoting sexual pleasure (n=42, 30.7%). Other common categories were entertainment apps (n=32, 23.4%), apps relating to sexual assault (n=19, 13.9%), and apps for sexual education and information (n=12, 8.8%) (Table 1). A total of 15 apps (11.0%) included any sexual health content, and 15 apps (11.0%) contained sexual assault or violence content. iTunes apps were more likely than Google Play apps to have apps for the purpose of sex aids and sexual exploration or sexual assault (P=.01). Most of the apps (n=107, 78.1%) did not contain any sexual health content; that is, information about STIs, STI testing, condom use or assistance locating condoms, or contraception. There was no statistically significant difference in sexual health content between iTunes and Google Play apps (P=.06). Among the 15 apps that contained sexual health content, 5 (33%) had both contraception and STI information, 4 (27%) contained contraception information, and 6 (40%) contained information about STI and condom use for STI prevention. Most of the apps containing sexual health content were from the sex education and information category (n=11, 73%). The remaining sexual health information-containing apps were distributed as follows: 2 sex aid/sexual exploration apps, 1 entertainment app, and 1 categorized as other.

Table 1. Sexual health content of sex-related apps (N=137).

Variable	Primary purpose of apps						Sexual health content	
	Sex aid/ exploration	Entertainment	Sex education/ information	Sexual assault	Others	None	Contraception/ STI[a] information	Sexual assault/ violence
Total apps, n (%)								
iTunes	33 (35.8)	21 (22.8)	8 (8.6)	16 (17.3)	14 (15.2)	68 (73.9)	10 (10.8)	14 (15.2)
Google Play	9 (20.0)	11 (24.4)	4 (8.8)	3 (6.6)	18 (40.0)	39 (86.6)	5 (11.1)	1 (2.2)
Apps containing STI or contraception information, n (%)								
iTunes	2 (2.1)	0 (0)	7 (7.6)	0 (0)	1 (1.0)	N/A[b]	N/A	N/A
Google Play	0 (0)	1 (2.2)	4 (8.8)	0 (0)	0 (0)	N/A	N/A	N/A
App store category, n (%)								
Entertainment	14 (53.8)	11 (42.3)	0 (0)	0 (0)	1 (3.8)	26 (100)	0 (0)	0 (0)
Games	3 (42.8)	3 (42.8)	0 (0)	1 (14.2)	0 (0)	6 (85.7)	0 (0)	1 (14.2)
Education	0 (0)	3 (20.0)	3 (20.0)	5 (33.3)	4 (26.6)	10 (66.6)	4 (26.6)	6 (40.0)
Books/ reference	1 (20.0)	1 (20.0)	1 (20.0)	2 (40.0)	0 (0)	3 (60.0)	1 (20.0)	1 (20.0)
Health and fitness	7 (35.0)	0 (0)	4 (20.0)	1 (5.0)	8 (40.0)	16 (80.0)	4 (20.0)	1 (5.0)
Lifestyle	13 (32.5)	11 (27.5)	2 (5.0)	4 (10.0)	10 (25.0)	34 (85.0)	5 (12.5)	2 (5.0)
Medical	0 (0)	1 (10.0)	2 (20.0)	1 (10.0)	6 (60.0)	8 (80.0)	1 (10.0)	1 (10.0)
Tools/ utilities	1 (20.0)	0 (0)	0 (0)	2 (40.0)	2 (40,0)	3 (60.0)	0 (0)	2 (40.0)
Others	2 (22.2)	2 (22,2)	0 (0)	2 (22.2)	3 (33.3)	9 (100)	0 (0)	1 (11.1)
User rating, n (%)								
Unrated	27 (31.7)	21 (24.7)	8 (9.4)	15 (17.6)	14 (16.4)	70 (82.3)	9 (10.5)	14 (16.4)
1–3.9	12 (41.3)	7 (24.1)	2 (6.8)	2 (6.8)	6 (20.6)	25 (86.2)	3 (10.3)	1 (3.4)
4–5	2 (8.6)	4 (17.3)	3 (13.0)	1 (4.3)	13 (56.5)	20 (86.9)	3 (13.0)	0 (0)

[a]STI: sexually transmissible infections.

[b]N/A: not applicable.

Of the 15 apps offering information about sexual assault, 5 (33%) had information regarding management after sexual assault, 5 (33%) had general information about sexual assault, 1 (7%) focused on identifying sexual assault victims, and the other 4 (27%) had information about sexual assault prevention. However, none of the apps with the primary purpose of sexual assault/violence contained any information about STIs or contraception.

We also recorded the number of downloads of apps from Google Play marketplace; this information is readily available within the marketplace and is displayed in a range (for example: between 1000 and 5000). We found that the 4 Android apps that contained sexual health information were downloaded less frequently than other sex aid or entertainment apps (the number of downloads is available in Multimedia Appendix 1).

Part 2: Dating Apps

We included the first 50 free dating apps from each store, then excluded 25 duplicates from the list (Multimedia Appendix 2). In the analysis we excluded the apps that were not available for download (n=5), that required in-app purchases to proceed to basic functions such as receiving messages and online chats (n=7), and apps that did not function after 3 attempts (n=6).

During initial review we realized that none of these apps targeted lesbian women, so we included the 3 top lesbian dating apps available in both the Apple iTunes store and the Google Play marketplace, found using the search term "lesbian dating." We included 60 apps in the study (Figure 2). Of these, 44 (73%) apps targeted heterosexual users, 9 (15%) targeted MSM, 3 (5%) targeted lesbians, and 4 (7%) were for group dating and finding partners for threesomes (Table 2).

XSL•FO

RenderX

Figure 1. Inclusion of sex-related apps.
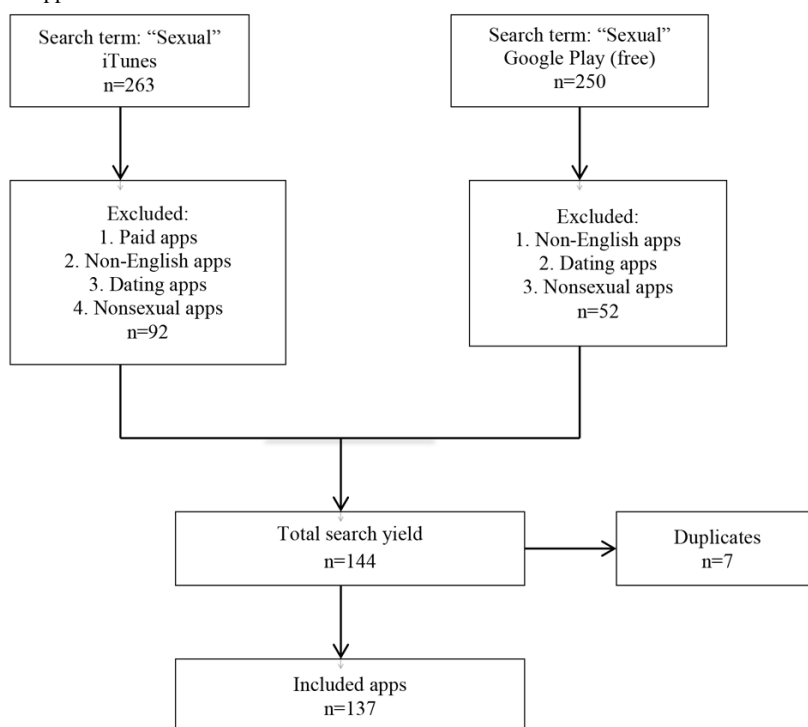
Table 2. Results of dating app review (N=60).

Variable	Target group				Sexual health content	Safe dating tips	Total
	Heterosexual	MSM[a]	Lesbian	Other[b]			
Total apps, n (%)							
iTunes	28 (68)	9 (22)	3 (7)	1 (2)	7 (17)	6 (15)	41 (100)
Google Play	16 (84)	0 (0)	0 (0)	3 (16)	2 (11)	1 (5)	19 (100)
App store category, n (%)							
Social networking	34 (71)	9 (19)	3 (6)	2 (4)	8 (17)	5 (10)	48 (100)
Communication	1 (100)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1 (100)
Entertainment	1 (100)	0 (0)	0 (0)	0 (0)	1 (100)	0 (0)	1 (100)
Lifestyle	8 (80)	0 (0)	0 (0)	2 (20)	0 (0)	1 (10)	10 (100)
User star rating, n (%)							
N/A[c]	22 (65)	8 (24)	3 (9)	1 (3)	5 (15)	5 (15)	34 (100)
1–3.9	8 (89)	0 (0)	0 (0)	1 (11)	1 (11)	0 (0)	9 (100)
4–5	14 (82)	1 (6)	0 (0)	2 (12)	3 (18)	2 (12)	17 (100)

[a]MSM: men who have sex with men.

[b]Other includes threesome or group dating.

[c]N/A: not available.

We found that 9 (15%) of the 60 apps included sexual health content. The sexual health content was displayed in the apps in four different ways (Textbox 2). The majority of the apps with sexual health content targeted MSM (7/9, 78%) (Table 3). Only 1 heterosexual app contained sexual health content (Table 3). None of the 3 lesbian-specific apps contained any sexual health or safe dating content (Table 3). We found safe dating tips in 7 apps (12%), which included information such as "do not disclose your true identity online," "always meet in public places," and "trust one's own instinct." All the apps with safe dating tips targeted heterosexual users. The availability of sexual health content and safe dating tips was related to the target group of the apps ($P<.001$).

Textbox 2. Sexual health content within dating apps.

- Preference of safe sex in users' profiles (always, depends, or never)
- Pop-up messages encouraging sexually transmissible infection (STI) testing
- STI status in users' profiles
- Links to or articles about STI information in apps or websites

Table 3. Availability of sexual health content and safe dating tips in dating and entertainment apps according to target group.

Target group	Content, n (%)			P value	Total no. of apps
	Sexual health content	Safe dating tips	None		
Heterosexual	1 (2)	7 (16)	36 (81)	<.001	44
MSM[a]	7 (78)	0 (0)	2 (22)		9
Lesbian	0 (0)	0 (0)	3 (100)		3
Other	1 (25)	0 (0)	3 (75)		4

[a]MSM: men who have sex with men.

Figure 2. Inclusion of dating apps.
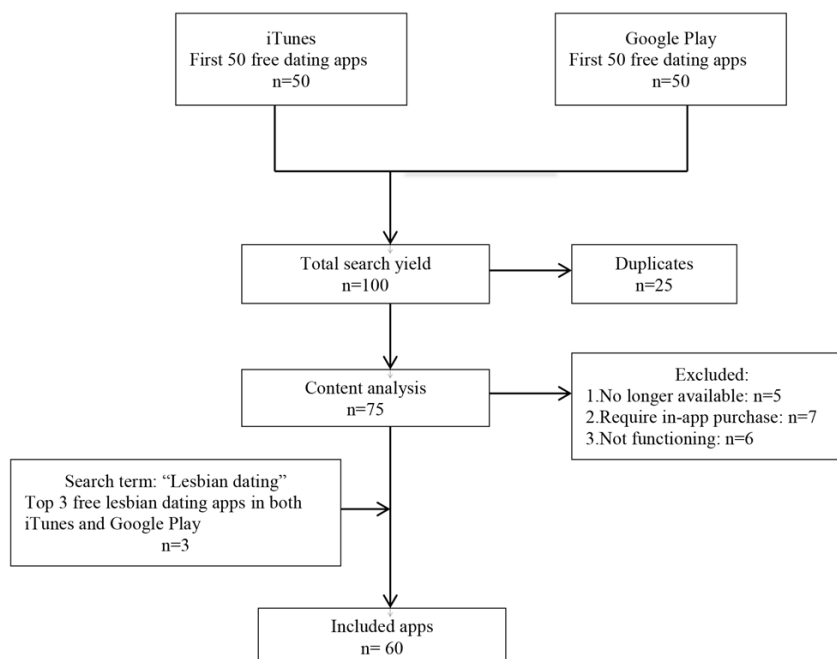
Discussion

Principal Findings

Our review identified a large number of smartphone apps that have sexual content or promote dating. We found that the majority of these sex-related and dating apps contained no information about sexual health promotion. We chose to review all apps found using our search terms, regardless of app store category, rather than focusing on apps that aim to provide sexual health information and education. The reason behind this decision was that previous research has shown that apps designed for educational purposes are infrequently downloaded and underused [23]. We made the same observation in our review. For example, *Sexually Transmitted STDs*, an app that provides information about STIs, including definition, transmission, symptoms, treatments, and prevention, had fewer than 1000 downloads to January 2015. On the other hand, some game apps were downloaded far more frequently; for example, *Bliss –The Game for Lovers* was downloaded more than 100,000 times from the Android Google Play marketplace (this app is also available in iTunes). This finding demonstrates the advantage of integrating sexual health information within popular entertainment apps in order to reach out to more people. We found that 73% of the apps with sexual health content had the primary purpose of education, which means that they might not be attractive to people who are not specifically seeking out sexual health information. Only 2 sex aid apps and 1 entertainment app contained sexual health content, demonstrating room for improvement. We hypothesize that exposing users to sexual health content while they are using these sex-related apps (presumably while they are thinking about

sex or during foreplay) might be a good way to remind them of safe sex practices. This hypothesis will need to be examined in future research.

In the first part of the study, we found that 19 sexual apps had the primary purpose of providing support and information for sexual assault or violence. These apps offered a variety of information about topics including mental health support after sexual assault, prevention of sexual assault, and general information regarding sexual assault. However, none mentioned the potential adverse sexual health outcomes faced by sexual assault victims. The risk of unwanted pregnancy and contracting STIs, and steps that can be taken to prevent or treat these, are critical in the aftercare of sexual assaults [25]. Many victims of sexual assault do not seek help from health professionals [25]; therefore, providing such information via smartphone apps might encourage victims to seek medical advice for STI and pregnancy prevention.

Our review of dating apps showed that very few included any sexual health content (9 of 60 dating apps reviewed, 15%). The majority of these apps were targeted at an MSM population (7 out of 9). These findings suggest that there is more focus on the sexual risks of MSM who use dating apps than other groups. MSM are disproportionately affected by HIV globally and are a key population for HIV infection and STI prevention [26-28]. These geosocial-networking apps can potentially function as an entry point for HIV intervention delivery, assisting health professionals to reach the key populations, particularly where populations are hidden or difficult to access [7,15]. Delivering sexual health interventions via dating apps is an important area to be addressed in prevention of HIV infection.

It is unfortunate that the opportunity to deliver messages to other groups via these apps is being missed. Only 1 of the 44 heterosexual dating apps reviewed had any sexual health content (STI status in users' profile). People who use dating apps seem likely to have more casual sexual partners than people who do not, which means higher risk of contracting STIs [13]. While sexual health content appears to be acceptable to MSM who use dating apps, its acceptability among other groups is unknown [15,24]. More work needs to be done to increase the sexual risk awareness of users of heterosexual dating apps.

Dating apps used four different modes to display sexual health content: pop-up messages, infection status in users' profiles, safe sex preferences in users' profiles, and blog posts or links to sexual health information in the apps or linked websites. Each of these messages has its own limitations in reminding users of their sexual risk. First, the frequency and timing of pop-up messages might influence users' acceptance. If the frequency is too high, it might desensitize users. Messages appearing during chats could cause annoyance, which could lead to users unsubscribing and turning to other apps that exclude these kinds of messages. Second, having HIV or STI status and safe sex preference on a profile can be a good way to assist users' partner filtering processes. Nevertheless, these messages are highly dependent on users' self-reports and their knowledge of infection status. These disclosures may also expose users to stigma and discrimination or cyberbullying [29]. In the 2 apps that enabled indication of preference for safe sex practice, the concept was

not defined. Moreover, since this information appears on users' profiles, users can decide to disclose the status or preference, or not. Third, in-app blog posts can be a good place to display information regarding STIs and HIV if these posts are updated frequently and the information provided is correct. On the other hand, having links to sexual health clinics in the websites rather than within the apps might be less effective, since users have to be actively looking for sexual health information and using the website at the same time to be exposed to these messages. This type of message is less likely to effectively remind users about their sexual risks. Further research is needed to understand the impact these messages have on users' behavior and health outcomes.

We identified another potential platform for intervention during the app review: the advertisement space within apps. Advertisements (ads) mostly exist in two forms: pop-up ads and ads that appear on the bottom of the screen; users can close pop-up ads, but they usually cannot remove bottom-of-the-screen ads. Once users click on the ads, they will usually be directed to a new page that contains more information about the product being advertised (most likely another paid app). Health promoters could purchase these ad spaces to display sexual health information or links. Some app developers sell ad spaces as pop-ups for advertisers to purchase. For example, Grindr sells mobile Web banner ads, which can link directly to advertisers' websites, emails, or mobile numbers [30]. These ads are sold as cost per thousand banner impressions, with the price ranging from USD $9 to $25 per thousand banner impressions for iPhone and Android devices. It is potentially a cost-effective way of promoting sexual health, as it is cheaper and more focused on target groups than traditional media ads. However, the limitation of this method is that users are usually encouraged to subscribe to premium membership (by paying a monthly fee or upgrading to the paid version of the apps) in order to avoid seeing the ads. Once users upgrade to the paid versions, they might no longer be exposed to sexual health information through this medium. More research is needed to evaluate the effectiveness of advertising through in-app ads and how to make sure all users receive the messages being advertised.

Research has shown that 80% of Internet users in the United States search online for health information, and that young people are gathering health information using mobile devices with increasing frequency, including sexual health information [31]. However, while new technologies, including smartphone apps, are used to facilitate health information seeking, health-related apps are infrequently downloaded and rarely used [32]. This suggests that, to promote sexual health through smartphone apps, researchers could partner with app developers in order to integrate sexual health promotion interventions in popular sex-related or dating apps [24]. Such partnerships will be difficult to form when the interests of the parties conflict. For dating app developers, sexual health content that reminds users of their sexual risk might be unattractive, as it could jeopardize their popularity among users [24]. However, it is evident that these apps provide novel opportunities to engage at-risk populations in sexual health interventions [7,8,11,15,24].

Several studies have suggested that young adults consider sexual health promotion via apps acceptable [10,15,24]. Sun et al found that approximately two-thirds of MSM were willing to receive sexual health-related information through apps, and 26% of them requested referrals for HIV and STI testing [15]. The willingness to participate in future HIV infection and STI prevention programs is even higher among MSM aged between 19 and 24 years [24]. Holloway et al found that 80% of young MSM recruited through Grindr expressed an interest in joining such programs, and 71% preferred to have the information delivered through smartphone apps [24].

Limitations and Strengths

Our study had some limitations. First, smartphone apps are changing rapidly, including their content, popularity, and even availability. The ranking of popularity varies over time; therefore, our search results might be different if repeated. Updates of the apps can change apps' features and functions, including the sexual health content that we looked for. For example, since our review, Tinder has agreed to provide information for STI testing locations [33]. Second, our categorization of pop-up sexual health content might have been inaccurate: we could have missed infrequent pop-up messages, or those appearing only around major events. We used only 1 device for each platform, which prevented us noticing variation in app function between devices (if any). We also did not identify any differences in the frequency or availability of sexual health content using different profiles. Third, the terms "sexual"

and "dating" used to search app stores for sex-related apps and dating apps may have constrained our search. Other terms such as "sex" or "networking" could be considered for future searches. Fourth, our search was limited to the Apple iTunes store and the Android Google Play store, and thus neglected apps from other smartphone operating systems (eg, Microsoft, Palm, Blackberry). However, this decision was justified by the fact that 96% of smartphone users worldwide use either Apple or Google operating systems [34].

Despite these limitations, this study is, to our knowledge, the first to review the inclusion of sexual health content within sexual and dating apps that are not primarily aimed at sex education. We are unsure how much influence these messages have on users. Further research in this field is needed to understand the effectiveness and efficiency of promoting sexual health through in-app messages.

Conclusions

The majority of sex-related and dating smartphone apps do not contain any sexual health content, with the exception of dating apps targeting MSM. Using smartphone apps to promote sexual health is a potentially important method of reaching at-risk populations. Due to the low rate of integration of sexual health information in dating apps and sex-related entertainment apps, we suggest that sexual health researchers work with app developers to promote sexual health within existing popular apps. Further investigation of the acceptability and effectiveness of sexual health content in sexual and dating apps is needed.

Conflicts of Interest

None declared.

Multimedia Appendix 1

List of sexual apps.

[XLSX File (Microsoft Excel File), 52KB - mhealth_v4i4e124_app1.xlsx]

Multimedia Appendix 2

List of dating apps.

[XLSX File (Microsoft Excel File), 37KB - mhealth_v4i4e124_app2.xlsx]

References

1. 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]
2. Nielsen. Mobile Millennials: Over 85% of Generation Y Owns Smartphones. New York, NY: The Nielsen Company; 2014. URL: http://www.nielsen.com/us/en/insights/news/2014/mobile-millennials-over-85-percent-of-generation-y-owns-smartphones.html [accessed 2015-08-31] [WebCite Cache ID 6bBh6BW7V]

3. Statistica. Cumulative Number of Apps Downloaded From the Google Play Android App Store as of July 2013. URL: http://www.statista.com/statistics/281106/number-of-android-app-downloads-from-google-play/ [accessed 2015-08-31] [WebCite Cache ID 6bBhmRR32]

4. Statistica. Cumulative Number of Apps Downloaded From the Apple App Store From July 2008 to June 2015. URL: http://www.statista.com/statistics/263794/number-of-downloads-from-the-apple-app-store/ [accessed 2015-08-31] [WebCite Cache ID 6bBhrE6Xl]

5. Shen N, Levitan M, Johnson A, Bender JL, Hamilton-Page M, Jadad AA, et al. Finding a depression app: a review and content analysis of the depression app marketplace. JMIR Mhealth Uhealth 2015;3(1):e16 [FREE Full text] [doi: 10.2196/mhealth.3713] [Medline: 25689790]

6. 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;1(1):e9 [FREE Full text] [doi: 10.2196/mhealth.2709] [Medline: 25100681]

7. Goedel WC, Duncan DT. Geosocial-networking app usage patterns of gay, bisexual, and other men who have sex with men: survey among users of Grindr, a mobile dating app. JMIR Public Health Surveill 2015;1(1):e4 [FREE Full text] [doi: 10.2196/publichealth.4353] [Medline: 27227127]

8. Rice E, Holloway IW, Winetrobe H, Rhoades H, Barman-Adhikari A, Gibbs J, et al. Sex risk among young men who have sex with men who use Grindr, a smartphone geosocial networking application. J AIDS Clinic Res 2012;01(S4). [doi: 10.4172/2155-6113.S4-005]

9. Lehmiller JJ, Ioerger M. Social networking smartphone applications and sexual health outcomes among men who have sex with men. PLoS One 2014;9(1):e86603 [FREE Full text] [doi: 10.1371/journal.pone.0086603] [Medline: 24466166]

10. Phillips G, Magnus M, Kuo I, Rawls A, Peterson J, Jia Y, et al. Use of geosocial networking (GSN) mobile phone applications to find men for sex by men who have sex with men (MSM) in Washington, DC. AIDS Behav 2014 Sep;18(9):1630-1637. [doi: 10.1007/s10461-014-0760-9] [Medline: 24682866]

11. Winetrobe H, Rice E, Bauermeister J, Petering R, Holloway IW. Associations of unprotected anal intercourse with Grindr-met partners among Grindr-using young men who have sex with men in Los Angeles. AIDS Care 2014;26(10):1303-1308. [doi: 10.1080/09540121.2014.911811] [Medline: 24754563]

12. Landovitz RJ, Tseng C, Weissman M, Haymer M, Mendenhall B, Rogers K, et al. Epidemiology, sexual risk behavior, and HIV prevention practices of men who have sex with men using GRINDR in Los Angeles, California. J Urban Health 2013 Aug;90(4):729-739 [FREE Full text] [doi: 10.1007/s11524-012-9766-7] [Medline: 22983721]

13. Lim M, Bowring A, Carrotte E, Hellard M. Sex, Drugs and Rock 'n' Roll 2015 Results. Melbourne, Australia: Burnet Institute; 2015. URL: https://http://www.burnet.edu.au/system/asset/file/1755/2015_Web_Summary__FINAL.pdf [accessed 2015-09-10] [WebCite Cache ID 6bR7WYTiA]

14. Grindr. Press & Media. Los Angeles, CA: Grindr LLC; 2015. URL: http://www.grindr.com/press/?lang=zh-hans [accessed 2016-10-27] [WebCite Cache ID 6lZK8u4u2]

15. Sun CJ, Stowers J, Miller C, Bachmann LH, Rhodes SD. Acceptability and feasibility of using established geosocial and sexual networking mobile applications to promote HIV and STD testing among men who have sex with men. AIDS Behav 2015 Mar;19(3):543-552. [doi: 10.1007/s10461-014-0942-5] [Medline: 25381563]

16. Couch D, Liamputtong P. Online dating and mating: the use of the internet to meet sexual partners. Qual Health Res 2008 Feb;18(2):268-279. [doi: 10.1177/1049732307312832] [Medline: 18216345]

17. Gold J, Pedrana AE, Sacks-Davis R, Hellard ME, Chang S, Howard S, et al. A systematic examination of the use of online social networking sites for sexual health promotion. BMC Public Health 2011;11:583 [FREE Full text] [doi: 10.1186/1471-2458-11-583] [Medline: 21777470]

18. Veale HJ, Sacks-Davis R, Weaver ER, Pedrana AE, Stoové MA, Hellard ME. The use of social networking platforms for sexual health promotion: identifying key strategies for successful user engagement. BMC Public Health 2015;15:85 [FREE Full text] [doi: 10.1186/s12889-015-1396-z] [Medline: 25884461]

19. Bailey J, Mann S, Wayal S, Abraham C, Murray E. Digital media interventions for sexual health promotion-opportunities and challenges: a great way to reach people, particularly those at increased risk of sexual ill health. BMJ 2015 Mar 03;350:h1099. [Medline: 25736806]

20. Carpenter KM, Stoner SA, Mikko AN, Dhanak LP, Parsons JT. Efficacy of a web-based intervention to reduce sexual risk in men who have sex with men. AIDS Behav 2010 Jun;14(3):549-557 [FREE Full text] [doi: 10.1007/s10461-009-9578-2] [Medline: 19499321]

21. Bowen AM, Williams ML, Daniel CM, Clayton S. Internet based HIV prevention research targeting rural MSM: feasibility, acceptability, and preliminary efficacy. J Behav Med 2008 Dec;31(6):463-477 [FREE Full text] [doi: 10.1007/s10865-008-9171-6] [Medline: 18770021]

22. Lim MS, Wright C, Hellard ME. The medium and the message: fitting sound health promotion methodology into 160 characters. JMIR Mhealth Uhealth 2014;2(4):e40 [FREE Full text] [doi: 10.2196/mhealth.3888] [Medline: 25367387]

23. Muessig KE, Pike EC, Legrand S, Hightow-Weidman LB. Mobile phone applications for the care and prevention of HIV and other sexually transmitted diseases: a review. J Med Internet Res 2013;15(1):e1 [FREE Full text] [doi: 10.2196/jmir.2301] [Medline: 23291245]

XSL•FO
RenderX

24.   Holloway IW, Rice E, Gibbs J, Winetrobe H, Dunlap S, Rhoades H. Acceptability of smartphone application-based HIV prevention among young men who have sex with men. AIDS Behav 2014 Feb;18(2):285-296 [FREE Full text] [doi: 10.1007/s10461-013-0671-1] [Medline: 24292281]

25.   Sumner SA, Mercy JA, Dahlberg LL, Hillis SD, Klevens J, Houry D. Violence in the United States: status, challenges, and opportunities. JAMA 2015 Aug 4;314(5):478-488 [FREE Full text] [doi: 10.1001/jama.2015.8371] [Medline: 26241599]

26.   Beyrer C, Baral SD, van Griensven F, Goodreau SM, Chariyalertsak S, Wirtz AL, et al. Global epidemiology of HIV infection in men who have sex with men. Lancet 2012 Jul 28;380(9839):367-377 [FREE Full text] [doi: 10.1016/S0140-6736(12)60821-6] [Medline: 22819660]

27.   Sullivan PS, Carballo-Diéguez A, Coates T, Goodreau SM, McGowan I, Sanders EJ, et al. Successes and challenges of HIV prevention in men who have sex with men. Lancet 2012 Jul 28;380(9839):388-399 [FREE Full text] [doi: 10.1016/S0140-6736(12)60955-6] [Medline: 22819659]

28.   Beyrer C, Sullivan PS, Sanchez J, Dowdy D, Altman D, Trapence G, et al. A call to action for comprehensive HIV services for men who have sex with men. Lancet 2012 Jul 28;380(9839):424-438 [FREE Full text] [doi: 10.1016/S0140-6736(12)61022-8] [Medline: 22819663]

29.   Reeders D. Scope Social Connectivity: Daniel's Story. URL: http://www.scopeproject.net/new-blog-2//daniels-story [accessed 2015-09-15] [WebCite Cache ID 6bY8YiVLA]

30.   Grindr Advertising. Los Angeles, CA: Grindr LLC; 2014. URL: https://s3.amazonaws.com/grindr_marketing/US+Media+Kit+05.2014.pdf [accessed 2016-08-24] [WebCite Cache ID 6jzjBB2a0]

31.   Amante DJ, Hogan TP, Pagoto SL, English TM, Lapane KL. Access to care and use of the Internet to search for health information: results from the US National Health Interview Survey. J Med Internet Res 2015;17(4):e106 [FREE Full text] [doi: 10.2196/jmir.4126] [Medline: 25925943]

32.   Wartella E, Rideout V, Zupancic H, Beaudoin-Ryan L, Lauricella A. Teens, health, and technology: a national survey. Media Commun 2016;4(3):13-23 [FREE Full text] [doi: 10.17645/mac.v4i3.515]

33.   White D. Tinder Adds STD Testing Center Locator to Dating App. New York, NY: Time Inc; 2016. URL: http://time.com/4190222/tinder-adds-std-testing-center-locator/ [accessed 2016-02-12] [WebCite Cache ID 6fEUbuDbH]

34.   Smartphone OS Market Share, Q1 2015. Framingham, MA: IDC Corporate USA; 2015. URL: http://www.idc.com/prodserv/smartphone-os-market-share.jsp [accessed 2015-08-31] [WebCite Cache ID 6bBhzPNfh]

Abbreviations

HIV: human immunodeficiency virus
MSM: men who have sex with men
STI: sexually transmissible infections

XSL•FO
RenderX

<u>Original Paper</u>

A Systematic Review of Apps using Mobile Criteria for Adolescent Pregnancy Prevention (mCAPP)

Elizabeth Chen[1*], MPH; Emily Rose Mangone[2,3*], MSc

[1]Department of Health Behavior, Gillings School of Global Public Health, The University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[2]International Health Division, Abt Associates, Bethesda, MD, United States

[3]Department of Health Policy and Management, Gillings School of Global Public Health, The University of North Carolina at Chapel Hill, Chapel Hill, NC, United States

[*]all authors contributed equally

Corresponding Author:
Emily Rose Mangone, MSc
International Health Division
Abt Associates
4550 Montgomery Avenue
Suite 800 North
Bethesda, MD, 20814
United States
Phone: 1 650 919 3414
Email: emilyrose.mangone@gmail.com

Abstract

Background: Adolescents in the United States and globally represent a high-risk population for unintended pregnancy, which leads to high social, economic, and health costs. Access to smartphone apps is rapidly increasing among youth, but little is known about the strategies that apps employ to prevent pregnancy among adolescents and young adults. Further, there are no guidelines on best practices for adolescent and young adult pregnancy prevention through mobile apps.

Objective: This review developed a preliminary evaluation framework for the assessment of mobile apps for adolescent and young adult pregnancy prevention and used this framework to assess available apps in the Apple App Store and Google Play that targeted adolescents and young adults with family planning and pregnancy prevention support.

Methods: We developed an assessment rubric called Mobile Criteria for Adolescent Pregnancy Prevention (mCAPP) for data extraction using evidence-based and promising best practices from the literature. mCAPP comprises 4 domains: (1) app characteristics, (2) user interface features, (3) adolescent pregnancy prevention best practices, and (4) general sexual and reproductive health (SRH) features. For inclusion in the review, apps that advertised pregnancy prevention services and explicitly mentioned youth, were in English, and were free were systematically identified in the Apple App Store and Google Play in 2015. Screening, data extraction, and 4 interrater reliability checks were conducted by 2 reviewers. Each app was assessed for 92 facets of the mCAPP checklist.

Results: Our search returned 4043 app descriptions in the Apple App Store (462) and Google Play (3581). After screening for inclusion criteria, 22 unique apps were included in our analysis. Included apps targeted teens in primarily developed countries, and the most common user interface features were clinic and health service locators. While app strengths included provision of SRH education, description of modern contraceptives, and some use of evidence-based adolescent best practices, gaps remain in the implementation of the majority of adolescent best practices and user interface features. Of the 8 best practices for teen pregnancy prevention operationalized through mCAPP, the most commonly implemented best practice was the provision of information on how to use contraceptives to prevent pregnancy (15/22), followed by provision of accurate information on pregnancy risk of sexual behaviors (13/22); information on SRH communication, negotiation, or refusal skills (10/22); and the use of persuasive language around contraceptive use (9/22).

Conclusions: The quality and scope of apps for adolescent pregnancy prevention varies, indicating that developers and researchers may need a supportive framework. mCAPP can help researchers and developers consider mobile-relevant evidence-based best practices for adolescent SRH as they develop teen pregnancy prevention apps. Given the novelty of the mobile approach, further research is needed on the impact of mCAPP criteria via mobile channels on adolescent health knowledge, behaviors, and outcomes.

XSL•FO
RenderX

Introduction

The prevention of unintended pregnancy among adolescents and young adults is a health priority in the United States and globally because of its substantial health, social, and economic impacts [1-3]. Factors cited as key components to preventing teen pregnancy include free access to information, access to reproductive health services, and sex education [4-6]. The technology that has the potential to connect adolescents to these interventions and services is rapidly proliferating. In the United States, smartphone ownership among teens has grown exponentially, from 37% in 2013 to 73% in 2015, and almost all teens (91%) go online from a mobile device at least occasionally [7-9]. Access to smartphones in the United States is not limited to the wealthy; 61% of teens in households with annual incomes of less than $30,000 have access to smartphones [7]. Additionally, black and Hispanic teens, groups that experience 2 to 3 times the rate of unintended pregnancy as white teens, have the same or higher access to smartphones compared to white teens (85%, 71%, and 71%, respectively) [7,10]. While smartphones are currently more common in developed country settings, globally, smartphone subscriptions are projected to increase from 2.1 billion subscriptions in 2015 to 6.1 billion in 2020, covering 70% of the world's population [11].

Mobile phone apps present a unique opportunity to connect teens to contraceptive information and sex education, behavior change interventions, and reproductive health services that could not be achieved previously. A recent review on apps for the prevention of unintended pregnancy in the general population found 218 apps that claimed to help prevent unintended pregnancy [12]. However, only 16 of these apps specifically targeted young adults, and the criteria used to evaluate these apps were developed for a broader population, without special consideration for adolescents [13].

Mobile health (mHealth), the use of mobile phones and other wireless technology in health care, is a burgeoning field within public health, and there is a need to standardize how mHealth interventions are developed, evaluated, and reported so that the field can advance. Recently, the mHealth Technical Evidence Review Group of the World Health Organization (WHO) developed the mHealth Evidence Reporting and Assessment (mERA) checklist to give guidance for reporting mHealth interventions, and a recent study implemented the mERA checklist to evaluate the reporting of mobile adolescent sexual and reproductive health interventions [14,15]. While adoption of mERA will lead to greater transparency and consistency in reporting across the broader field of mHealth, there is a need to develop nuanced and evidence-based checklists for specific mHealth approaches, such as adolescent pregnancy prevention and the promotion of sexual and reproductive health through mobile channels. Recognizing this gap, we build on the

methodology and findings of a previous review of mobile apps for pregnancy prevention and identify evidence-based best practices for adolescent and young adult pregnancy prevention from the literature to develop and implement a preliminary framework for mobile Criteria for Adolescent Pregnancy Prevention (mCAPP) [12].

Using mCAPP, we conducted a systematic analysis of apps explicitly targeting adolescents and young adults and posed the following research questions:

1. What are the characteristics of teen pregnancy prevention apps currently on the market?

2. What features are included in the user interface of these apps?

3. What teen pregnancy prevention best practices are used?

4. What sexual and reproductive health features and content are offered?

Our review presents the answers to these questions and offers mCAPP as a foundation for future research in this field. Finally, we offer recommendations to app developers and researchers developing mobile phone apps for the prevention of unintended pregnancy among adolescents and young adults.

Methods

Codebook and Mobile Criteria for Adolescent Pregnancy Prevention Checklist Development

The codebook and data extraction form were developed iteratively between June and August 2015 and were informed by literature on mHealth, eHealth (electronic health information and communication technology, of which mHealth is a subset characterized by its mobile channels), and evidence-based adolescent reproductive health best practices (cited individually below). We established 4 domains: general app information, user interface, teen pregnancy prevention best practices for mobile apps, and general sexual and reproductive health features and content. Data extraction points were created in each domain, and data were collected and analyzed in Excel (Microsoft Corp). The codebook included 92 data extraction points for every app. To create a user-friendly mCAPP checklist (see Multimedia Appendix 1) for general use, we condensed some of the separate binary data extraction points (eg, inclusion of 7 race/ethnicity categories for images) into an easily digestible, 2-page format that remains organized around the 4 domains from our codebook.

Search and Screening Strategy

In July 2015, we developed a set of 52 search terms informed by 3 previous systematic reviews on adolescent sexual health by combining 4 age-targeted terms, "adolescent," "teen," "young adult," and "youth," with 13 pregnancy prevention terms: "pregnancy prevention," "sex education," "sexual negotiation,"

"family planning," "sexual health," "reproductive health," "abstinence," "sexual communication," "sexual decision making," "sexuality," "condom," "contraception," and "birth control" [12,16,17]. Each of these 52 terms was searched separately by both authors in the Apple App Store and the Android Google Play store, and the number of results per search

term was documented. Since Android and iOS (Apple) phones make up over 95% of the worldwide smartphone market share, the authors chose to review apps only in the relevant two app stores [18]. All app descriptions were screened for inclusion and exclusion criteria just as abstracts are initially screened in systematic reviews (Textbox 1).

Textbox 1. Inclusion and exclusion criteria.

Inclusion criteria:

- App advertises and includes a component of pregnancy prevention, including sexual decision making information, contraception information that explicitly notes pregnancy prevention, or communication/negotiation information.

- App explicitly targets "adolescents," "teens," "young adults," or "youth"or "students" in title or app store description.

- App is in English.

- App is free.

Exclusion criteria:

- App is primarily for the purpose of facilitating or tracking an established pregnancy.

- App is primarily a birth control reminder or fertility tracker with minimal additional information.

- App is primarily a locator (for clinics, condoms, etc).

- App is primarily information about hours, locations, and services provided by brick-and-mortar centers or clinics.

- App is for a specific event such as a conference, march, or rally.

- App primarily addresses infertility or in vitro fertilization.

- App primarily promotes pro-life viewpoints or opinions with no component of pregnancy prevention.

- App primarily targets adults.

- App is primarily a condom game that is not explicitly for educational purposes.

- App includes screenshots or descriptions that are not in English (even if English is listed as a language), is in beta testing, contains pranks, or is dysfunctional.

- Google Play app is a duplicate of a relevant app in the Apple App Store (Google Play app excluded, Apple App Store app included).

Apps that were primarily fertility trackers, birth control reminders, or locators were excluded because they were evaluated thoroughly in a recent systematic review and have too narrow of a purpose to benefit from an evaluation that takes a holistic approach [12]. Apps that were available in both the Apple App Store and the Google Play store were removed from the Google Play group and only analyzed as Apple App Store apps to avoid duplication. Full apps were downloaded for data abstraction.

Relevance Review and Data Extraction

Two rounds of interrater reliability (IRR) calculations were conducted during the initial screening of app descriptions. Both authors documented the number of results they had for each search term and then compared the list of apps that met the inclusion criteria from each search term. IRR was 99% in both round 1 and round 2. IRR scores were high because of the low sensitivity of the Apple App Store and Google Play search engines and the resulting high volume of unrelated apps. Relevant apps were downloaded using an Android phone and an iPhone, and data extraction occurred in September 2015. Two additional rounds of IRR checks were done for the 92 data extraction points for 22% (8/37) of the downloaded apps, with IRR scores of 90% in round 3 and 96% in round 4.

Characteristics of Apps

To classify apps and describe the general characteristics of apps available for prevention of unintended pregnancy among adolescents, we documented data from the app stores such as developer name, number of average stars, date last updated, app store category, and age rating; evaluated the race/ethnicity featured in the app images; determined which gender the app targeted; documented whether the app was affiliated with a specific religion; evaluated app credibility; evaluated app abortion stance; and evaluated the primary purpose of app (education, linkage to care, counseling support). We included characteristics that teens may use to determine credibility of source, peer use and approval, and fit with personal need or demographics [19]. There were 25 data extraction points in this section in the codebook.

User Interface Features

To describe the design, interactivity, and engagement of included apps, we reviewed apps for 15 useful interface features. These features were informed by the literature on teen and young adult user interface preferences and marketing as well as criteria for evaluating health websites (Textbox 2) [19-21].

We also evaluated apps for the presence of 4 undesirable interface components (Textbox 3).

Textbox 2. Positive user interface features.

☐ Global positioning system (GPS) capabilities

☐ Clinic or service locators

☐ Contraceptive locators

☐ Customizable look

☐ Appointment scheduling

☐ Public, forum, or social network communication

☐ Direct communication (chat, text, call)

☐ Push notifications

☐ Internal search function

☐ Videos, films, or movies

☐ Gamification elements

　○ Quizzes

　○ Direct manipulation

☐ Decision aid

☐ Other notable features

Textbox 3. Negative user interface features.

☐ Faulty element (eg, broken links, blank pages, or broken/unintelligible English)

☐ App crashed

☐ Advertising

☐ App required purchase to use

Teen Pregnancy Prevention Best Practices for Apps

We were most interested in teen pregnancy prevention best practices relevant for mobile platforms, specifically apps. Because there are no current best practices for mHealth or eHealth interventions for teen pregnancy prevention, we consulted experts in the field and reviewed and operationalized sets of best practices in the peer-reviewed literature for in-person teen pregnancy prevention interventions that could reasonably be expected to succeed via this new mode of intervention delivery. Best practices that were specific to traditional classroom-based or other in-person teen pregnancy interventions and outside of the control of app developers were excluded. Upon review of the literature, we developed the following list of best practices as an evaluation framework for current apps (Textbox 4) [22-26]. As with general teen pregnancy interventions, it is ideal for teen pregnancy apps to use all 8 of these best practices.

Textbox 4. Best practices for app-based mHealth interventions for teen pregnancy prevention.

☐ Deliver and consistently reinforce persuasive communication about abstaining from sexual activity [22-24]

☐ Deliver and consistently reinforce persuasive communication about using condoms or other forms of contraception when sexually active [22-24]

☐ Be based on theoretical approaches that have been demonstrated to influence other health-related behavior and identify specific important sexual antecedents to be targeted [22,23,25,26]

☐ Provide clear, accurate information about the risk of pregnancy due to sexual activity [22,23]

☐ Provide clear, accurate information and skill-building exercises on how to use contraceptives to prevent unwanted pregnancy [22,23]

☐ Provide skill-building exercises or practice with sexual communication, negotiation, and refusal [22,24]

☐ Provide activities designed to engage users, personalize or internalize information, and provide tailored feedback [22,24]

☐ Target information for racial and ethnic subgroups of adolescents (eg, Hispanic females) [25,26]

Textbox 5. Promising practices for adolescent pregnancy prevention.

☐ Address whether parental consent is required for sexual and reproductive health services [27-30]

☐ Encourage parental communication [31]

☐ Offer peer sexual and reproductive health stories or peer counseling [32-34]

We also found 3 practices that, while not currently hailed as best practices in the literature, were cited by multiple sources as promising practices given the specific challenges that adolescents face in the context of reproductive health (Textbox 5).

General Sexual and Reproductive Health Features and Content

WHO defines sexual and reproductive health (SRH) as a state of physical, emotional, mental, and social well-being in all matters relating to sexuality and the reproductive system and its processes. WHO adds that this state "implies that people are able to have a satisfying and safe sex life and that they have the capacity to reproduce and the freedom to decide if, when, and how often to do so" [35]. For our final research question about general SRH features and content, we organized the data into 2 subsections: one that included 7 SRH features relevant for all ages (Textbox 6) and one that focused on the provision of contraceptive information (Textbox 7). Both subsections were adapted from the SRH evaluation criteria developed in the earlier review [12].

In the contraceptive information subsection, we reviewed for the inclusion of 6 modern contraceptive methods, their effectiveness, and how to use or length of effectiveness of the specific method. We also reviewed for inclusion of related information. There were 37 data extraction points in this section in the codebook.

Textbox 6. Sexual and reproductive health features relevant for all ages.

☐ State that app is not a replacement for professional medical advice

☐ Mention the confidentiality or privacy of the app or services accessed through the app

☐ Address the cost of SRH services

☐ Describe or counsel on abusive relationships or intimate partner violence

☐ Describe or counsel on alcohol or substance abuse

☐ Refer for pregnancy testing

☐ Provide information about or refer for abortion services

Textbox 7. Sexual and reproductive health contraceptive information.

☐ Information on condoms, pills, injection, implant, intrauterine device (IUD), emergency contraception, withdrawal, or other method

☐ Information on where the user can get contraceptives

☐ Information/counseling about contraceptive side effects and risks

☐ Information/counseling on side effect management

☐ Information/counseling on switching contraceptive methods

☐ Information/counseling on dual protection from pregnancy and sexually transmitted infections (STIs) including HIV

☐ Information/counseling on STIs and STI testing

☐ Notable misinformation or bad practices in the app (such as shaming or scare tactics)

Results

Search Results

Our search returned 4043 app descriptions from the Apple App Store (462) and Google Play (3581). Of these, 22 unique apps met inclusion criteria and were included in the review (Figure 1). Because of the poor sensitivity of the app store search engines and the hundreds of returned apps that were unrelated to reproductive health as well as the fact that it is not possible to export or copy information from the Apple App Store and Google Play, it was not feasible to document all reasons for exclusion at the app description level.

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram [36] for Apple App Store and Google Play search results.

^aDupes: Duplicate apps
^bNo preg prevent: No pregnancy prevention features offered
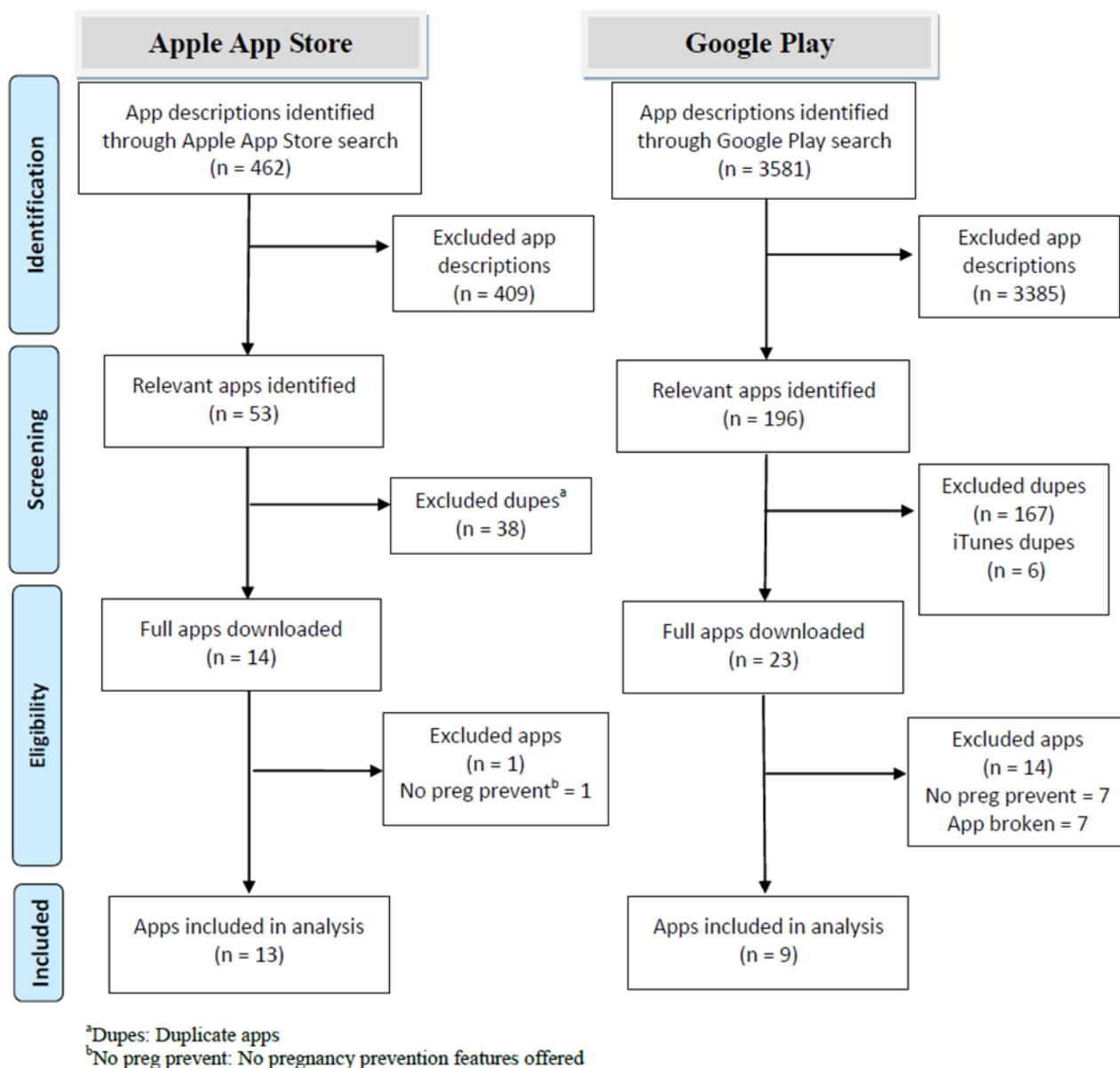
Characteristics of Apps

Of the 22 included apps, 8 were available through the Apple App Store only, 9 were available through Google Play only, and 5 were available through both app stores. The 22 included apps were listed under a variety of app store categories: Health & Fitness (8/22), Education (5/22), Lifestyle (6/22), Reference (1/22), Medical (1/22), and Entertainment (1/22). The majority of the apps targeted adolescents in the United States (8/22) or the United Kingdom (7/22). The other apps focused on adolescents in Kenya (1/22), Africa (1/22), Pakistan (1/22), and Canada (1/22). Only 3 apps did not have a discernable geographic focus. A total of 13 apps included images of teens or young adults, and of these, 4 apps included images of white teens only, 2 apps included images of black teens only, and 7 apps included images of teens from more than one race. Almost all (21/22) apps provided information that was relevant for both males and females; 1 app specifically targeted females. Almost half (10/22) of the apps were downloaded less than 1000 times,

4 apps were downloaded 1000 to 10,000 times, and 8 apps did not have this information. See Multimedia Appendix 2 for general characteristics of included apps.

A total of 17 of the apps cited credible information from a reputable public health source (ie, Centers for Disease Control and Prevention or other health organization). The source of information in 5 apps was unclear or not credible. Many (15/22) of the apps had been updated in the past 12 months; all were updated within the past 24 months. We also evaluated each app's abortion stance: 15 apps were categorized as pro-choice, 2 were categorized as pro-life, and 5 were categorized as unclear or no information. No apps were found to have an affiliation with a religion. Finally, we assessed each app's apparent purpose(s). Categories were not mutually exclusive, and the overwhelming majority of the apps provided SRH education (19/22). Many also provided linkage to care (13/22) and counseling or support (12/22).

User Interface Features

In evaluating the apps for desirable user features, we found that the most common features included in the apps were clinic and service locators (12/22) (see Table 1). Less than half (9/22) had global positioning system (GPS) capabilities; 7 of these 9 apps located contraceptives. Less than a quarter of the included apps offered entertainment, gamification, or communication features (5/22, 5/22, and 4/22, respectively). No apps allowed users to customize their design or schedule appointments.

With regard to the undesirable user interface features that we evaluated, we encountered faulty elements with 5 apps, advertisements with 3 apps, and crashing with 1 app. None of the apps required in-app purchases.

Teen Pregnancy Prevention Best Practices for Apps

Of the 8 best practices for teen pregnancy prevention that we operationalized for mHealth interventions, the most commonly implemented best practice was the provision of information on how to use contraceptives to prevent pregnancy (15/22); followed by provision of accurate information on pregnancy risk of sexual behaviors (13/22); information on SRH communication, negotiation, or refusal skills (10/22); and the use of persuasive language around contraceptive use (9/22) (see Table 2). Among the 22 apps, *my choice by PPT* included 5 of the 8 best practices and one promising practice (parental consent). *Safe Sex Tips* also had 5 best practices but no promising practices. *Get S.M.A.R.T.* had 4 best practices and 2 promising practices. A total of 3 apps did not include any of the best or promising practices. No apps referenced a theoretical approach or encouraged parental communication, only 2 apps included targeting for racial/ethnic groups, and only 3 apps included persuasive language on abstinence.

General Sexual and Reproductive Health Features and Contraceptive Content

Of the 7 general SRH features, 6 apps included 4 or more (Table 3). *my choice by PPT* again led with the highest number (6/7)

of included SRH features. A total of 3 apps, *ASK My Body App, Healthwise*, and *ICAH*, did not include any of the SRH features. The most frequently offered SRH feature was the description of abortion options or counseling (12/22), followed by pregnancy testing referral, a statement of confidentiality, and information about abusive relationships (10/22 each). The least frequently provided SRH feature was a note that the app was not a replacement for medical advice (4/22) and information about alcohol or substance abuse (5/22).

Almost all apps (20/22) provided some information about male condoms as a way to prevent unintended pregnancy, and 17 apps mentioned condoms in the context of dual protection (from pregnancy and STIs). Only half (10/20) of apps that mentioned condoms also provided any information about how to use condoms. Emergency contraception was mentioned in 15 apps, followed by oral contraception (14/22), the intrauterine device (IUD) (14/22), and other methods (14/22). One-half to two-thirds of the apps that mentioned the condom, IUD, injection, and implant also described the effectiveness of the method at preventing pregnancy. Almost all of the apps that described long-acting methods, including the IUD, injection, and implant, also described the length of duration of the methods (13/14, 11/12, and 9/10, respectively). Almost all (14/15) apps that provided information about emergency contraception also provided some information about how to use it. Most (12/14) apps that provided information about oral contraception provided information about how to use it.

A total of 14 apps provided information on where to get contraceptives, and 10 apps provided information about side effects of using contraceptives. However, only 3 of these apps provided information about side effect management and switching methods. Most apps (19/22) provided information about STIs, and 15 provided information about getting tested for STIs.

Table 1. Desirable user interface features found in apps.

App name	GPS[a]	CLI[b]	CON[c]	CUS[d]	SCH[e]	FOR[f]	DIR[g]	PUSH[h]	SEA[i]	VID[j]	GAM[k]	DA[l]	Total
Get S.M.A.R.T.	✓	✓				✓	✓	✓	✓	✓	✓	✓	9
gPower	✓	✓	✓				✓		✓	✓	✓	✓	8
my choice by PPT	✓	✓	✓					✓			✓	✓	6
Teens in NYC	✓	✓	✓						✓	✓			5
Love Matters		✓				✓		✓				✓	4
NeedTayKnow	✓	✓				✓						✓	4
Your Choice Your Voice	✓	✓	✓									✓	4
CaSH 2 U	✓	✓									✓		3
Kent C Card	✓	✓	✓										3
The Choice—It's kind of a big deal						✓			✓	✓			3
Girls Incorporated of Lynn		✓	✓										2
OC Teens Mobile		✓										✓	2
SexPositive										✓	✓		2
SafeSex101		✓	✓										2
The Real Deal	✓											✓	2
aSk UK							✓						1
ICAH								✓					1
ASK My Body App													0
HealthWise													0
My Sex Doctor													0
My Sex Doctor Lite													0
Safe Sex Tips													0
Total	9	12	7	0	0	4	3	4	4	5	5	8	

[a]GPS: Global positioning system

[b]CLI: Clinic or service locator

[c]CON: Contraceptive locator

[d]CUS: Customizable look

[e]SCH: Appointment scheduler

[f]FOR: Forum or social network

[g]DIR: Direct communication (text, call, or chat)

[h]PUSH: Push notifications

[i]SEA: Search function

[j]VID: Videos, films, or movies

[k]GAM: Gamification elements

[l]DA: Decision aid

XSL•FO

RenderX

Table 2. Adolescent pregnancy prevention best practices found in apps.

App name	ABS[a]	CON[b]	THY[c]	PRI[d]	CI[e]	NEG[f]	PER[g]	TAR[h]	PAR[i]	COM[j]	PS[k]	Total
Get S.M.A.R.T.		✓		✓	✓	✓			✓		✓	6
my choice by PPT		✓		✓	✓	✓	✓		✓			6
gPower		✓		✓	✓		✓		✓			5
Safe Sex Tips		✓		✓	✓	✓	✓					5
The Choice—It's kind of a big deal	✓			✓		✓	✓				✓	5
Your Choice Your Voice	✓	✓			✓		✓		✓			5
Love Matters				✓	✓	✓		✓				4
OC Teens Mobile	✓			✓	✓	✓						4
SexPositive					✓	✓	✓				✓	4
The Real Deal		✓		✓	✓	✓						4
ASK My Body App					✓		✓	✓				3
aSk UK		✓		✓	✓							3
My Sex Doctor				✓	✓	✓						3
My Sex Doctor Lite				✓	✓	✓						3
NeedTayKnow		✓		✓	✓							3
Teens in NYC					✓				✓		✓	3
Girls Incorporated of Lynn		✓		✓								2
CaSH 2 U									✓			1
ICAH											✓	1
HealthWise												0
Kent C Card												0
SafeSex101												0
Total	3	9	0	13	15	10	7	2	6	0	5	

[a]ABS: Pro-abstinence messaging

[b]CON: Pro-contraception messaging

[c]THY: Noted a theoretical approach

[d]PRI: Pregnancy risk information

[e]CI: Contraceptive use information

[f]NEG: Communication, negotiation, and refusal skill information

[g]PER: Personalization

[h]TAR: Targeting of information for ethnic or racial groups

[i]PAR: Parental consent for sexual and reproductive health (SRH) services information

[j]COM: Parental communication encouragement

[k]PS: Peer stories or peer counseling

Table 3. Sexual and reproductive health features found in apps.

App name	Sexual and reproductive health features							
	DIS[a]	CON[b]	COS[c]	VIO[d]	SUB[e]	TES[f]	ABO[g]	Total
my choice by PPT	✓	✓	✓	✓		✓	✓	6
aSk UK		✓	✓	✓	✓		✓	5
Get S.M.A.R.T.		✓	✓	✓		✓	✓	5
NeedTayKnow		✓	✓	✓	✓		✓	5
Your Choice Your Voice		✓	✓	✓		✓	✓	5
CaSH 2 U		✓	✓			✓	✓	4
Love Matters				✓		✓	✓	3
My Sex Doctor				✓		✓	✓	3
My Sex Doctor Lite				✓		✓	✓	3
OC Teens Mobile	✓			✓	✓			3
SafeSex101			✓			✓	✓	3
The Choice—It's kind of a big deal		✓		✓	✓			3
The Real Deal	✓					✓	✓	3
Girls Incorporated of Lynn						✓	✓	2
gPower	✓		✓					2
Kent C Card		✓	✓					2
Safe Sex Tips					✓			1
SexPositive		✓						1
Teens in NYC		✓						1
ASK My Body App								0
HealthWise								0
ICAH								0
Total	4	10	9	10	5	10	12	

[a]DIS: Medical advice disclaimer

[b]CON: Statement of confidentiality

[c]COS: Service cost information

[d]VIO: Violence or abuse information

[e]SUB: Substance abuse information

[f]TES: Pregnancy testing referral available

[g]ABO: Abortion counseling or information

Discussion

Principal Findings

The 22 reviewed apps employed a variety of strategies to assist teens and adolescents in preventing unintended pregnancies. Overall strengths of included apps were the provision of SRH education, descriptions of at least 1 modern contraceptive method, and the use of 4 or more of the adolescent pregnancy prevention best or promising practices. It was also promising that 10 of the 22 apps provided information or counseling around SRH communication, negotiation, and refusal skills.

As a group, the apps also had weaknesses. Only half included persuasive messages, indicating that many apps were not explicitly advocating for behavior change. Only 2 of the apps (*Love Matters* and *ASK My Body App*) included targeted information for a specific ethnic or racial group of adolescents. Additionally, none of the apps explicitly cited theory, so there is no evidence of an established change strategy in the development of the apps. *Safe Sex Tips* is the only app that included 6 of the 8 best practices we evaluated, but this app is also one that was categorized as having unclear or no credibility. Moreover, there were 5 apps that did not include any of the best practices that we evaluated. Given that there were only 22 apps included in this study, it is concerning that 5 apps do not offer any of the best practices for teen pregnancy prevention and 3 apps do not offer any best or promising practices. These findings indicate that current best practices are not being implemented in available apps for teen pregnancy prevention, possibly because there are no guidelines for developers in this field.

XSL·FO
RenderX

In terms of user interface, one approach that was underused with this demographic was gamification. Only 5 of the 22 apps employed gamification elements. Encouragingly, these 5 apps included promising features for engagement. For example, *SexPositive* featured an interactive spinning wheel that matched body parts (eg, mouth, vagina) with different body parts or objects (eg, penis, teddy bear) and then provided information on the resulting risk of sexually transmitted infections, recommended safer sex practices, and communication and advice. While using games for teen pregnancy prevention is not entirely new, using a mobile phone instead of a computer or console is [36-39]. Given that 59% of girls and 84% of boys ages 13 to 17 years play video games online or on their phones, gamification may be a promising strategy to explore for future teen pregnancy prevention apps [40].

One factor that may limit the impact of the 22 teen pregnancy prevention apps was that 19 of them were location- or organization-specific. For example, *SafeSex 101* was developed for students at the University of California, Los Angeles. The app listed information on local resources and nearby health care providers and guided students on how to access birth control given the university's specific insurance plans. While these apps may provide a lot of relevant information, adolescents searching for these apps may assume that these apps do not apply to them simply because they are not part of the app's target demographic. Because of GPS, apps can provide global and local services based on the user's location. More apps are needed that follow best practices and include resources for national and even international audiences.

Limitations

There are limitations to this review. Because there is not a list of best practices for mHealth-based teen pregnancy prevention interventions, we synthesized and adopted existing best practices from the literature with a critical eye for the strengths and limitations of mobile phones, creating mCAPP. This ambitious 92-item checklist represents a novel and important foundation for developing and evaluating adolescent sexual and reproductive health interventions on a mobile platform. However, the validity and effectiveness of mCAPP as a cohesive set of criteria has not yet been evaluated for this platform, and

mCAPP will require further study to understand its impact in practice. As it stands, this checklist is intended to help developers and evaluators alike consider best practices for adolescent SRH in the context of mobile apps where previously there was little guidance. Future studies are needed to determine the impact of the mCAPP criteria and whether there are synergistic effects between different criteria that lead to improved outcomes for adolescents.

Another limitation is that while we sought to assess whether apps were based on theory, we only evaluated whether a theory was explicitly cited. Therefore, some apps may have had theoretical underpinnings that went undocumented. One technical limitation was that we tried to only evaluate content available in the app (assuming that data plans and Internet connection may not always be available) and not consider secondary content (YouTube videos, external links). However, it was occasionally difficult to make this distinction when content was embedded in the app. When in doubt, we reviewed inclusively. A final limitation is that by the time this article is published, there will almost certainly be more apps available at the Apple App Store and Google Play that would fit our inclusion criteria and propose innovative strategies to address adolescent pregnancy on this platform.

Conclusion

In this quickly evolving field and market, we must assess the strategies that are used to prevent unintended pregnancy among youth through mobile phones and lay groundwork for more effective interventions that emphasize behavior change as well as knowledge dissemination. We put forth mCAPP as a preliminary framework that can guide development and evaluation efforts for developers and researchers alike. Additional research should be conducted to assess the impact of apps that leverage mCAPP strategies compared to apps that do not on reproductive and sexual health knowledge, behaviors, and outcomes of adolescents. With more and more adolescents accessing smartphones, our hope is that mCAPP can serve as a framework and catalyst for new mHealth interventions that can revolutionize mobile health, starting with adolescent pregnancy prevention.

Conflicts of Interest

None declared.

XSL•FO

RenderX

Multimedia Appendix 1

Mobile Criteria for Adolescent Pregnancy Prevention checklist.

[PDF File (Adobe PDF File), 582KB - mhealth_v4i4e122_app1.pdf]

Multimedia Appendix 2

Overview of apps included in review.

[PDF File (Adobe PDF File), 39KB - mhealth_v4i4e122_app2.pdf]

References

1. Healthy People 2020. Family Planning. 2016. URL: https://www.healthypeople.gov/2020/topics-objectives/topic/family-planning [accessed 2016-01-06] [WebCite Cache ID 6lSXIFnhp]
2. Family Planning 2020. About us.. Washington DC: United Nations Foundation; 2015. URL: http://www.familyplanning2020.org/microsite/about-us [accessed 2016-02-01] [WebCite Cache ID 6lSXUgSXg]
3. World Health Organization. Adolescent pregnancy. 2014. URL: http://www.who.int/mediacentre/factsheets/fs364/en/ [accessed 2015-12-01] [WebCite Cache ID 6lSXj8wfW]
4. Advocates for Youth. Teen pregnancy prevention. 2008. URL: http://www.advocatesforyouth.org/teen-pregnancy-prevention [accessed 2016-01-07] [WebCite Cache ID 6lSXnLAyV]
5. Planned Parenthood Federation of America. Reducing teenage pregnancy. 2013. URL: https://www.plannedparenthood.org/files/6813/9611/7632/Reducing_Teen_Pregnancy.pdf [accessed 2016-01-07] [WebCite Cache ID 6lSXqHD3s]
6. World Health Organization. Guidelines on preventing early pregnancy and poor reproductive outcomes among adolescents in developing countries. 2015. URL: http://www.who.int/maternal_child_adolescent/documents/preventing_early_pregnancy/en/ [accessed 2016-01-05] [WebCite Cache ID 6lSXueh08]
7. Pew Research Center. 73% of teens have access to a smartphone; 15% have only a basic phone. Washington, DC: Pew Internet and American Life Project; 2015. URL: http://www.pewinternet.org/2015/04/09/teens-social-media-technology-2015/pi_2015-04-09_teensandtech_06/ [accessed 2016-01-07] [WebCite Cache ID 6lSY5Y4Jr]
8. Madden M, Lenhart A, Duggan M, Cortesi S, Gasser U. Teens and technology 2013. Washington, DC: Pew Internet and American Life Project; 2013. URL: http://www.pewinternet.org/files/old-media//Files/Reports/2013/PIP_TeensandTechnology2013.pdf [accessed 2016-03-01] [WebCite Cache ID 6lSYD2G1I]
9. Lenhart A. Mobile access shifts social media use and other online activities. Washington, DC: Pew Internet and American Life Project; 2015. URL: http://www.pewinternet.org/files/2015/04/PI_TeensandTech_Update2015_0409151.pdf [accessed 2016-01-07] [WebCite Cache ID 6lSYIBfkJ]
10. Martin J, Hamilton B, Osterman M, Curtin S, Matthews TJ. Births: final data for 2013. Natl Vital Stat Rep 2015 Jan 15;64(1):1-65 [FREE Full text] [Medline: 25603115]
11. Ericsson. Mobility Report: 70 percent of world's population using smartphones by 2020. 2015 Jun 03. URL: http://www.ericsson.com/news/1925907 [accessed 2015-12-15] [WebCite Cache ID 6lSYZH0gF]
12. Mangone ER, Lebrun V, Muessig KE. Mobile phone apps for the prevention of unintended pregnancy: a systematic review and content analysis. JMIR Mhealth Uhealth 2016;4(1):e6 [FREE Full text] [doi: 10.2196/mhealth.4846] [Medline: 26787311]
13. Taylor D, James EA. An evidence-based guideline for unintended pregnancy prevention. J Obstet Gynecol Neonatal Nurs 2011;40(6):782-793 [FREE Full text] [doi: 10.1111/j.1552-6909.2011.01296.x] [Medline: 22092349]
14. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, WHO mHealth Technical Evidence Review Group. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. BMJ 2016;352:i1174. [Medline: 26988021]
15. 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]
16. Guse K, Levine D, Martins S, Lira A, Gaarde J, Westmorland W, et al. Interventions using new digital media to improve adolescent sexual health: a systematic review. J Adolesc Health 2012 Dec;51(6):535-543. [doi: 10.1016/j.jadohealth.2012.03.014] [Medline: 23174462]
17. Jones K, Eathington P, Baldwin K, Sipsma H. The impact of health education transmitted via social media or text messaging on adolescent and young adult risky sexual behavior: a systematic review of the literature. Sex Transm Dis 2014 Jul;41(7):413-419. [doi: 10.1097/OLQ.0000000000000146] [Medline: 24922099]
18. International Data Corporation. Smartphone OS marketshare 2016 Q2. 2015. URL: http://www.idc.com/prodserv/smartphone-os-market-share.jsp [accessed 2016-09-01] [WebCite Cache ID 6lSZ1FPBc]
19. 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]
20. Loranger H, Nielsen J. Teenage usability: designing teen-targeted websites. Nielsen Norman Group. 2013. URL: https://www.nngroup.com/articles/usability-of-websites-for-teenagers/ [accessed 2016-01-08] [WebCite Cache ID 6lSZBFsH6]

XSL•FO
RenderX

21. Kim P, Eng TR, Deering MJ, Maxfield A. Published criteria for evaluating health related web sites: review. BMJ 1999 Mar 6;318(7184):647-649 [FREE Full text] [Medline: 10066209]

22. Kirby D. Emerging answers: Research findings on programs to reduce teen pregnancy (summary). Am J Health Educ 2001;32(6):348-355.

23. Kirby DB, Laris BA, Rolleri LA. Sex and HIV education programs: their impact on sexual behaviors of young people throughout the world. J Adolesc Health 2007 Mar;40(3):206-217. [doi: 10.1016/j.jadohealth.2006.11.143] [Medline: 17321420]

24. Suellentrop K. What works 2011-2012: curriculum-based programs that help prevent teen pregnancy. Washington, DC: National Campaign to Prevent Teen and Unplanned Pregnancy; 2011. URL: https://thenationalcampaign.org/sites/default/files/resource-primary-download/WhatWorks.pdf [accessed 2015-07-15] [WebCite Cache ID 6lSaKuYgk]

25. Sales JM, Milhausen RR, Diclemente RJ. A decade in review: building on the experiences of past adolescent STI/HIV interventions to optimise future prevention efforts. Sex Transm Infect 2006 Dec;82(6):431-436 [FREE Full text] [doi: 10.1136/sti.2005.018002] [Medline: 17151029]

26. Sales J, DiClemente R. Adolescent STI/HIV prevention programs: what works for teens.Research Facts and Findings. 2010. URL: http://www.actforyouth.net/resources/rf/rf_sti_0510.pdf [accessed 2016-10-25] [WebCite Cache ID 6lWYOlhdL]

27. Guttmacher Institute. State policies in brief. New York, NY; 2016. URL: https://www.guttmacher.org/state-policy/explore/overview-minors-consent-law [accessed 2016-01-08] [WebCite Cache ID 6lSaUZRIk]

28. Girma S, Paton D. Does parental consent for birth control affect underage pregnancy rates? The case of Texas. Demography 2013 Dec;50(6):2105-2128. [doi: 10.1007/s13524-013-0225-1] [Medline: 23765262]

29. Caal S, Guzman L, Berger A, Ramos M, Golub E. "Because you're on birth control, it automatically makes you promiscuous or something": Latina women's perceptions of parental approval to use reproductive health care. J Adolesc Health 2013 Nov;53(5):617-622. [doi: 10.1016/j.jadohealth.2013.05.003] [Medline: 23768462]

30. Hopkins K, White K, Linkin F, Hubert C, Grossman D, Potter JE. Women's experiences seeking publicly funded family planning services in Texas. Perspect Sex Reprod Health 2015 Jun;47(2):63-70 [FREE Full text] [doi: 10.1363/47e2815] [Medline: 25639913]

31. Widman L, Choukas-Bradley S, Noar SM, Nesi J, Garrett K. Parent-adolescent sexual communication and adolescent safer sex behavior: a meta-analysis. JAMA Pediatr 2016 Jan;170(1):52-61. [doi: 10.1001/jamapediatrics.2015.2731] [Medline: 26524189]

32. Jennings JM, Howard S, Perotte CL. Effects of a school-based sexuality education program on peer educators: the Teen PEP model. Health Educ Res 2014 Apr;29(2):319-329 [FREE Full text] [doi: 10.1093/her/cyt153] [Medline: 24488649]

33. Kapadia F, Frye V, Bonner S, Emmanuel PJ, Samples CL, Latka MH. Perceived peer safer sex norms and sexual risk behaviors among substance-using Latino adolescents. AIDS Educ Prev 2012 Feb;24(1):27-40. [doi: 10.1521/aeap.2012.24.1.27] [Medline: 22339143]

34. Layzer C, Rosapep L, Barr S. A peer education program: delivering highly reliable sexual health promotion messages in schools. J Adolesc Health 2014 Mar;54(3 Suppl):S70-S77 [FREE Full text] [doi: 10.1016/j.jadohealth.2013.12.023] [Medline: 24560080]

35. World Health Organization Department of Reproductive Health and Research. Sexual and reproductive health core competencies in primary care. 2011. URL: http://www.who.int/reproductivehealth/publications/health_systems/9789241501002/en/ [accessed 2016-01-07] [WebCite Cache ID 6lSarjIR9]

36. Paperny DM, Starn JR. Adolescent pregnancy prevention by health education computer games: computer-assisted instruction of knowledge and attitudes. Pediatrics 1989 May;83(5):742-752. [Medline: 2654867]

37. Roberto AJ, Zimmerman RS, Carlyle KE, Abner EL. A computer-based approach to preventing pregnancy, STD, and HIV in rural adolescents. J Health Commun 2007;12(1):53-76. [doi: 10.1080/10810730601096622] [Medline: 17365349]

38. Shegog R, Markham C, Peskin M, Dancel M, Coton C, Tortolero S. "It's your game": an innovative multimedia virtual world to prevent HIV/STI and pregnancy in middle school youth. Stud Health Technol Inform 2007;129(Pt 2):983-987. [Medline: 17911862]

39. Thomas R, Cahill J, Santilli L. Using an interactive computer game to increase skill and self-efficacy regarding safer sex negotiation: field test results. Health Educ Behav 1997 Feb;24(1):71-86. [Medline: 9112099]

40. Lenhart A. Teens, social media, and technology: overview 2015. Washington, DC: Pew Internet and American Life Project; 2015. URL: http://www.pewinternet.org/2015/04/09/teens-social-media-technology-2015/ [accessed 2016-01-07] [WebCite Cache ID 6lSbC8Lt1]

Abbreviations

eHealth: electronic health information and communication technology
GPS: global positioning system
IRR: interrater reliability
IUD: interuterine device
mCAPP: Mobile Criteria for Adolescent Pregnancy Prevention

XSL•FO
RenderX

mERA: mHealth Evidence Reporting and Assessment
mHealth: mobile health
SRH: sexual and reproductive health
STI: sexually transmitted infection
WHO: World Health Organization

XSL•FO
RenderX

Original Paper

Mobile Phone Apps to Improve Medication Adherence: A Systematic Stepwise Process to Identify High-Quality Apps

Karla Santo[1,2], MD; Sarah S Richtering[1], MD; John Chalmers[1,2], MD, PhD; Aravinda Thiagalingam[1,2,3,4], MD, PhD; Clara K Chow[1,2,3,5], MD, PhD; Julie Redfern[1,2], PhD

[1]The George Institute for Global Health, University of Sydney, Sydney, Australia

[2]Sydney Medical School, University of Sydney, Sydney, Australia

[3]Cardiology Department, Westmead Hospital, Sydney, Australia

[4]Westmead Institute for Medical Research, Sydney, Australia

[5]Charles Perkins Centre Westmead, Sydney, Australia

Corresponding Author:
Karla Santo, MD
The George Institute for Global Health
University of Sydney
Level 10, King George V Building, 83-117 Missenden Rd, Camperdown, NSW
Sydney, 2050
Australia
Phone: 61 280524620
Fax: 61 280524502
Email: ksanto@georgeinstitute.org.au

Abstract

Background: There are a growing number of mobile phone apps available to support people in taking their medications and to improve medication adherence. However, little is known about how these apps differ in terms of features, quality, and effectiveness.

Objective: We aimed to systematically review the medication reminder apps available in the Australian iTunes store and Google Play to assess their features and their quality in order to identify high-quality apps.

Methods: This review was conducted in a similar manner to a systematic review by using a stepwise approach that included (1) a search strategy; (2) eligibility assessment; (3) app selection process through an initial screening of all retrieved apps and full app review of the included apps; (4) data extraction using a predefined set of features considered important or desirable in medication reminder apps; (5) analysis by classifying the apps as basic and advanced medication reminder apps and scoring and ranking them; and (6) a quality assessment by using the Mobile App Rating Scale (MARS), a reliable tool to assess mobile health apps.

Results: We identified 272 medication reminder apps, of which 152 were found only in Google Play, 87 only in iTunes, and 33 in both app stores. Apps found in Google Play had more customer reviews, higher star ratings, and lower cost compared with apps in iTunes. Only 109 apps were available for free and 124 were recently updated in 2015 or 2016. Overall, the median number of features per app was 3.0 (interquartile range 4.0) and only 18 apps had ≥9 of the 17 desirable features. The most common features were flexible scheduling that was present in 56.3% (153/272) of the included apps, medication tracking history in 54.8% (149/272), snooze option in 34.9% (95/272), and visual aids in 32.4% (88/272). We classified 54.8% (149/272) of the included apps as advanced medication reminder apps and 45.2% (123/272) as basic medication reminder apps. The advanced apps had a higher number of features per app compared with the basic apps. Using the MARS instrument, we were able to identify high-quality apps that were rated as being very interesting and entertaining, highly interactive and customizable, intuitive, and easy to use and to navigate as well as having a high level of visual appeal and good-quality information.

Conclusions: Many medication reminder apps are available in the app stores; however, the majority of them did not have many of the desirable features and were, therefore, considered low quality. Through a systematic stepwise process, we were able to identify high-quality apps to be tested in a future study that will provide evidence on the use of medication reminder apps to improve medication adherence.

Introduction

Nonadherence to long-term therapies in chronic diseases is a global concern highlighted by the World Health Organization report in 2003 [1]. Medication nonadherence is associated with increased risk of morbidity [2], mortality [3], and costs [4]; therefore, there is a need for effective interventions to improve adherence. It is known that current interventions provide inconsistent results in improving adherence [5]. In recent years, the growing mobile phone ownership [6] has made mobile phones a promising tool to deliver health care interventions. Furthermore, there has been an increasing interest in using mobile phones as a tool to improve medication adherence.

Reminders sent via text messages have been shown to improve adherence in chronic diseases [7]. The recent growth in mobile phone subscriptions [6], however, has spawned an exponential increase in the number of health-related mobile phone apps available in the app stores, including those dedicated to improve medication-taking behavior. These medication adherence apps have many features, including reminders, that may help patients take their medication correctly and avoid medication errors and, hence, could address known barriers to adherence [8], especially for patients with high pill burden and complex regimens, such as patients with cardiovascular diseases.

Despite this plethora of medication adherence apps, there is a lack of information on how they differ, how many and which features they have, their overall quality, and whether they are effective. Previous reviews have identified available medication adherence–related apps and described the relevant features present in these apps [9-11]. However, these reviews only provided a descriptive analysis of the available apps and their features without a deeper quality assessment. The aim of this research was to describe a systematic and stepwise process to identify high-quality medication reminder apps by identifying and reviewing the current available apps and their features and to assess the apps' quality by using a reliable quality assessment tool for mobile health apps.

Methods

Design

This review was conducted in a similar manner to a systematic review by using a stepwise approach that included a search strategy, prespecified eligibility criteria, app selection through an initial screening of all retrieved apps and full app review of the included apps, data extraction and analysis, and quality assessment of selected apps using a reliable quality assessment tool for mobile health apps.

Search Strategy

The search was conducted in the main online app stores, iTunes (Apple Inc, Australia) and Google Play (Google Inc, Australia), that have more than 2 million apps available for download [12]. The apps available in these app stores are compatible with any mobile phone that uses the leading operating systems in Australia, iOS and Android systems, which together account for 97% of the Australian mobile phone market [13]. We, therefore, searched the Australian iTunes and Google Play app stores between December 10, 2015, and December 20, 2015, using 8 search terms that during the preliminary searches had the best performance in retrieving the apps of interest for this review. The search terms used were *medication reminder*, *medication pill reminder*, *pill reminder*, *meds reminder*, *medication tracker*, *medication management*, *Rx*, and *medication*.

Eligibility Criteria

Apps were eligible to be included in this review if they met all the following inclusion criteria: (1) apps that aimed to support medication self-management, (2) apps capable of generating scheduled reminders for medication-taking behavior, and (3) apps that were in English.

As we aimed to include apps that could be used by a large number of patients or individuals, we excluded apps that were restricted to a specific group of individuals or type of medication, with the exception of apps related to cardiovascular diseases that have the highest prevalence worldwide. Therefore, apps were excluded from the review if they (1) generated general reminders not specifically designed for medication reminders (eg, general calendars and alarms); (2) focused on one medication (eg, contraception); (3) focused on individuals with only one medical condition, except cardiovascular diseases (eg, asthma); (4) focused on one specific group of people (eg, seniors); (5) only listed medication prescribed without sending reminders; (6) generated reminders for medication refill or expiration date without daily reminders for medication adherence; (7) were designed for ordering medication refills online (eg, pharmacy-owned apps); (8) were owned by health care services targeting only their own patients (eg, hospitals and family practices); (9) focused on general health, fitness, lifestyle, and well-being; (10) only provided medication information (eg, medication dosing information and side effects); and (11) lacked enough information to determine eligibility.

App Selection Process

One author (KS) carried out the app store searches. Information about the apps retrieved from the search was entered into a predesigned electronic spreadsheet developed for this review. The information entered included name of the app, name of app developer, cost, app store or stores in which the app was available, and the search term or terms that retrieved the app. Apps available in both app stores were only entered once in the electronic spreadsheet, including same apps that had slightly different names or app developers' name in the 2 app stores. Apps that had 2 versions in the same app store, for example, a free or lite version and a paid or pro version, were entered separately in the spreadsheet as the features may differ in each version.

All apps retrieved from the search were screened for eligibility by 1 reviewer (KS). The screening process consisted of reviewing the information available about the app characteristics on the product list in each app store, including written information, pictures, and videos. Additional apps found during the data extraction process were also added to the initial spreadsheet and screened for eligibility. Apps were included in this review if they met all the inclusion criteria, and the reasons for exclusion were recorded. A full detailed review of the apps included in this review was performed.

Data Extraction and Analysis

Before data extraction, a set of features considered important or desirable in medication reminder apps was developed for this review based on previous reviews [9-11] and on a panel consensus (KS, CKC, and JR). The set of important features consisted of 3 practical and 17 functionality features (Table 1). To assess the features present in the apps, the information available about the app from the app store, including the written description, photos, and videos, was extracted. If an app was available in both app stores, the information available in both app stores was combined for the review. In addition, information about the app or the app developer provided on its own website was also extracted and assessed. The information extracted included name of the app, name of the app developer, app store in which the app was available, star rating, number of reviews, cost, and the last date updated, as well as the practical and functionality features. For each practical and functionality feature, the reviewer determined if the criteria were present or absent. If the presence of a feature could not be ascertained by evaluating the different sources of information described above, the feature was considered to be absent in the app.

Table 1. Practical and functionality features' description.

Features	Rationale or description
Practical features	
Available in both app stores	Allows the app to be used by individuals who own mobile phones that use the leading operating systems (iOS and Android).
Available for free without ads	A full free version of the app without advertising for third-party products is likely to be used by a large number of individuals.
Updated in 2015 or 2016	A recent update ensures ongoing technical support to fix any software issues.
Functionality features	
Medication tracking history	Ability to record and track taken and missed doses.
Snooze option	Ability to snooze the reminder for a predetermined period of time, for example, 15 minutes.
Flexible scheduling	Ability to schedule reminders to occur on a nondaily or monthly basis or every X days, or ability to schedule medications with stop dates.
Medication database	Availability of a medication database that allows the user to search and select a medication from the database.
Refill reminders	Ability to set reminders to the date when medication needs to be refilled.
Reminders with no connectivity	No Internet connection required for the reminders to function.
Data exporting and sharing	Ability to export and share the medication information to a third party, for example, family member or health care provider.
Multiple users support	Ability to generate medication reminders on different medications for more than 1 user, for example, family members.
Notification for other people	Availability of an option to alert other people about when to take their medication or when missed doses are registered.
Data security	The app developer ensures data security, for example, data are only stored in the user's device or statement of HIPAA[a] compliance.
Data privacy: password protection	Password option to access the app.
Multilingual	Availability of languages other than English.
Time zone support	Ability to change time zones to ensure medication is taken at the right time when traveling.
Adherence rewards	Availability of a feature that rewards the patient when the medication is taken on schedule, for example, awarding points for each medication taken that could be redeemed into vouchers.
Adherence statistics and charts	Availability of statistics and charts describing medication usage trends and adherence rates.
Customizable alert sounds	Availability of different types of notification sounds.
Visual aids	Availability of icons (eg, tablet, syringe, drops) or ability to add a picture to provide visual clues (eg, to ensure correct medication is taken).

[a]HIPAA: Health Insurance Portability and Accountability Act.

Extracted data were entered into an electronic spreadsheet and were analyzed using IBM SPSS version 22.0 (IBM Corporation). App characteristics and features were summarized as means or medians for continuous data and as frequencies and proportions for categorical data. We further analyzed the included apps by (1) classifying them as basic or advanced medication reminder apps and (2) ranking the apps using a scoring system developed for this review.

Classification of Apps

On the basis of the presence or absence of the medication tracking history feature, each app was classified into the following categories that were previously described by Stawarz et al [10]: (1) basic medication reminder apps and (2) advanced medication reminder apps.

The first type of app offers basic features to support prospective memory, by providing daily, simple, timed reminders to reinforce medication-taking behavior with no further interactivity. This basic app acts similarly to an alarm or a text message by showing a reminder on the mobile phone screen at a set time every day.

The second type of app offers not only the same basic features to support prospective memory, by providing the same daily, simple, timed reminders to reinforce medication-taking behavior, but also additional features to support retrospective memory, by having the ability to track the medications taken or missed, as well as having more customizable and interactive features, such as adherence statistics, time zone support, data sharing, and multiple user support.

App Ranking

The apps in each of the categories described above were ranked based on a scoring system developed for this review by assessing the number of practical and functionality features detailed in Table 1. To rank the apps, we calculated a score for each app by adding 2 points for each practical feature present in the app and 1 point for each functionality feature. The practical features were given a higher weight in the scoring system as these features were considered important to ensure that the apps are available for a large number of people. The total number of points possible was therefore 23, where a maximum of 6 points were available for practical features and a maximum of 17 points were available for functionality features. The apps were then ranked from highest to lowest, where those with the most practical and functional features were ranked highest.

Quality Assessment Using the Mobile App Rating Scale

As it was not feasible to download all the apps included in this review, we decided to select the top 5 scoring apps in each category (basic and advanced medication reminder apps) for further assessment using the Mobile App Rating Scale (MARS) [14]. The MARS tool is a 23-item scale developed by researchers to assess the quality of mobile health apps. The MARS instrument provides a deeper evaluation of the app quality by testing the app thoroughly for 10 minutes and grading the app in several domains, including user engagement, functionality, aesthetics, information, and app subjective quality. Each item was scored using a 5-point scale (1-Inadequate,

2-Poor, 3-Acceptable, 4-Good, 5-Excellent). For each domain, we calculated a mean score that ranged from 0 to 1, where a score of 0 would mean inadequate quality and a score of 1 would mean excellent quality.

In the selection process for app download, if more than 5 apps had the same scores, the apps were selected for download using the following predefined hierarchy: apps available in both app stores, apps for free, and apps updated in 2015 or 2016. Two reviewers (KS and SR) downloaded and independently tested the selected apps using the MARS instrument. The reviewers were trained to use the MARS instrument by watching an online tutorial to ensure that both reviewers used the tool in the same manner. One reviewer (KS) assessed the apps using an iOS device and the other reviewer (SR) used an Android device. In addition, the reviewers were required to check if all the functionality features described in the app store were, in fact, present in the app. Any disagreement was discussed and resolved by consensus.

Results

Search

The process of identification and inclusion of apps is outlined in Figure 1. A total of 1471 apps were screened for eligibility and 1199 apps were excluded for several reasons presented in Figure 1. A total of 1042 apps were found only in Google Play, 351 only in iTunes, and 73 in both app stores. Another 5 apps were identified during the data extraction process and were also screened for eligibility. After exclusions, 272 apps were eventually included in this review.

General Characteristics of Included Apps

Of the 272 apps included, 55.9% (152/272) were found only in Google Play, 32.0% (87/272) only in iTunes, and 12.1% (33/272) in both app stores. A total of 34 apps had 2 versions: a free or lite version and a paid or pro version. In Google Play, 91.9% (170/185) of the apps were reviewed by customers who gave a star rating to the apps (star ratings range from 0 to 5 stars). The median star rating for Google Play apps was 3.9 (interquartile range, IQR, 0.82) with a minimum of 1.0 star given to 2 apps and a maximum of 5.0 stars given to 17 apps. The median number of reviews per app was 22.5 (IQR 78.75), with a minimum of 1 review for 12 apps and a maximum of 98,179 reviews for 1 app (Medisafe app). In terms of cost, 20.5% (38/185) of the apps found in Google Play required a payment for download at a median cost of Aus $1.88 (IQR 1.63) and a range of Aus $0.99 to Aus $5.70.

In the iTunes store, only 34.2% (41/120) of the apps were reviewed by customers who gave a star rating to the apps (star ratings range from 0 to 5 stars). The median star rating for iTunes apps was 3.0 (IQR 2.75) with a minimum of 1.0 star given to 9 apps and a maximum of 5.0 stars given to 5 apps. The median number of reviews per app was 2.0 (IQR 9.5), with a minimum of 1 review for 11 apps and a maximum of 153 reviews for 1 app (Medisafe app). In terms of cost, 48.3% (58/120) of the apps found in iTunes required a payment for download at a median cost of Aus $2.99 (IQR 3.00) and a range of Aus $1.49 to Aus $42.99.

**Figure 1.** Flowchart of selection of included apps.
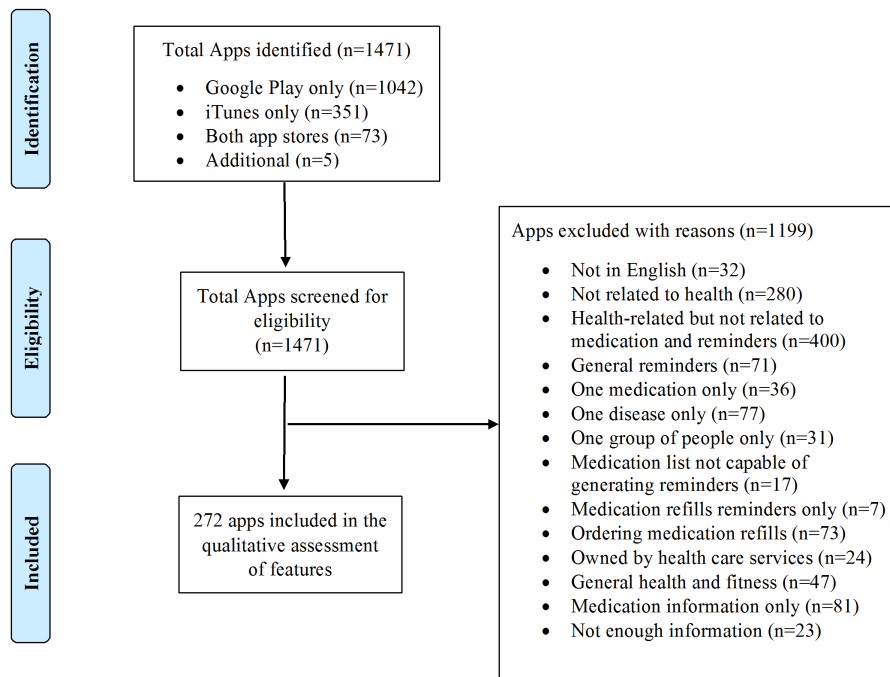
Features of Included Apps

In terms of the practical features, as stated above, 12.1% (33/272) of the apps were available in both app stores. In addition, 40.1% (109/272) were fully available for free without third-party advertisement or in-app purchases and 45.6% (124/272) were recently updated in 2015 or 2016. In terms of functionality, the median number of features per app was 3.0 (IQR 4.0) and only 6.6% (18/272) of the apps had at least 9 of

the 17 desirable features. Flexible scheduling and medication tracking history were the only 2 features present in more than half of the apps. Other common functionality features were snooze option, visual aids, customizable alert sounds, multiple users support, data exporting and sharing, and languages other than English, which were present in around a third of the apps. All the other functionality features were present in less than a quarter of the apps (Figure 2).

Figure 2. Results of functionality criteria assessment.
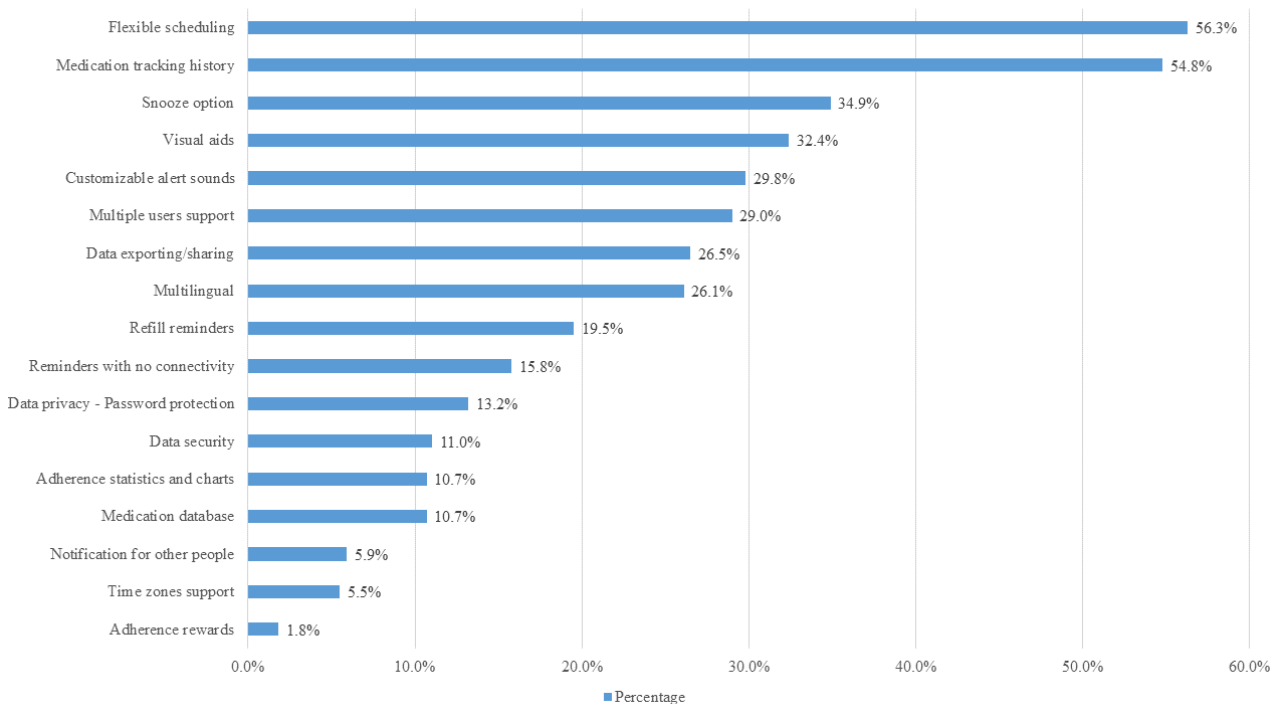
Classification and Ranking of Included Apps

In total, 54.8% (149/272) of the included apps were classified as advanced medication reminder apps as they had the ability to track taken and missed doses, while the other 45.2% (123/272) were classified as basic medication reminder apps. In terms of functionality features according to the classification group, the advanced apps had more than double the number of features compared with the basic apps, having a median of 5 (IQR 4) and 2 (IQR 2) features per app, respectively. Among the advanced apps, the apps with the highest number of functionality features were Medisafe and Pill Reminder by

Drugs.com with 14 features, while among the basic apps it was AlarMeds reminder app with 7 features.

Regarding the ranking of the apps, Medisafe was ranked number 1 among the advanced medication reminder apps, achieving 20 out of a maximum of 23 points (Table 2). The median score among the advanced apps was 7 (IQR 5), having a score range of 1 to 20 points. Among the basic medication reminder apps, My heart, my life was the highest-scoring app, achieving 9 points. The median app score in this category was 4 (IQR 3), having a score range of 0 to 9. The practical and functionality features in the top-ranking apps among the advanced and basic medication adherence apps are presented in Tables 2 and 3.

Table 2. Rank and score of and practical features present in the top advanced and basic medication reminder apps.

Rank	App names	Score	Available in both app stores	Full version available for free	Updated in 2015 or 2016
Advanced apps					
1	Medisafe	20	✓[a]	✓	✓
2	Dosecast	15	✓		✓
3	MyMeds	15	✓		✓
4	CareZone	14	✓	✓	✓
5	My Pillbox	14	✓		✓
6	MedicineList+[b]	14	✓	✓	✓
Basic apps					
1	My heart, my life	9	✓	✓	✓
2	MediWare[c]	8	✓	✓	✓
3	MyMedManager	8	✓	✓	
4	Pill Reminder (Aplicativos Legais)[d]	8	✓		✓

[a]The symbol ✓ means that the feature was present in the app when tested in both iOS and Android devices.

[b]Initially evaluated as a basic medication reminder app but after download classified as an advanced app as the medication tracking history feature was present.

[c]Could not be assessed owing to crashes and technical issues in both iOS and Android devices.

[d]Only assessed on an iOS device because of technical problems in the Android device.

Table 3. Functionality features present in the top advanced and basic medication reminder apps.

App names	Functionality features[a]																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Advanced apps																	
Medisafe	✓[b]	✓	✓	✓	✓	✓	✓	✓	✓			✓	✓		✓	✓	✓
Dosecast	✓	✓	✓		✓	✓	✓	✓		✓						✓	✓
MyMeds	✓	✓		✓	✓	✓		✓		✓	✓			✓	✓		✓
CareZone	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓			✓		✓		✓
My Pillbox	✓	✓	✓		✓	✓	✓	✓			✓					✓	✓
MedicineList+[c]	✓	✓	✓	✓		✓	✓	✓		✓							
Basic apps																	
My heart, my life				✓	✓					✓	✓						✓
MediWare[d]																	
MyMedManager		✓	✓			✓										✓	
Pill Reminder (Aplicativos Legais)[e]																	

[a]See Table 1 for the 17 functionality features.

[b]The symbol ✓ means that the feature was present in the app when tested in both iOS and Android devices.

[c]Initially evaluated as a basic medication reminder app but after download classified as an advanced app as the medication tracking history feature was present.

[d]Could not be assessed owing to crashes and technical issues in both iOS and Android devices.

[e]Only assessed on an iOS device because of technical problems in the Android device.

Quality Assessment Using the Mobile App Rating Scale

Medisafe was the highest-scoring app of the 10 apps assessed using the MARS instrument; as it had the highest scores in the engagement and aesthetics domains because it was found to be interesting, entertaining, highly interactive, and customizable and to have a high level of visual appeal (Table 4). In addition, Medisafe had the maximum score in the subjective quality section, meaning that the reviewers would use this app regularly and recommend it to others. Medisafe was also the only app rated as having some evidence supporting its effectiveness in nonrandomized studies. The My heart, my life app had the maximum score in the functionality domain as it was intuitive, was easy to use and to navigate, and did not present any technical issues during use. The MedicineList+ app had the highest score in the information domain as it provided high-quality information from a credible source called National Prescribing Service MedicineWise Australia, which is an independent not-for-profit organization.

Table 4. The Mobile App Rating Scale mean scores assessed by domains.

App names	Mean scores by domains[a]					
	Engagement	Functionality	Aesthetics	Information	Subjective quality	MARS[b] total score
Advanced apps						
Medisafe	0.94	0.90	0.93	0.83	1.00	0.92
MedicineList+[c]	0.74	0.90	0.87	0.93	0.78	0.84
CareZone	0.78	0.90	0.93	0.84	0.70	0.83
My Pillbox	0.76	0.83	0.77	0.68	0.53	0.71
Dosecast	0.56	0.90	0.80	0.80	0.55	0.70
MyMeds	0.52	0.75	0.87	0.55	0.48	0.63
Basic apps						
My heart, my life	0.60	1.00	0.83	0.82	0.58	0.77
MyMedManager	0.50	0.83	0.70	0.63	0.33	0.60
Pill Reminder (Aplicativos Legais)[d]	0.36	0.80	0.60	0.33	0.25	0.47
MediWare[e]	0.00	0.00	0.00	0.00	0.00	0.00

[a]Mean score ranges from 0 to 1, where a score of 0 means inadequate quality and a score of 5 means excellent quality.

[b]MARS: Mobile App Rating Scale.

[c]Initially evaluated as a basic medication reminder app but after download classified as an advanced app as the medication tracking history feature was present.

[d]Only assessed on an iOS device because of technical problems in the Android device.

[e]Could not be assessed owing to crashes and technical issues in both iOS and Android devices.

On the basis of the MARS assessment, Medisafe app was evaluated as the best app currently available in the app stores overall and among the advanced medication reminder apps, while My heart, my life was the best available app among the basic medication reminder apps. MedicineList+ app was classified as a basic medication reminder app in the initial assessment of features during the data extraction process; however, after the app was downloaded for quality assessment using the MARS tool, it was reclassified as an advanced app as this app had the ability to record and track taken and missed doses.

Discussion

Principal Findings

This review documents a systematic stepwise process to identify high-quality medication reminder apps. A comprehensive search identified 272 medication reminder apps, of which only a small number of apps were available in both app stores. Importantly, less than half of the apps were fully available at no cost and have been recently updated. In addition, the average number of desirable features per app was low and only a very small number of apps had more than half of these important features, and, therefore, the majority of apps were considered low quality. About half of them were classified as advanced medication reminder apps, while the other half was classified as basic apps. As expected, advanced apps had a higher average number of features and higher scores compared with basic apps. Through an in-depth quality assessment using a reliable tool, high-quality medication reminder apps were identified.

Previous reviews have also attempted to evaluate the availability of apps related to medication adherence in the app stores and their features. Similar to our results, other authors have found a large number of medication reminder apps available in the app stores with only a small number being available in more than one app store [9-11,15]. In addition, Bailey et al [11] also found that approximately half of the apps had a medication tracking history feature. However, these reviews only provided a descriptive analysis of the available apps and their features without a deeper quality assessment. Furthermore, these reviews are dated and were performed using different app stores in the United States and United Kingdom. It is important to mention that the app stores in different countries have different apps available, as the choice of countries in which the apps are available is determined by the app developers [16,17]. In this review, among the 10 apps selected for download, 2 apps (My heart, my life and MedicineList+) are available only in Australia as they were developed by Australian not-for-profit organizations.

It is important to highlight that the current mobile health app market is poorly regulated. The app stores provide guidelines about restricted content, privacy and security of the data, and monetization of the apps; however, these guidelines are not a quality control assessment of the available apps. Recently, the US Food and Drug Administration (FDA) released a guidance document stating which type of mobile medical apps will be subject to their regulation [18]. However, at this stage, apps to promote medication adherence will not be within the FDA

regulation oversight, as these apps are not intended to provide diagnosis or treatment recommendations.

The lack of quality control and assessment makes it difficult for individuals to choose and even for health professionals to recommend high-quality apps to their patients. In this work, the described process to identify high-quality apps might be useful to guide others researchers, clinicians, and stakeholders on how to assess mobile health apps. First, a systematic and comprehensive search of the app stores is needed to ensure that all available apps are identified. Second, inclusion and exclusion criteria should be predefined to select apps that are appropriate for a target population. Third, the features that are considered essential and desirable to the apps being evaluated should be predetermined so that higher-quality apps can be identified. Finally, download and testing of selected apps using a quality assessment tool can confirm whether these apps are of high quality.

This review is not without limitations. The search conducted in the Australian app stores might have restricted the results of this review as some of the selected apps are only found in Australia. However, one of the aims of this review was to detail a process for use by future researchers who might need to identify and assess apps in their region and area of interest. In addition, the rigorous eligibility criteria might have resulted in the exclusion of some good-quality medication reminder apps that might be suitable for specific groups of patients. We also acknowledge our inability to download and assess all the included apps; however, we believe that this limitation did not compromise our results. Although we were able to identify high-quality medication reminder apps, currently there is no evidence that these apps are effective in improving medication adherence. To fill this gap in knowledge, our team is designing a study (including a qualitative component) to test the high-quality apps identified in this review. If proven effective, medication reminder apps can have an impact on clinical practice as they can be used as an additional tool among other strategies to improve adherence.

Conclusions

In the current technology-driven world, apps have been gaining space in our everyday lives. Health apps, including medication reminder apps, are becoming more and more popular and are a promising tool to improve people's health. In this review, we found a large number of medication reminder apps available in the app stores; however, majority of them were considered low quality. Through a systematic stepwise process, we were able to identify high-quality apps to be tested in a future study that will provide evidence on the use of medication reminder apps to improve medication adherence.

Conflicts of Interest

None declared.

References

1. World Health Organization. Adherence to long-term therapies: evidence for action. Switzerland: WHO Library Cataloguing-in-Publication Data; 2003. URL: http://apps.who.int/iris/bitstream/10665/42682/1/9241545992.pdf [accessed 2016-09-28] [WebCite Cache ID 6kqutIExA]
2. Chowdhury R, Khan H, Heydon E, Shroufi A, Fahimi S, Moore C, et al. Adherence to cardiovascular therapy: a meta-analysis of prevalence and clinical consequences. Eur Heart J 2013 Oct;34(38):2940-2948 [FREE Full text] [doi: 10.1093/eurheartj/eht295] [Medline: 23907142]
3. Rasmussen JN, Chong A, Alter DA. Relationship between adherence to evidence-based pharmacotherapy and long-term mortality after acute myocardial infarction. JAMA 2007 Jan 10;297(2):177-186. [doi: 10.1001/jama.297.2.177] [Medline: 17213401]
4. Roebuck MC, Liberman JN, Gemmill-Toyama M, Brennan TA. Medication adherence leads to lower health care use and costs despite increased drug spending. Health Aff (Millwood) 2011 Jan;30(1):91-99 [FREE Full text] [doi: 10.1377/hlthaff.2009.1087] [Medline: 21209444]
5. Nieuwlaat R, Wilczynski N, Navarro T, Hobson N, Jeffery R, Keepanasseril A, et al. Interventions for enhancing medication adherence. Cochrane Database Syst Rev 2014;11:CD000011. [doi: 10.1002/14651858.CD000011.pub4] [Medline: 25412402]
6. International Telecommunication Union. Measuring the Information Society Report 2014. URL: http://www.itu.int/en/ITU-D/Statistics/Documents/publications/mis2014/MIS2014_without_Annex_4.pdf [accessed 2016-09-28] [WebCite Cache ID 6kqtdeLwz]
7. Thakkar J, Kurup R, Laba T, Santo K, Thiagalingam A, Rodgers A, et al. Mobile Telephone Text Messaging for Medication Adherence in Chronic Disease: A Meta-analysis. JAMA Intern Med 2016 Mar;176(3):340-349. [doi: 10.1001/jamainternmed.2015.7667] [Medline: 26831740]

XSL·FO

RenderX

8.   Osterberg L, Blaschke T. Adherence to medication. N Engl J Med 2005 Aug 4;353(5):487-497. [doi: 10.1056/NEJMra050100] [Medline: 16079372]

9.   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]

10.   Stawarz K, Cox A, Blandford A. Don't forget your pill!: designing effective medication reminder apps that support users' daily routines. 2014 Presented at: 32nd annual Association for Computing Machinery (ACM) conference on Human factors in computing systems; April 26 - May 01, 2014; Toronto, Ontario, Canada p. 2269-2278. [doi: 10.1145/2556288.2557079]

11.   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]

12.   Statista Inc. Number of apps available in leading app stores as of June 2016. 2016 Jun. URL: http://www.statista.com/statistics/276623/number-of-apps-available-in-leading-app-stores/ [accessed 2016-09-15] [WebCite Cache ID 6kqsVd38B]

13.   Snakk Blog - Australian mobile advertising and media. 3 Ways to Understand Australian Smartphone Usage. 2016 Aug 2. URL: http://blog.snakkmedia.com/3-ways-to-understand-australian-smartphone-usage/ [accessed 2016-09-28] [WebCite Cache ID 6mIcLdRXy]

14.   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]

15.   van Kerkhof LW, van der Laar CW, de Jong C, Weda M, Hegger I. Characterization of Apps and Other e-Tools for Medication Use: Insights Into Possible Benefits and Risks. JMIR Mhealth Uhealth 2016 Apr 06;4(2):e34 [FREE Full text] [doi: 10.2196/mhealth.4149] [Medline: 27052946]

16.   Google Inc. The Secrets to App Success on Google Play. 2015. URL: https://commondatastorage.googleapis.com/androiddevelopers/shareables/distribute/secrets_play/v2/web/secrets_to_app_success_v2_en.pdf [accessed 2016-10-30] [WebCite Cache ID 6ldWJPk86]

17.   Apple Inc. iTunes Connect Developer Guide - Changing the Territories Rights Information. 2016. URL: https://developer.apple.com/library/content/documentation/LanguagesUtilities/Conceptual/iTunesConnect_Guide/Chapters/ChangingAppStatus.html#//apple_ref/doc/uid/TP40011225-CH30-SW1 [accessed 2016-10-30] [WebCite Cache ID 6ldWhAIrB]

18.   Food and Drug Administration. Mobile medical applications - Guidance for Industry Food Drug Administration staff. 2015. URL: http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM263366.pdf [accessed 2016-09-28] [WebCite Cache ID 6kqtH9qBw]

Abbreviations

FDA: US Food and Drug Administration
IQR: interquartile range
MARS: Mobile App Rating Scale

XSL•FO
RenderX

Original Paper

Smartphone Apps for Measuring Human Health and Climate Change Co-Benefits: A Comparison and Quality Rating of Available Apps

Rachel K Sullivan[1], BSc; Samantha Marsh[1], PhD; Jakob Halvarsson[2]; Michelle Holdsworth[3], RNutr, RD, PhD; Wilma Waterlander[1], PhD; Maartje P Poelman[4], PhD; Jennifer Ann Salmond[5], PhD; Hayley Christian[6], PhD; Lenny SC Koh[7], PhD; Janet E Cade[8], RNutr, PhD; John C Spence[9], PhD; Alistair Woodward[10], PhD; Ralph Maddison[1,11], PhD[‡]

[1]National Institute for Health Innovation, School of Population Health, The University of Auckland, Auckland, New Zealand

[2]Faculty of Medicine, Linköping University, Linköping, Sweden

[3]School of Health and Related Research, Section of Public Health, University of Sheffield, Sheffield, United Kingdom

[4]Department of Human Geography and Spatial Planning, Utrecht University, Utrecht, Netherlands

[5]School of Environment, Faculty of Science, The University of Auckland, Auckland, New Zealand

[6]School of Population Health and Telethon Kids Institute, The University of Western Australia, Perth, Australia

[7]Centre for Energy, Environment and Sustainability, Management School, University of Sheffield, Sheffield, United Kingdom

[8]Nutritional Epidemiology Group, Schools of Food Science and Nutrition, University of Leeds, Leeds, United Kingdom

[9]Faculty of Physical Education and Recreation, University of Alberta, Edmonton, AB, Canada

[10]Epidemiology and Biostatistics, School of Population Health, The University of Auckland, Auckland, New Zealand

[11]Institute for Physical Activity and Nutrition, Deakin University, Victoria, Australia

[‡]CHANGE-IT Research Collaboration

Corresponding Author:
Ralph Maddison, PhD
Institute for Physical Activity and Nutrition
Deakin University
Locked Bag 20000
Geelong
Victoria,
Australia
Phone: 61 3 52271100
Fax: 61 3 52271100
Email: ralph.maddison@deakin.edu.au

Abstract

Background: Climate change and the burden of noncommunicable diseases are major global challenges. Opportunities exist to investigate health and climate change co-benefits through a shift from motorized to active transport (walking and cycling) and a shift in dietary patterns away from a globalized diet to reduced consumption of meat and energy dense foods. Given the ubiquitous use and proliferation of smartphone apps, an opportunity exists to use this technology to capture individual travel and dietary behavior and the associated impact on the environment and health.

Objective: The objective of the study is to identify, describe the features, and rate the quality of existing smartphone apps which capture personal travel and dietary behavior and simultaneously estimate the carbon cost and potential health consequences of these actions.

Methods: The Google Play and Apple App Stores were searched between October 19 and November 6, 2015, and a secondary Google search using the apps filter was conducted between August 8 and September 18, 2016. Eligible apps were required to estimate the carbon cost of personal behaviors with the potential to include features to maximize health outcomes. The quality of included apps was assessed by 2 researchers using the Mobile Application Rating Scale (MARS).

Results: Out of 7213 results, 40 apps were identified and rated. Multiple travel-related apps were identified, however no apps solely focused on the carbon impact or health consequences of dietary behavior. None of the rated apps provided sufficient information on the health consequences of travel and dietary behavior. Some apps included features to maximize participant

XSL·FO

RenderX

engagement and encourage behavior change towards reduced greenhouse gas emissions. Most apps were rated as acceptable quality as determined by the MARS; 1 was of poor quality and 10 apps were of good quality. Interrater reliability of the 2 evaluators was excellent (ICC=0.94, 95% CI 0.87-0.97).

Conclusions: Existing apps capturing travel and dietary behavior and the associated health and environmental impact are of mixed quality. Most apps do not include all desirable features or provide sufficient health information. Further research is needed to determine the potential of smartphone apps to evoke behavior change resulting in climate change and health co-benefits.

Introduction

Reducing the impact of climate change on our planet and reducing the burden of noncommunicable diseases (NCDs) are major global challenges [1,2]. A strong link exists between climate change and public health, which can be demonstrated by measuring greenhouse gas (GHG) emissions. For example, increased use of motorized transport leads to increased GHG emissions, thereby contributing to climate change as well as reducing physical activity levels, which has been linked to the development of many NCDs [3-7]. Furthermore, our current food production system is one of the most important contributors to global GHG emissions, where certain dietary changes (eg, eating less meat and fewer calories) can benefit the environment and public health, especially for adult populations in high-income countries [3,8-10]. The unique connection between these 2 issues presents an opportunity to make changes which have co-benefits to health and the environment. Transport and diet are 2 areas in which behavioral change may offer the greatest potential for reducing the impact of both climate change and NCDs [5,7].

It is estimated that transport accounts for approximately 22% of the world's energy-related carbon emissions due to increased reliance on motorized transport within both developed and developing countries [11-14]. While motorized transport may be time-efficient in the modern world, it contributes to the problem of physical inactivity and sedentary behavior, which have been identified as major behavioral risk factors contributing to many NCDs and their determinants, causing approximately 3.2 million deaths per year [3,4]. The World Health Organization predicts that 7.7% of the total mortality risk within high income countries is attributed to physical inactivity alone, with a further 8.4% associated with being overweight or obese [15]. However, the prevalence of physical inactivity and the associated burden of chronic disease could be lowered with small changes to individual travel behavior, such as reduced vehicle use and increased active travel (ie, walking or cycling) within urban areas [5,7,12-14,16-18]. Such changes also reduce the impact of climate change by lowering GHG emissions.

Agriculture accounts for approximately 31% of global GHG emissions with most emissions coming from livestock production, including methane from ruminant digestion, nitric oxide from fertilizer use, and carbon dioxide from deforestation/felled vegetation and fossil fuel use [6,19]. Livestock production is expected to rise substantially over the coming years to meet rapidly growing demands driven by population growth, economic growth, and urbanization within low and middle income countries, thereby exacerbating the effects of climate change [20]. In addition to this, globalization has seen increases in diets characterized by high energy (calories), saturated fat, free sugars, and salt, commonly associated with obesity and the development of many NCDs such as type 2 diabetes, ischemic heart disease, and some cancers [4,10,21]. Not only are these diets detrimental to health, they tend to focus on a narrow range of food crops, which increases the vulnerability of food supply to diseases, pests, and weather extremes that may arise with climate change, thereby also threatening food security [2]. Therefore, a universal shift away from a globalized diet toward reduced consumption of animal products and energy dense foods not only has the potential to reduce the burden of NCDs [4,8,9], it has also been predicted to significantly reduce GHG emissions due to a reduction in livestock production and associated emissions as well as increased land availability suitable for growing alternative food or the regrowth of native vegetation [22].

To date, studies of ways to reduce GHG emissions within transport and agricultural sectors have mainly focused on high-level modulations of potential policy changes [5,6,12,14,16] or technological and managerial approaches such as improving productivity, restoring soil carbon, optimizing nutrient use and fertilizers, improving livestock diets, and better management of waste [10,19,20]. However, little research has been done to examine the ways in which personal behavior affects both GHG emissions and the incidence of NCDs [17,19,23]. Although at the individual level the net effect of changing lifestyles (eg, changing 2 daily commutes per week from driving to cycling) might be relatively small, the impact at the population level could be substantial [14,16]. For example, studies show that if 10% of Canadians who are currently inactive or sedentary swap to active commuting by walking instead of driving and sit less, it could result in cost savings to the health care system of CaD $2.6 billion (US $2.0 billion) by 2040 and a cumulative CaD $7.5 billion (US $5.7 billion) boost to the Canadian gross domestic product [24]. Modeled data from New Zealand also showed that shifting 5% of vehicle kilometers to cycling would reduce vehicle travel by approximately 223 million kilometers each year and reduce transport-related GHG emissions by 0.4%, along with reduced NCD-related mortality [7].

XSL•FO
RenderX

A clear need exists for cost effective, minimal intervention strategies to promote and motivate individual behavior change. These strategies should focus on simple clear steps such as "leave the car at home today" [7,16] and "eat a vegetarian meal" [21,25]. Individual knowledge of small simple steps that shrink personal carbon footprints and reduce disease risk may increase peoples' awareness and willingness to engage with large-scale problems such as climate change and obesogenic environments, which otherwise appear impossibly daunting and remote [26]. It is therefore important to address individual level behavior in addition to top-down approaches assessing the impact of policy changes and managerial or technological interventions. Such a bottom-up approach may prove to be more successful in addressing both health and environmental issues.

Traditional Web-based carbon calculators or surveys, requiring memory recall and manual input of personal behaviors, often result in nonuse, low adherence, and therefore inaccurate or incomplete data [27,28]. However, given the ubiquitous use and availability of smartphones, potential exists to use smartphone apps to capture personal travel and dietary behaviors, while simultaneously estimating the carbon cost and health benefits of these actions through new matrices and integrative carbon-health computation methodology.

Smartphones have advantages for collecting travel data compared with traditional surveys or Global Positioning System (GPS) loggers because as they are always taken with the user and have multiple inbuilt sensors such as GPS and accelerometers allowing automatic and continuous data collection, thereby creating a vast source of data ("big data")

for immediate data analytics, feedback, and decision making. For example, these data can be used to determine the frequency and duration of motorized transport and active travel and thereby estimate personal carbon emissions. For diet, technologies exist to assess food consumption and purchasing habits [29-31], which could be adapted to estimate the associated carbon cost.

To identify relevant smartphone apps which estimate the carbon cost and health impact of personal travel and dietary behavior and possibly promote changes in these behaviors, we reviewed free and paid smartphone apps to describe their features and rate the quality of these apps against a valid quality rating tool.

Methods

Search Strategy

A list of smartphone apps for both iOS and Android operating systems was compiled between October 19, 2015, and September 18, 2016. The Apple App Store (version 12.1.3) and the Google Play Store (version 6.9.21, using the apps filter) were searched between October 19 and November 6, 2015. A secondary Google search using the apps filter was conducted between August 8 and September 18, 2016, to identify any apps that may not have been available on the New Zealand app stores. This search was carried out using Mozilla Firefox (version 45.1.0) ensuring the researcher was logged out of their personal account and searching the web (rather than limiting the search to New Zealand) for both free and paid apps. Within the app stores, independent search phrases were used to identify relevant apps (Table 1). The method of searching used in Google to identify apps is shown in Table 2.

Table 1. Search terms used and results generated within Google Play store and Apple App Store to identify diet and transport-related apps (each search statement represents a separate search).

Search term	Results App Store	Results Google Play	Downloaded
Carbon footprint	71	233	16
Carbon calculator	39	16	2
Carbon AND food	3	124	1
Carbon AND diet	0	44	0
Environment AND food	0	250	0
Environment AND diet	1	250	1
GHG[a] AND food	0	9	0
GHG AND diet	0	11	0
GHG calculator	0	9	0
CO_2[b] AND food	0	42	2
CO_2 AND diet	0	11	0
CO_2 calculator	36	114	1
CO_2 emissions AND diet	0	4	0
Emissions food	1	148	1
Emissions diet	0	27	0
Greenhouse gas AND food	0	33	0
Greenhouse gas AND diet	0	65	0
CO_2 tracker	12	109	0
Emissions calculator	17	250	1
Greenhouse gas calculator	1	25	0
Travel carbon emissions	8	86	7
Transport carbon emissions	0	132	2
Travel carbon footprint	7	62	2
Transport carbon footprint	0	74	1
Greenhouse gas emissions AND travel	0	53	0
Greenhouse gas emissions AND transport	0	66	1
GHG travel	0	21	0
GHG transport	0	28	0
CO_2 emissions travel	13	72	4
CO_2 emissions transport	1	112	1
Travel CO_2	100	74	4
Transport CO_2	100	121	0
Carbon AND travel	3	126	1
Carbon AND transport	0	138	0
Active transport CO_2	0	5	0
Sustainable transport	4	202	5
Sustainable transport CO_2	0	12	0
Sustainable travel	27	250	0
Sustainable mobility	3	118	0
Transport type	100	249	0

Search term	Results App Store	Results Google Play	Downloaded
Transport type AND carbon emissions	0	25	0
Transport type AND carbon footprint	0	14	0
Transport diary	3	128	1
Active transport	2	250	2
Active transport AND CO_2	0	6	0
Green transport	43	20	0
Commute CO_2	1	17	0
Commute carbon emissions	1	19	0

[a]GHG: greenhouse gas

[b]CO_2: carbon dioxide

Table 2. Search terms used for secondary Google search using the apps filter.

Search #	Search terms	Results	Downloaded
1	Carbon emissions AND health	60	5
	Carbon emissions AND travel	93	3
	Carbon emissions AND food	51	0
	Carbon emissions AND diet	5	0
	Search c & d	16	0
	Search a & b & e	15	0
2	Carbon emissions AND food AND travel	33	1
3	Health AND carbon emissions AND diet AND activity	19	1
4	Carbon footprint AND health	62	0
	Carbon footprint AND diet	18	0
	Carbon footprint AND food	70	2
	Carbon footprint AND travel	87	1
	Carbon footprint AND transport	63	1
	Search d & e	59	0
	Search a & f	18	0
	Search a & c	43	0
	Search g & h	5	0
5	CO_2[a] AND health	96	2
	CO_2 AND travel	78	1
	CO_2 AND food	102	0
	CO_2 AND diet	43	0
	Search a & b & d	11	0
6	Greenhouse gas emissions AND travel	32	1
	Greenhouse gas emissions AND transport	57	0
	Greenhouse gas emissions AND food	16	0
	Greenhouse gas emissions AND diet	6	0
	Greenhouse gas emissions AND health	24	0
	Search a & b & c & d & e	5	0

[a]CO_2: carbon dioxide

Inclusion Criteria

Potentially relevant apps were identified after careful consideration of titles and descriptions. Apps included in this review were in English, captured individual level travel or dietary behavior, and estimated the associated carbon cost, potentially including features to maximize health and reduce the burden of NCDs.

Due to the limitations (such as nonuse, low adherence, incompleteness, and inaccuracy) associated with manual input of personal travel behavior using carbon calculators or surveys, we searched for travel-based apps that captured personal travel data automatically and reported the associated carbon impact. However, auto-geolocation used in these apps has several limitations such as significant battery power consumption and privacy concerns, reducing its acceptability to users [28]. The majority of travel-based apps currently available are not fully automated but require users to manually initiate and terminate trip tracking or require completely manual data entry. Therefore, to be comprehensive in our search, travel-based apps were included if they recorded personal travel behavior automatically or with a manual start/stop and reported the associated carbon impact or required manual input to compare the carbon cost of different travel modes to help users make more sustainable travel choices.

Inclusion criteria for dietary apps were less specific. As dietary behavior cannot be easily recorded automatically, we searched for any apps that recorded some aspect of an individual's dietary behavior in the estimation of personal GHG emissions, with potential to include information and/or tips to reduce diet-related emissions and improve health.

Exclusion Criteria

Simple carbon calculators for transport emissions (ie, those requiring input of distance, type of fuel, and fuel consumption or engine efficiency) were excluded due to the limitations associated with manual data capture. Apps tracking travel location but not transport methods (ie, GPS tracking apps) were also excluded as they were unable to determine personal GHG emissions. Carpooling or taxi rideshare apps and health/fitness/"lifelog" style apps were also excluded if they did not include any reference to personal carbon emissions.

Out of the possible 7213 apps returned in searches, 28 potentially relevant transport-based apps and 34 dietary apps were downloaded. Following download, apps were reviewed to identify relevance for inclusion in the final sample. Of the 40 relevant apps, 11 transport-based apps and 4 dietary-based apps were offered on both iPhone and Android operating systems. In these cases, both versions were downloaded to check for consistency across both operating systems.

App Quality

Apps were rated for quality using the Mobile Application Rating Scale (MARS). This scale was developed at the Queensland University of Technology, Australia, following a comprehensive review of Web- and app-related quality rating scores [32]. The MARS includes 4 domains, engagement, functionality, aesthetics, and information, and provides an overall mean score

of the 4 domains. A separate subjective score assesses the user's overall satisfaction, and an app specific score assesses the app's ability to produce changes in the user's awareness, knowledge, understanding, attitudes, and behavior.

All apps identified during the app store searches were reviewed and rated by 2 researchers (RS and JH) on either an iPhone 5S or an iPad for iOS and a Samsung Galaxy S5 Mini or Samsung Galaxy Note II for Android. Apps identified during the secondary Google search were rated by 1 researcher (RS). Apps that were consistent across both operating systems were only rated on iOS; however, if differences in layout or content were detected both Android and iOS versions were rated.

Before rating, online training videos [33] were reviewed to ensure correct use of the MARS scale, and the modified scales were tested using 3 carbon calculators not included in this review. Travel apps were typically used for at least 2 days to record travel activity, and dietary apps were used for a minimum of 15 minutes prior to rating.

Analysis

All analysis was undertaken using SPSS Statistics version 21 (IBM Corp). Descriptive scores were calculated from the MARS scale. Independent sample, equal variance t tests identified the significance of any differences between travel and dietary apps or free and paid apps. Interrater reliability for the scores of the 2 researchers was calculated using the intraclass correlation coefficient (ICC). A 2-way mixed, absolute agreement, average measures model estimated the reliability of average measures between the 2 researchers.

Results

Search

The procedure for identifying relevant apps for inclusion is shown in Figure 1. Of the apps downloaded, 8 travel-based apps and 14 dietary apps were excluded from further analysis. Out of 22 apps, 14 apps (64%) were irrelevant for our purposes (ie, did not include dietary information, did not capture individual level behavior/purely educational, or did not include carbon emissions data), 6 apps (27%) were restricted to certain geographical locations or countries, and 2 apps (9%) did not work (ie, crashing or having problems contacting the activation server). The final sample included 40 apps (20 travel apps and 20 dietary apps).

App Content

Transport Apps

The majority (14/20, 70%) of travel-based apps recorded individual transport behavior with the aim of making users more aware of their travel-related carbon footprint and encouraging more sustainable transport. Additionally, 3 apps (15%) focused on encouraging cycling behavior (Bike da firma, Bikes vs Cars, Cycling 365), 1 app recorded cycling with a fitness/training focus but included personal emissions information (Bike Companion), and 2 apps allowed comparison of different transport modes to make sustainable choices (Green Travel Choice, TripGo). The majority of apps (13/20, 65%) captured multiple transport modes and the associated carbon impact,

XSL·FO

RenderX

while 7 apps recorded only 1 mode of transportation such as bicycle, car, bus, or airplane travel (see Multimedia Appendix 1). Although 7 apps reported the calories burned during active transportation, no other health information was included, and no apps directly mentioned the benefit of active transport in reducing sedentary behavior and improving health outcomes. Only 4 apps were completely automated (recorded trip information in the background without having to start and stop trip tracking manually).

Diet Apps

Although 1 app focused on the emissions related to the transportation of food (Food Miles Footprint), the majority of dietary apps (19/20, 95%) were simple carbon calculators or surveys that attempted to estimate an individual's carbon footprint based on inputs from multiple behavioral categories. However, only 8 of these apps (8/19, 42%) displayed the emissions result of food separately from other behaviors (such as household energy use). Dietary inputs typically focused on consumption and purchasing habits as well as information on food transport (eg, local, imported) and farming methods (eg, organic, nonorganic). Meat consumption was the most common dietary habit captured (17/19, 89%), with inputs ranging from simple multiple choice answers such as "no meat," "some meat," or "a lot of meat" to the quantity of meat consumed or purchased over a period of time. Only 4 of these apps allowed users to specify the type of meat bought or consumed (Carbon Footprinter, EcoChallenge, MathTappers: Carbon Choices, SustainableI). Fish and dairy consumption were also commonly considered. A total of 8 apps (8/19, 42%) included only a broad representation of dietary behavior with 1 or 2 questions addressing either dietary lifestyle (eg, vegan, vegetarian, meat eater), total food consumption or food transport (eg, local, imported). However, 4 apps captured more detailed information about the consumption of a wide range of food groups such as meat, fish, dairy, rice, bread/cereals/grains, and fruits and vegetables. Only 1 app captured coffee or alcohol consumption; 6 apps captured other information such as food packaging or farming method (eg, organic, nonorganic).

A total of 7 dietary apps (35%) also included information to enhance the user's knowledge of how dietary habits affect personal GHG emissions, however only 2 apps briefly stated the co-benefits for health and the environment by changing dietary and travel behaviors (Green Plaza and Oroeco).

Figure 1. Flow chart of the systematic process for determining apps for inclusion in the final sample.
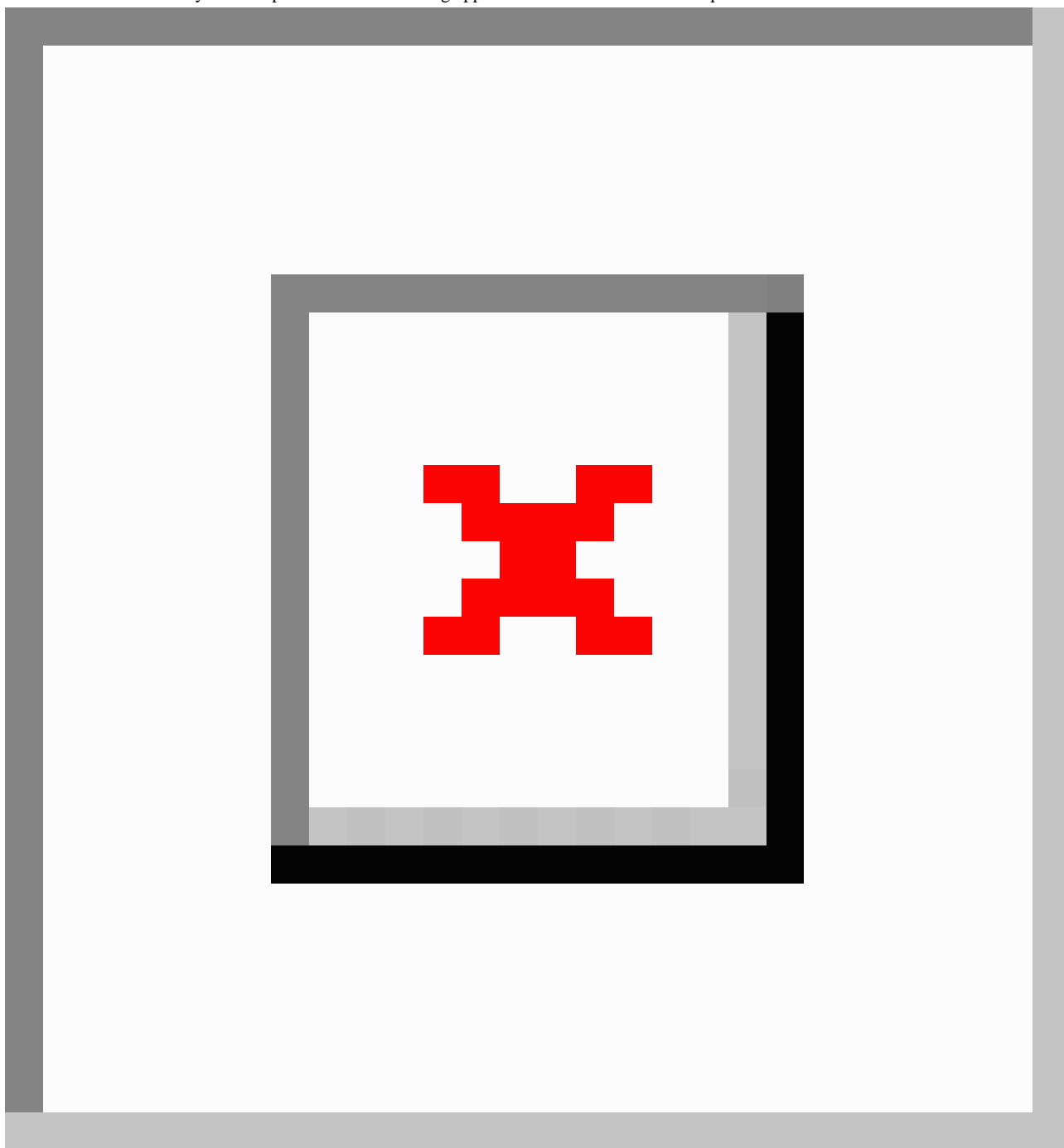
Table 3. App name, developer, type of app, method of data capture, cost, and availability of transport-related apps.

App name	Developer	Type	Data capture	Cost $NZ	Availability
Bikes vs Cars	Fredrik Gertten (WG Film)	Travel + CO_2[a]	Start/stop	Free	iPhone & Android
Bike Companion	Karlheinz Agsteiner	Travel + CO_2+ kcal burned[b]	Automatic	Free	Android only
Bike da firma	Bike da firma LTDA-ME	Travel + CO_2+ kcal burned	Start/stop	Free	iPhone & Android
Changers	Blacksquared GmbH	Travel + CO_2	Automatic	Free	iPhone & Android
Commute Greener	Pocketweb GmbH	Travel + CO_2+ kcal burned	Manual	Free	iPhone & Android
Cycling 365	SRM—Societa' Reti e Mobilita' Srl	Travel + CO_2	Start/stop	Free	iPhone & Android
Ecolife	Pipat Apiruktanakorn	Travel + CO_2+ kcal burned	Start/stop	Free	iPhone & Android
Eco Via	Gregory Carpentier	Travel + CO_2	Start/stop	Free	Android only
EMission	Kalyanaraman Shankari	Travel + CO_2+ kcal burned	Automatic	Free	iPhone & Android
Electrip	Volvo Bussar AB	Travel + CO_2	Start/stop	Free	iPhone & iPad
FuelGood	Energy Saving Trust	Travel + CO_2	Start/stop	Free	iPhone & Android
Greener Mile	Patrick Hardie	Travel + CO_2	Start/stop	Free	iPhone only
Green Steps	Alkemy Lab	Travel + CO_2	Start/stop	Free	iPhone & Android
Green Travel Choice	PocketWeb Ltd	Comparison of travel modes	Manual	$2.59	iPhone & iPad
modalyzer	Innovationszentrum fuer Mobilitaet und gesellschaftlichen Wandel (InnoZ) GmbH	Travel + CO_2	Automatic	Free	iPhone & Android
My Carbon	International College, KMITL	Travel + CO_2	Start/stop	Free	Android only
My Open Road	My Open Road Corp	Travel + CO_2+ kcal burned	Start/stop	Free	iPhone & Android
Singapore G1 Live Green	Balanced Consultancy	Travel + CO_2+ kcal burned	Start/stop	Free	Android only
TripGo	SkedGo Pty Ltd	Comparison of travel modes	Manual	Free	iPhone & iPad
Vapourz	Oliver Wilson	Travel + CO_2	Start/stop	Free	iPhone only

[a]Travel + CO_2 refers to apps that record personal travel and give an estimation of resulting carbon dioxide emissions or savings.

[b]Travel +CO_2 + kcal burned refers to apps that record personal travel, the associated carbon impact, and kcal burned during active travel.

XSL•FO
RenderX

Table 4. App name, developer, type of app, method of data capture, cost, and availability of diet-related apps.

App name	Developer	Type	Data capture	Cost $NZ	Availability
AGE Carbon Calculator	Palacegroup	Carbon Calculator	Multiple choice	Free	Android only
CarbonBuster	Dunman Secondary	Carbon Calculator/educational	Multiple choice	Free	iPhone Only
Carbon Footprinter	Luhui Yan	Carbon Calculator	Numerical input	Free	iPhone Only
CarbonSins	Team Maple Bangalore India	Carbon Calculator	Multiple choice	Free	Android Only
Count Carbon	Don Kershaw	Carbon Calculator	Multiple choice	Free	Android Only
CO_2 Emission Calculator	FerviDroid	Carbon Calculator	Numerical input	Free	Android only
CO_2 Footprint	Ship Shape	Carbon Calculator	Multiple choice	Free	iPhone & iPad
EcoChallenge	Raureif GmbH	Carbon Calculator/educational/behavior change	Multiple choice	Free	iPhone Only
eco footprint	Max Gontar	Carbon Calculator	Multiple choice	Free	Android Only
Eco Life Hacks	Anako Dev	Carbon Calculator/educational	Multiple choice	Free	Android Only
ecological footprint	Talents & Treasures, Lda	Carbon Calculator	Multiple choice	Free	Android Only
Food Miles Footprint	BW15 Apps	Calculates emissions due to transportation of food	Numerical/data input	Free	iPhone Only
Green Plaza	Webdunia.com	Carbon Calculator/educational + health	Multiple choice	$1.29	iPhone & iPad
GreenYou	ITAnyplace	Carbon Calculator	Numerical input	iOS–free; Android–$1.29	iPhone & Android
Lotus Greens Carbon Calculator	Lotus Greens Developers Pvt Ltd	Carbon Calculator	Multiple choice/scales	Free	Android Only
Math Tappers: Carbon Choices	HeavyLifters Network Ltd.	Carbon Calculator/educational	Numerical input	Free	iPhone Only
Oroeco	Oroeco Mobile	Carbon Calculator/educational/behavior change + health	Multiple choice/numerical input/scales	Free	iPhone & Android
Residential CO_2 Alert	Jacques Joosten (Apes apps)	Carbon Calculator	Multiple choice	Free	iPhone Only
SustainableI	DNV GL Business Assurance UK	Carbon Calculator/educational/behavior change	Numerical input	Free	iPhone & Android
VES CO_2 Tool HD	Veolia Environmental Services	Carbon Calculator/educational	Numerical input/multiple choice	Free	iPhone & iPad

Figure 2. Typical features or information included in travel- or dietary-based apps.

Content and features of included apps
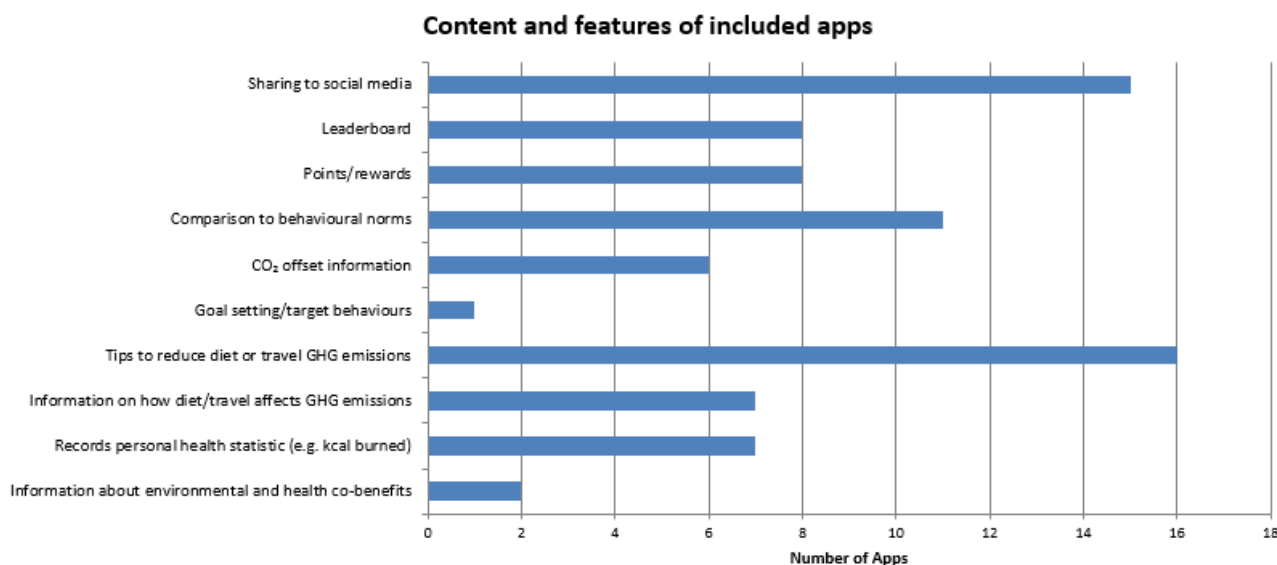
App Features and Behavioral Change Techniques

Several apps included interactive features to enhance user engagement (Figure 2, Multimedia Appendix 1). Sharing to social media such as Facebook was a feature included in 14 apps (8 transport apps and 6 dietary apps), and 1 dietary app (Carbon Footprinter) allowed sharing to Asian social media such as Sina Weibo. Gamificaton features were incorporated into 9 apps, including reward points and badges for sustainable actions or a leader board where users could compete against others.

Several apps also included features to encourage behavioral change. One app included goal setting, where users set targets for the amount of financial or carbon emissions savings they wanted to make or the calories they wished to burn. A total of 5 travel apps included comparison of different transport modes to help users make more sustainable transport choices, and 16 apps (12 dietary, 4 travel) included specific tips to help users reduce personal GHG emissions. However, the tips included in travel-based apps were typically more implicit and harder to find compared with dietary apps (eg, in the help section of the app). Finally, 11 apps (9 dietary, 2 travel) compared personal emissions with behavioral norms. The majority of these apps (8/11, 73%) used visual comparisons to help user's easily identify problematic behavior.

App Quality

In total, 40 apps were rated. One app was rated on both iOS and Android operating systems due to slight differences (My Open Road), and 2 apps were only rated on Android due to issues with functionality on iOS (Green Steps and Eco Life Hacks).

Most apps were rated as acceptable quality (score of 2.50-3.49 out of 5) as determined by the MARS overall mean score; however, 1 travel app was rated as poor quality (score of 1.50-2.49 out of 5), and 7 travel apps and 3 dietary apps were rated as good quality (score of 3.50-4.49 out of 5, see Multimedia Appendix 2). On average, travel apps received higher subjective satisfaction scores compared to dietary apps

(P=.014, 95% CI 0.13-1.09), and although travel apps tended towards higher overall mean scores (P=.08, 95% CI −0.04 to 0.54) and app specific scores (P=.15, 95% CI −0.14 to 0.85) compared to dietary apps, they were not significantly different. Travel apps tended to be more engaging (P=.03, 95% CI 0.06-0.87) and aesthetically pleasing (P=.05, 95% CI −0.003 to 0.88) compared to dietary-based apps, although functionality (P=.63, 95% CI −0.51 to 0.31) and information scores (P=.15, 95% CI −0.08 to 0.53) were similar between the 2 groups. Free apps were rated slightly higher than paid apps; however, the difference was not significant (P=.21, 95% CI −0.21 to 0.92).

Interrater reliability was excellent for the overall mean app quality scores (ICC=0.94, 95% CI 0.87-0.97), the app-specific scores (ICC=0.94, 95% CI 0.87-0.97), and the subjective satisfaction scores (ICC=0.96, 95% CI 0.91-0.98).

Discussion

Principal Findings

This study sought to review existing smartphone apps that collect data on individual transport and dietary-related behaviors and estimate personal carbon emissions with a view to determining suitable preexisting tools for assessing changes in behaviors that have health and environment co-benefits. This is the first review to describe features of these types of apps and rate their quality against a valid quality rating tool. Although we found multiple apps for capturing either travel or dietary behavior, we found no single app that fitted inclusion criteria for capturing both behaviors simultaneously, and only 2 apps mentioned the potential co-benefits to health and the environment by changing behaviors. Overall, apps were of mixed quality and few included all the features of being fully automated, including health information, and providing personalized feedback to participants with strategies to make changes to their behavior.

Motorized transport is one of the fastest rising sources of GHG emissions among energy-using sectors and is predicted to rise by 80% between 2007 and 2030 [12,14]. Not only does this

pose a grave risk to the environment, it also increases the disease burden associated with sedentary lifestyles [7,12-14]. Strategies to reduce travel-related GHG emissions, such as moving toward more active travel (walking or cycling instead of driving), have demonstrated a reduction in the risk of cardiovascular disease, type 2 diabetes, stroke, dementia, depression, and cancer as a result of improved physical activity levels [7,12,13,16,17]. Moreover, reductions in air pollution associated with alternative transport provide further health benefits by lowering the risk of respiratory disease, cardiovascular disease, and lung cancer [13]. While increased use of active travel may result in small increases in road traffic injuries, it has been repeatedly shown that the potential health benefits heavily outweigh the risks of injury, and risks can be reduced with the implementation of appropriate programs and policies [5,7,13,16].

Even small individual-level changes to travel mode can result in meaningful population-level health benefits [7,13,14,16]. A recent study has estimated that changing just 5% of vehicle travel in Adelaide, South Australia, to cycling results in annual carbon dioxide reductions of 191,313 tons per year, reduces particulate matter and air pollution by 8.5%, and could save 155 deaths and 1991 disability-adjusted life-years associated with chronic disease [13]. However, the largest benefit to both public health and the environment results from the combination of significant increases in active travel and reduced reliance on motorized travel [12,13,17]. Changing 40% of vehicle travel to alternative modes, such as public transportation and cycling, is estimated to potentially reduce the total disease burden attributed to physical inactivity by 55% [13]. Despite the abundance of evidence showing the potential health benefits of small changes to travel mode, none of the reviewed transport apps provided adequate information regarding the impact of personal travel behavior on health and the risk of NCDs (other than the number of calories burned during active travel).

Although the environmental effects of motorized transport may be widely acknowledged, individuals may be less aware of the indirect effects of individual dietary consumption on climate change. Agricultural activities and deforestation contribute more to global GHG emissions than transport, thereby exacerbating climate change and its subsequent effects on health, including food yields [6,19,21,22]. Despite the significant impact that individual dietary behavior may have on climate change, we found no apps that were solely focused on estimating personal emissions attributed to dietary behavior. Diet-related apps included in this review were mostly carbon calculators that addressed multiple behavioral categories. Although some apps attempted to measure the environmental impact of dietary behavior in more depth, few apps adequately captured all aspects contributing to personal diet-related emissions. Kim and Neff [34] reported similar findings with only 25% of all carbon calculators including a diet component, of which most addressed only 1 diet-related behavior. The limited scope of dietary behaviors captured by apps included in this review may be attributed to the trade-off between the accuracy of measurement and the burden of significant manual data input.

Livestock production alone accounts for 18% of global GHG emissions, and global meat and dairy consumption (and therefore livestock production) are predicted to double from 229 million

tons in 1999-2001 to 465 million tons in 2050 [21]. Therefore, meat consumption was appropriately one of the primary considerations of most diet-related apps in this review, with some also considering milk or dairy consumption. Ruminant animals (eg, cattle, sheep, and goats) account for the majority of livestock's GHG emissions due to significant methane emissions from enteric fermentation compared to monogastric animals (eg, pigs and poultry) [25]. However, only 4 of the included dietary apps allowed users to specify the type of meat consumed. Rice consumption was considered by 4 of the included apps, as its cultivation also contributes to methane emissions, which are more damaging to the environment than carbon emissions [19]. Additionally, since the majority of emissions "beyond the farm gate" come from the transportation of food [19], some apps addressed the food source, namely whether it was imported or grown locally.

Changes to dietary behavior such as reducing the intake of meat and dairy products are necessary to achieve meaningful reductions in food-related GHG emissions [6,19,21]. However, these changes also have potential health benefits, especially within high-income countries where dietary excess and the increase in availability of animal-based foods high in saturated fat contribute to the burden of NCDs [6,8,9,19,21]. Furthermore, if GHG emissions from livestock production could be curtailed, a reduction in red meat consumption in high-income countries could in the short term allow for slight increases in very low-income countries, addressing issues of mal- and undernutrition, thereby providing health benefits for all [6,21]. However, in the face of large projected increases in meat consumption within the developing world, not only will changes be necessary in high-income countries, but low and middle-income countries will have to moderate their intakes per capita [19]. Only 2 of the dietary apps included in this review made reference to the potential health benefits of dietary changes. The Oroeco app provided an explanation of why a reduction in the consumption of animal products is not only good for the environment but also reduces the risk of obesity and associated NCDs (heart disease, diabetes, and many cancers), while the Green Plaza app only mentioned that these changes are good for your health, your budget, and the planet.

Only 10 out of 40 apps (25%) included in this review explicitly encouraged behavior change by rewarding sustainable behavior or providing clear instruction to change behavior (eg, try walking, cycling, or using public transport instead of driving). The remaining apps were more implicit, focusing on creating awareness of the impact of current behavior with some providing information, tips or sustainable alternatives with potential to encourage behavior change. Of the reviewed apps, dietary apps typically included more tips to help users reduce emissions and were more likely to provide clarity regarding behavioral norms compared with travel apps. However, although a few apps provided many good quality tips or recommendations to mitigate personal emissions (Oroeco, VES CO2 Tool, Eco Life Hacks, SustainableI, Carbon Buster), most apps provided very few or no tips with no explanation or background information, potentially reducing their effectiveness in changing users' attitudes and ultimately behavior. Apps that rated highest typically included features to maximize user interaction and

engagement such as sharing to social media (Facebook) and elements of gamification (leader board, competition/comparison to other users, or rewards) in addition to producing reasonable estimates of personal emissions, providing good quality information and strategies to mitigate personal GHG emissions. Technology employing behavior change techniques such as self-monitoring, feedback, comparison with behavioral norms, peer influence/social networking, or gamification has been promising in evoking changes in attitudes and behavior in both travel and health contexts [35-38]. Future research is needed to determine which features or techniques within smartphone apps are required to trigger and sustain behavior change, resulting in climate change and health co-benefits.

Previous research has identified that carbon calculators tend to be meaningless unless represented in a more tangible form such as comparison to social norms/target behaviors or showing the direct environmental impact of individual actions [39]. Although the majority of apps included in this review provided emissions results in tons, kilograms, or carbon dioxide equivalents annually, which may be difficult for users to interpret, several apps used more effective communication methods. The Changers and VES CO2 Tool HD apps incorporated real-world representations of carbon dioxide emissions or savings (eg, "this is equivalent to one cycle of laundry at 60°C or watching TV for a year"), 8 apps showed clear comparison to behavioral norms (ie, clear visual representation), 2 apps represented emissions as the number of planets required to sustain behavior if all humans had the same impact, and 5 apps represented emissions in terms of the number of trees required to achieve carbon neutrality.

Strengths and Limitations

A strength of this review is that, to our knowledge, it is the first study of its kind to synthesize available smartphone apps for capturing personal travel and dietary behavior with an emphasis on GHG emissions and NCDs. It also used a validated quality rating tool and described the presence or absence of behavior change techniques required to initiate and maintain travel and dietary behavior. This study, however, was limited to apps for Android or iOS devices available at the time of the search and therefore does not include apps for Windows phones or other technologies. Furthermore, new apps are constantly being developed and existing apps improved, thus our review is relevant to the versions available at this time.

Future Recommendations

Future apps should incorporate information and features to promote the health benefits of strategies to reduce personal GHG emissions. This could potentially result in greater behavioral change by appealing to the user's self-interest in achieving optimal personal health in addition to enabling contributions to wider global challenges such as climate change [39]. Potential exists to include health monitoring data (eg, heart rate, arterial oxygen saturation) by linking with other devices, thereby guiding users towards improved health and wellbeing while simultaneously reducing personal GHG emissions. More information could be incorporated to ensure at-risk populations are still meeting dietary guidelines despite changes in diet. For example, recommendations to reduce meat consumption could

be paired with information about alternative protein and iron sources, possibly including recipes, especially for menstruating women, high intensity athletes, or anyone cooking for growing children. Apps should also simultaneously and automatically track all forms of physical activity and travel, giving a clear indication of health gain from active travel.

In addition, potential exists to include a broader scope of dietary and transport behaviors to more accurately and dynamically capture personal carbon emissions. Diet-based apps should ideally capture all aspects of food-related emissions from farm to waste, for all foods consumed (especially high-impact foods such as animal products and rice), and specify the type of meat consumed (eg, red meat from ruminants vs poultry or pork). Travel apps have potential for greater integration with multimodal transportation methods such as plane, rail, bus, car, motorcycle, and boat/ferry travel.

Studies suggest that participants value accuracy in tracked data [40] but view manual data entry as burdensome and may forget to record smaller travel trips if required to do so [27]. Therefore future apps should capture data continuously and ubiquitously, with minimal manual participant input. However, thought should be given to the trade-off between data accuracy, using continuous automatic GPS tracking, and the preservation of limited battery energy. App users are very aware of battery power consumption and may delete apps with high battery use, deeming them unnecessary [28,40]. Therefore, the inconvenience of frequent battery recharge may outweigh any benefit of using auto-geolocation. Furthermore, care should be taken to protect the user's privacy by granting control of the tracking feature such as enabling the user to set "sensitive areas" or "sensitive times" [28].

The ability for apps to deliver personalized advice or information "on the go" has been identified as a valued feature among app users [40]. However, personal carbon calculators may produce abstract scores, making them difficult to interpret and understand. Furthermore, the inconsistencies in carbon calculator scores and unclear methodology makes them difficult to compare, standardize, and benchmark [34,41]. Future apps should consider the use of more effective ways to communicate personal emissions data in a way that users will understand. Future work should also include improved visibility and transparency of data sources and development of new metrics and computation methodology that provide insightful and valuable information on people's personal carbon emissions and potential health co-benefits associated with dietary and physical activity-related behaviors.

Finally, future apps could include more techniques and interactive features (such as sharing to social media and app communities) to maximize participant engagement, promote sustained use, and ultimately evoke behavior change.

Conclusion

This review revealed multiple apps for capturing either dietary or travel behavior and estimating the associated carbon cost. However, we found no single app that adequately captured both behaviors simultaneously and addressed the potential co-benefits for the environment and health by changing these behaviors.

Overall, existing apps are of mixed quality, and none included all of the features of being fully automated, providing adequate health information and personalised feedback to participants with strategies to make changes to their behavior.

The ubiquity of smartphone use, their inbuilt sensors (accelerometers, GPS), and their computational power makes them an ideal device for collecting personal data on travel and dietary behaviors as well as personal carbon emissions and using these data to provide "just-in-time" interventions to evoke behavior change. Future research is needed to determine the potential of such approaches to change behavior.

Acknowledgments

Conflicts of Interest

None declared.

Multimedia Appendix 1

Content and features of all included apps.

[XLSX File (Microsoft Excel File), 20KB - mhealth_v4i4e135_app1.xlsx]

Multimedia Appendix 2

Mobile Application Rating Scale (MARS) quality scores of the 2 evaluators for all included apps.

[PDF File (Adobe PDF File), 75KB - mhealth_v4i4e135_app2.pdf]

References

1. Woodward A, Smith KR, Campbell-Lendrum D, Chadee DD, Honda Y, Liu Q, et al. Climate change and health: on the latest IPCC report. Lancet 2014 Apr 5;383(9924):1185-1189. [doi: 10.1016/S0140-6736(14)60576-6] [Medline: 24703554]

2. McMichael AJ. Globalization, climate change, and human health. N Engl J Med 2013 Apr 4;368(14):1335-1343. [doi: 10.1056/NEJMra1109341] [Medline: 23550671]

3. Mendis S, Armstrong T, Bettcher D, Branca F, Lauer J, Mace C, et al. Global status report on noncommunicable diseases. Switzerland: World Health Organization; 2015. URL: http://apps.who.int/iris/bitstream/10665/148114/1/9789241564854_eng.pdf?ua=1 [accessed 2016-04-14] [WebCite Cache ID 6gm3Kp3ZF]

4. Marmot M, Atinmo T, Byers T, Chen J, Hirohata T, Jackson A, et al. Food, nutrition, physical activity, and the prevention of cancer: a global perspective. Washington, DC: World Cancer Research Fund/American Institute for Cancer Research; 2007. URL: http://discovery.ucl.ac.uk/4841/1/4841.pdf [accessed 2016-04-15] [WebCite Cache ID 6gm51MT2L]

5. Haines A, McMichael AJ, Smith KR, Roberts I, Woodcock J, Markandya A, et al. Public health benefits of strategies to reduce greenhouse-gas emissions: overview and implications for policy makers. Lancet 2009 Dec 19;374(9707):2104-2114. [doi: 10.1016/S0140-6736(09)61759-1] [Medline: 19942281]

6. Haines A, Smith KR, Anderson D, Epstein PR, McMichael AJ, Roberts I, et al. Policies for accelerating access to clean energy, improving health, advancing development, and mitigating climate change. Lancet 2007 Oct 6;370(9594):1264-1281. [doi: 10.1016/S0140-6736(07)61257-4] [Medline: 17868819]

7. Lindsay G, Macmillan A, Woodward A. Moving urban trips from cars to bicycles: impact on health and emissions. Aust N Z J Public Health 2011 Feb;35(1):54-60. [doi: 10.1111/j.1753-6405.2010.00621.x] [Medline: 21299701]

8. Li D, Siriamornpun S, Wahlqvist ML, Mann NJ, Sinclair AJ. Lean meat and heart health. Asia Pac J Clin Nutr 2005;14(2):113-119 [FREE Full text] [Medline: 15927927]

9. Taylor EF, Burley VJ, Greenwood DC, Cade JE. Meat consumption and risk of breast cancer in the UK Women's Cohort Study. Br J Cancer 2007 Apr 10;96(7):1139-1146 [FREE Full text] [doi: 10.1038/sj.bjc.6603689] [Medline: 17406351]

10. Friel S, Dangour AD, Garnett T, Lock K, Chalabi Z, Roberts I, et al. Public health benefits of strategies to reduce greenhouse-gas emissions: food and agriculture. Lancet 2009 Dec 12;374(9706):2016-2025. [doi: 10.1016/S0140-6736(09)61753-0] [Medline: 19942280]

11.   Friel S, Bowen K, Campbell-Lendrum D, Frumkin H, McMichael AJ, Rasanathan K. Climate change, noncommunicable diseases, and development: the relationships and common policy opportunities. Annu Rev Public Health 2011;32:133-147. [doi: 10.1146/annurev-publhealth-071910-140612] [Medline: 21091194]

12.   Woodcock J, Edwards P, Tonne C, Armstrong BG, Ashiru O, Banister D, et al. Public health benefits of strategies to reduce greenhouse-gas emissions: urban land transport. Lancet 2009 Dec 5;374(9705):1930-1943. [doi: 10.1016/S0140-6736(09)61714-1] [Medline: 19942277]

13.   Xia T, Nitschke M, Zhang Y, Shah P, Crabb S, Hansen A. Traffic-related air pollution and health co-benefits of alternative transport in Adelaide, South Australia. Environ Int 2015 Jan;74:281-290. [doi: 10.1016/j.envint.2014.10.004] [Medline: 25454245]

14.   Shaw C, Hales S, Howden-Chapman P, Edwards R. Health co-benefits of climate change mitigation policies in the transport sector. Nature Climate change 2014 May 28;4(6):427-433. [doi: 10.1038/NCLIMATE2247]

15.   World Health Organization. Global health risks: mortality and burden of disease attributable to selected major risks. 2009. URL: http://apps.who.int/iris/bitstream/10665/44203/1/9789241563871_eng.pdf [accessed 2016-12-07] [WebCite Cache ID 6masPub10]

16.   Macmillan A, Connor J, Witten K, Kearns R, Rees D, Woodward A. The societal costs and benefits of commuter bicycling: simulating the effects of specific policies using system dynamics modeling. Environ Health Perspect 2014 Apr;122(4):335-344 [FREE Full text] [doi: 10.1289/ehp.1307250] [Medline: 24496244]

17.   Woodcock J, Givoni M, Morgan AS. Health impact modelling of active travel visions for England and Wales using an Integrated Transport and Health Impact Modelling Tool (ITHIM). PLoS One 2013;8(1):e51462 [FREE Full text] [doi: 10.1371/journal.pone.0051462] [Medline: 23326315]

18.   Maizlish N, Woodcock J, Co S, Ostro B, Fanai A, Fairley D. Health cobenefits and transportation-related reductions in greenhouse gas emissions in the San Francisco Bay area. Am J Public Health 2013 Apr;103(4):703-709. [doi: 10.2105/AJPH.2012.300939] [Medline: 23409903]

19.   Garnett T. Where are the best opportunities for reducing greenhouse gas emissions in the food system (including the food chain)? Food Policy 2011 Jan;36:S23-S32. [doi: 10.1016/j.foodpol.2010.10.010]

20.   Steinfeld H, Gerber P, Wassenaar T, Castel V, de Haan C. Livestock's Long Shadow: Environmental Issues and Options. Rome: Food & Agriculture Organisation of the United Nations; 2006.

21.   McMichael AJ, Powles JW, Butler CD, Uauy R. Food, livestock production, energy, climate change, and health. Lancet 2007 Oct 6;370(9594):1253-1263. [doi: 10.1016/S0140-6736(07)61256-2] [Medline: 17868818]

22.   Stehfest E, Bouwman L, van Vuuren DP, den Elzen MGJ, Eickhout B, Kabat P. Climate benefits of changing diet. Climatic Change 2009 Feb 4;95(1-2):83-102. [doi: 10.1007/s10584-008-9534-6]

23.   Sovacool BK. Diversity: energy studies need social science. Nature 2014 Jul 31;511(7511):529-530. [doi: 10.1038/511529a] [Medline: 25079540]

24.   Bounajm F, Dinh T, Thériault L. Alberta Centre for Active Living. 2015 Feb. Moving ahead: The economic impact of reducing physical inactivity and sedentary behaviour URL: https://www.centre4activeliving.ca/media/filer_public/5b/2f/5b2f73fb-cc21-46de-836d-dd49b5357ea0/2015-feb-sedentary.pdf [accessed 2016-12-12] [WebCite Cache ID 6mgfIXyHU]

25.   Opio C, Gerber P, Mottet A, Falcucci A, Tempio G, MacLeod M, et al. Greenhouse gas emissions from ruminant supply chains: a global life cycle assessment. Rome: Food and Agriculture Organization of the United Nations; 2013. URL: http://www.fao.org/docrep/018/i3461e/i3461e.pdf [accessed 2016-12-08] [WebCite Cache ID 6matl7eBb]

26.   Lorenzoni I, Nicholson-Cole S, Whitmarsh L. Barriers perceived to engaging with climate change among the UK public and their policy implications. Global Environmental Change 2007 Aug;17(3-4):445-459. [doi: 10.1016/j.gloenvcha.2007.01.004]

27.   Wolf J, Oliveira M, Thompson M. Impact of underreporting on mileage and travel time estimates: results from global positioning system-enhanced household travel survey. J Transport Res Board 2003 Jan;1854:189-198. [doi: 10.3141/1854-21]

28.   Liu H, Chen C, Fan Y. Apps and battery efficient technologies for smartphone-based travel data collection: state of the art. 2016 Presented at: Transportation Research Board 95th Annual Meeting; 2016; Washington.

29.   Sharp DB, Allman-Farinelli M. Feasibility and validity of mobile phones to assess dietary intake. Nutrition 2014;30(11-12):1257-1266. [doi: 10.1016/j.nut.2014.02.020] [Medline: 24976425]

30.   Chen J, Cade JE, Allman-Farinelli M. The most popular smartphone apps for weight loss: a quality assessment. JMIR Mhealth Uhealth 2015;3(4):e104 [FREE Full text] [doi: 10.2196/mhealth.4334] [Medline: 26678569]

31.   Carter MC, Albar SA, Morris MA, Mulla UZ, Hancock N, Evans CE, et al. Development of a UK online 24-hour dietary assessment tool: myfood24. Nutrients 2015 Jun;7(6):4016-4032 [FREE Full text] [doi: 10.3390/nu7064016] [Medline: 26024292]

32.   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]

33.   Stoyan Stoyanov. Youtube. [video file]. 2016 Jun 14. MARS training video URL: https://www.youtube.com/watch?v=25vBwJQIOcE [accessed 2016-12-13] [WebCite Cache ID 6mi8tkcEB]

34.   Kim B, Neff R. Measurement and communication of greenhouse gas emissions from US food consumption via carbon calculators. Ecological Economics 2009;69(1):186-196.

35.   Froehlich J, Dillahunt T, Klasnja P, Mankoff J, Consolvo S, Harrison B, et al. UbiGreen: investigating a mobile tool for tracking and supporting green transportation habits. 2009 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; 2009 Apr 4; Boston p. 1043-1052. [doi: 10.1145/1518701.1518861]

36.   Jariyasunant J, Abou-Zeid M, Carrel A, Ekambaram V, Gaker D, Sengupta R, et al. Quantified traveler: travel feedback meets the cloud to change behavior. J Intelligent Transport Syst 2014 Sep 25;19(2):109-124. [doi: 10.1080/15472450.2013.856714]

37.   Mankoff J, Matthews D, Fussell S, Johnson M. Leveraging social networks to motivate individuals to reduce their ecological footprints. 2007 Presented at: 47th Hawaii International Conference on System Sciences; 2007; Hawaii p. 87. [doi: 10.1109/HICSS.2007.325]

38.   Taniguchi A, Suzuki H, Fujii S. Mobility management in Japan: its development and meta-analysis of travel feedback programs. J Transport Res Board 2007 Dec;2021:100-109. [doi: 10.3141/2021-12]

39.   Chatterton TJ, Coulter A, Musselwhite C, Lyons G, Clegg S. Understanding how transport choices are affected by the environment and health: views expressed in a study on the use of carbon calculators. Public Health 2009 Jan;123(1):e45-e49. [doi: 10.1016/j.puhe.2008.10.022] [Medline: 19101000]

40.   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]

41.   Padgett JP, Steinemann AC, Clarke JH, Vandenbergh MP. A comparison of carbon calculators. Environmental Impact Assessment Review 2008 Feb;28(2-3):106-115. [doi: 10.1016/j.eiar.2007.08.001]

Abbreviations

CO2: carbon dioxide
GHG emissions: greenhouse gas emissions
GPS: Global Positioning System
ICC: intraclass correlation coefficient
MARS: Mobile Application Rating Scale
NCD: noncommunicable disease

<u>Original Paper</u>

Engaging Gatekeeper-Stakeholders in Development of a Mobile Health Intervention to Improve Medication Adherence Among African American and Pacific Islander Elderly Patients With Hypertension

Hamed Yazdanshenas[1], MD; Mohsen Bazargan[2], PhD; Loretta Jones[3], MA, ThD; May Vawer[4], RN; Todd B Seto[4], MD, MPH; Summer Farooq[5], MS; Deborah A Taira[6], ScD

[1]College of Medicine, Departments of Family Medicine and Orthopedic Surgery, Charles R Drew University of Medicine and Science/University of California, Los Angeles (UCLA), Los Angeles, CA, United States

[2]College of Medicine, Department of Family Medicine, Charles R Drew University of Medicine and Science/University of California, Los Angeles (UCLA), Los Angeles, CA, United States

[3]Department of Community Engagement, Charles R Drew University of Medicine and Science, Los Angeles, CA, United States

[4]The Queen's Medical Center and the John A Burns School of Medicine, University of Hawaii, Honolulu, HI, United States

[5]College of Medicine, Department of Family Medicine, Charles R Drew University of Medicine and Science, Los Angeles, CA, United States

[6]Daniel K Inouye College of Pharmacy, University of Hawai'i at Hilo, Honolulu, HI, United States

Corresponding Author:
Hamed Yazdanshenas, MD
College of Medicine
Departments of Family Medicine and Orthopedic Surgery
Charles R Drew University of Medicine and Science/University of California, Los Angeles (UCLA)
1748 East, LSRN Bldg, # N150, 118th St
Los Angeles, CA, 90059
United States
Phone: 1 3233573452
Email: Yazdanshenas@ucla.edu

Abstract

Background: Approximately 70 million people in the United States have hypertension. Although antihypertensive therapy can reduce the morbidity and mortality associated with hypertension, often patients do not take their medication as prescribed.

Objective: The goal of this study was to better understand issues affecting the acceptability and usability of mobile health technology (mHealth) to improve medication adherence for elderly African American and Native Hawaiian and Pacific Islander patients with hypertension.

Methods: In-depth interviews were conducted with 20 gatekeeper-stakeholders using targeted open-ended questions. Interviews were deidentified, transcribed, organized, and coded manually by two independent coders. Analysis of patient interviews used largely a deductive approach because the targeted open-ended interview questions were designed to explore issues specific to the design and acceptability of a mHealth intervention for seniors.

Results: A number of similar themes regarding elements of a successful intervention emerged from our two groups of African American and Native Hawaiian and Pacific Islander gatekeeper-stakeholders. First was the need to teach participants both about the importance of adherence to antihypertensive medications. Second, was the use of mobile phones for messaging and patients need to be able to access ongoing technical support. Third, messaging needs to be short and simple, but personalized, and to come from someone the participant trusts and with whom they have a connection. There were some differences between groups. For instance, there was a strong sentiment among the African American group that the church be involved and that the intervention begin with group workshops, whereas the Native Hawaiian and Pacific Islander group seemed to believe that the teaching could occur on a one-to-one basis with the health care provider.

Conclusions: Information from our gatekeeper-stakeholder (key informant) interviews suggests that the design of a mHealth intervention to improve adherence to antihypertensives among the elderly could be very similar for African Americans and Native Hawaiian and Pacific Islanders. The main difference might be in the way in which the program is initiated (possibly through

XSL•FO

RenderX

church-based workshops for African Americans and by individual providers for Native Hawaiian and Pacific Islanders). Another difference might be who sends the messages with African Americans wanting someone outside the health care system, but Native Hawaiian and Pacific Islanders preferring a provider.

Introduction

Hypertension is the most common condition seen in primary care and may lead to myocardial infarction, stroke, renal failure, and death if not treated appropriately. In 2011, the cost burden in the United States associated with hypertension was estimated at US $46 billion in health care services, medications, and missed days of work [1]. Although advances in pharmacotherapy have decreased morbidity and improved life expectancy, poor medication adherence has offset some of these gains.

Racial and ethnic minority populations bear a disproportionate burden of hypertension and its sequelae. The prevalence of high blood pressure in African Americans is the highest in the world with approximately 32% of African Americans having hypertension. Estimates from the National Health Interview Survey in 2007 found that Native Hawaiians and Pacific Islanders had the second highest rates of any ethnic group in the United States at 29% [2-5]. Despite evidence that appropriate pharmacological treatment reduces morbidity and mortality across all populations [6-8], racial and ethnic minorities, including African Americans and Native Hawaiian and Pacific Islanders, often underutilize antihypertensive medications [9-11].

Even among people who see their providers regularly, 50% are nonadherent to prescribed medications after 6 months [12]. This nonadherence results in thousands of premature deaths, hospitalizations, and increased health care costs [13,14]. Of all medication-related hospital admissions, 33% to 69% are due to poor medication adherence, with approximate health care costs of US $100 billion annually [15,16]. Greater adherence to medications for chronic conditions including hypertension has been associated with higher medication costs, but net overall reduction in total health care costs [17].

Efforts to improve health-related behaviors and provider-patient communication play an increasingly important role in the management of patients with hypertension. Mobile health technology (mHealth) has been shown to be a feasible, well-accepted, and cost-effective means of achieving sustained gains in health-related behaviors, including physical activity, medication adherence, and blood pressure control [18-26]. The use of mHealth offers a platform for innovative approaches to health enhancement and disease management at low cost with widespread applicability, while promoting self-efficacy and autonomous regulation. Importantly, mHealth may be particularly relevant for populations with limited access to health care, including those who live in rural communities, and among racial and ethnic minorities. Examples of mHealth include text-based reminders, electronic delivery of educational materials, and use of mobile devices to monitor physical activity, medication adherence, and blood pressure.

The goal of this study was to better understand issues affecting the acceptability and usability of mobile health technology to improve medication adherence for elderly African American and Native Hawaiian and Pacific Islander patients with hypertension.

Methods

We conducted semistructured (30 minute) interviews of gatekeeper-stakeholders (N=20) who were health care providers or community leaders and were familiar with issues surrounding medication adherence among elderly African American (n=10) or Native Hawaiian and Pacific Islander (n=10). Potential participants asked to participate in this study were connected to the geriatric population. In addition, attempt was made to invite potential participants who were older. All gatekeeper-stakeholders who were approached agreed to participate. Established qualitative approaches were used to analyze and to interpret data [27,28]. This study was reviewed and approved by the Committee on Human Subjects for the two study sites at The Queen's Medical Center in Honolulu, HI, and at Charles Drew University in Los Angeles, CA.

Interviews

Because of the paucity of information about the acceptability of mHealth for African American and Native Hawaiian and Pacific Islander seniors, we chose a qualitative approach to identify themes that could be further explored in future quantitative studies. Two members of the research team conducted in-depth interviews with targeted open-ended questions and follow-up probes to elicit further clarification (Textbox 1). The interview covered the following key topics: (1) major concerns that each ethnic elderly population may encounter related to hypertension medication adherence; (2) ethnic differences, perceptions, and barriers to using mHealth interventions for medication self-management in older adults with hypertension; (3) desired daily frequency of mHealth interventions as well as the assessment of content to ensure cultural and linguistic appropriateness; (4) most effective mode of delivery of mHealth interventions to increase medication adherence; and (5) ways to increase older adults' interest and awareness about using mHealth interventions to improve compliance with their antihypertensive medications. Gatekeeper-stakeholders received a US $30 gift card for their participation to cover their time.

XSL•FO
RenderX

Textbox 1. Main questions from the semistructured key informant interviews.

Do you think most seniors in your community have a cell phone? How about a smartphone?

How effective do you think it would be to use a cell phone to remind seniors in your community about taking their blood pressure medication on time?

What would be the best way to phrase the text reminder/question?

From what source do you think seniors in your community would like to receive the reminder to take their blood pressure medication (eg, physician office, nurse, other family members...)?

Whom else do you think should receive the reminders that can help the senior population in your community to take their blood pressure on time?

Are there any cultural considerations?

Analysis

Interviews were deidentified and transcribed by members of the research team using a combination of direct quotes, paraphrasing, and summarization. They were then organized and coded manually by two independent coders. Analysis of patient interviews used largely a deductive approach because the targeted open-ended interview questions were designed to explore issues specific to the design and acceptability of a mHealth intervention for seniors. We also used inductive methods to identify additional themes that emerged during the interviews. Two primary members of the research team performed the first- and second-level coding. Input from the other members of the team was obtained to confer about the codes, quotes, and interpretation of quotes.

Results

Characteristics of Gatekeeper-Stakeholders

A convenience sample of community-based gatekeeper-stakeholders familiar with either the African American (n=10) or Native Hawaiian and Pacific Islander (n=10) elderly community were recruited for the interviews. Overall, the mean age of the gatekeeper-stakeholders (N=20) was 59.4 (SD 13.3) years. All but four participants were 50 years and older. All participants were connected to the geriatric population. The sample included founder or chief executive officer of a nongovernmental organization, community advocates, church leaders, and retired or active health care providers, including physicians, nurses, pharmacist, and medical assistant. Overall, 70% (14/20) of participants were female. All participants had a mobile phone, although one participant had a mobile phone with only a phone feature. All gatekeeper-stakeholders with mobile phones used their phones to access the Internet, but only four said they downloaded apps with their mobile phone.

Elderly Adults Use of Mobile Phones

When asked if they believe most elderly adults have mobile phones (including smartphones and cell phones), most gatekeeper-stakeholders said that having and using mobile phones by older adults was probably related to age. A common sentiment was that patients older than age 70 or 75 years might not have mobile phones or, if they do, they only use them for emergencies or to talk to relatives:

> *They just use cellphone for emergency purposes (making and receiving calls), but for other features for using cellphone (like sending text and taking picture), they don't have interest. The majority of individuals are retirees, age makes a difference; 50-year-olds are more willing to learn all that phones have to offer, once you get to 70, that changes.* [P12, African American]

A few also mentioned education level of the patient as an indication of whether they would have a mobile phone:

> *Some of the elderly who are more professional and just retired, they are very savvy with social media and stuff so they are using it in a comprehensive manner. Those who are more progressive and professional, they are texting, Facebooking, Instagramming. Those who maybe don't have a high school diploma (regular people in the community), worked in blue-collar jobs all their lives, and are just about to retire, they just use it for calling friends and relatives.* [P15, African American]

The general sentiment among gatekeeper-stakeholders was that lack of a mobile phone would not be a barrier, but that older participants would need a lot of training and support.

Elements Necessary for Success

There were a number of themes that emerged from both groups and some that arose only from either the African American or the Native Hawaiian and Pacific Islander gatekeeper-stakeholders (Textbox 2). The most frequently mentioned themes for both groups were (1) need to personalize the content; (2) teach about the importance of medication adherence and support people in using their mobile phone; (3) show caring, connection, and trust; (4) keep it simple; and (5) include music. Other suggestions by both groups were to include a picture of the medication pill, to enable people to help one another, and to ask nicely (eg, say "please").

Textbox 2. Themes for an effective mHealth intervention by group.

Both groups

- Personalized

- Teach, remind, follow up, be available to answer questions

- Caring/love/connection/trust

- Short/simple/easy to understand

- Include music

- Picture of pill

- Enable people to help one another, be contagious

- Ask nicely/say please/include positive affirmations

African Americans

- Workshop to teach about mobile phone

- Involve church

- Subsidize/incentivize

- Assistance from family member; voice/photo of family member

- Include quotes from scripture, authors, the President; tips on hobbies; slang/jokes

Native Hawaiians/Pacific Islanders

- Cannot be just out of the blue/need plan

- Patient choice/control/involvement/partnership

- Consider where they are at, readiness

- Start by reducing the number of medications if possible

Elements Wanted by Both Groups

Personalize the Content

Virtually every gatekeeper-stakeholder spoke of the need to personalize the intervention to the individual's needs:

> *Whatever can be personalized to the patient the better it will be and at the same time can't be in isolation, needs to be part of the bigger plan with the doctor-team...* [P9, Native Hawaiian/Pacific Islander]

Teach and Support

Another very common theme was the need to teach, not only about the importance of taking medications as prescribed, but on the use of mobile phones in general. There was also a strong sentiment expressed by many that there would need to be ongoing support. Specific quotes from gatekeeper-stakeholders included:

> *...Need to talk story about what and why medications being ordered, who to call with questions, follow up to see how it's going, how they are doing.* [P1, Native Hawaiian/Pacific Islander]

> *You can't give it to them and say "BYE," then you'll never hear from them for life. Educate them, follow up, and then follow up after the follow-up.* [P14, African American]

Importance of Caring, Connection, and Trust

Another common theme expressed by gatekeeper-stakeholders in both groups was the need to have the messaging come from someone with whom they have a connection and trust:

> *A person for them to connect with—so they know them—have worked with, learned what to do—so when they get their texts (to take medications) they know who it's from [you] it connects them and it's not just this machine that's texting them.* [P5, Native Hawaiian/Pacific Islander]

> *Love is important. People talk about it with their social groups. Use words like "I care about you" or "I want you to be around" or "we love you." You can go to Hallmark and see the greeting cards they have. Use a more common, less formal vernacular, like misspelling words. "Wazzup?" Personalize it. Use their names, first name and a picture if possible. Hallmark greeting card language is a suggestion. Maybe using some slang or jokes in the text reminders.* [P19, African American]

Keep it Simple

Regarding the nature of the messaging, the gatekeeper-stakeholders consistently emphasized the need to keep it short and sweet, simple, and direct:

Yeah and not annoying like telemarketing—wasting our time...pause...short and sweet shows a flash of caring. [P4, Native Hawaiian/Pacific Islander]

Shorter. People tend to want to read less. People have visual or literacy problems as they get older. Maybe once in a while, use a longer one to break things up (every fifth or tenth text). [P19, African American]

Include Music

Without prompting, many gatekeeper-stakeholders expressed the idea of including music as part of the messaging:

Yeah—if the doctor talks to you about "the reminder to take pills" coming and it's gonna be three beeps...or...one song...something they can look forward to from the doctor...I'd want one song as my "beep": Bob Marley, Three Little Birds, something can smile about, connect to. [P7, Native Hawaiian/Pacific Islander]

Use a sound but it is going to be gospel music, a very popular gospel song. Pretty soon, it is going to get stuck in their head, you want to use a very popular gospel song. [P15, African American]

Elements Wanted by Native Hawaiian and Pacific Islanders

Several themes were unique to the Native Hawaiian and Pacific Islander community. These included (1) need for a plan, (2) patient partnership and engagement, (3) reduction of medications if possible, and (4) consider level of readiness.

Need a Plan

One common sentiment among the Native Hawaiian and Pacific Islander group was the need to have a plan that the patient helps to develop:

Needs to be part of a bigger coordinated plan of care it can become useless...quickly...if no feedback step incorporated with the text... [P9, Native Hawaiian/Pacific Islander]

Patient Partnership

Gatekeeper-stakeholders also stated that patients need to be active partners in decisions and interventions need to adapt to their needs:

It's gonna be different for everybody—not one size—one set fits all—need them (Patient) feedback and commitment to whatever is decided with a process in place to make changes if needed. Depends—midmorning always a good start—but needs to be personalized and agreed upon with the patient. [P3, Native Hawaiian/Pacific Islander]

Reduction of Medications at Start (if Possible)

Several Native Hawaiian and Pacific Islander gatekeeper-stakeholders brought up the idea that patients want to be on fewer medications, if possible:

Patients don't want more pills—they want less, they want to feel better, they want to try anything before

taking more blood pressure pills. [P8, Native Hawaiian/Pacific Islander]

Elements Wanted by African Americans

Some ideas for a successful intervention that were only mentioned by the African American gatekeeper-stakeholders included (1) having workshops to train, (2) involving the church, (3) incentivizing participation, (4) involvement of family, and (5) quotations from scripture or famous authors or information on hobbies.

Workshop

Most African American gatekeeper-stakeholders said having group workshops at the church or community center would be a good way to start the intervention:

The majority of individuals are retirees and they are just sitting at the center to pass the day. They will come to workshops because it will be something new to them, they will participate. They will gain the knowledge of how to go through the cell phone, what to do, how to do. That would be at the top of the agenda, an introductory course. [P12, African American]

They come in groups, they are not going to come individually. We need to tell them "let me show you how to do this" and then they will do it. [P13, African American]

Workshops (including church workshops) can help improve access and increase knowledge of mobile technology among elderly populations. Have a representative (poster person) describe the benefits of research performed on them. [P12, African American]

Involvement of Church

African American gatekeeper-stakeholders emphasized the need to involve the church in health improvement efforts:

In the black community, you need to somehow link your workshop to the churches and talk about the importance of using mobile for this purpose. Also, you need to identify those gatekeepers who had a good relation with seniors and also pastors and make them involved with your workshops in churches. You might want to have a primary evaluation about each senior's knowledge and education background about using cellphone and then have your workshops on different levels. [P15, African American]

Subsidize/Incentivize

Many African American gatekeeper-stakeholders also mentioned the benefits of financial incentives to participate:

Regarding the teacher of this workshop, I would say that this age population group love church. The senior citizen center and some schools are also could be wonderful to have the workshops. One of the incentives to come is providing food at the workshops. Since most group of this population has fixed income, they would appreciate change in their hands (like $30

is excellent). Also, this age group is the Target and Walmarters. Giving them a gift card [to] Target would be excellent. [P16, African American]

Involvement of Family

Many of the African American gatekeeper-stakeholders recommended involving the family in delivering the messaging and incorporating their voices and photos:

You could ask them when you have an appointment to bring one of their children or grandchildren in so that they could have a partner, who they'd love to be in touch with, someone that they love. Preferably that person would be comfortable with technology so that they can help with keeping said person on his or her medication schedule. [P17, African American]

You also could send the picture of the medication or even a picture of his/her son/daughter...as a reminder to them. [P12, African American]

Quotes from Scripture/Authors and Information on Hobbies

In addition, many African American gatekeeper-stakeholders suggested incorporating quotations from scripture or famous authors into the messaging:

In their first interview, ask them about their favorite scripture (if they have one). Ask them about their favorite writer and then you can put some sentence from that writer. For instance my favorite activist is Marian Wright Edelman and her favorite motto is "Service is the rent we pay for being. It is the very purpose of life, and not something you do in your spare time." I would love seeing that every morning. If you want, ask them about their hobby and interest for phrasing your text, you have to make sure they are comfortable with it and know why you are asking this question and how it connects to your study. You need to ask them in a very nice way, such as "We would like you to wake up or to go to rest with a thought that you find interesting or pleasant. What would be the most interesting in having us say to you

or send to you?" and then have the list of those kind of things: their hobby, their favorite writer, and their favorite scripture. [P16, African American]

Be careful using faith-based statements straight from the Bible. It may feel like an intrusion into their faith. Possibly use statements from nonreligious influential people: Maya Angelou, Martin Luther King, Jr...Start with a phrase of gratitude like "God woke me this morning," but I do not think it is necessary to get into faith. Have messages come from patients, people that are good at adherence, like why they take their medications every day. There is peer credibility. [P17, African American]

Barriers to Success

In discussing whether a mHealth intervention would be successful, gatekeeper-stakeholders mentioned several barriers to success (Textbox 3). The main themes in this area were (1) there are other reasons for nonadherence (side effects, distrust, and depression), (2) they do not always have their mobile phone with them or do not have it on unless talking to family or in an emergency, and (3) they do not like text messages or need eye-to-eye contact. Several gatekeeper-stakeholders mentioned the fact that there are many reasons patients do not take their medications. This intervention will primarily help those who do not take their medication due to forgetfulness, but does not address other causes of nonadherence.

In addition, the Native Hawaiian and Pacific Islander gatekeeper-stakeholders mentioned some other barriers, including that we need to be specific about medications because they may not be sure which pill we are talking about if we just say "blood pressure medication" because pills are constantly changing. Another issue that was brought up was that the mobile phone keys are small so it may be difficult for participants to text responses. Finally, someone mentioned that it is likely that even if patients become more adherent initially, they may lose interest in the intervention and stop participating. This raises a question about the sustainability of the impact of the intervention.

Textbox 3. Barriers to a successful mHealth intervention by group.

Both groups

- Other reasons for nonadherence (side effects, distrust, depression)

- Do not carry mobile phone or have it off unless talking to family or in an emergency only

- Do not like text messages; need eye-to-eye contact

- Keys are small; text hard to read

African Americans

None stated from this group

Native Hawaiians/Pacific Islanders

- Not sure which pill is which; pills constantly changing

- Do not want to sound like telemarketing / be annoying

- People may lose interest and stop participating

XSL•FO

RenderX

Here are some of the statements that the gatekeeper-stakeholders made regarding barriers:

> *Depression could be a barrier to people using medications. People have individualized reasons why they do not adhere to medication.* [P17, African American]

> *...but gotta remember people don't take their medications for a lot of reasons—it's not just that they forget—so a daily text for those with other reasons might not work. They don't believe it helps, don't trust the doctor, don't want to question the doctor, and just don't take the pill or don't even fill the prescription...social barriers in terms of homelessness, money, financial.* [P6, Native Hawaiian/Pacific Islander]

> *I need eye-to-eye contact, face-to-face contact. I am not text savvy. If you don't have time to call them, you can send a prerecorded message to them with a voice of their close friends or relatives (instead of having a real person call).* [P15, African American]

An African American gatekeeper-stakeholder mentioned the elderly may be scared of new technology:

> *Overall, I think seniors are scared of technology, but if people that they trust persuade them, then they will make the transition.* [P17, African American]

In general, even though gatekeeper-stakeholders discussed barriers, they seemed to think that these barriers could be overcome, particularly if the connection, trust, and caring were present.

Phrasing

When asked what specific phrasing we should use, gatekeeper-stakeholders suggested:

> *"It's time again...Mrs/Mr..., it's time for your morning pill. Just a friendly reminder. It's time to take your pill. It's time to take your BP medicine. Even if you feel ok, remember to take your BP medicine. Taking your BP medicine as recommended is important for your health. It's time to take yours. It's time for your BP medicine. How about a short walk afterwards if you are able and it's ok with your doctor. It's that time again. Remember to take your BP medicine. Keep your BP medicine in a place that helps you remember to take it. It's time now."* [P11, African American]

> *Good morning. This is just a reminder that it's time for you to take your blood pressure medicine."* We have to be community-friendly so we have to use their terms. We cannot use "hypertension" we need to say "high blood pressure." Giving them a directive is offensive. Make it like you are talking to them.* [P13, African American]

In the morning:

> *"Good morning, Chauncy, sleepy head. It's time to take your blood pressure medication."* Tell them the name of the medication because they could take the wrong thing at the wrong time. Tell them any important instruction for taking the medication...Tell them you do have to text back this time—send "yes" or "oops" if you took it or not.* [P14, African American]

Cultural Considerations

Cultural considerations varied by group. Among the African American group, the main consideration was lack of trust in the medical system. As a result, many suggested that the messaging should come from someone outside the medical/research community, such as a family member, friends, pastor, or someone else they trusted (Textbox 4), whereas the Native Hawaiian and Pacific Islander group mainly suggested that providers send the messages.

Textbox 4. Who should send the reminders and who should receive them?

Who Should Send the Reminder?

Both groups

- Doctor/provider

- Pharmacist

- Patient choice

African Americans

- Family member/friend

- Study personnel

- Pastor

- Community liaison

- Peer

Native Hawaiians/Pacific Islanders

- Whoever is most engaged and taught them program

- Knows patient and medical history

Who Else, Besides the Patient, Should Receive the Reminder?

Both groups

- Family/spouse/companion

- Caregiver / senior housing

- Patient choice

- Pharmacist

African Americans

- Peer with hypertension

- Doctor

Native Hawaiians/Pacific Islanders

None stated by this group

Some older people may think it is rude to use the first name, like it is disrespectful and want to be called by Mrs/Mr and their last name (because they are coming out of serious racism where they called boy and things like that). They might say, "You don't know me, don't call me by my name." But, the younger people, they don't care. "Good morning Mrs/Mr so and so." I would go with that, since it's safe. [P17, African American]

If a voice doesn't sound African American...I don't trust it. If it sounds European, I'm like those people are talking again...It has to be culturally appropriate. (P18, African American)

People of color in general still have a mistrust of the medical system. Most people score moderate to extreme distrust of the medical system and they are around 40; that mistrust might increase as they get older. Something culturally important must increase trust in people who are doing this particular project. Remove the mistrusted person and replace with the family member or church leader. Community or family

liaison is critical to patients trusting in the project. [P19, African American]

For the Native Hawaiian and Pacific Islander group, trust of the system did not appear to be as big an issue. Most, however, expressed the need to deliver the intervention with warmth, caring, and aloha from someone with whom they have a connection:

I think warmth—aloha—so they know that we care for them about how they're doing and not just grabbing them to do something—it has to matter and what they do has to matter—a text can't be cultural—it has to...come from the sender that it matters that they matter. [P2, Native Hawaiian/Pacific Islander]

Yeah important to always show aloha, we do at the clinic always and that can work in the app too—personalized to what's important to them we can personalize it together. [P5, Native Hawaiian/Pacific Islander]

Although both the African American and Native Hawaiian and Pacific Islander groups expressed trust as an issue, it seemed to be a more important consideration in the African American community. The Native Hawaiian and Pacific Islander gatekeeper-stakeholders emphasized the need for aloha and warmth.

Discussion

Overall, the response from gatekeeper-stakeholders was very positive toward the potential for a mHealth intervention to improve adherence to antihypertensives among elderly adults. This acceptability of mHealth is consistent with a prior study of a culturally diverse low-income population at two primary care clinics in California [29]. In this study, 86% of respondents said they were interested in using mHealth to improve their health. Our study differs in that we focus on medication adherence and acceptance of mHealth in elderly African Americans and Native Hawaiian and Pacific Islanders.

A number of similar themes regarding what would be needed to make this intervention successful emerged from our two groups of African American and Native Hawaiian and Pacific Islander gatekeeper-stakeholders). First was the need to teach participants both about the importance of adherence to antihypertensive medications and the use of mobile phones for messaging. Second, gatekeeper-stakeholders in both communities expressed a strong need for participant access to ongoing "technical support" and "checking in" to make sure everything was going okay. Therefore, it is important that interventional studies that target older adults and intend to use mobile phones provide culturally appropriate trainings for these populations. We were repeatedly reminded by our participants that older adults are very interested in and excited about learning how to effectively operate their mobile phone. They also advised that such trainings should be provided in a friendly environment, adjusted to their level of understating and needs, and be delivered by a peer educator. In addition, they noted that trainings should continue during the course of interventions and repeated frequently to enhance their usage.

Third, almost all gatekeeper-stakeholders mentioned the need to personalize the intervention to the needs of the participant in terms of frequency of text messages, who should be sending the text message, what it should contain (eg, quotes from scripture). Fourth, the messaging needs to come from someone the participant trusts and with whom they have a connection. Fifth, the messaging needs be caring and warm, not cold and impersonal. Finally, short, simple, and direct, but personalized, is the best type of messaging.

However, African American and Native Hawaiian and Pacific Islander gatekeeper-stakeholders did tend to differ on some thoughts regarding the intervention. For instance, there was a strong sentiment among the African American group that the church be involved and that the intervention begin with group workshops, whereas the Native Hawaiian and Pacific Islander group seemed to believe that the teaching could occur on a one-to-one basis with the health care provider (eg, physician, pharmacist, community health worker). Along the same lines, the African American group was also much more likely to recommend including verse from scripture or quotes from famous authors in the messaging, as well as incorporating tips on hobbies and recipes. Moreover, a distrust of the medical system was more commonly expressed among the African American gatekeeper-stakeholders, leading to the suggestion that the messaging originate from someone outside the medical system, such as a family member, pastor, or peer.

In conclusion, our study provides the first qualitative assessment of perspectives on use of mHealth technology to improve medication adherence among elderly African Americans and Native Hawaiian and Pacific Islanders. Successful interventions need to be personalized to meet patient needs, come from a place of caring from a trusted person, and be short and simple to understand.

Conflicts of Interest

None declared.

References

1. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, American Heart Association Statistics Committee and Stroke Statistics Subcommittee. Heart disease and stroke statistics--2015 update: a report from the American Heart Association. Circulation 2015 Jan 27;131(4):e29-e322. [doi: 10.1161/CIR.0000000000000152] [Medline: 25520374]
2. Yazdanshenas H, Bazargan M, Orum G, Loni L, Mahabadi N, Husaini B. Prescribing patterns in the treatment of hypertension among underserved African American elderly. Ethn Dis 2014;24(4):431-437 [FREE Full text] [doi: 10.1177/0898264315592600] [Medline: 25417425]

XSL•FO
RenderX

3. Moy KL, Sallis JF, David KJ. Health indicators of Native Hawaiian and Pacific Islanders in the United States. J Community Health 2010 Feb;35(1):81-92 [FREE Full text] [doi: 10.1007/s10900-009-9194-0] [Medline: 19856087]

4. Juarez DT, Davis JW, Brady SK, Chung RS. Prevalence of heart disease and its risk factors related to age in Asians, Pacific Islanders, and Whites in Hawai'i. J Health Care Poor Underserved 2012 Aug;23(3):1000-1010. [doi: 10.1353/hpu.2012.0103] [Medline: 24212153]

5. Havas S, Sherwin R. Putting it all together: summary of the NHLBI Workshop on the Epidemiology of Hypertension in Hispanic American, Native American, and Asian/Pacific Islander American Populations. Public Health Rep 1996;111 Suppl 2:77-79 [FREE Full text] [Medline: 8898784]

6. Stannard D. Summaries of nursing care-related systematic reviews from the Cochrane Library: pharmacotherapy for mild hypertension. J Cardiovasc Nurs 2014;29(1):10-11. [doi: 10.1097/JCN.0b013e31829eae67] [Medline: 24309431]

7. Frieden TR, King SM, Wright JS. Protocol-based treatment of hypertension: a critical step on the pathway to progress. JAMA 2014 Jan 1;311(1):21-22. [doi: 10.1001/jama.2013.282615] [Medline: 24231925]

8. Dave GJ, Bibeau DL, Schulz MR, Aronson RE, Ivanov LL, Black A, et al. Predictors of uncontrolled hypertension in the Stroke Belt. J Clin Hypertens (Greenwich) 2013 Aug;15(8):562-569 [FREE Full text] [doi: 10.1111/jch.12122] [Medline: 23889718]

9. Zhao B, Jose PO, Pu J, Chung S, Ancheta IB, Fortmann SP, et al. Racial/ethnic differences in hypertension prevalence, treatment, and control for outpatients in northern California 2010-2012. Am J Hypertens 2015 May;28(5):631-639. [doi: 10.1093/ajh/hpu189] [Medline: 25352230]

10. Taira DA, Gelber RP, Davis J, Gronley K, Chung RS, Seto TB. Antihypertensive adherence and drug class among Asian Pacific Americans. Ethn Health 2007 Jun;12(3):265-281. [doi: 10.1080/13557850701234955] [Medline: 17454100]

11. Juarez DT, Williams AE, Chen C, Daida YG, Tanaka SK, Trinacty CM, et al. Factors affecting medication adherence trajectories for patients with heart failure. Am J Manag Care 2015 Mar;21(3):e197-e205 [FREE Full text] [Medline: 26014307]

12. Sabaté E. Adherence to Long-Term Therapies: Evidence for Action. Geneva: World Health Organization; 2003.

13. Col N, Fanale JE, Kronholm P. The role of medication noncompliance and adverse drug reactions in hospitalizations of the elderly. Arch Intern Med 1990 Apr;150(4):841-845. [Medline: 2327844]

14. Berg J, Dischler J, Wagner D, Raia J, Palmer-Shevlin N. Medication compliance: a healthcare problem. Ann Pharmacother 1993 Sep;27(9 Suppl):S1-S24. [Medline: 8400462]

15. Sokol MC, McGuigan KA, Verbrugge RR, Epstein RS. Impact of medication adherence on hospitalization risk and healthcare cost. Med Care 2005 Jun;43(6):521-530. [Medline: 15908846]

16. Osterberg L, Blaschke T. Adherence to medication. N Engl J Med 2005 Aug 4;353(5):487-497. [doi: 10.1056/NEJMra050100] [Medline: 16079372]

17. Roebuck MC, Liberman JN, Gemmill-Toyama M, Brennan TA. Medication adherence leads to lower health care use and costs despite increased drug spending. Health Aff (Millwood) 2011 Jan;30(1):91-99 [FREE Full text] [doi: 10.1377/hlthaff.2009.1087] [Medline: 21209444]

18. Rosen MI, Rigsby MO, Salahi JT, Ryan CE, Cramer JA. Electronic monitoring and counseling to improve medication adherence. Behav Res Ther 2004 Apr;42(4):409-422. [doi: 10.1016/S0005-7967(03)00149-9] [Medline: 14998735]

19. McGillicuddy JW, Gregoski MJ, Weiland AK, Rock RA, Brunner-Jackson BM, Patel SK, et al. Mobile health medication adherence and blood pressure control in renal transplant recipients: a proof-of-concept randomized controlled trial. JMIR Res Protoc 2013 Aug;2(2):e32 [FREE Full text] [doi: 10.2196/resprot.2633] [Medline: 24004517]

20. Matsuyama JR, Mason BJ, Jue SG. Pharmacists' interventions using an electronic medication-event monitoring device's adherence data versus pill counts. Ann Pharmacother 1993;27(7-8):851-855. [Medline: 8364262]

21. Mann D, Kudesia V, Reddy S, Weng M, Imler D, Quintiliani L. Development of DASH Mobile: a mHealth lifestyle change intervention for the management of hypertension. Stud Health Technol Inform 2013;192:973. [Medline: 23920747]

22. Logan AG. Transforming hypertension management using mobile health technology for telemonitoring and self-care support. Can J Cardiol 2013 May;29(5):579-585. [doi: 10.1016/j.cjca.2013.02.024] [Medline: 23618506]

23. Green BB, Cook AJ, Ralston JD, Fishman PA, Catz SL, Carlson J, et al. Effectiveness of home blood pressure monitoring, Web communication, and pharmacist care on hypertension control: a randomized controlled trial. JAMA 2008 Jun 25;299(24):2857-2867 [FREE Full text] [doi: 10.1001/jama.299.24.2857] [Medline: 18577730]

24. Feinn R, Tennen H, Cramer J, Kranzler HR. Measurement and prediction of medication compliance in problem drinkers. Alcohol Clin Exp Res 2003 Aug;27(8):1286-1292. [doi: 10.1097/01.ALC.0000080670.59386.6E] [Medline: 12966323]

25. de Bruin M, Hospers H, van den Borne HW, Kok G, Prins J. Theory- and evidence-based intervention to improve adherence to antiretroviral therapy among HIV-infected patients in the Netherlands: a pilot study. AIDS Patient Care STDS 2005 Jun;19(6):384-394. [doi: 10.1089/apc.2005.19.384] [Medline: 15989434]

26. Brath H, Morak J, Kästenbauer T, Modre-Osprian R, Strohner-Kästenbauer H, Schwarz M, et al. Mobile health (mHealth) based medication adherence measurement - a pilot trial using electronic blisters in diabetes patients. Br J Clin Pharmacol 2013 Sep;76 Suppl 1:47-55 [FREE Full text] [doi: 10.1111/bcp.12184] [Medline: 24007452]

27. Smith J, Firth J. Qualitative data analysis: the framework approach. Nurse Res 2011;18(2):52-62. [doi: 10.7748/nr2011.01.18.2.52.c8284] [Medline: 21319484]

XSL·FO
RenderX

28.   Bradley EH, Curry LA, Devers KJ. Qualitative data analysis for health services research: developing taxonomy, themes, and theory. Health Serv Res 2007 Aug;42(4):1758-1772 [FREE Full text] [doi: 10.1111/j.1475-6773.2006.00684.x] [Medline: 17286625]

29.   Ramirez V, Johnson E, Gonzalez C, Ramirez V, Rubino B, Rossetti G. Assessing the use of mobile health technology by patients: an observational study in primary care clinics. JMIR Mhealth Uhealth 2016 Apr;4(2):e41 [FREE Full text] [doi: 10.2196/mhealth.4928] [Medline: 27095507]

XSL•FO

RenderX

Original Paper

The Mobile Phone Affinity Scale: Enhancement and Refinement

Beth C Bock[1,2,3*], PhD; Ryan Lantini[1*], MA; Herpreet Thind[4*], PhD; Kristen Walaska[1*], BS; Rochelle K Rosen[1,3], PhD; Joseph L Fava[1*], PhD; Nancy P Barnett[5*], PhD; Lori AJ Scott-Sheldon[1,2,3*], PhD

[1]Centers for Behavioral and Preventive Medicine, The Miriam Hospital, Providence, RI, United States

[2]Department of Psychiatry and Human Behavior, Alpert School of Medicine, Brown University, Providence, RI, United States

[3]Department of Behavioral and Social Sciences, School of Public Health, Brown University, Providence, RI, United States

[4]Department of Public Health, University of Massachusetts Lowell, Lowell, MA, United States

[5]Department of Behavioral and Social Sciences, Center for Alcohol and Addiction Studies, School of Public Health, Brown University, Providence, RI, United States

[*]these authors contributed equally

Corresponding Author:
Beth C Bock, PhD
Centers for Behavioral and Preventive Medicine
The Miriam Hospital
CORO Building, Suite 309
164 Summit Avenue
Providence, RI, 02906
United States
Phone: 1 401 793 8020
Fax: 1 401 793 8059
Email: Bbock@lifespan.org

Abstract

Background: Existing instruments that assess individuals' relationships with mobile phones tend to focus on negative constructs such as addiction or dependence, and appear to assume that high mobile phone use reflects pathology. Mobile phones can be beneficial for health behavior change, disease management, work productivity, and social connections, so there is a need for an instrument that provides a more balanced assessment of the various aspects of individuals' relationships with mobile phones.

Objective: The purpose of this research was to develop, revise, and validate the Mobile Phone Affinity Scale, a multi-scale instrument designed to assess key factors associated with mobile phone use.

Methods: Participants (N=1058, mean age 33) were recruited from Amazon Mechanical Turk between March and April of 2016 to complete a survey that assessed participants' mobile phone attitudes and use, anxious and depressive symptoms, and resilience.

Results: Confirmatory factor analysis supported a 6-factor model. The final measure consisted of 24 items, with 4 items on each of 6 factors: Connectedness, Productivity, Empowerment, Anxious Attachment, Addiction, and Continuous Use. The subscales demonstrated strong internal consistency (Cronbach alpha range=0.76-0.88, mean 0.83), and high item factor loadings (range=0.57-0.87, mean 0.75). Tests for validity further demonstrated support for the individual subscales.

Conclusions: Mobile phone affinity may have an important impact in the development and effectiveness of mobile health interventions, and continued research is needed to assess its predictive ability in health behavior change interventions delivered via mobile phones.

Introduction

Mobile phones have shown promise as an effective delivery tool for health behavior change and disease management [1-6], and may be particularly well suited for interventions designed for young adults and adolescents. This population uses texting, apps, and other phone-based applications regularly, a trend that is particularly strong among younger age groups [7]. When developing an intervention that is delivered through mobile devices, it is important to consider how an individual uses his/her mobile phone, as mobile phone use may influence the receptivity to, and ultimately the efficacy of, mobile health (mHealth) programs and interventions [8,9].

There are several published measures that assess the use of technology, including excessive mobile phone use and Internet addiction [10-13]. These measures are largely derived from the addictions and psychopathology literature, and are intended to measure problematic use of technology within a conceptual framework of use-as-pathology [13-15]. Problematic psychological constructs that have been linked to mobile phone use include impulsivity [11], depression [14], and anxiety [14]. However, mobile phones can serve many positive functions. For example, many apps now exist to help people track health behaviors (eg, exercise, weight loss) and manage medical conditions, including diabetes and asthma [16-18]. Mobile phones also serve positive functions, including increasing efficiency and productivity at work, and improving connectivity to social networks, family, and friends [19,20]. However, to our knowledge no measures assess mobile phone and technology use that include items reflective of these positive elements. Therefore, the goals of this study were to expand, revise, and validate the psychometric properties of the Mobile Phone Affinity Scale (MPAS) based on a version that we previously developed in a study of community college students [21]. The current study employs a national cross-sectional survey of adults living in the United States to identify important constructs related to mobile phone use, develop scales to measure these constructs, evaluate the internal consistency of the constructs, and establish the validity of the newly revised instrument.

Methods

Questionnaire Development

In the initial development of the MPAS, factors that may be associated with mobile phone use (and more broadly technology use) were identified by conducting a search of the relevant literature using PubMed. Identified constructs included social connectivity, dependence, addiction, mood, and continuous use [21]. To better assess positive functions of mobile phone use, the same search procedures were used in this study to expand the MPAS to include constructs related to empowerment, safety, and usefulness in the domains of social and personal use, and work and school-related use. Individuals from our research team independently wrote a series of items for each of these three additional constructs in English at a sixth grade reading level. The entire research team reviewed the items to determine face validity. Any confusing or ambiguous items were edited and duplicates were deleted. Instructions and response format were

also reviewed for clarity. The resulting instrument contained 57 items measuring 7 constructs, with 6-to-9 items per construct, and used a Likert-type scale response format. This initial draft of the revised MPAS was then pretested with eight adults to confirm item clarity and comprehension before administering the measure to a larger sample.

Psychometric Testing

To test the MPAS, we conducted a national cross-sectional survey between March 25 and April 1, 2016, to assess respondents' mobile phone use, attitudes toward their mobile phone, current mood, demographic information, and other characteristics. Participants were registered users (*workers*) of Amazon Mechanical Turk (MTurk). We used MTurk to recruit for this survey since research has shown it to be a fast, inexpensive, reliable, and useful approach to collect data from a large and ethnically diverse sample [22-26].

To participate in the study, workers were redirected through the MTurk website to our project survey website, which presented detailed information about the study and an informed consent document. Interested workers who indicated their consent were then linked to the screening questions to assess eligibility. Workers were eligible to participate in the study if they were 18 years of age or older, lived in the United States, could read fluently in English, and owned a mobile phone. If eligible, participants were then directed to the full survey, which asked questions about their mobile phone use, attitudes toward their mobile phone, their current mood, and demographic information.

The online survey was managed using a secure Web-based application, Research Electronic Data Capture [27], hosted in our institution's Information Services department. Participants were compensated US $1 upon completing the survey. No identifying information was asked of participants, thus keeping their responses anonymous. The informed consent process, assessment of eligibility, and completion of the surveys took an average of 17 minutes to complete. This study was approved by the Institutional Review Board at The Miriam Hospital.

Measures

Mobile Phone Affinity Scale

The initial MPAS consisted of 57 statements about mobile phone use. Participants were asked to report how true each statement was for them, using a 5-point Likert-type response format (1=not at all true to 5=extremely true).

Other Measures

In addition to the MPAS, participants responded to demographic questions regarding age, gender, race, ethnicity, education, employment status, and marital status. Participants also responded to questions about their mobile phone (ie, *is it a smartphone?*), the types of activities and apps used on the mobile phone, and the frequency with which they used text messaging.

Previous research has reported increased levels of anxiety, depression, and impulsivity associated with problematic mobile phone use [11,14], and since the MPAS was designed to assess both negative and positive constructs associated with mobile

phone use, we included measures of these constructs to validate MPAS subscales. These measures were administered as part of the online questionnaire and included: (1) *symptoms of anxiety*, which were assessed using the 20-item State-Trait Anxiety Inventory (STAI) [28] which is scored on a 4-point scale (1=almost never to 4=almost always), with higher scores indicating higher levels of anxiety; (2) *depressive symptoms*, which were measured using the 10-item Centers for Epidemiological Studies Depression Scale (CESD-10) [29] scored on a 4-point scale (0=rarely/none of the time to 4=most/all of the time); and (3) *impulsiveness*, which was assessed using the 30-item Barratt Impulsiveness Scale (BIS-11) [30] scored on a 4-point scale (1=rarely/never to 4=almost always/always), with higher scores indicating greater impulsiveness. In addition, to differentiate subscales within the MPAS that we hoped would assess more positive and/or beneficial applications of mobile phone use, we chose *psychological resilience* as a positive psychological construct. We assessed this parameter using the 6-item Brief Resilience Scale (BRS) [31] scored on 5-point scale (1=strongly disagree to 5=strongly agree), with higher scores indicating greater psychological resilience.

Statistical Analyses

First, we used summary statistics (means, standard deviations, and frequencies) to describe the sample characteristics and measures for the entire sample. Second, a preliminary analysis was conducted to examine the MPAS item characteristics, and psychometric analyses were then conducted using confirmatory factor analysis (CFA) in Mplus Version 7.3 [32]. Model fit was evaluated based on the minimum fit function chi-square statistic, the comparative fit index (CFI; [33]), the nonnormed or Tucker-Lewis index (TLI; [34]), the root means square error of approximation (RMSEA; [35]), the standardized root mean square residual (SRMR; [36]), and the model chi-square. Respecifications to the model were guided by theory, in combination with modification indices that are part of the statistical output, and designed to produce a brief measure of mobile phone attachment that is suitable for use in field research. Third, Cronbach alpha [37], a measure of internal consistency reliability, was estimated for each subscale. Finally, to test for concurrent validity of the final measure, we examined the association between MPAS subscales and measures of anxiety, depression, and impulsivity.

Results

Participants

Of the 1241 MTurk workers who were redirected to our survey website, 1.05% (13/1241) never completed the informed consent process, 5.32% (66/1241) of the potential participants completed the informed consent process but were deemed to be ineligible for the study, and 8.30% (103/1241) of the participants did not complete the survey. The data for one additional participant was corrupted and removed from the analyses.

Our analyses were restricted to the 1058 participants who agreed to participate in the study and completed all aspects of the survey. Participants were predominately white (877/1058; 82.89%). Men accounted for 49.91% (528/1058) of the participants, and women comprised 50.09% (530/1058) of the sample. Participants were between 18 and 87 years of age (mean 32.5, standard deviation 10.3). Table 1 provides the complete demographic information for the sample.

Table 1. Demographics for the overall sample (N=1058).

Sample Characteristic	Proportion, % (n)
Sex	
Male	49.91 (528)
Female	50.09 (530)
Race	
White	82.89 (877)
Black	6.14 (65)
Asian	5.01 (53)
Native Hawaiian	0.47 (5)
Pacific Islander	0.28 (3)
Other	1.51 (16)
Multiple	3.69 (39)
Ethnicity	
Hispanic	10.02 (106)
Non-Hispanic	89.98 (952)
Marital status	
Single	51.13 (541)
Married	39.89 (422)
Separated	1.61 (17)
Divorced	6.62 (70)
Widowed	0.76 (8)
Census region	
Northeast	18.53 (196)
South	39.13 (414)
Midwest	21.17 (224)
West	20.70 (219)
Pacific	0.47 (5)
Education	
High school or less	10.87 (115)
Some college	40.36 (427)
College degree or above	48.68 (515)
Student	
Yes, full-time	14.37 (152)
Yes, part-time	6.24 (66)
No	79.30 (839)
Employed	
Yes, full-time	57.09 (604)
Yes, part-time	20.42 (216)
No	22.40 (237)
Income	
<$25,000	17.77 (188)
$25,000-$49,999	32.42 (343)
$50,000-$99,999	34.50 (365)

Sample Characteristic	Proportion, % (n)
>$100,000	11.53 (122)
Missing	3.78 (40)
Living arrangement	
Alone	18.34 (194)
Spouse/partner	30.91 (327)
Adult roommate	9.36 (99)
Parents	10.40 (110)
Child(ren)	3.78 (40)
Spouse/partner and child(ren)	22.59 (239)
Multiple	4.63 (49)

Dimensional and Internal Validity Analyses

Preliminary item level analyses were conducted to examine individual item means, standard deviations, skew, and kurtosis in each of the 57 items, and the results were judged adequate to proceed with the dimensional analysis. An initial CFA model was fit using the full 57-item variable set and specified 7 correlated factors, with each of the items only allowed to load and be freely estimated on its hypothesized factor. This initial model fit poorly ($\chi^2_{1,518}$=9375.0, RMSEA=.07, CFA=.79, TLI=.78, SRMR=.071). The model fit was improved by removing items with poor loadings (<.4), very high cross-factor error correlations, or potentially high cross-factor loading, which indicated a complex item. One further adjustment included collapsing the initially posited separate factors of safety and empowerment into a single factor, based on the modification indices and conceptual similarity of the item content.

The final model (χ^2_{237}=1100.9, RMSEA=.059, CFA=.941, TLI=.931, SRMR=.042) represented a parsimonious and balanced solution with 6 correlated factors, each measured by 4 items, creating a final measure consisting of 24 items

(Multimedia Appendix 1). This 24-variable solution fit very well based on alternative criteria suggested by Hu and Bentler [38] and did not use any superfluous adjustments, such as freeing error covariance parameters or allowing variables to load on additional factors, to achieve the final improved model fit. Individual item loadings were high for all items on their respective factors (range=0.57-0.87, mean 0.75; see Table 2), and internal consistency reliability was also very good for each factor, as measured by Cronbach alpha (range=0.76-0.88, mean 0.83; see Table 2). The disattenuated correlations for the latent constructs are presented in Table 3.

Concurrent validity analyses examined correlations between each MPAS subscale and measures of depressive symptoms (CESD-10), anxiety (STAI), resilience (BRS), and impulsiveness (BIS). Depressive symptoms, anxiety, and impulsivity were significantly correlated with *Anxious Attachment*, *Addiction*, and *Continuous Use* subscales, but not with the *Connectedness*, *Productivity*, or *Empowerment* subscales. Resilience was significantly negatively correlated with scores on the *Addiction* and *Anxious Attachment* subscales and positively correlated with the *Productivity* subscale (Table 4).

Table 2. Confirmatory factor analyses, item loadings, and Cronbach coefficient alpha for 6 correlated factors that were modeled using the whole sample.

Parameter	Connectedness	Productivity	Empowerment	Anxious Attachment	Addiction	Continuous Use
My phone helps me keep track of my social life	.76					
When it comes to my health or social life, my phone is my personal assistant	.71					
My phone helps me stay close to my friends and family	.70					
My phone makes it easy to cancel plans with others	.60					
My phone helps me to be more organized at work/school		.74				
I use my phone to connect with my co-workers or other students		.75				
My phone is necessary for work/school		.72				
My phone helps me stay up-to-date with work/school activities		.87				
Having my phone with me makes it easier to leave a risky situation			.73			
I feel in control when I have my phone with me			.81			
My phone gives me a sense of security			.84			
I feel safe when I have my phone with me			.83			
I feel anxious if I don't have my phone with me				.73		
I feel isolated without my phone				.79		
I feel dependent on my phone				.80		
Without my mobile phone, I feel out of touch with the world				.82		
I find myself occupied on my phone even when I'm with other people					.75	
I find myself occupied with my phone when I should be doing other things					.84	
I find myself engaged with my mobile phone for longer periods of time than I intended					.85	
I would get more work done if I spent less time on my phone					.68	
I read/send text messages from my mobile phone, when I am at work or in class, that are not related to what I am doing						.57
I use my phone all day						.74
I am never bored if I have my phone with me						.61
I rely on my phone 24/7						.79
Cronbach alpha	.78	.85	.88	.86	.86	.76

XSL•FO
RenderX

Table 3. Disattenuated correlations among 6 latent constructs from confirmatory factor analyses (upper-right triangle), with the raw summated scale score correlations (lower-left triangle).

	Connectedness	Productivity	Empowerment	Anxious Attachment	Addiction	Continuous Use
Connectedness	–	.74	.82	.77	.66	.81
Productivity	.60	–	.55	.53	.46	.65
Empowerment	.69	.48	–	.78	.47	.69
Anxious Attachment	.63	.45	.68	–	.72	.84
Addiction	.54	.40	.41	.61	–	.78
Continuous Use	.65	.55	.57	.68	.67	–

Table 4. Correlations for complete sample among the 6 Mobile Phone Affinity Scale (MPAS) scores and Centers for Epidemiologic Studies Depression Scale (CESD), State-Trait Anxiety Inventory (STAI), Barratt Impulsiveness Scale (BIS-11), and Brief Resilience Scale (BRS).

	Depression (CESD-10)		Anxiety (STAI)		Impulsivity (BIS-11)		Resilience (BRS)	
	r^a	P value	r^a	P value	r^a	P value	r^a	P value
Connectedness	.03	.36	-.00006	.99	.01	.81	.06	.07
Productivity	.02	.56	-.04	.20	-.04	.21	.09	<.001
Empowerment	.02	.51	.03	.39	-.01	.64	.00	.95
Anxious Attachment	.13	<.001	.15	<.001	.13	<.001	-.12	<.001
Addiction	.25	<.001	.18	<.001	.29	<.001	-.15	<.001
Continuous Use	.10	<.001	.08	.01	.14	<.001	-.03	.28

[a]r=Pearson product-moment correlation coefficient

Discussion

This study developed, revised, and validated the MPAS, a multi-dimensional survey instrument with strong internal consistency and high item factor loadings across 6 distinct subscales. The MPAS provides a measure of an individual's relationship to his/her mobile phone with positive (*Connectedness*, *Productivity*, *Empowerment*), negative (*Anxious Attachment*, *Addiction*), and neutral (*Continuous Use*) valences. Results showed that the subscales we conceived of as positive (eg, *Connectedness*) were not correlated with depressive symptoms, anxiety, or impulsivity, while both negative subscales (*Anxious Attachment*, *Addiction*) were correlated with these features. Personal resilience, a positive characteristic, was significantly and positively correlated with affinity for the mobile phone for *Productivity* purposes and negatively correlated with both negative subscales. This finding differentiates the MPAS from other instruments in that both impulsivity and anxiety have been shown to be associated with mobile phone use in studies of mobile phone addiction or dependence [39-41]. However, the *Connectedness* subscale was not correlated with resilience, which is reasonable given that personal relationships with others can help promote personal resilience [42,43], but can also place a burden on individuals [44]. Thus, no clear positive or negative association with resilience would be expected among those with high affinity for using their mobile phone for social connections.

In this revision of the MPAS, we have succeeded in our goal of creating a scale that represents both positive and negative aspects of increased use of mobile technology. Using the previous 4-subscale version, most scales were correlated with anxiety-related and depressive symptoms [21], while a much clearer differentiation was observed in the present 6-subscale version. In the current revised version, only *Continuous Use*, *Anxious Attachment,* and *Addiction* were associated with depressive symptoms, while use of mobile phones for *Connectedness*, *Productivity,* and *Empowerment* were not. Furthermore, in the current version of the MPAS, the *Continuous Use* subscale is correlated with anxiety, but less strongly than the *Anxious Attachment* or *Addiction* subscales, suggesting that some level of anxiety can be a positive, functional trait, by enhancing attentiveness to important things (eg, making sure to bring your phone with you or ensuring the battery is charged).

Limitations

Ultimately, we anticipate the value of the MPAS will be to provide predictive value for mHealth interventions, necessitating the assessment of whether scores on the MPAS (or any of its subscales) relate significantly to outcomes of interventions. However, this instrument was not developed in the context of an intervention study, and it is a limitation of this study that we did not assess other factors (eg, intentions to change health behaviors or outcome expectations) that may have served as a proxy for intervention outcomes. A second important limitation relates to the nature of the study sample. Although MTurk

workers tend to be racially and ethnically diverse, they are likely to be more technologically savvy than most mobile phone users, and may not be representative of the overall US population in that regard. Thus, it is possible that the mean value of any individual MPAS subscale may differ between the MTurk study sample and other US population samples, but the high internal consistency reliability of each subscale, and the associated high item loadings on each subscale, indicate that the instrument has strong internal validity. We expect that this instrument will prove valuable for its intended purposes when used in other adult samples.

Future Directions

Additional work is needed to examine whether scores on the MPAS or its subscales are predictive of uptake, maintenance, and successful outcomes among individuals who are interacting with a behavioral health intervention delivered (in whole or in part) through mobile technology. It is our hope that the MPAS developed in this study may prove to be a useful indicator of the quality of the individual's relationship with their mobile phone, and may comprise an important element in understanding the efficacy of mHealth interventions and programs.

There is tremendous growth in technologies that can provide health care and behavioral health interventions through mobile channels, such as apps and text messaging [16-18]. While the qualities of both the intervention and the delivery technology are important, in our quest to understand the effects that these interventions have on behaviors, it is also important to understand the relationship that the individual has with their mobile phone. mHealth involves not only health behaviors, but also those behaviors and attitudes relevant to interacting with mobile technology, and interacting with other people through mobile devices. Given the immediate and reciprocal nature of both health behavior and the behavior of interacting with a mobile device, thought leaders have suggested that our current behavioral health theories may be inadequate, particularly as mHealth interventions become increasingly interactive and adaptive [45]. Several papers have called for an expansion of our understanding of how interacting with a mobile device impacts health behavior [46,47]. Research is needed that can contribute to new theories regarding the interaction between mobile technology use, mHealth interventions, and behavior change.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Final 24-item Mobile Phone Affinity Scale.

[PDF File (Adobe PDF File), 46KB - mhealth_v4i4e134_app1.pdf]

References

1. 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]
2. 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]
3. Finitsis DJ, Pellowski JA, Johnson BT. Text message intervention designs to promote adherence to antiretroviral therapy (ART): a meta-analysis of randomized controlled trials. PLoS One 2014;9(2):e88166 [FREE Full text] [doi: 10.1371/journal.pone.0088166] [Medline: 24505411]
4. 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]
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. Free C, Phillips G, Watson L, Galli L, Felix L, Edwards P, et al. The effectiveness of mobile-health technologies to improve health care service delivery processes: a systematic review and meta-analysis. PLoS Med 2013 Jan;10(1):e1001363 [FREE Full text] [doi: 10.1371/journal.pmed.1001363] [Medline: 23458994]
7. 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;3(1):e22 [FREE Full text] [doi: 10.2196/mhealth.3779] [Medline: 25714907]
8. Duggan M. Cell phone activities. Pew Research Center Internet & American Life Project. Washington, DC; 2013. URL: http://www.pewinternet.org/2013/09/19/cell-phone-activities-2013/ [accessed 2016-08-23] [WebCite Cache ID 6jytbp1AY]

9.   Duggan M. Pew Research Center Internet & American Life Project. Mobile messaging and social media. Pew Research Center Internet & American Life Project. Washington, DC; 2015. URL: http://www.pewinternet.org/2015/08/19/mobile-messaging-and-social-media-2015/ [accessed 2016-08-23] [WebCite Cache ID 6jytp07LY]

10. Bianchi A, Phillips JG. Psychological predictors of problem mobile phone use. Cyberpsychol Behav 2005 Feb;8(1):39-51. [doi: 10.1089/cpb.2005.8.39] [Medline: 15738692]

11. Billieux J. Problematic use of the mobile phone: a literature review and a pathways model. CPSR 2012 Oct 01;8(4):299-307. [doi: 10.2174/157340012803520522]

12. Merlo LJ, Stone AM, Bibbey A. Measuring problematic mobile phone use: development and preliminary psychometric properties of the PUMP scale. J Addict 2013;2013:912807 [FREE Full text] [doi: 10.1155/2013/912807] [Medline: 24826371]

13. Jenaro C, Flores N, Gómez-Vela M, González-Gil F, Caballo C. Problematic internet and cell-phone use: psychological, behavioral, and health correlates. Addict Res Theory 2009 Jul 11;15(3):309-320. [doi: 10.1080/16066350701350247]

14. Thomée S, Härenstam A, Hagberg M. Mobile phone use and stress, sleep disturbances, and symptoms of depression among young adults--a prospective cohort study. BMC Public Health 2011;11:66 [FREE Full text] [doi: 10.1186/1471-2458-11-66] [Medline: 21281471]

15. Augner C, Hacker GW. Associations between problematic mobile phone use and psychological parameters in young adults. Int J Public Health 2012 Apr;57(2):437-441. [doi: 10.1007/s00038-011-0234-z] [Medline: 21290162]

16. 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]

17. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: a systematic review. Telemed J E Health 2009 Apr;15(3):231-240. [doi: 10.1089/tmj.2008.0099] [Medline: 19382860]

18. Peterson A. Improving type 1 diabetes management with mobile tools: a systematic review. J Diabetes Sci Technol 2014 Jul;8(4):859-864 [FREE Full text] [doi: 10.1177/1932296814529885] [Medline: 24876414]

19. Walsh SP, White KM, Cox S, Young RM. Keeping in constant touch: the predictors of young Australians' mobile phone involvement. Comput Human Behav 2011 Jan;27(1):333-342. [doi: 10.1016/j.chb.2010.08.011]

20. Walsh S, White KM, McD Young R. Needing to connect: the effect of self and others on young people's involvement with their mobile phones. Aus J Psych 2010 Dec;62(4):194-203. [doi: 10.1080/00049530903567229]

21. Bock B, Thind H, Fava J, Walaska K, Barnett N, Rosen R, et al. Development of the mobile phone attachment scale. 2016 Jan 05 Presented at: 49th Hawaii International Conference on System Sciences (HICSS); January 2016; Kauai, Hawaii p. 5-8. [doi: 10.1109/HICSS.2016.424]

22. Paolacci G, Chandler J. Inside the Turk: understanding Mechanical Turk as a participant pool. Curr Dir Psychol Sci 2014 Jun 03;23(3):184-188. [doi: 10.1177/0963721414531598]

23. Paolacci G, Chandler J, Ipeirotis P. Running experiments on Amazon Mechanical Turk. Judgm Decis Mak 2010;5(5):411-419.

24. Buhrmester M, Kwang T, Gosling SD. Amazon's Mechanical Turk: a new source of inexpensive, yet high-quality, data? Perspect Psychol Sci 2011 Jan;6(1):3-5. [doi: 10.1177/1745691610393980] [Medline: 26162106]

25. Berinsky AJ, Huber GA, Lenz GS. Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk. Political Analysis 2012 Mar 02;20(3):351-368. [doi: 10.1093/pan/mpr057]

26. Mason W, Suri S. Conducting behavioral research on Amazon's Mechanical Turk. Behav Res Methods 2012 Mar;44(1):1-23. [doi: 10.3758/s13428-011-0124-6] [Medline: 21717266]

27. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009 Apr;42(2):377-381 [FREE Full text] [doi: 10.1016/j.jbi.2008.08.010] [Medline: 18929686]

28. Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. Manual for the state-trait anxiety inventory. Palo Alto, CA: Consulting Psychologists Press; 1983.

29. Andresen EM, Malmgren JA, Carter WB, Patrick DL. Screening for depression in well older adults: evaluation of a short form of the CES-D (Center for Epidemiologic Studies Depression Scale). Am J Prev Med 1994 Mar;10(2):77-84. [Medline: 8037935]

30. Patton JH, Stanford MS, Barratt ES. Factor structure of the Barratt impulsiveness scale. J Clin Psychol 1995 Nov;51(6):768-774. [Medline: 8778124]

31. Smith BW, Dalen J, Wiggins K, Tooley E, Christopher P, Bernard J. The brief resilience scale: assessing the ability to bounce back. Int J Behav Med 2008;15(3):194-200. [doi: 10.1080/10705500802222972] [Medline: 18696313]

32. Muthén L, Muthén B. Mplus User's Guide. Seventh Edition. Los Angeles, CA: Muthén & Muthén; 1998.

33. Bentler PM. Comparative fit indexes in structural models. Psychol Bull 1990 Mar;107(2):238-246. [Medline: 2320703]

34. Tucker LR, Lewis C. A reliability coefficient for maximum likelihood factor analysis. Psychometrika 1973 Mar;38(1):1-10. [doi: 10.1007/BF02291170]

35. Steiger JH. Structural model evaluation and modification: an interval estimation approach. Multivariate Behav Res 1990 Apr 01;25(2):173-180. [doi: 10.1207/s15327906mbr2502_4] [Medline: 26794479]

36. Bentler P. EQS: a structural equations program manual. Los Angeles, CA: BMDP Statistical Software; 1989.

37. Cronbach LJ. Coefficient alpha and the internal structure of tests. Psychometrika 1951 Sep;16(3):297-334. [doi: 10.1007/BF02310555]

38.   Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999 Jan;6(1):1-55. [doi: 10.1080/10705519909540118]

39.   Billieux J, Van der Linden M, Rochat L. The role of impulsivity in actual and problematic use of the mobile phone. Appl Cognit Psychol 2008 Dec;22(9):1195-1210. [doi: 10.1002/acp.1429]

40.   Choliz M. Mobile-phone addiction in adolescence: the test of mobile phone dependence (TMD). Prog Health Sci 2012;2(1):33-44.

41.   King AL, Valença AM, Silva AC, Sancassiani F, Machado S, Nardi AE. "Nomophobia": impact of cell phone use interfering with symptoms and emotions of individuals with panic disorder compared with a control group. Clin Pract Epidemiol Ment Health 2014;10:28-35 [FREE Full text] [doi: 10.2174/1745017901410010028] [Medline: 24669231]

42.   Kawachi I, Berkman LF. Social ties and mental health. J Urban Health 2001 Sep;78(3):458-467 [FREE Full text] [doi: 10.1093/jurban/78.3.458] [Medline: 11564849]

43.   van Kessel G. The ability of older people to overcome adversity: a review of the resilience concept. Geriatr Nurs 2013;34(2):122-127. [doi: 10.1016/j.gerinurse.2012.12.011] [Medline: 23332474]

44.   Pemberton R, Fuller Tyszkiewicz MD. Factors contributing to depressive mood states in everyday life: a systematic review. J Affect Disord 2016 Aug;200:103-110. [doi: 10.1016/j.jad.2016.04.023] [Medline: 27131503]

45.   Fiordelli M, Diviani N, Schulz PJ. Mapping mHealth research: a decade of evolution. J Med Internet Res 2013;15(5):e95 [FREE Full text] [doi: 10.2196/jmir.2430] [Medline: 23697600]

46.   de la Vega R, Miró J. mHealth: a strategic field without a solid scientific soul. A systematic review of pain-related apps. PLoS One 2014;9(7):e101312 [FREE Full text] [doi: 10.1371/journal.pone.0101312] [Medline: 24999983]

47.   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]

Abbreviations

BIS-11: Barratt Impulsiveness Scale
BRS: Brief Resilience Scale
CESD-10: Centers for Epidemiologic Studies Depression Scale
CFA: confirmatory factor analysis
CFI: Comparative Fit Index
mHealth: mobile health
MPAS: Mobile Phone Affinity Scale
MTurk: Amazon Mechanical Turk Service
RMSEA: root mean square error of approximation
SRMR: standardized root mean square residual
STAI: State-Trait Anxiety Inventory
TLI: Tucker-Lewis Index

XSL•FO
RenderX

Original Paper

Evaluation of Diet-Related Infographics on Pinterest for Use of Behavior Change Theories: A Content Analysis

Jessica L Wilkinson[1], MPH; Kate Strickling[1], MPH; Hannah E Payne[1], MPH; Kayla C Jensen[1], RD, MS; Joshua H West[1], MPH, PhD

Computational Health Science Research Group, Department of Health Science, Brigham Young University, Provo, UT, United States

Corresponding Author:
Jessica L Wilkinson, MPH
Computational Health Science Research Group
Department of Health Science
Brigham Young University
4002 LSB
Provo, UT
United States
Phone: 1 801 319 1777
Fax: 1 801 422 0273
Email: jltwilkinson@gmail.com

Abstract

Background: There is increasing interest in Pinterest as a method of disseminating health information. However, it is unclear whether the health information promoted on Pinterest is evidence-based or incorporates behavior change theory.

Objectives: The objective of the study was to determine the presence of health behavior theory (HBT) constructs in pins found on Pinterest and assess the relationship between various pin characteristics and the likelihood of inclusion of HBT.

Methods: A content analysis was conducted on pins collected from Pinterest identified with the search terms "nutrition infographic" and "healthy eating infographic." The coding rubric included HBT constructs, pin characteristics, and visual communication tools. Each HBT construct was coded as present or not present (yes=1, no=0). A total theory score was calculated by summing the values for each of the 9 constructs (range 0-9). Adjusted regression analysis was used to identify factors associated with the inclusion of health behavior change theory in pins ($P<.05$).

Results: The mean total theory score was 2.03 (SD 1.2). Perceived benefits were present most often (170/236, 72%), followed by behavioral capability (123/238, 51.7%) and perceived severity (79/236, 33.5%). The construct that appeared the least was self-regulation/self-control (2/237, 0.8%). Pin characteristics associated with the inclusion of HBT included a large amount of text ($P=.01$), photographs of real people ($P=.001$), cartoon pictures of food ($P=.01$), and the presence of references ($P=.001$). The number of repins ($P=.04$), likes ($P=.01$), and comments ($P=.01$) were positively associated with the inclusion of HBT.

Conclusions: These findings suggest that current Pinterest infographics targeting healthy eating contain few HBT elements. Health professionals and organizations should create and disseminate infographics that contain more elements of HBT to better influence healthy eating behavior. This may be accomplished by creating pins that use both text and images of people and food in order to portray elements of HBT and convey nutritional information.

Introduction

Chronic disease is a major public health concern. In the United States, heart disease, cancer, and stroke cause over 50% of all deaths [1]; globally, this figure is 60% [2]. While the etiology of obesity and related chronic diseases is multifactorial, poor diet is one of the most prominent risk factors [3]. The quality of the average American diet has decreased over the past several decades as sugar, salt, and trans fat consumption increased continually, resulting in an increased average caloric intake [4-6]. The Centers for Disease Control and Prevention reports that Americans are consuming a median of 1.8 servings of

vegetables and 1.1 servings of fruit per day compared to the 2 to 3 servings of vegetables and 2 servings of fruit recommended [7,8]. Promoting dietary modifications is one of the most important chronic disease prevention strategies [9]. The benefits of improved diet are not limited to disease prevention but also include positive health outcomes such as increased life expectancy, stable body weight, and improved mental health [10-12].

In an attempt to alter these unhealthy eating trends, health professionals are increasingly using the Internet to improve dietary behaviors in populations [13,14]. The Internet, including social media, has become a means of communicating health information [15]; in 2008, more than half of adult patients reported searching online for health information, and searching for health information was the most popular online activity for adults after email and general searches [16]. Pinterest, a social media platform that allows users to share content through photos and pinboards, is one site that has become a repository of health information, both formal and informal. A total of 40% of daily pinners (using Pinterest at least 1 time daily) and 25% of active pinners (using Pinterest at least 1 time monthly) consider Pinterest their "go-to" source for health information; furthermore, 84% of daily pinners reported trying something new once a week or more because of something they saw on Pinterest [17]. Additionally, more people with more diversity are using Pinterest; 67% of users are under age 40 years and 82% are female [17]. Minority group membership is growing with over half of users having joined in the past year [17]. Infographics are a common category of pins shared on Pinterest. Infographics are data visualization tools that aim to communicate information through elements such as graphs and images. Infographics have become increasingly popular in education due to their ability to present complex data in a simple and clear manner and are used by many public health professionals to disseminate health information [18-20]; social media sites like Pinterest may be a useful medium for health educators to share infographics efficiently to large numbers of people.

Pinterest is beginning to impact health educators and the way they share information [21]. Indeed, it is emerging as a tool not just among health professionals [22] but also among many health organizations [23,24]. However, despite the increasing interest in Pinterest as a method of disseminating health information, it is unclear whether the health information promoted on Pinterest is evidence-based or promotes behavior change. This is concerning provided that nutrition and diet information found on the Internet may largely be inaccurate; Hirasawa et al [25] found in a content analysis of online nutrition searches that most advice followed recommended nutrition guidelines only partially. Researchers have conducted content analyses of other health information promoted on social media [26-28] but to date there has been little research about health information found on Pinterest. In one of the few studies of health-related pins on Pinterest, Paige et al [29] reported that chronic obstructive pulmonary disease (COPD) pins incorporated significant levels of verbal persuasion and social modeling and may be useful as a health communication tool for COPD patients [29]. It is uncertain whether diet and nutrition pins are similarly appropriate for distribution by health professionals.

Pinterest may be a useful tool to disseminate information about dietary behaviors, but there is no research about the content of nutrition and diet pins. In particular, it is of interest whether they contain constructs of health behavior theory (HBT), as health promotion materials containing more elements of HBT have been demonstrated to be more effective in changing behavior [30-32]. Specifically, constructs from social cognitive theory [33-35] and health belief model [36-38] have been related to improved dietary practices. Moreover, it may be that approaches to changing dietary practices should incorporate both theories simultaneously [39]. Theory is used to assist a practitioner in organizing information along certain principles believed to change behavior [39]. In the case of Pinterest, then, theory could inform the content of an infographic in order to be most impactful at changing the end user's behavior. The purpose of this study was to identify a typical sample of infographics and determine the extent to which HBT was integrated. Drawing upon previous research of health technologies we expected HBT to be only marginally represented. Secondarily, we explored the relationship between various pin characteristics and the inclusion of HBT. Whereas there was no previous research to guide this analysis, we provided these data to assist in formulating future research questions and for practitioners wishing to identify the infographics most likely to contain HBT.

Methods

Overview

The study protocols met the exemption criteria of the university's institutional review board. No human subjects were involved in this study and only existing, publicly available data were collected for analysis.

This content analysis evaluated HBT in healthy eating infographic pins selected from Pinterest. Two public health graduate students trained in HBT and nutrition coded the pins to evaluate the extent to which constructs of the health belief model and social cognitive theory, 2 major theories of behavior change, were included. KJ, JLW, and KS are graduate students studying HBT. As it relates to nutritional qualifications, KJ is a registered dietitian nutritionist and was responsible for the nutrition-specific codes in the instrument. JLW has a BS in nutritional science, and KS works professionally as a nutrition coach.

Pin Selection

The sample was collected from Pinterest in September and October 2015. The study authors created a new Pinterest user account so that no search history would influence the search results. The study sample was identified using the following terms independently entered into the main Pinterest search bar: "nutrition infographic" and "healthy eating infographic." The first 250 eligible pins that were returned for each term were saved for analysis, resulting in 500 initial pins. Eligible pins included English-language pins that addressed some aspect of nutrition or healthy eating. Examples of this included pins that

explained the benefits of vitamins and minerals or provided tips for a healthy diet. Examples of ineligible pins included those addressing physical fitness or the use of food as a beauty product. An Excel spreadsheet (Microsoft Corp) was used to save the URL for each pin and to record the number of repins, likes, and comments. Finally, duplicate pins were deleted resulting in a final sample of 238.

The methodology for identifying pins was adapted from previous eHealth content analysis research [23,40-42]. Whereas previous content analyses of pins identified study samples through boards and then filtered samples by board popularity [29], it was not feasible in the current study because there are too many diet and nutrition boards available on Pinterest and they cannot be sorted by popularity.

Coding Procedures and Measurements

The researchers coded each pin in the study sample using an HBT-based coding rubric adapted from previous content analysis studies [28,40]. The rubric was adapted to be relevant for pins.

JHW is a senior health communications researcher and trained the other study authors in content analysis research during biweekly sessions over the course of 4 months. First, all authors met to define and reach a common understanding of all the study variables. Second, a coding instrument was developed and pilot tested by jointly coding pins and then resolving any discrepancies in codes. As part of the biweekly training sessions and in response to discrepancies in codes, which were either resolved until agreement was reached or they were removed from the instrument, JHW trained the other study authors to be able to identify the HBT constructs that were measured in this study. Third, the authors revised the instrument to remove coding options for which agreement could not be achieved or that were otherwise determined to be not applicable. Lastly, interrater reliability was established using a subset of the study sample.

The data were coded into an electronic spreadsheet and then exported for analyses. The coding rubric included 4 primary categories: (1) pin characteristics, (2) visual communication tools, (3) health belief model constructs, and (4) social cognitive theory constructs.

Pin characteristics included the URL and the number of repins, likes, and comments for each pin. The variables repins, likes, and comments were not normally distributed and underwent square root transformations for analyses. Pin affiliation/author (business, government, or individual) and the pin category of healthy eating (macronutrient, micronutrient, disease management, portion control/weight management, other) were recorded. Each pin was also coded with respect to whether or not it addressed a particular diet trend, such as paleo (eating only whole unprocessed foods) or avoiding genetically modified foods. Lastly, each pin was assessed to determine whether or not the coder would recommend the pin for use in promoting healthy eating. The coders were uniquely qualified to evaluate this aspect of pins.

The visual communication aspects of each pin were coded. These characteristics included the amount of text in the pin (no text, text light [covering <50% of the pin], or text heavy [covering >50% of the pin]); whether or not there was a person depicted in the pin (yes/no, and if yes, whether or not it was a photo of a real person or a cartoon); whether or not food was depicted in the pin (yes/no, and if yes, whether or not it was a picture of real food or a cartoon); and finally, the dominant colors of the pin (vibrant colors, muted colors, or black and white).

Constructs from social cognitive theory and health belief model were coded. Each HBT construct was coded as present or not present (yes=1, no=0).

Analysis

To ensure intercoder reliability, 2 coders evaluated a common 5% (26/500) of the study sample, which is considered adequate in cases of a large sample [43]. A Cohen's kappa coefficient of .6 was calculated, which is categorized as good agreement and is an acceptable level of intercoder reliability [44].

All analyses were conducted using Stata version 12 (StataCorp). Descriptive statistics were computed and summarized in aggregate. A total theory score was calculated by summing the values for each of the 9 construct codes (range 0-9). Combining multiple constructs to form a total theory score has been done previously [45-47] and stems from the notion that greater total amounts of theory may be most effective at changing complex diet-related behaviors [39]. Total theory scores were not normally distributed and a square root transformation was used to normalize this variable. Adjusted regression analysis was used to identify factors associated with the inclusion of HBT in pins.

Results

Sample Characteristics

A total of 238 infographic pins from Pinterest were analyzed. Characteristics of the pins are described in Table 1. Of these pins, 98.7% of the infographics were affiliated with a business or individual; the remainder were affiliated with a government organization. Healthy eating infographics related to either disease management or portion control constituted 32.6% (77/236) of the pins. The majority of pins were focused on very specific nutrition topics, such as lists of superfoods, health foods for babies, and nutritional content of specific fruits and vegetables. Only 33.8% (80/237) of the infographics depicted a person, and in 53.4% (126/236) of the infographics, food was depicted using cartoon figures. Under half (100/229, 43.7%) of the infographics cited references, and 34.8% (82/236) of the infographics received a recommendation by nutrition professionals. Other descriptive information is presented in Table 1. Engagement metrics, including the average number of comments, repins, and likes, are presented in Table 2.

Table 1. Pin characteristics.

Variables	n^a (%)
Pin affiliation[b]	
Business	198 (83.2)
Individual	3 (15.6)
Government	37 (1.3)
Pin category	
Macronutrients[c]	25 (10.6)
Micronutrients[d]	36 (15.3)
Disease management[e]	30 (12.7)
Portion control/weight management[f]	47 (19.9)
Other[g]	98 (41.5)
Text[h]	
Text heavy (>50%)	109 (46.2)
Text light (<50%)	127 (53.8)
Color[i]	
Vibrant	155 (65.7)
Muted	78 (33.1)
Black and white	1 (0.4)
Person depicted	
No person depicted	157 (66.2)
Cartoon	66 (27.9)
Photograph	14 (5.9)
Food Depicted	
No food depicted	15 (6.4)
Cartoon	126 (53.4)
Photograph	95 (40.3)
References[j]	100 (43.7)
Professional recommendation[k]	82 (34.8)

[a]Not all categories in every variable will sum to 238 due to some instances of missing data.

[b]Pin affiliation: who authored the pin.

[c]Macronutrients: carbohydrates, proteins, fats.

[d]Micronutrients: vitamins and minerals.

[e]Disease management: cancer, obesity, arthritis, cardiovascular disease, etc.

[f]Portion control/weight management: identifying and promoting healthy portion sizes.

[g]Other: included highly specific nutrition topics including top 10 healthiest foods, lists of superfoods, harms of specific foods, and nutritional content of specific fruits and vegetables.

[h]Text: pins were categorized as text heavy if more than 50% of the infographic contained text and text light if less than 50% of the infographic contained text.

[i]Color: pins were categorized as vibrant colored, muted in color, or black and white.

[j]References: presence of references or citations.

[k]Professional recommendation: whether or not the pin provided accurate information and advice to recommend to clients.

Table 2. Engagement metrics.[a]

Engagement metrics	Average	SD	Range
Comments	2.35	4.37	0-30
Repins	1318	2267	3-15,848
Likes	215	440	0-5027

[a]Variables were not normally distributed and underwent transformations for analyses. The mean and SD are presented here in raw form for interpretation.

Presence of Specific Health Behavior Theory Constructs

The prevalence of each construct [48] is presented in Table 3. Perceived benefits were present most often (170/236, 72%), followed by behavioral capability (123/238, 51.7%) and perceived severity (79/236, 33.5%). The construct that appeared the least was self-regulation/self-control (2/237, 0.8%).

Health Behavior Theory Scores

Table 4 illustrates a summary of HBT scores according to pin affiliation. The mean total theory score was 2.03 (SD 1.2) out of a possible 9. Pins authored by an individual had the lowest mean score at 1.67, pins from nongovernment businesses had an average score of 2.09, and pins created by government organizations had the highest average score at 2.75. The pin with the highest total theory score received an HBT score of 6.

Engagement Metrics and Health Behavior Theory

The number of repins (P=.04), likes (P=.01), and comments (P=.01) were positively associated with the inclusion of HBT. Including HBT constructs in an infographic was associated with greater user interaction through repins, likes, and comments (Table 5).

Table 3. Prevalence of health behavior theory constructs among pins.

Construct	Description [48]	n (%)
Perceived benefits	Belief about the potential positive aspects of a health action	170 (72.0)
Perceived barriers	Belief about the potential negative aspects of a health action	37 (15.6)
Perceived susceptibility	Belief about getting a disease or condition	38 (16.1)
Perceived severity	Belief about the seriousness of a condition or the consequences of leaving it un-addressed	79 (33.5)
Self-efficacy	Belief that one can achieve the behavior required to execute the outcome	11 (4.7)
Self-regulation/control	Controlling oneself through self-monitoring, goal-setting, feedback, self-reward, self-instruction, and enlistment of social support	2 (0.8)
Behavioral capability	Providing tools, resources, or environmental changes that make new behaviors easier to perform	123 (51.7)
Observational learning/modeling	Beliefs based on observing similar individuals or role models perform a new behavior	3 (1.3)
Subjective norm	An individual's perception of social norms or his/her peers' beliefs about a behavior. A function of an individual's normative beliefs and motivation to comply with beliefs	17 (7.2)

Table 4. Summary of health behavior theory scores.[a]

	Mean	SD	Frequency
Overall	2.03	1.20	238
Business	2.09	1.21	189
Government	2.67	0.577	3
Individual	1.67	1.12	36

[a]Variables were not normally distributed and underwent transformations for analyses. The mean and SD are presented here in raw form for interpretation.

XSL•FO

RenderX

Table 5. Correlation between pin engagement metrics and health behavior theory score.

	Total theory		Likes		Repins	
	Coefficient	*P* value	Coefficient	*P* value	Coefficient	*P* value
Repins	0.13	.04	0.95	<.001	1.00	—
Likes	0.16	.01	1.00	—	0.95	<.001
Comments	0.17	.01	0.66	<.001	0.64	<.001

Pin Characteristic and Health Behavior Theory

Shown in Table 6, several characteristics were positively associated with a higher HBT score in adjusted regression analyses after controlling for other factors. Compared to a small amount of text, a large amount of text (*P*=.01) was associated with higher HBT score, as were photographs of real people

(*P*=.001) (compared to no photographs of real people), cartoon pictures of food (*P*=.01) (compared to photographs of real food), and the presence of references (*P*=.001) (compared to no references). Lastly, when compared to infographics that were not coded as recommendable for professional use, infographics that received this professional recommendation had higher HBT scores (*P*=.001).

Table 6. Regression analysis of pin characteristic and presence of health behavior theory.

Theory square root	Coefficient	SE	*t*	*P* value	95% CI
Professional recommendation	.197	.0577	3.4	.001	.0827 to .3104
Text heavy	.141	.0547	2.58	.011	.0334 to .2491
Color	−.110	.0578	−1.89	.060	−.2237 to .0044
Photograph of person	.196	.0581	3.38	.001	.0816 to .3107
Cartoon of food	.132	.0551	2.40	.017	.0235 to .2407
References	.184	.0567	3.25	.001	.0727 to .2964

Discussion

Principal Findings

The purpose of this study was to determine the level of integration of HBT in nutrition infographics on Pinterest. Secondarily, we identified factors associated with its inclusion. The secondary interest was largely exploratory and was intended to inform future research and efforts that might use Pinterest to promote healthy eating. Overall, it was found that HBT constructs are integrated into nutrition pins only minimally. The low levels of HBT are not surprising, as pins can be created and shared by lay parties who may not have training in HBT or know why its inclusion may improve pin impact on behavior. This may be a recurring problem for health professionals wishing to adopt and use social media and other health technology in their health promotion efforts. Payne et al [40] and Cowan et al [45] reported similarly low levels of HBT in their respective content analyses of physical activity mobile apps and recommended increased collaboration between experts in HBT and app developers. Likewise, it may be beneficial for health professionals to partner with infographic developers to create pins. There is a rich literature base supporting the connection between HBT and healthy dietary practices [33-39]. Future research could test the impact on behavior provided that practitioners and developers were to work collaboratively to accomplish a more complete integration of HBT.

Interestingly, there were very few pins created by government organizations such as Let's Move! and the US Department of Agriculture, although they had the highest average HBT scores. It seems plausible that pins created by government health

organizations would contain higher levels of HBT because individuals working in these settings may be more likely to have training in a related discipline. In general, however, these results may indicate that businesses and individuals who create nutrition infographics for Pinterest lack the training to effectively incorporate HBT into social media campaigns. While research elsewhere indicates that health professionals currently use and understand social media in vocational roles only minimally [47,49,50], the creation of accurate nutrition infographics that include HBT by health professionals may allow Pinterest to be used as an effective health promotion tool.

The majority of pins were focused on very specific nutrition topics, such as lists of superfoods, health foods for babies, and nutritional content of specific fruits and vegetables. The second most common infographic category was portion control and weight management. It is promising that the latter category was a relatively large percentage of the sample, especially considering that portion sizes have dramatically increased over the past several decades [51] and this trend is a major contributor to the global obesity epidemic [52]. However, the high percentage of very specific nutrition topics (eg, the nutritional content of a banana or a list of healthy foods for babies) not necessarily related to preventive health behaviors such as portion control may be problematic, as providing highly specific nutrition information may not be an effective way to change health outcomes.

While Pinterest is dominated by young middle- to upper-middle-class white females, Pinterest use grew significantly among individuals living in rural locations, those with an annual salary of less than $30,000 per year, and those

aged 50 years and older from 2013 to 2014 [53]. These populations are at significant risk for poor dietary behaviors and obesity in the United States [54,55], and it is not clear to what extent current infographics appeal to these populations. The lack of references incorporated in this sample is also troubling, especially as infographics containing references were more likely to contain HBT. It may also be that pins without citations are less likely to be evidence-based, an important attribute in public health interventions [56]. Health organizations creating infographics for Pinterest may want to consider tailoring pins for high-risk populations and incorporating information from reliable references to improve the accuracy of nutrition information. Including references at the bottom of infographics in a smaller font can improve credibility without hindering visual appeal.

The most common HBT constructs in this sample were perceived benefits, behavioral capability, and perceived severity. The remaining 6 constructs appeared in 15% or less of the sample, with self-regulation and control being the least common. It is concerning that self-efficacy was one of the least incorporated constructs because self-efficacy has been shown to be a significant predictor of behavior change [48] and is positively associated with better chronic disease management [29,57]. It is unsurprising that the most common HBT constructs were related to disseminating knowledge and general information; pins are largely noninteractive, especially compared to other health technology tools (apps, videos, etc), and it may be difficult to incorporate more complex HBT constructs. Public health researchers debate whether education-heavy health interventions are effective at changing behavior [58]. However, some researchers indicate that there is an association between increased nutrition knowledge and dietary behaviors, including eating smaller portions, eating foods with fewer calories, and using nutritional labels more effectively [59-61]. Additionally, given that 84% of those who access Pinterest daily report being inspired to try something new once a week or more, health information distributed through Pinterest may have the potential to influence behavior [17]. As theory-based interventions are reported to be more successful at changing behavior than those that do not contain HBT [30], Pinterest infographics containing HBT may prove to be an effective health promotion tool, although further research on the impact of infographics to change health behaviors is needed.

Factors positively associated with HBT included the presence of heavy amounts of text, a photograph of a real person, or a cartoon image of food. This suggests that depicting HBT constructs on Pinterest is more likely to be accomplished by incorporating a combination of text, people, and images rather than text alone. Indeed, research has demonstrated that images, rather than text, are the most desirable way to communicate information [62]. Less than half of nutrition pins analyzed depicted a person and even fewer included a photograph of a real person. This is not ideal, considering that realistic pictures that resonate with users are more effective at disseminating health information [29]. Including more realistic images of people engaging in healthy eating behaviors in infographics

may be more likely to promote behavior change. A study with this focus might be warranted to test the impact on behavior. These findings might also guide a practitioner's selection of images for social distribution if they perceive that such pins are more likely to include HBT.

Social media engagement is a key performance indicator that links social media usage to action [63] and can be categorized as low, medium, or high engagement. In the context of Pinterest, the number of "likes" on each pin can be considered low engagement (ie, users acknowledge or agree with content), while repins and comments can be considered medium engagement (ie, users create or share content) [29]. High engagement refers to actual offline participation and is not measurable by Pinterest engagement metrics. Although engagement with nutrition pins in the form of comments was relatively minimal in this study, findings suggest that the commenting on, repinning, and liking of nutrition-related infographics is more likely when HBT constructs are depicted. Future research could test the impact of including HBT in pins on the likelihood of offline action.

Limitations

The researchers only assessed the image and did not evaluate the websites the pin linked to. Additionally, there are many duplicates of each pin shared on Pinterest, so engagement metrics may be spread among pins. While the researchers made note of information accuracy, it was not the primary focus of the study. A more rigorous analysis of content accuracy could be done to determine accuracy of messages.

Conclusions

Promoting dietary modifications is an important public health strategy for preventing chronic disease. Pinterest as a social media platform has the potential to communicate health information and influence healthy eating behavior through infographics. However, current Pinterest infographics targeting healthy eating contain few HBT elements.

It is recommended that health professionals and organizations create and disseminate infographics that contain more elements of HBT to better influence healthy eating behavior and be more effective in changing behavior [30-32]. Pins should be tailored for high-risk populations and incorporate information from reliable references. This may be accomplished by including a combination of images and texts and portraying HBT constructs. In so doing, individuals (or the public) may obtain the information and skills necessary to eat healthy and prevent chronic disease. Including HBT in pins could foster increased user engagement and result in a greater likelihood of offline action.

This study may also help dietitians, public health workers, and health educators make an informed decision about whether or not to recommend Pinterest as a health information source. Because many health professionals use infographics to disseminate health information [18,19], they should take into account the accuracy of infographics on Pinterest and have a realistic expectation about whether or not pins are effective in producing behavior change.

Conflicts of Interest

None declared.

References

1. Heron M, Hoyert D, Murphy S, Xu J, Kochanek K, Tejada-Vera B. Deaths: final data for 2006. Centers For Disease Control And Prevention, National Center For Health Statistics, National Vital Statistics System. 2009 Apr 17. URL: https://www.cdc.gov/nchs/data/nvsr/nvsr57/nvsr57_14.pdf [accessed 2016-11-30] [WebCite Cache ID 6j4vGaAZH]

2. World Health Organization. Chronic Diseases and Health Promotion. 2015. URL: http://www.who.int/chp/en/ [accessed 2016-07-17] [WebCite Cache ID 6j4vO43JL]

3. Bauer UE, Briss PA, Goodman RA, Bowman BA. Prevention of chronic disease in the 21st century: elimination of the leading preventable causes of premature death and disability in the USA. Lancet 2014 Jul 5;384(9937):45-52. [doi: 10.1016/S0140-6736(14)60648-6] [Medline: 24996589]

4. Putnam J, Allshouse J, Kantor L. US per capita food supply trends: more calories, refined carbohydrates, and fats. Food Review. 2002. URL: http://ageconsearch.umn.edu/bitstream/234624/2/frvol25i3a.pdf [accessed 2016-11-30] [WebCite Cache ID 6j4vSiQud]

5. Grotto D, Zied E. The Standard American Diet and its relationship to the health status of Americans. Nutr Clin Pract 2010 Dec;25(6):603-612. [doi: 10.1177/0884533610386234] [Medline: 21139124]

6. Johnson RK, Appel LJ, Brands M, Howard BV, Lefevre M, Lustig RH, American Heart Association Nutrition Committee of the Council on Nutrition, Physical Activity, and Metabolism and the Council on EpidemiologyPrevention. Dietary sugars intake and cardiovascular health: a scientific statement from the American Heart Association. Circulation 2009 Sep 15;120(11):1011-1020 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.109.192627] [Medline: 19704096]

7. Center for Chronic Disease Prevention and Health Promotion. State Indicator Report on Fruits and Vegetables. 2013. URL: http://www.cdc.gov/nutrition/downloads/state-indicator-report-fruits-vegetables-2013.pdf [accessed 2016-07-17] [WebCite Cache ID 6j4veWb2g]

8. Moore LV, Thompson FE. Adults Meeting Fruit and Vegetable Intake Recommendations—United States, 2013. Atlanta, GA: Centers for Disease Control and Prevention; 2015 Jul 10. URL: http://www.cdc.gov/mmwr/preview/mmwrhtml/mm6426a1.htm [accessed 2016-11-30] [WebCite Cache ID 6j4w8BG1P]

9. Bazzano LA, Serdula MK, Liu S. Dietary intake of fruits and vegetables and risk of cardiovascular disease. Curr Atheroscler Rep 2003 Nov;5(6):492-499. [Medline: 14525683]

10. Guenther P, Fungwe T, Lucas R, Hiza H, Juan W, Lino M. Diet Quality of Low-Income and Higher Income Americans in 2003-04 as Measured by the Healthy Eating Index—2005. US Department of Agriculture. 2008 Dec.. URL: http://www.cnpp.usda.gov/sites/default/files/nutrition_insights_uploads/Insight42.pdf [accessed 2016-11-30] [WebCite Cache ID 6j4wFi0WW]

11. Trichopoulou A, Vasilopoulou E. Mediterranean diet and longevity. Br J Nutr 2000 Dec;84 Suppl 2:S205-S209. [Medline: 11242471]

12. Philippou C, Andreou E, Menelaou N, Hajigeorgiou P, Papandreou D. Effects of diet and exercise in 337 overweight/obese adults. Hippokratia 2012 Jan;16(1):46-50 [FREE Full text] [Medline: 23930057]

13. Frenn M, Malin S, Brown RL, Greer Y, Fox J, Greer J, et al. Changing the tide: an Internet/video exercise and low-fat diet intervention with middle-school students. Appl Nurs Res 2005 Feb;18(1):13-21. [doi: 10.1016/j.apnr.2004.04.003] [Medline: 15812731]

14. Williams G, Hamm MP, Shulhan J, Vandermeer B, Hartling L. Social media interventions for diet and exercise behaviours: a systematic review and meta-analysis of randomised controlled trials. BMJ Open 2014 Feb;4(2):e003926 [FREE Full text] [doi: 10.1136/bmjopen-2013-003926] [Medline: 24525388]

15. Moorhead SA, Hazlett DE, Harrison L, Carroll JK, Irwin A, Hoving C. A new dimension of health care: systematic review of the uses, benefits, and limitations of social media for health communication. J Med Internet Res 2013;15(4):e85 [FREE Full text] [doi: 10.2196/jmir.1933] [Medline: 23615206]

16. Korda H, Itani Z. Harnessing social media for health promotion and behavior change. Health Promot Pract 2013 Jan;14(1):15-23. [doi: 10.1177/1524839911405850] [Medline: 21558472]

17. Ahalogy. Pinterest Media Consumption Study. 2015. URL: https://www.ahalogy.com/blog/2015-pinterest-media-consumption-study [accessed 2016-07-07] [WebCite Cache ID 6ipYuiwQl]

18. Kibar PN, Akkoyunlu B. A new approach to equip students with visual literacy skills: Use of infographics in education. 2014 Presented at: European Conference on Information Literacy; 2014; Dubrovnik, Croatia p. 456-465.

19. Late M. Graphic shows how public health saves lives, costs. Nation's Health 2012;42(9).

20. Stones C, Gent M. "If The Guardian can do it, we should be able to do it!": examining public health infographic strategies used by public health professionals. 2015 Presented at: Proceedings of the 3rd European Conference on Design4Health; 2015; Sheffield, England.

21. Hansen K, Nowlan G, Winter C. Pinterest as a tool: applications in academic libraries and higher education. Partnership 2012 Dec 05;7(2). [doi: 10.21083/partnership.v7i2.2011]

XSL·FO

RenderX

22. Turck CJ, Silva MA, Tremblay SR, Sachse SL. A preliminary study of health care professionals preferences for infographics versus conventional abstracts for communicating the results of clinical research. J Cont Educ Health Prof 2014;34:S36-S38. [doi: 10.1002/chp.21232]

23. Lister C, Royne M, Payne HE, Cannon B, Hanson C, Barnes M. The laugh model: reframing and rebranding public health through social media. Am J Public Health 2015 Nov;105(11):2245-2251. [doi: 10.2105/AJPH.2015.302669] [Medline: 26378824]

24. Davis County Health Department. Healthy info from Davis Co. 2015. URL: https://www.pinterest.com/daviscohealth/ [accessed 2016-07-07] [WebCite Cache ID 6ipYxfjbd]

25. Hirasawa R, Yachi Y, Yoshizawa S, Horikawa C, Heianza Y, Sugawara A, et al. Quality and accuracy of Internet information concerning a healthy diet. Int J Food Sci Nutr 2013 Dec;64(8):1007-1013. [doi: 10.3109/09637486.2013.812620] [Medline: 23863089]

26. De la Torre-Díez I, Díaz-Pernas FJ, Antón-Rodríguez M. A content analysis of chronic diseases social groups on Facebook and Twitter. Telemed J E Health 2012;18(6):404-408. [doi: 10.1089/tmj.2011.0227] [Medline: 22650380]

27. Lister C, Brutsch E, Johnson A, Boyer C, Hall P, West J. It gets better: a content analysis of health behavior theory in anti-bullying YouTube videos. Int J Health 2013;1(2). [doi: 10.14419/ijh.v1i2.1002]

28. West JH, Lister C, Perry JA, Church JL, Vance DL. Share videos, not pills: a content analysis of prescription drug videos on YouTube for presence of behavior change theory. Health 2014;06(11):1255-1262. [doi: 10.4236/health.2014.611154]

29. Paige SR, Stellefson M, Chaney BH, Alber JM. Pinterest as a resource for health information on chronic obstructive pulmonary disease (COPD): a social media content analysis. Am J Health Educ 2015 Jul 06;46(4):241-251. [doi: 10.1080/19325037.2015.1044586]

30. Glanz K, Lewis FM, Rimer BK. Health Behavior and Health Education: Theory, Research, and Practice. San Francisco: Jossey-Bass; 1997.

31. Lopez LM, Tolley EE, Grimes DA, Chen-Mok M. Theory-based interventions for contraception. Cochrane Database Syst Rev 2011;CD007249(3). [doi: 10.1002/14651858.CD007249.pub3] [Medline: 21412901]

32. Noar SM, Black HG, Pierce LB. Efficacy of computer technology-based HIV prevention interventions: a meta-analysis. AIDS 2009 Jan 2;23(1):107-115. [doi: 10.1097/QAD.0b013e32831c5500] [Medline: 19050392]

33. Anderson ES, Winett RA, Wojcik JR. Self-regulation, self-efficacy, outcome expectations, and social support: social cognitive theory and nutrition behavior. Ann Behav Med 2007;34(3):304-312. [doi: 10.1080/08836610701677659] [Medline: 18020940]

34. Brebbia CA, Miralles i Garcia JL, Lee KA, Bai Y, Wunderlich SM. A pilot study for plant the seed: a nutrition education program using local food environment to put theory into action. Int J SDP 2016;11(6):1028-1036. [doi: 10.2495/SDP-V11-N6-1028-1036]

35. Muturi NW, Kidd T, Khan T, Kattelmann K, Zies S, Lindshield E, et al. An examination of factors associated with self-efficacy for food choice and healthy eating among low-income adolescents in three U.S. states. Front Commun 2016 Sep 21;1. [doi: 10.3389/fcomm.2016.00006]

36. Deshpande S, Basil M, Basil D. Factors influencing healthy eating habits among college students: an application of the health belief model. Health Mark Q 2009;26(2):145-164. [doi: 10.1080/07359680802619834] [Medline: 19408181]

37. Abood DA, Black DR, Feral D. Nutrition education worksite intervention for university staff: application of the health belief model. J Nutr Educ Behav 2003;35(5):260-267. [Medline: 14521826]

38. Gillibrand R, Stevenson J. The extended health belief model applied to the experience of diabetes in young people. Br J Health Psychol 2006 Feb;11(Pt 1):155-169. [doi: 10.1348/135910705X39485] [Medline: 16480561]

39. Achterberg C, Miller C. Is one theory better than another in nutrition education? A viewpoint: more is better. J Nutr Educ Behav 2004;36(1):40-42. [Medline: 14756981]

40. Payne HE, Moxley VB, MacDonald E. Health behavior theory in physical activity game apps: a content analysis. JMIR Serious Games 2015;3(2):e4 [FREE Full text] [doi: 10.2196/games.4187] [Medline: 26168926]

41. Briones R, Nan X, Madden K, Waks L. When vaccines go viral: an analysis of HPV vaccine coverage on YouTube. Health Commun 2012;27(5):478-485. [doi: 10.1080/10410236.2011.610258] [Medline: 22029723]

42. Bromberg JE, Augustson EM, Backinger CL. Portrayal of smokeless tobacco in YouTube videos. Nicotine Tob Res 2012 Apr;14(4):455-462 [FREE Full text] [doi: 10.1093/ntr/ntr235] [Medline: 22080585]

43. Kaid L, Wadsworth A. Content analysis. In: Measurement of Communication Behavior. New York: Longman; 1989.

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

45. 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]

46. West JH, Hall PC, Arredondo V, Berrett B, Guerra B, Farrell J. Health behavior theories in diet apps. J Consumer Health Internet 2013 Jan;17(1):10-24. [doi: 10.1080/15398285.2013.756343]

47. Victor Arredondo HP. Use and acceptance of social media among community health workers. J Community Med Health Educ 2015;05(03). [doi: 10.4172/2161-0711.1000354]

48.   Glanz K, Rimer B, Viswanath K. Health behavior and health education: theory, research, and practice. San Francisco: Jossey-Bass; 2008.

49.   Keller B, Labrique A, Jain KM, Pekosz A, Levine O. Mind the gap: social media engagement by public health researchers. J Med Internet Res 2014;16(1):e8 [FREE Full text] [doi: 10.2196/jmir.2982] [Medline: 24425670]

50.   Thackeray R, Neiger BL, Smith AK, Van Wagenen SB. Adoption and use of social media among public health departments. BMC Public Health 2012;12:242 [FREE Full text] [doi: 10.1186/1471-2458-12-242] [Medline: 22449137]

51.   Nielsen SJ, Popkin BM. Patterns and trends in food portion sizes, 1977-1998. JAMA 2003;289(4):450-453. [Medline: 12533124]

52.   Young LR, Nestle M. The contribution of expanding portion sizes to the US obesity epidemic. Am J Public Health 2002 Feb;92(2):246-249. [Medline: 11818300]

53.   Duggan M, Ellison N, Lampe C, Lenhart A, Madden M. Demographics of key social networking platforms. Washington, DC: Pew Internet and American Life Project; 2015 Jan 09 URL: http://www.pewinternet.org/files/2015/01/PI_SocialMediaUpdate20144.pdf [accessed 2016-11-30] [WebCite Cache ID 6j4wmQRh2]

54.   Befort CA, Nazir N, Perri MG. Prevalence of obesity among adults from rural and urban areas of the United States: findings from NHANES (2005-2008). J Rural Health 2012;28(4):392-397 [FREE Full text] [doi: 10.1111/j.1748-0361.2012.00411.x] [Medline: 23083085]

55.   Ogden CL, Carroll MD, Fryar CD, Flegal KM. Prevalence of obesity among adults and youth: United States, 2011-2014. NCHS Data Brief 2015 Nov(219):1-8 [FREE Full text] [Medline: 26633046]

56.   Brownson R, Fielding JE, Maylahn CM. Evidence-based public health: a fundamental concept for public health practice. Annu Rev Public Health 2009;30:175-201. [doi: 10.1146/annurev.pu.30.031709.100001] [Medline: 19296775]

57.   Farrell K, Wicks MN, Martin JC. Chronic disease self-management improved with enhanced self-efficacy. Clin Nurs Res 2004 Nov;13(4):289-308. [doi: 10.1177/1054773804267878] [Medline: 15448281]

58.   Tones K, Tilford S. Health Education: Effectiveness, Efficiency and Equity. London: Chapman & Hall; 1994.

59.   Laz TH, Rahman M, Pohlmeier AM, Berenson AB. Level of nutrition knowledge and its association with weight loss behaviors among low-income reproductive-age women. J Community Health 2015 Jun;40(3):542-548 [FREE Full text] [doi: 10.1007/s10900-014-9969-9] [Medline: 25394404]

60.   Bonaccio M, Costanzo S, Olivieri M, Donati MB, Iacoviello L, Bonanni A, et al. Nutrition knowledge is associated with higher adherence to Mediterranean diet and lower prevalence of obesity. Results from the Moli-sani study. Appetite 2013 Sep;68:139-146. [doi: 10.1016/j.appet.2013.04.026] [Medline: 23665233]

61.   Miller LMS, Cassady DL. The effects of nutrition knowledge on food label use. A review of the literature. Appetite 2015 Sep;92:207-216 [FREE Full text] [doi: 10.1016/j.appet.2015.05.029] [Medline: 26025086]

62.   Katz MG, Kripalani S, Weiss BD. Use of pictorial aids in medication instructions: a review of the literature. Am J Health Syst Pharm 2006 Dec 1;63(23):2391-2397. [doi: 10.2146/ajhp060162] [Medline: 17106013]

63.   Neiger BL, Thackeray R, Van Wagenen SA, Hanson CL, West JH, Barnes MD, et al. Use of social media in health promotion: purposes, key performance indicators, and evaluation metrics. Health Promot Pract 2012 Mar;13(2):159-164. [doi: 10.1177/1524839911433467] [Medline: 22382491]

Abbreviations

COPD: chronic obstructive pulmonary disease
HBT: health behavior theory

XSL•FO
RenderX

XSL•FO

RenderX

Original Paper

Design and Testing of the Safety Agenda Mobile App for Managing Health Care Managers' Patient Safety Responsibilities

José Joaquín Mira[1,2], PhD; Irene Carrillo[2], M Psych (Clin); Cesar Fernandez[3], PhD; Maria Asuncion Vicente[3], PhD; Mercedes Guilabert[2], MPH, PhD

[1]Alicante-Sant Joan Health District, Consellería Sanitat, Alicante, Spain

[2]Health Psychology Department, Miguel Hernández University, Elche, Spain

[3]Systems Engineering and Automation Department, Miguel Hernández University, Elche, Spain

Corresponding Author:
José Joaquín Mira, PhD
Alicante-Sant Joan Health District, Consellería Sanitat
Hospital-Plá Health Center c/ Hermanos López Osaba s/n
Alicante, 03013
Spain
Phone: 34 606433599
Fax: 34 966658984
Email: jose.mira@umh.es

Abstract

Background: Adverse events are a reality in clinical practice. Reducing the prevalence of preventable adverse events by stemming their causes requires health managers' engagement.

Objective: The objective of our study was to develop an app for mobile phones and tablets that would provide managers with an overview of their responsibilities in matters of patient safety and would help them manage interventions that are expected to be carried out throughout the year.

Methods: The Safety Agenda Mobile App (SAMA) was designed based on standardized regulations and reviews of studies about health managers' roles in patient safety. A total of 7 managers used a beta version of SAMA for 2 months and then they assessed and proposed improvements in its design. Their experience permitted redesigning SAMA, improving functions and navigation. A total of 74 Spanish health managers tried out the revised version of SAMA. After 4 months, their assessment was requested in a voluntary and anonymous manner.

Results: SAMA is an iOS app that includes 37 predefined tasks that are the responsibility of health managers. Health managers can adapt these tasks to their schedule, add new ones, and share them with their team. SAMA menus are structured in 4 main areas: information, registry, task list, and settings. Of the 74 users who tested SAMA, 64 (86%) users provided a positive assessment of SAMA characteristics and utility. Over an 11-month period, 238 users downloaded SAMA. This mobile app has obtained the AppSaludable (HealthyApp) Quality Seal.

Conclusions: SAMA includes a set of activities that are expected to be carried out by health managers in matters of patient safety and contributes toward improving the awareness of their responsibilities in matters of safety.

Introduction

Patient Safety

Patients do not expect to be harmed in the course of treatment to recover their health. Professionals also do not expect this result. However, over the course of administering health care, incidents do occur unexpectedly and involuntarily that cause harm to patients (adverse events) [1]. Latent errors within the organization and clinical errors are considered the most common causes of avoidable adverse events [2].

In the Organisation for Economic Co-operation and Development's member countries, 9% of hospitalized patients [1] suffer an adverse event, whereas studies in primary care

centers have identified a rate of adverse events of less than 2% for all consultations [3,4]. At hospitals and primary care centers, 18% and 7% of patients, respectively, experienced more than 1 adverse event [4,5]. Almost half of hospital adverse events were preventable [1]. In developing countries, the prevalence of adverse events in hospitals oscillates around 10.5%, whereas the prevalence of adverse events in ambulatory care is around 5% [6].

The total cost of a lack of patient safety (hospitals, primary care centers, and medication errors in patients) can reach 5.6% [7] of the total health expenditure. In Europe, the cost of adverse events in hospitals has been calculated to vary between 1.5% and 2.4% of the total health expenditure [8,9].

Adverse events in European countries cause 3.5 million disability-adjusted life-years every year, of which 1.5 million are likely due to preventable adverse events. These findings have recently been corroborated (TB Agbabiaka, unpublished data, 2016).

Responsibility of Managers on Patient Safety

Reducing the frequency of preventable adverse events is a challenge for health care organizations. However, sometime patients suffer an adverse event. When it occurs, healthcare professionals should seek: first, that the same patient does not suffer more than 1 adverse event over the course of treatment; and second, that the same adverse event is not repeated. To achieve this, the health manager's role is crucial, although the organizations must count on involvement by all their professionals [10]. Root cause analysis, critical incident analysis, and incident simulations are the most useful techniques for investigating what happened [11-13]. A Web-based tool named BACRA (based on root cause analysis, in Spanish Basado en Análisis Causa-RAíz) has been developed to involve frontline health care professionals and middle managers in implementing solutions to prevent recurrence of these incidents [14].

A culture of safety at health care centers is a factor that contributes to reduced risks for patients [15]. However, studies [11] have suggested that many times managers do not spend enough time and do not outline a clear hospital strategy on quality and safety. They do not exert enough effort to shape a positive safety culture, lack awareness of what activities they are expected to do, and do not always ensure an appropriate information system to help clinicians improve their practice [12].

Although health managers have the responsibility of advocating a proactive culture of patient safety and defining a safety framework in health care organizations, there are no available tools to assist with raising the awareness of managers with respect to their responsibilities and roles in patient safety.

The objective of our study was to develop an app for mobile phones and tablets that would provide managers at hospitals and primary care centers with an overview of their responsibilities in matters of patient safety and would help them manage and record the interventions that are expected to be carried out throughout a year in order to properly manage the risks inherent to the health care that patients receive.

Methods

SAMA development

Study design of an app to involve hospital managers and staff on an annual patient safety set of actions. Figure 1 shows the steps followed in the development of the Safety Agenda Mobile App (SAMA).

Information Sources

A literature review, the Spanish standard Una Norma Española (UNE) 179003:2013 of Risk Management for Patient Safety [13], and a qualitative consensus technique were used to collect information about managers' responsibilities on patient safety.

The narrative review of the scientific and gray literature identified a set of interventions (activities) on matters of safety that are the responsibility of managers and should be carried out throughout the year. To do this, first a search was conducted using MEDLINE for review studies on the role and responsibilities of managers in patient safety from 2000 to 2015 with a combination of the words *leader* and *managers* with the following descriptors: "patient safety" and "adverse events." Using the Google meta-search engine, websites that could provide further information on the role and responsibilities of managers with respect to patient safety were analyzed.

This review yielded a selection of studies as information sources. The study by White et al [16] was used to describe the role of managers in patient safety. Rodrigues et al [17] described the reach of the studies and the interventions in patient safety. Research by Clarke et al [18], White et al [19], and Parand et al [11] were the main sources of information for identifying the tasks and responsibilities of managers, along with studies by Van Gerven et al [20] and Mira et al [21] concerning the role of managers in reducing the impact of adverse events on professionals and on the very reputations of health care institutions.

Second, the Spanish standard UNE 179003:2013 of Risk Management for Patient Safety [13] describes a set of appropriate and expected patient safety guidelines that ensure that health care risks are effectively managed. The UNE 179003:2013 forces to deploy operating procedures aimed at reducing the incidence of adverse events from a risk management perspective. This standard was used to elaborate a list of the activities that managers are responsible for was completed.

Figure 1. Overview of the steps involved in Safety Agenda Mobile App (SAMA) design and development. UNE: Una Norma Española.
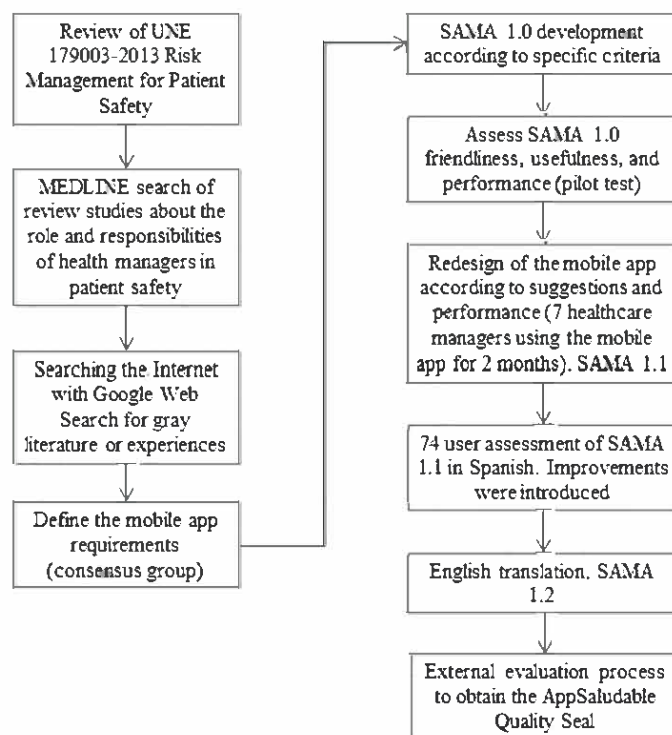
Third, a consensus group was formed in April 2015 with voluntary participation by 6 patient safety experts. They responded to what characteristics and functions a mobile app is expected to have with the aim of involving managers as leaders on matters of patient safety. The consensus group participants reviewed and prescribed the set of tasks managers are responsible for and also established the requirements that the mobile app should fulfill.

On the basis of all this information, version 1.0 (beta) of the mobile app was designed.

Assessment of the Safety Agenda Mobile App

To test the operation of SAMA 1.0 and compile ideas about improving this first version, a group of 7 managers from 6 regional health services in Spain were asked to use the app for 2 months. In this initial evaluation, they considered the ease of access and use, the suitability of its contents, the feasibility of the proposed activities, and usefulness of its different functions. This experience permitted redesigning SAMA, and resulted in version 1.1, with improvements implemented in its functions and in the navigation of various menus. This new version of the mobile app, once the improvements were made, was uploaded to the App Store and disseminated in health manager forums within Spain.

In June 2015, SAMA 1.1 became available at the App Store and interested users could download it [22]. This version included 4 nonmandatory questions about utility, strengths, and improvements needed. Four months subsequent to the SAMA launch, the system asked its users to express their opinions about the mobile app; they were also able to propose suggestions for improvement in terms of content, design, and navigation.

SAMA was evaluated externally to assure the appropriateness of the design, quality and safety information provided, confidentiality and privacy, and services provided by the App. This is part of the Quality and Safety Strategy for Spanish health mobile apps. The apps are blind reviewed in a set of 31 criteria before obtaining this AppSaludable (HealthyApp) Quality Seal. The HealthyApp Quality Seal is part of a validation process conducted by an international quality institution in Spain [23].

Results

SAMA characteristics

SAMA 1.0 was organized in general content blocks, called activities, and in a set of specific tasks. Table 1 shows the scope of expected SAMA activities. A total of 37 tasks were extracted from the information sources used. They are organized in 8 activities: (1) identification and analysis of risk management processes; (2) analysis of outcomes and risk management monitoring; (3) training actions in patient safety; (4) communication, information, and documentation; (5) consequences of adverse effects; (6) audits; (7) positive safety culture; and (8) management agreements on patient safety aims.

Table 1. Activities and subcategories in which the Safety Agenda Mobile App's tasks are organized (N=37).

Activities and subcategories	Tasks n (%)
Activity 1. Identification and analysis of risk management processes	6 (16)
Risk management processes	1 (17)
Resources	1 (17)
Study of adverse event frequency	1 (17)
Risk analysis	2 (33)
Clinical sessions	1 (17)
Activity 2. Analysis of outcomes and risk management at center	9 (24)
RCA[a] outcomes	1 (11)
Notification: safety committee	2 (22)
Notification: management committee	2 (22)
Anonymous notifications	3 (33)
Effectiveness tracking	1 (11)
Activity 3. Training actions in patient safety	6 (16)
Training revision	3 (50)
New staff incorporations	2 (33)
Effectiveness of training actions	1 (17)
Activity 4. Communication, information, and documentation	3 (8)
Communication plan	2 (67)
Documentation	1 (33)
Activity 5. Consequences of adverse events	7 (19)
Information to relatives	1 (14)
Victim compensation report	1 (14)
Insurance policy report	1 (14)
Impact on second victim	2 (29)
Impact on third victim	2 (29)
Activity 6. Audits	2 (5)
Revision and planning of the internal and external audit program	2 (100)
Activity 7. Positive safety culture	2 (5)
Foster safety culture within the organization	2 (100)
Activity 8. Management agreements on patient safety aims	2 (5)
Management agreements	2 (100)

[a]RCA: root cause analysis.

The recommended periodicity of most predefined tasks was annual (n=28), but there were some that were recommended to be executed every 3 (n=1), 4 (n=1), or 6 months (n=7).

The requirements to be fulfilled by SAMA were defined by the consensus group and included the following:

- A list of activities that managers at health centers are responsible for, based on the UNE 179003:2013 standard and criteria suggested by scientific documents and expert participants.

- The possibility of personalizing the activities throughout the year, adapting all interventions to each user's work plan.
- Associating text to each activity as a way to remind, summarize tasks, etc.
- A report of activities generated automatically, which permits monitoring the degree of compliance of the activities for which managers are responsible.
- User-friendly interface. If it is not, users will not continue using the app.
- Highly configurable. Not all hospitals and not all managers follow the same procedures on safety issues.

- High level of privacy. Data entered into the app are extremely sensitive, so they should be kept locally within the device and never uploaded to external servers.
- Access by log-in and password to keep data secure.
- Capability of gathering (anonymous, nonsensitive) data via surveys. Concretely, an initial survey before using the app for the first time and a second survey after several months of use.

Safety Agenda Mobile App Design and Redesigns

After analyzing the 2 main mobile platforms (Android and iOS), iOS was selected because of various reasons: first, the iPhone is the most common device possessed by prospective app users (management staff); second, fragmentation is much lower in iOS and, thus, it is easier to ensure that the app will work properly on all devices. Besides, the operating system user update rate is higher compared with Android, so there is no need to develop an app capable of working properly on old versions of the operating system. Finally, all iOS apps follow a strict review process by Apple (this is not the case for Android apps). This review process assures app quality indirectly (ie, there are no memory losses, etc).

After the pilot tests by initial users (N=7), the following modifications were suggested: first, more customization options, such as the capability of creating new tasks and groups of tasks other than those recommended by the UNE standard and that

are shown by default in the app; second, a way for the user to add text notes to certain tasks; third, functionality for backing up data; and, finally, slight changes in usability (mainly for task scheduling). All suggested modifications were carried out, and a second version of the app (version 1.1) was released. SAMA was structured in 4 main areas (see Figure 2):

1. *Info area*, which provides information about the app and the project.

2. *Log-in area*, which is responsible for access control and is also the area where users are asked to fill out surveys.

3. *Task list area*, the main area of the app, which shows the list of tasks ordered by immediacy and allows the user to update such tasks when there are new schedules or when new meetings have been carried out. Besides, the user can generate reports and can also create backup copies from this area.

4. *Setup area*, from where it is possible to customize the app: the user can enable or disable tasks or activities (groups of tasks), and they can also add tasks of their own.

Info Area

The screen that appears when the app is run for the first time is shown in Figure 3 (left). From this screen, the user can access information about the app and the project, both as a PDF document [24] (Figure 3, right) and as a video tutorial [25].
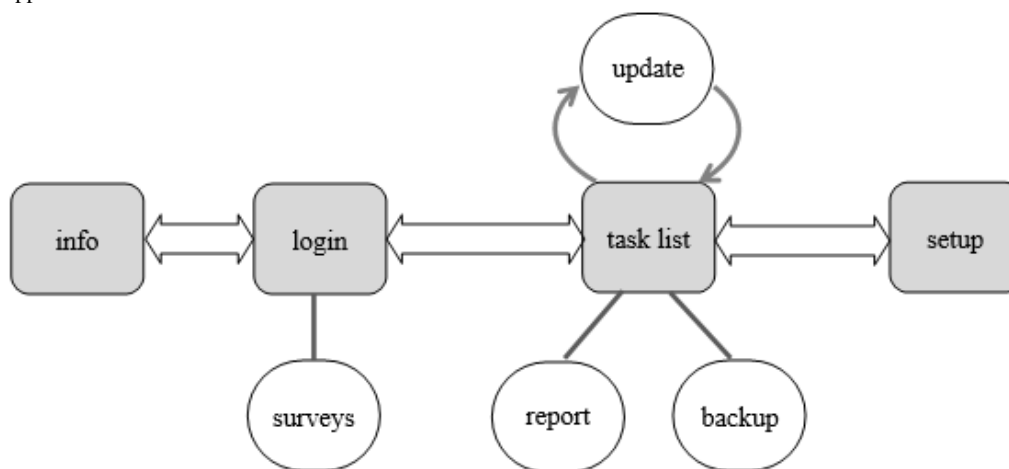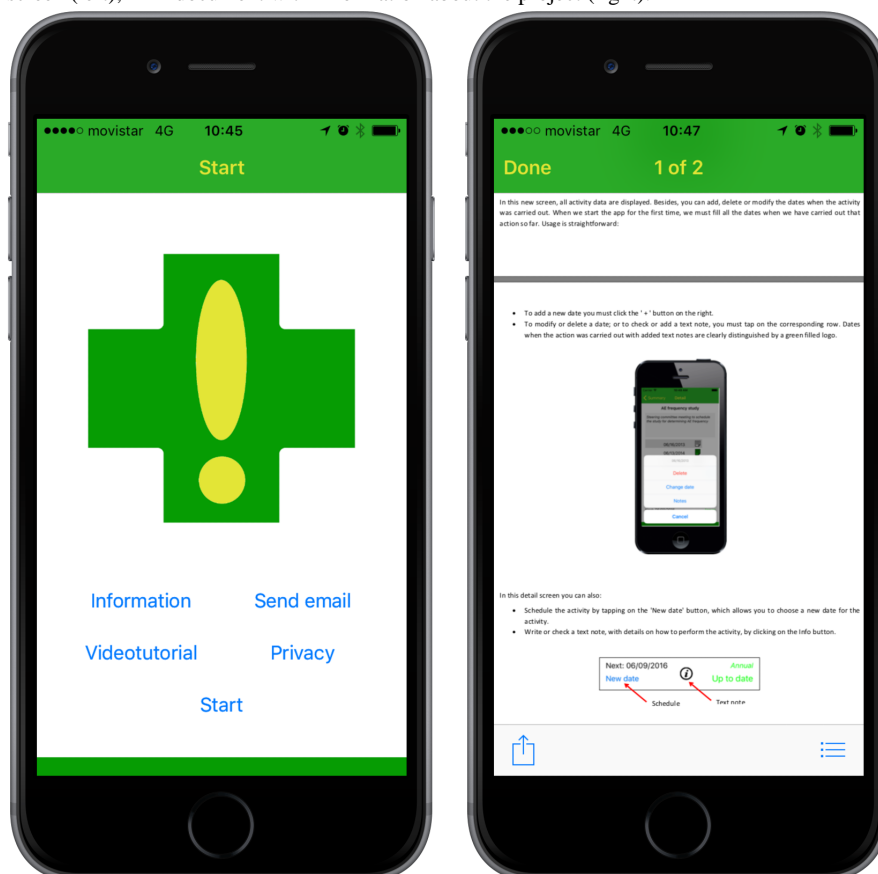
Figure 2. Mobile app structure.

Figure 3. Info area. First screen (left); PDF document with information about the project (right).


Log-In Area

Access to the app is restricted by username and password.

Task List Area

The main screen in this area shows the list of tasks ordered by immediacy. Every task has a certain periodicity, so it has to be carried out at regular times. There are 3 possible states for a task: pending (ie, its due date has arrived and it has not been completed or scheduled), scheduled (ie, its due date has been postponed by the user), and up-to-date (ie, its due date has not yet arrived). Figure 4 (left) shows this screen.

Tapping one of the tasks takes you to a detailed screen where you can update the task status: you can schedule the task for a later date, you can mark this task as completed on a certain date, and you can also add additional text information. Figure 4 (center) shows the detail screen.

Additionally, there are 2 more actions that can be performed from the task list screen: you can generate a PDF report on the status of all tasks, like Figure 4 (right) shows, and you can also create a backup copy of your data.

Figure 4. Task list area. Task list (left), detail list (center), and report on status (right).
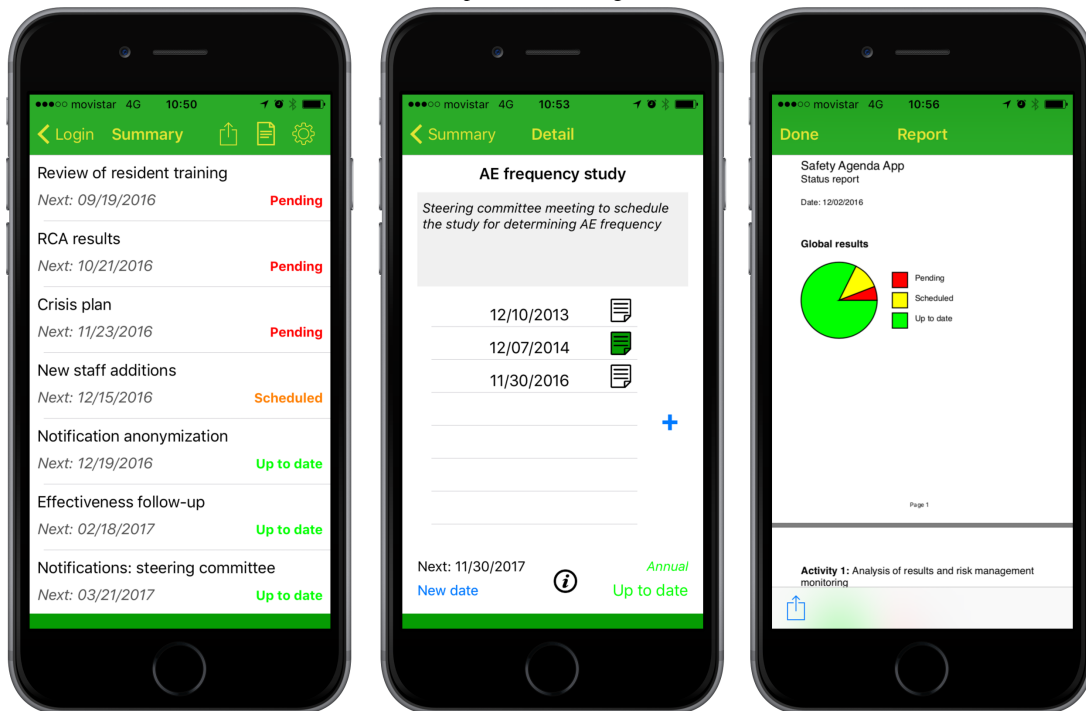
Setup Area

App customizations are performed in the setup area. Basically, users can enable or disable tasks depending on users' hospital or primary care needs and also create their own custom tasks. Tasks are grouped into activities and sections. Figure 5 (left) shows the first setup screen, where users can enable or disable activities (groups of tasks) and create their own activities. Figure 5 (center) shows that you can edit these activities just by swiping the row to the left. Finally, Figure 5 (right) shows the second setup screen, or detail view, where you can enable or disable each task independently; you can also create your own tasks.

Once the user has been identified, the list of tasks included in the agenda appears along with their status, which can be pending (when the task has yet to be done within the recommended implementation period), up-to-date (when the date of a finished task is introduced that is within the recommended implementation period), and programmed (by introducing an upcoming date to finish a task that was either pending or up-to-date). Each task is previously introduced into the system with a recommend schedule for completion (four times a year, three times, twice, or once a year). Touching any specific task makes its description and its compliance record appear, and the system permits users to add the last date that the task was completed, thereby modifying its status. The app has an edit screen that permits users to either hide or show specific tasks or groups of tasks (activities). SAMA permits downloading and emailing of compliance records.

Safety Agenda Mobile App 1.1 Assessment

Between June and September 2015, a total of 102 users downloaded the mobile app, and of these, 74 (72%) registered and assessed SAMA 1.1. Among those registered, it was more frequent that the management team of the health institution analyzed the risks to patient safety and shared its conclusions with the quality and safety commissions, and instructors (Table 2). As for its main strengths, these users identified that SAMA 1.1 constitutes a reminding system of the responsibilities of patient safety managers at the same time as it emerged as a verification list of safety tasks. The aspects needing improvement, according to user opinions, had to do with its design (user-friendliness, background color, font size) and with additional functions such as personalizing the content of the compliance reports, prioritizing tasks, synchronizing them with calendar applications, and expanding the task status scale (Table 3).

Version 1.2, which is about to be released, supports the English language (versions 1.0 and 1.1 supported only Spanish). SAMA 1.2 is also available at the App Store. A total of 238 users downloaded the mobile app till June 2016, either the version in Spanish, SAMA 1.1, or the version in English, SAMA 1.2.

As the current version of the app (version 1.1) only supports the Spanish language, most downloads correspond to Spain (167/238, 70.2%). However, the number of downloads in other countries suggest that version 1.2 (which supports the English language and is about to be released) will be accepted well worldwide (71/238, 29.8%).

On November 24, 2016, SAMA has been accredited receiving the AppSaludable (HealthyApp) Quality Seal [26]. This mean that SAMA meets a set of criteria about design, functionality, safety and quality and could be used in the health context to provide managers with an overview of their responsibilities in matters of patient safety.

Figure 5. Setup area. Activity list (left), deleting or editing your own activities (center), and detail screen with the task list for a certain activity (right).
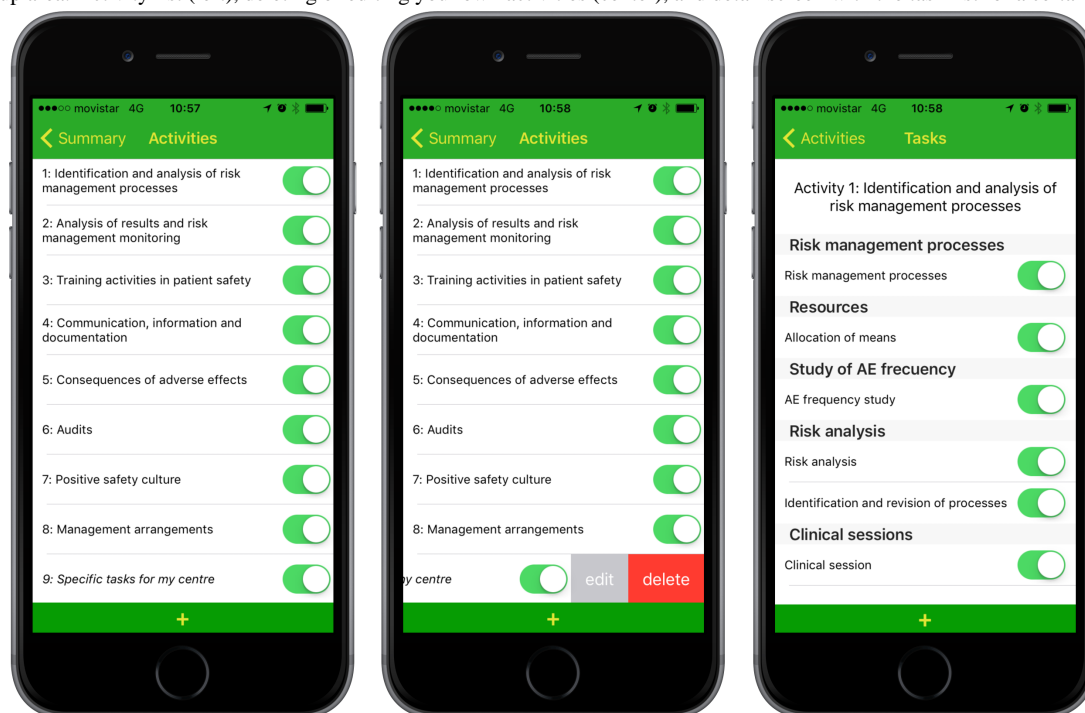
Table 2. Focus and management activity in patient safety (N=74).

Users experiences (quantitative data)	Yes (%)	No (%)
The most frequent causes of adverse events are periodically reviewed with the steering committee.	54 (73)	20 (27)
The management team dedicates at least one session every 4 months to analyze the risks to patient safety and shares the conclusions with the quality and safety commissions, and instructors.	41 (55)	33 (45)
SAMA[a] 1.1 assessed positively on its usefulness and navigation facilities	64 (87)	10 (13)

[a]SAMA: Safety Agenda Mobile App.

Table 3. Safety Agenda Mobile App 1.1 strengths and areas for improvement identified by users (N=74).

Users experiences (qualitative data)	n
Strengths	
Reminds managers about tasks	5
Facilitates a checklist of safety tasks	3
Clear action scheme	2
Easy to use	2
Very intuitive	1
Areas for improvement	
Trial design	1
Small font size	1
Wallpaper color cannot be chosen	1
Ability to print the report without the graphic summaries	1
Ability to classify items to comply with prioritizing the tasks	1
Ability to transfer dates to Outlook or other calendar schedules	1
It could consider other degrees of the status of each task (eg, not started, started, in progress, final phase, finished, halted)	1

Discussion

Principal Findings

SAMA was designed for the managing team at a health center to become aware of all the activities in matters of patient safety that are their responsibility and to facilitate their involvement directing these activities that are planned throughout the year. The agenda's format allows programming the tasks in a manner consistent with the center's patient safety plan and personalizing the tasks according to such plan and the annual forecasts. Moreover, SAMA permits sharing activities with the center's middle management and editing a report of the activities carried out that reflect the degree of commitment by the management to patient safety. SAMA includes a set of activities that are expected to be carried out by health managers in matters of patient safety and contributes toward improving the awareness of their responsibilities in matters of safety.

SAMA Utility

The review by Parand et al [11] showed that when managers incorporate quality and safety objectives into their work agendas, health care center results improve. SAMA adopts these results as a premise and goes further, incorporating the responsibilities in matters of patient safety that the UNE 179003:2013 standard incorporates in the management of risks inherent in health intervention into the design and promotion of a culture of safety at the center, to the review of outcome indicators, etc.

Not having a vision of the tasks that must be performed on matters of safety is one of the main problems for managers—for this, SAMA provides a dynamic, flexible, and trustworthy structure that helps management staff conceptualize their responsibilities in matters of safety and include them in their work agendas. Another problem is not having sufficient time for all the tasks that are expected to be carried out throughout a year in safety matters. Time management, the delegation of tasks, and appropriate input of all the activities expected of managers in matters of safety are the ways in which SAMA helps them overcome the identified gaps.

Progressively, we count on evidence of the effectiveness of the proposed interventions to reduce the occurrence and impact of adverse events. The projects of Bacteremia Zero (in Spain, it achieved in 1 year a reduction of some 3800 cases of bacteremia associated with central venous catheter in the intensive care unit or ICU [27]) and Pneumonia Zero (a reduction of pneumonia associated with mechanical ventilation in ICUs to 6 per 1000 mechanical ventilation days in 1 year [28]) are just 2 examples. These actions cannot be an option; instead, the responsibility of managers is to do whatever possible to incorporate them in their centers. SAMA is designed to be customized and permits introducing those actions that are considered effective and necessary.

Conclusion

The culture of safety of health care organizations has been related with the outcome of the risk management that affects the number of adverse events patients suffer [29]. Regrettably, a people approach is still preferred over a system approach, and we forget that a safe system is that which reduces the probability of professionals committing errors [30]. SAMA, just like other tools such as those designed for reporting systems [31], can contribute to implementing a positive and solid culture of safety in this case because it facilitates a set of activities managers are responsible for in order to accomplish proper management of the risks inherent to health care.

Limitations

The current SAMA design is limited to use by only iOS devices. The effect SAMA may have on the center's culture of safety has not been analyzed, nor whether its use affects the risk management for patient safety to any extent.

Future Studies

This study requires future research to determine the impact on patient safety indicators resulting from continuous use of SAMA at health care centers. New app updates could incorporate modifications based on user comments, and a version for Android operating systems could even be built.

Conflicts of Interest

None declared.

References

1. de Vries EN, Ramrattan MA, Smorenburg SM, Gouma DJ, Boermeester MA. The incidence and nature of in-hospital adverse events: a systematic review. Qual Saf Health Care 2008 Jun;17(3):216-223 [FREE Full text] [doi: 10.1136/qshc.2007.023622] [Medline: 18519629]
2. Seys D, Wu AW, Van Gerven E, Vleugels A, Euwema M, Panella M, et al. Health care professionals as second victims after adverse events: a systematic review. Eval Health Prof 2013 Jun;36(2):135-162. [doi: 10.1177/0163278712458918] [Medline: 22976126]

XSL•FO

RenderX

3. Tsang C, Bottle A, Majeed A, Aylin P. Adverse events recorded in English primary care: observational study using the General Practice Research Database. Br J Gen Pract 2013 Aug;63(613):e534-e542 [FREE Full text] [doi: 10.3399/bjgp13X670660] [Medline: 23972194]

4. Aranaz-Andrés JM, Aibar C, Limón R, Mira JJ, Vitaller J, Agra Y, et al. A study of the prevalence of adverse events in primary healthcare in Spain. Eur J Public Health 2012 Dec;22(6):921-925 [FREE Full text] [doi: 10.1093/eurpub/ckr168] [Medline: 23180803]

5. Allué N, Chiarello P, Bernal DE, Castells X, Giraldo P, Martínez N, et al. [Assessing the economic impact of adverse events in Spanish hospitals by using administrative data]. Gac Sanit 2014;28(1):48-54 [FREE Full text] [doi: 10.1016/j.gaceta.2013.06.004] [Medline: 24309522]

6. Montserrat-Capella D, Suárez M, Ortiz L, Mira JJ, Duarte HG, Reveiz L. Frequency of ambulatory care adverse events in Latin American countries: the AMBEAS/PAHO cohort study. Int J Qual Health Care 2015 Feb;27(1):52-59 [FREE Full text] [doi: 10.1093/intqhc/mzu100] [Medline: 25609774]

7. Antoñanzas VF. [Non safety costs in the Spanish health care system]. Rev Esp Salud Publica 2013 Jun;87(3):283-292 [FREE Full text] [doi: 10.4321/S1135-57272013000300008] [Medline: 23892680]

8. Hoonhout LH, de Bruijne MC, Wagner C, Zegers M, Waaijman R, Spreeuwenberg P, et al. Direct medical costs of adverse events in Dutch hospitals. BMC Health Serv Res 2009 Feb;9:27 [FREE Full text] [doi: 10.1186/1472-6963-9-27] [Medline: 19203365]

9. Allué N, Chiarello P, Bernal DE, Castells X, Giraldo P, Martínez N, et al. [Assessing the economic impact of adverse events in Spanish hospitals by using administrative data]. Gac Sanit 2014;28(1):48-54 [FREE Full text] [doi: 10.1016/j.gaceta.2013.06.004] [Medline: 24309522]

10. Etchegaray JM, Thomas EJ. Engaging Employees: The Importance of High-Performance Work Systems for Patient Safety. J Patient Saf 2015 Dec;11(4):221-227. [doi: 10.1097/PTS.0000000000000076] [Medline: 24522214]

11. Parand A, Dopson S, Renz A, Vincent C. The role of hospital managers in quality and patient safety: a systematic review. BMJ Open 2014 Sep;4(9):e005055 [FREE Full text] [doi: 10.1136/bmjopen-2014-005055] [Medline: 25192876]

12. Mira JJ, Carrillo I, Lorenzo S, Ferrús L, Silvestre C, Pérez-Pérez P, Research Group on Second and Third Victims. The aftermath of adverse events in Spanish primary care and hospital health professionals. BMC Health Serv Res 2015 Apr;15:151 [FREE Full text] [doi: 10.1186/s12913-015-0790-7] [Medline: 25886369]

13. AENOR. Madrid: AENOR; 2013. Health services: Risk management for patient safety URL: http://www.aenor.es/aenor/normas/normas/fichanorma.asp?tipo=N&codigo=N0051018#.WDoJTXdDnq1

14. Carrillo I, Mira JJ, Vicente MA, Fernandez C, Guilabert M, Ferrús L, et al. Design and Testing of BACRA, a Web-Based Tool for Middle Managers at Health Care Facilities to Lead the Search for Solutions to Patient Safety Incidents. J Med Internet Res 2016 Sep 27;18(9):e257 [FREE Full text] [doi: 10.2196/jmir.5942] [Medline: 27678308]

15. McFadden KL, Stock GN, Gowen CR. Leadership, safety climate, and continuous quality improvement: impact on process quality and patient safety. Health Care Manage Rev 2015 Mar;40(1):24-34. [doi: 10.1097/HMR.0000000000000006] [Medline: 24566246]

16. White AA, Brock DM, McCotter PI, Hofeldt R, Edrees HH, Wu AW, et al. Risk managers' descriptions of programs to support second victims after adverse events. J Healthc Risk Manag 2015;34(4):30-40 [FREE Full text] [doi: 10.1002/jhrm.21169] [Medline: 25891288]

17. Rodrigues SP, van Eck NJ, Waltman L, Jansen FW. Mapping patient safety: a large-scale literature review using bibliometric visualisation techniques. BMJ Open 2014 Mar;4(3):e004468 [FREE Full text] [doi: 10.1136/bmjopen-2013-004468] [Medline: 24625640]

18. Clarke JR, Lerner JC, Marella W. The role for leaders of health care organizations in patient safety. Am J Med Qual 2007 Oct;22(5):311-318. [doi: 10.1177/1062860607304743] [Medline: 17804390]

19. White AA, Waterman AD, McCotter P, Boyle DJ, Gallagher TH. Supporting health care workers after medical error: considerations for healthcare leaders. J Clin Outcomes Manag 2008 May;15(5):240-247 [FREE Full text]

20. Gerven EV, Seys D, Panella M, Sermeus W, Euwema M, Federico F, et al. Involvement of health-care professionals in an adverse event: the role of management in supporting their workforce. Pol Arch Med Wewn 2014 Apr;124(6):313-320 [FREE Full text] [Medline: 24781784]

21. Mira JJ, Lorenzo S, Carrillo I, Ferrús L, Pérez-Pérez P, Iglesias F, Research Group on Second and Third Victims. Interventions in health organisations to reduce the impact of adverse events in second and third victims. BMC Health Serv Res 2015 Aug;15:341 [FREE Full text] [doi: 10.1186/s12913-015-0994-x] [Medline: 26297015]

22. Appandabout SL. Safety Agenda Mobile App (SAMA). iTunes Apple. 2016. URL: https://itunes.apple.com/es/app/safety-agenda-mobile-app/id1008897674?l=en&mt=8 [WebCite Cache ID 6kr1MjKEs]

23. Calidad. AppSaludable Quality Seal. 2016. URL: http://www.calidadappsalud.com/en/distintivo-appsaludable/ [accessed 2016-11-26] [WebCite Cache ID 6mHtMlP13]

24. Mira JJ, Carrillo I, Fernández C, Vicente MA, Guilabert M. Safety Agenda Mobile App. Appandabout. 2016. URL: http://appandabout.es/sama/info_safetyagenda.pdf [accessed 2016-09-28] [WebCite Cache ID 6kr0RyIJY]

XSL•FO

RenderX

25.   Mira JJ, Carrillo I, Fernández C, Vicente MA, Guilabert M. Video tutorial: How to use the Safety Agenda Mobile App (SAMA). Youtube. 2016. URL: https://www.youtube.com/watch?v=d5-8NacXwZw [accessed 2016-09-28] [WebCite Cache ID 6kr11VEzx]

26.   Andalusian Agency for Health Care Quality. Safety Agenda Mobile App (SAMA). 2016. URL: http://www.calidadappsalud.com/distintivo/info?app=safety-agenda-mobile-sama [accessed 2016-11-26] [WebCite Cache ID 6mHtqYDbE]

27.   Palomar M, Álvarez-Lerma F, Riera A, Díaz MT, Torres F, Agra Y, Bacteremia Zero Working Group. Impact of a national multimodal intervention to prevent catheter-related bloodstream infection in the ICU: the Spanish experience. Crit Care Med 2013 Oct;41(10):2364-2372. [doi: 10.1097/CCM.0b013e3182923622] [Medline: 23939352]

28.   Álvarez-Lerma F, Álvarez J, Añón JM, Arias S, García R, Gordo F, et al. Informe final del proyecto Neumonía Zero. En: X Congreso Panamericano e Ibérico de Medicina Crítica y Terapia Intensiva. Med Intensiva 2014;38(Espec Cong):1-441.

29.   Wakefield JG, McLaws M, Whitby M, Patton L. Patient safety culture: factors that influence clinician involvement in patient safety behaviours. Qual Saf Health Care 2010 Dec;19(6):585-591. [doi: 10.1136/qshc.2008.030700] [Medline: 20724390]

30.   Panella M, Leigheb F, Rinaldi C, Donnarumma C, Tozzi Q, Di Stanislao F. [Defensive Medicine: Defensive Medicine: Overview of the literature]. Ig Sanita Pubbl 2015 Jun;71(3):335-351. [Medline: 26241515]

31.   Elliott P, Martin D, Neville D. Electronic clinical safety reporting system: a benefits evaluation. JMIR Med Inform 2014 Jun;2(1):e12 [FREE Full text] [doi: 10.2196/medinform.3316] [Medline: 25600569]

Abbreviations

ICU: intensive care unit
SAMA: Safety Agenda Mobile App
UNE: Una Norma Española

Review

Feasibility and Effectiveness of Using Wearable Activity Trackers in Youth: A Systematic Review

Nicola D Ridgers[1], BSc (Hons), MSc, PhD; Melitta A McNarry[2], BSc (Hons), PhD; Kelly A Mackintosh[2], BSc (Hons), MSc, PhD

[1]Deakin University, Geelong, Australia, Institute for Physical Activity and Nutrition (IPAN), School of Exercise and Nutrition Sciences
[2]Applied Sports Science Technology and Medicine Research Centre (A-STEM), College of Engineering, Swansea University, Swansea, United Kingdom

Corresponding Author:
Nicola D Ridgers, BSc (Hons), MSc, PhD
Institute for Physical Activity and Nutrition (IPAN)
School of Exercise and Nutrition Sciences
Deakin University
221 Burwood Highway
Burwood, 3125
Australia
Phone: 61 0392446718
Fax: 61 0392446017
Email: nicky.ridgers@deakin.edu.au

Abstract

Background: The proliferation and popularity of wearable activity trackers (eg, Fitbit, Jawbone, Misfit) may present an opportunity to integrate such technology into physical activity interventions. While several systematic reviews have reported intervention effects of using wearable activity trackers on adults' physical activity levels, none to date have focused specifically on children and adolescents.

Objective: The aim of this review was to examine the effectiveness of wearable activity trackers as a tool for increasing children's and adolescents' physical activity levels. We also examined the feasibility of using such technology in younger populations (age range 5-19 years).

Methods: We conducted a systematic search of 5 electronic databases, reference lists, and personal archives to identify articles published up until August 2016 that met the inclusion criteria. Articles were included if they (1) specifically examined the use of a wearable device within an intervention or a feasibility study; (2) included participants aged 5-19 years old; (3) had a measure of physical activity as an outcome variable for intervention studies; (4) reported process data concerning the feasibility of the device in feasibility studies; and (5) were published in English. Data were analyzed in August 2016.

Results: In total, we identified and analyzed 5 studies (3 intervention, 2 feasibility). Intervention delivery ranged from 19 days to 3 months, with only 1 study using a randomized controlled trial design. Wearable activity trackers were typically combined with other intervention approaches such as goal setting and researcher feedback. While intervention effects were generally positive, the reported differences were largely nonsignificant. The feasibility studies indicated that monitor comfort and design and feedback features were important factors to children and adolescents.

Conclusions: There is a paucity of research concerning the effectiveness and feasibility of wearable activity trackers as a tool for increasing children's and adolescents' physical activity levels. While there are some preliminary data to suggest these devices may have the potential to increase activity levels through self-monitoring and goal setting in the short term, more research is needed to establish longer-term effects on behavior.

KEYWORDS

XSL•FO
RenderX

Introduction

Physical inactivity is a global pandemic and has been identified as the fourth leading cause of death worldwide [1]. Regular physical activity plays a critical role in preventing precursors to metabolic and cardiovascular ill health in children [2], providing numerous health benefits during childhood that persist into adulthood [3]. Such health benefits include protective effects on bone health, as well as positive effects on fitness, body fat, and blood pressure [4]. Several countries (eg, the United States, United Kingdom, and Australia) recommend that children should engage in at least 60 minutes of moderate- to vigorous-intensity physical activity (MVPA) every day to benefit health [5-7]. However, the majority of children and adolescents (defined as youth hereinafter) do not meet these recommended levels and are therefore not sufficiently active to accrue the associated health benefits [8-10]. Since physical inactivity is a major, yet modifiable, risk factor for the burden of disease, there is a need for effective, preventive interventions that aim to increase physical activity levels in this population.

Self-monitoring has been identified as an effective behavior change technique that has been used in behavioral interventions targeting increases in physical activity levels [11]. Indeed, self-monitoring and feedback are fundamental to increasing awareness of individual physical activity levels, which is particularly important given that youth are unlikely to change their behavior if they do not know how active they actually are and how this translates to government guidelines. Specifically, Corder and colleagues found that approximately 60% of inactive adolescents thought that they met physical activity guidelines [12], suggesting that they may see no need to change their behavior, despite the associated health benefits. Traditionally, hip-worn pedometers have been used to increase individuals' awareness of their physical activity [13]; however, participants are required to record their activity at the end of each day, which can be burdensome for them [13]. In recent years, there has been increasing interest in emerging technologies and wearable sensors as self-monitoring tools for promoting physical activity levels [14]. The proliferation of wearable activity trackers, as well as their growing commercial availability, popularity, and widespread adoption [15], presents an opportunity to integrate such technologies into physical activity interventions. While ownership data are not available for youth, it is estimated that 10% and 20% of US and Australian adults, respectively, own some form of wearable technology [15,16]. An integral component of wearable activity trackers, such as Fitbit and Jawbone, is the automation of real-time physical activity tracking [17]. The wireless syncing of such devices to Web- or app-based profiles not only negates the burden of manual data entry, but also enables the wearer to self-monitor against physical activity recommendations or set goals [14,17].

To date, physical activity research has primarily focused on establishing the validity and reliability of wearable activity trackers for measuring a range of outcomes, including steps and sleep [18]. In comparison, little is known about the feasibility and effectiveness of these devices as a tool for increasing physical activity levels, whether in isolation or in combination with other strategies. A recent review reported that there was

some initial evidence that wearable activity trackers can increase physical activity levels, though only studies conducted in adult populations were included [19]. Given that engagement with technology is a highly valued behavior for youth and plays an important role in different domains of their lives (eg, education, socialization, and entertainment; [20]), there is a need to establish whether wearable activity trackers are feasible and effective in changing physical activity levels in this population. Such information is important for informing future physical activity interventions and has the potential to contribute to the development of public health guidance concerning the role of these tools in physical activity and health promotion practice.

The aim of this review, therefore, was to examine the effectiveness of wearable activity trackers as a tool for increasing children's and adolescents' physical activity levels. We also examined the feasibility of using such technology in youth populations (defined as those in the age range 5-19 years).

Methods

We conducted a systematic literature search in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [21]. We searched 5 electronic databases (PubMed, Web of Science, SPORTDiscus, Scopus, and ProQuest Central). Search strategies for the different databases included the following search strings in four main areas: wearable activity trackers (electronic track*, electronic activ*, electronic monitor, electronic fitness track*, wearable device, wearable act*, wearable track*, consumer wearable, Fitbit, SenseWear, Jawbone, Nike Fuelband, PAM), population (child*, adolescent, youth), study design (intervention, trial, feasibility), and outcome variable (physical activity, energy expenditure, fitness, exercise). The full search strategies for each database are presented in Multimedia Appendix 1. Articles that had been published in peer-reviewed journals or conference proceedings were considered for review. We did not include abstracts, dissertations, systematic reviews, and case studies. In addition to electronic searches, we also searched our personal collections and the bibliographies of retrieved studies. This is a commonly used approach for the identification of additional relevant studies for potential inclusion in systematic reviews [22].

For the purpose of this review, we defined wearable activity trackers as an electronic device with the following features: was designed to be worn on the user's body; uses accelerometers, altimeters, or other sensors to track the wearer's movements or biometric data, or both; and can provide feedback, beyond the display of basic activity count information, via the monitor display or through a partnering app to elicit continual self-monitoring of activity behavior [19,23]. To be included in the review, studies were required to (1) specifically examine the use of a wearable device within an intervention or a feasibility study; (2) include participants aged 5-19 years old; (3) have a measure of physical activity as an outcome variable for intervention studies; (4) report process data concerning the feasibility of the device in feasibility studies; (5) be published between the start date of each electronic database and August 2016; and (6) be published in English. We excluded studies that

XSL•FO

RenderX

reported study protocols, used mobile phones rather than a wearable activity tracker, used mobile phone or tablet apps without an accompanying wearable activity tracker, or only used the wearable activity tracker to evaluate an intervention (eg, worn at baseline and postintervention). In the event that a wearable activity tracker was used in conjunction with other tools, such as Facebook to share activity tips, studies were eligible for inclusion if physical activity data were reported or the feasibility of the device was reported separately. Studies that implemented the use of such technologies in clinical populations were eligible for inclusion if the focus was on using a wearable device to increase physical activity levels. When studies were still in press or were an advanced publication ahead of print but had a unique digital object identifier, they were eligible for inclusion. Conference proceedings were eligible for review due to the potential for such devices to be examined in different disciplines (eg, computer science) where such outputs are often more reputable than journal articles.

All authors independently assessed the results obtained from the initial literature search. Articles were screened in 4 steps: first, we removed duplicates, then screened the title, abstract, and full text. We then screened the articles based on the inclusion and exclusion criteria outlined above. If we could not determine suitability during the screening of the title and abstract, we accessed full-text articles and compared them against the inclusion criteria. Any disagreements were settled by discussion between the authors. We then extracted the following data using a standardized form for each study that met the inclusion criteria: country of study, study design, sample characteristics (eg, sample size, age), wearable device used, and results. The first author extracted the data, which were checked by the remaining authors (MAM, KAM). We then undertook a narrative review of the included studies.

The second and last authors (MAM, KAM) independently assessed the risk of bias in the intervention studies that met the inclusion criteria. We adapted the criteria for assessing risk of bias from the *Methods Guide for Comparative Effectiveness Reviews* [24] and previous reviews in similar areas [19]. We identified 8 criteria as being important to this review: (1) participants were allocated randomly; (2) an adequate proportion of participants had complete data for the outcome variable (ie, no more than 20% of data were missing); (3) data were analyzed according to group allocated; (4) the study population was representative of the population of interest; (5) the timing of outcome assessments was similar in all groups; (6) the study reported the validity of the device used (either data were provided in the article or there was an appropriate reference to the original study); (7) the study reported the reliability of the device used (either data were provided in the article or there was an appropriate reference to the original study); and (8) the study was conducted independently of the manufacturer of the device used. We assessed only the last 3 risk-of-bias criteria for feasibility studies given the nature of such study designs. Each criterion was scored as "yes" (1), "no" (0), or "unsure" (?).

Results

We screened and analyzed data in August 2016. Through the systematic search, we initially identified 259 articles, then identified 1 additional article through other sources. Of these, 5 were included in the review (Figure 1): 3 were intervention studies [25-27] and focused primarily on increasing physical activity levels across the whole day [26,27] or during recess [25], and 2 studies were classed as feasibility studies [28,29]. Of the intervention studies, 2 [26,27] also included process measures about the device used. The majority of the studies (n=4) focused on children and were conducted in the United States [25,27-29]. The main brand of wearable device used was Fitbit [25,27,28]. Table 1 reports the characteristics of each included study. Figure 2 shows the different wearable devices used in the included studies, and Table 2 provides an overview of the features of these devices.

Table 3 reports the risk of bias for each study. All were conducted independently of the manufacturers of the device(s) used, while 2 studies reported the reliability and validity of the devices used. Of the intervention studies conducted, only 1 used a randomized controlled trial design [26].

Table 1. Summary of included studies on the effectiveness and feasibility of wearable activity trackers in youth (chronological order by study design).

Study	Country	Participants	Type of study	Study design and description	Device examined	Outcomes assessed
Slootmaker, 2010 [26]	Nether-lands	Adolescents (13-17 years old), 32 boys, 55 girls (15.1 years) at baseline	Intervention plus feasibility component	Least active group of youth recruited. Randomized to intervention or control. 3-month Web-based intervention combining self-monitoring, goal setting, device and PAM COACH.	PAM and PAM COACH	Physical activity: Activity Questionnaire for Adolescents and Adults (self-report). Time spent in SED[a], LPA[b], MPA[c], and VPA[d]. Process measures: evaluation of PAM and PAM coach.
Hayes, 2015 [25]	USA	6 grade-3 girls (aged 8 years old) from 1 school; intact social group	Intervention	Recess intervention (22 sessions in total). Fitbit used to self-monitor physical activity levels against set goals. Tangible rewards provided if goals met.	Fitbit (model not reported)	Steps/recess. MVPA[e] (min) during recess.
Hooke, 2016 [27]	USA	16 children (5 boys, 11 girls) aged mean 8.7, SD 3.1 years; participants receiving a cycle of maintenance chemotherapy for lymphoblastic leukemia	Intervention plus feasibility component	Used for 17 days before and 5 days after a corticosteroid pulse. Step goal tailored based on data and daily feedback against goal provided (either to increase or maintain physical activity). Goal set in Fitbit website by study nurse for participants to track progress.	Fitbit One	Steps/day. Feasibility component included ease of recruitment, ease of use and enjoyment of Fitbit, and days of wear.
Schaefer, 2014 [29]	USA	24 children (11 boys, 13 girls) aged 7-10 years (mean 8.9, SD 1.3 years)	Feasibility	Each child wore a different monitor for 1 week (4 weeks total). Underwent structured interview about each device and then summary (exit) interview at the end, with child and parents interviewed separately.	Actical SenseWear Polar Active Polar heart rate monitor	Frequency of removal, reasons for removal, enjoyment, comfort of use, favorite/least favorite device characteristics. Devices also ranked in terms of most and least favorite.
Schaefer, 2016 [28]	USA	34 children (22 boys, 12 girls) 11-12 years old (mean age 12.6 years); attending a low-socioeconomic-status school	Feasibility	6-month feasibility study. Initially asked to wear devices during after-school program, which then increased to daily wear.	Fitbit One	Fitbit data (ie, steps). Interviews examining experiences of using the Fitbit.

[a]SED: sedentary time.

[b]LPA: light-intensity physical activity.

[c]MPA: moderate-intensity physical activity.

[d]VPA: vigorous-intensity physical activity.

[e]MVPA: moderate- to vigorous-intensity physical activity.

XSL•FO

RenderX

Table 2. Overview of features of wearable devices used in the included studies on the effectiveness and feasibility of wearable activity trackers in youth.

Device	Location worn	Main measures	Device display	Compatibility	Sensors	Memory	Waterproof
Fitbit One	Waist	Steps, stairs, distance, calories, sleep	Yes	Personal computer, iOS, Android, Windows	Accelerometer (3 axis), altimeter	Up to 23 days	No
PAM	Waist	Physical activity score	Yes	Personal computer	Accelerometer (3 axis)	Not reported	No
SenseWear	Upper arm	Physical activity, energy expenditure, steps, sleep	No (optional display required)	Personal computer	Accelerometer (3 axis), heat flux, galvanic skin response, skin temperature, near-body ambient temperature	Up to 34 days	No
Actical	Wrist, waist, ankle	Physical activity, energy expenditure, steps	No	Personal computer	Accelerometer (omnidirectional)	Up to 194 days	Yes
Polar Active	Wrist	Physical activity, steps, calories, sleep	Yes	Personal computer, iOS, Android	Accelerometer (3 axis)	21 days (activity diary)	Yes
Polar heart rate monitor	Chest	Heart rate, calories	No	Personal computer	Heart rate	Up to 30 hours	Yes

Table 3. Risk-of-bias results[a] in studies on the effectiveness and feasibility of wearable activity trackers in youth.

Study	Random allocation	Minimal missing data	Analyzed in group allocated	Representative sampling	Timing of outcome assessments	Reliability of device	Validity of device	Independence from device manufacturer
Hayes, 2015 [25]	0	0	1	0	1	1	1	1
Hooke, 2016 [27]	0	1	1	0	1	1	1	1
Schaefer, 2014 [29]	N/A[b]	N/A	N/A	N/A	N/A	0	0	1
Schaefer, 2016 [28]	N/A	N/A	N/A	N/A	N/A	0	0	1
Slootmaker, 2010 [26]	1	1	1	1	1	0	0	1

[a]Scored as follows: 1 = yes; 0 = no.

[b]N/A: not applicable.

Figure 1. Flow diagram of screening process and results.
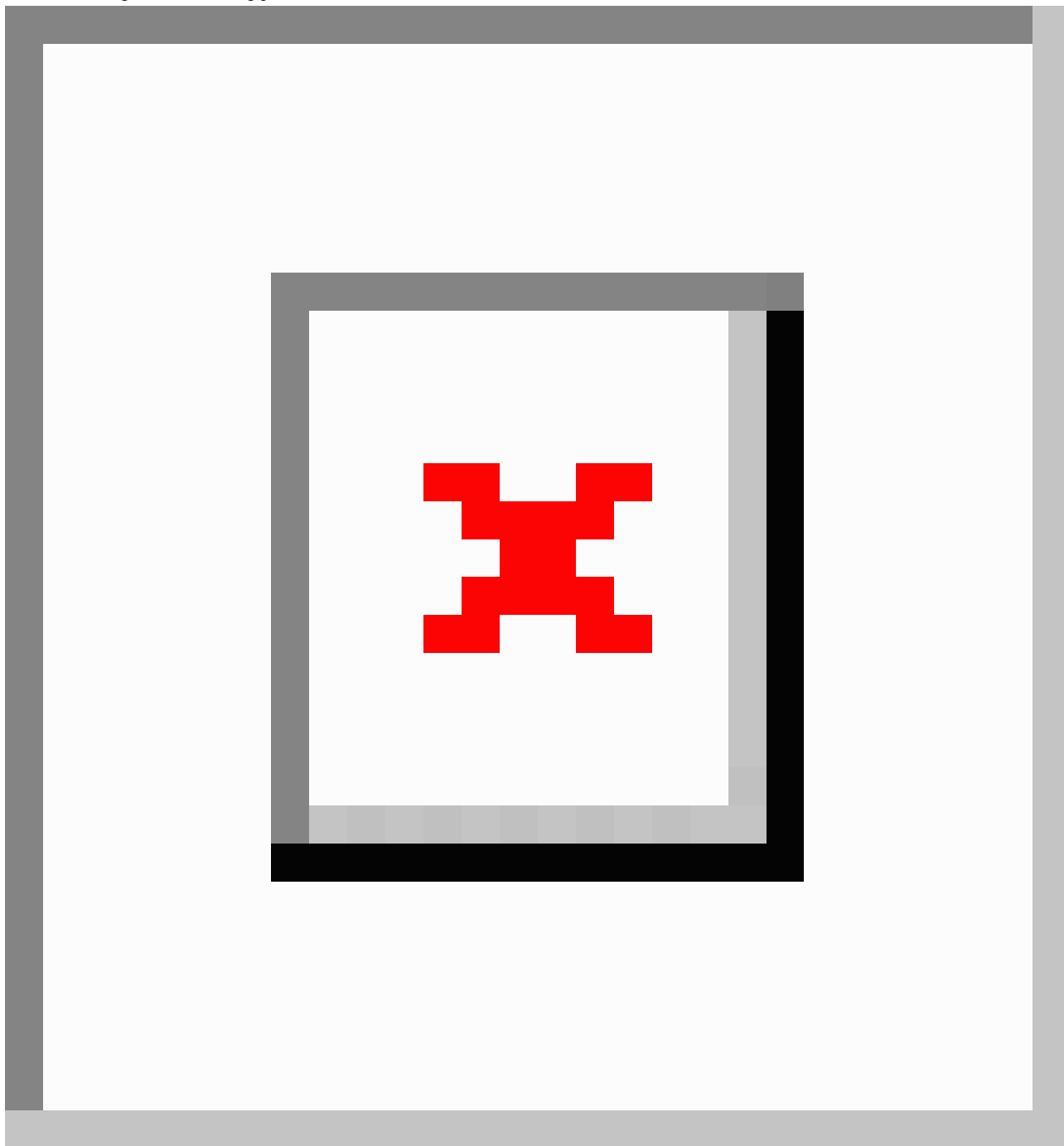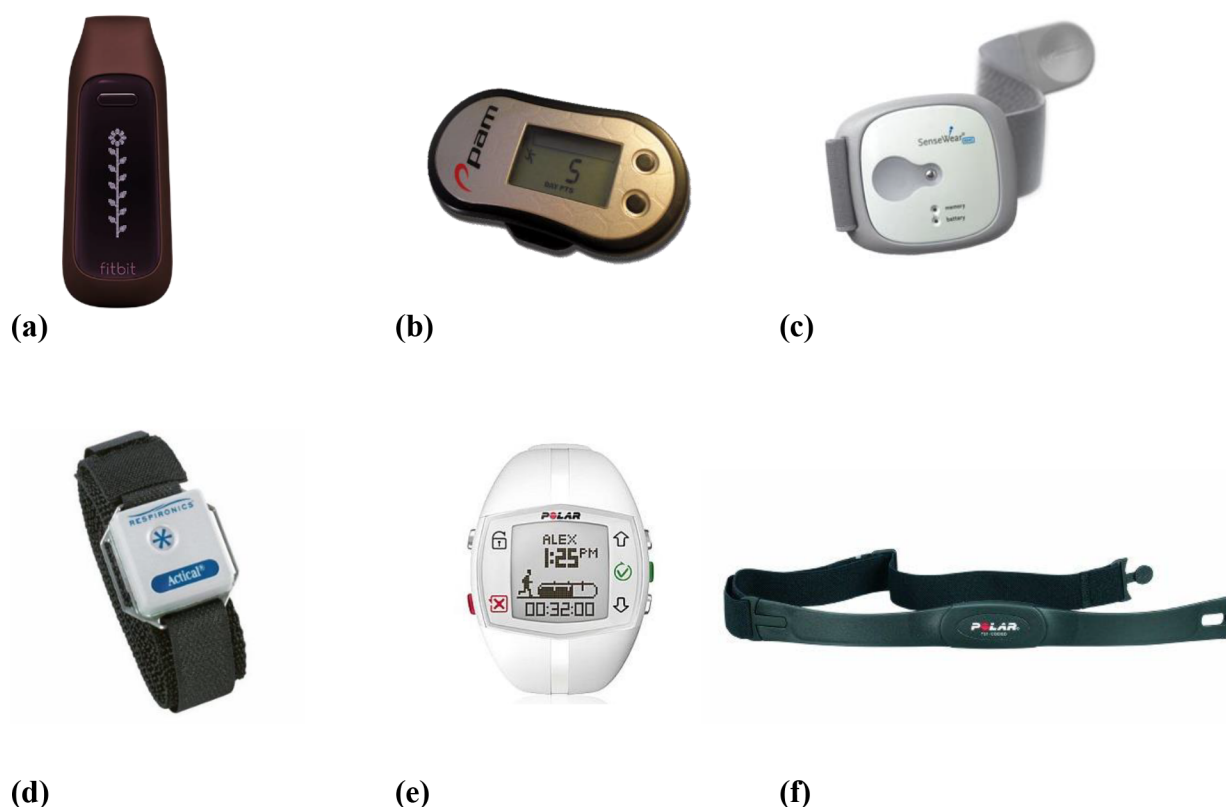
Figure 2. Wearable devices used in the included studies: (a) Fitbit One, (b) the PAM (new model shown), (c) SenseWear, (d) Actical, (e) Polar Active, and (f) Polar heart rate monitor.

(a)   **(b)**   **(c)**

(d)   **(e)**   **(f)**

Intervention Studies

In the Netherlands, Slootmaker and colleagues [26] investigated the effectiveness of the PAM monitor in combination with PAM COACH for increasing the physical activity levels of adolescents (aged 13-17 years; n=87). Activity data from the PAM device were uploaded to PAM COACH, which was a Web-based system for self-monitoring activity levels and setting goals. At the conclusion of the 3-month intervention, girls in the intervention group increased their self-reported weekly moderate physical activity relative to controls (equated to ~59 min/day), though this was not evident at 8 months postintervention. No intervention effects were observed for boys after 3 months, though self-reported sedentary time was lower at 8 months relative to boys in the control group (~257 min/day). Greater adherence to the study (eg, log-in and upload frequency to PAM COACH) was not associated with greater physical activity. Minimal attrition was observed, with 78% and 91% of participants providing follow-up data at 3 and 8 months postintervention, respectively. Overall, the PAM was viewed positively by participants, and 65% reported regular wear, though monitor loss (12%) and damage (7%) may have influenced the results.

Hayes and Van Camp [25] used the Fitbit (model not reported) as a tool for increasing 6 grade 3 (8 years old) girls' physical activity levels during school recess. After baseline data were collected from 7 recess periods, girls were provided with step goals (increments based on baseline data) for 7 further recess periods and were encouraged to self-monitor their steps against these goals. Following this, data were then collected for a further

7 periods (no step goals provided). The project culminated in a final intervention session where 3 goals were given, and a tangible reward (eg, a small toy) was provided based on the goal(s) achieved. The number of steps taken increased by 47% from baseline (1326 steps) to intervention (1956 steps; contribution of 18% to daily step recommendations [30]), while the proportion of time spent in MVPA increased from 4% to 25%, which equates to 5 minutes of MVPA during recess, or a contribution of 8% to daily recommendations [5]. Without the use of the Fitbit to self-monitor recess activity, steps taken and MVPA decreased to initial baseline levels. Basic process evaluation measures suggested that data were lost due to syncing issues, particularly during later recess sessions.

Hooke and colleagues [27] examined the efficacy of the Fitbit One to promote physical activity in clinical settings. A total of 16 children (mean age 8.7, SD 3.1 years) with acute lymphoblastic leukemia were provided with a Fitbit to wear for 17 days prior to and 5 days during a corticosteroid pulse. Monitoring over 3 days was used to identify baseline activity levels. Step goals were then tailored for each participant by a research nurse, and daily feedback was provided against these goals. No significant increases in daily steps were recorded, though there was a trend for steps to increase from week 1 to week 2 (average of 269 steps/day; 2% of daily step guidelines [30]) but to decrease from week 2 to week 3 (average of 307 steps/day). Process evaluation indicated that participants and their families were able to use the Fitbit One and the accompanying website, that they liked the Fitbit, and that data were available for 92% of measured days.

Feasibility Studies

Schaefer and colleagues conducted 2 feasibility studies in US primary school-age children [28,29]. In the first, 24 children each wore 4 activity monitors (Actical, SenseWear, Polar Active, and Polar heart rate monitor) separately for 1 week [29]. Of these activity monitors, the Polar Active and SenseWear met the definition of a wearable activity tracker. Following each week of wear, children and their parents were interviewed about their experiences of using the monitors. Overall, the Polar Active was the most popular monitor, with its comfort and feedback features (including a clock function) noted. It was used for 98% of the total time. In comparison, the SenseWear was the least popular, largely due to its placement on the arm (uncomfortable, embarrassing). This also corresponded with its lack of use (28% of the total time). No reactivity to wear was reported for the Polar Active.

In their second study, Schaefer and colleagues examined the feasibility of the Fitbit One in children aged 11-12 years attending a school located in an area of socioeconomic disadvantage [28]. Initially, 24 children were provided with a Fitbit to monitor their activity levels during an after-school program. After several weeks (approximately 1 month), children were provided with a Fitbit to wear every day for 5 months. On average, children accumulated 8406 steps/day. The number of steps taken increased during the monitoring phase, but not significantly. On average, 58 days of data were collected from each participant during the study, of which 19 days were considered to be valid days (overall use of 15%). Only 2 participants were still using their Fitbit at the end of the study (8%). Interview data indicated that, while the range of functions were well received, the design of the Fitbit One was unpopular and it was easy to forget to wear. Some children reported using the monitor to try to change behavior, while others (mainly boys) used it to compete against each other. One of the biggest barriers to use was the children's ability to sync and access their data outside of the after-school program.

Discussion

Principal Findings

This systematic, narrative review evaluated the effectiveness of wearable activity trackers as a tool for increasing children's and adolescents' physical activity levels. We also examined the feasibility of using wearable activity trackers in this population. Overall, there is a dearth of studies that have reported the use of wearable activity trackers in youth populations to increase activity levels. We identified 3 intervention studies, with 1 implemented in school-age children [25], 1 in a clinical population [27], and 1 with adolescents [26], the latter being the only study to use a randomized controlled trial design. There was some evidence to suggest that wearable activity trackers may have the potential to increase youth activity levels, with increases in physical activity compared with baseline or a control group observed. In addition, there was some evidence that wearable activity trackers were viewed positively by youth and they enjoyed wearing them. However, given that the studies had numerous methodological shortcomings and the majority of the reported differences were largely nonsignificant, it is

clear that further research using rigorous and well-designed methodologies is needed to establish the effectiveness of these devices for increasing youth activity levels.

Intervention Effects

The limited intervention effects observed in studies included in this review may be attributable to several factors. First, the length of the interventions implemented ranged from 19 days to 3 months. It is possible that the shorter intervention periods were not sufficient to change behavior, a notion supported by a recent review that highlighted that behavioral interventions of longer durations (≥6 months) had greater success in changing physical activity levels [31]. There is a clear need for studies using longer measurement periods to examine the effectiveness of these devices in youth. Second, it is not clear whether the interventions were grounded on behavioral theories, which are critical for intervention effectiveness [32]. Third, 2 of the studies may have been underpowered to detect a change in physical activity due to the smaller number of children recruited [25,27]. Fourth, intervention effects were not examined using validated objective monitors (eg, accelerometers). While several studies used data from the wearable activity tracker, these devices have not been validated for assessing physical activity outcomes in youth to date [18]. This could be viewed as a limitation of these included studies. However, this may be less of an issue if the focus of an intervention study is to use the device as a tool to facilitate behavior change rather than to evaluate the outcome (ie, validity and reliability should be established prior to such use). In the only study to report significant effects, self-reported physical activity data were collected using a questionnaire with low validity [26]. There is a clear need for longer-term studies using randomized controlled trials that are grounded on behavioral theories to identify the effectiveness of wearable activity trackers for changing youth physical activity behavior.

Self-Monitoring Using Wearable Activity Trackers

A common feature of the identified intervention studies was that the wearable activity tracker was used to self-monitor physical activity in combination with different intervention approaches [17]. Unsurprisingly, these intervention approaches largely consisted of goal setting, identified as an effective behavior change technique [11] to enhance physical activity levels [33]. However, there was some variability in who set the goals. In the interventions with children, the researchers set the goals based on baseline values obtained, and then provided regular support [27] or rewards [25] in relation to reaching the goals. In contrast, the adolescents had an opportunity to set their own activity goal and then received tailored advice to achieve these based on their preferred activities [26]. While research has shown that assigned goals are as effective as self-set goals, provided the reason for the goal is given [34], it is unclear whether the included studies provided this information to participants. It is important for future research to establish how youth engage with wearable activity trackers (eg, is the frequency of self-monitoring mediated by the activity goal source?), as this will provide critical insights into how these devices can be integrated into future interventions and strategies for engaging children and adolescents in the behavior change process.

Sustainability of Wearable Activity Trackers Over Time

Arguably, one of the biggest concerns regarding wearable activity trackers is whether individuals sustain their engagement with the technology over time [16]. Research has suggested that approximately one-third of US adults stop using their wearable activity tracker after 6 months [16], with expectation mismatch (ie, the technology doesn't do what was expected) a commonly cited reason [35]. While it is difficult to draw any conclusions about this due to the small number of studies identified and the variability in the length of time the wearable activity trackers were worn, there is some preliminary evidence to suggest that youth might regularly use the technology to self-monitor their physical activity levels when the technology is integrated into an intervention [26,27], but sustained use may not be observed in the longer term when these devices are simply provided to youth to use [28]. Interestingly, Slootmaker and colleagues found that the frequency that adolescents uploaded their data to PAM COACH was not associated with physical activity, yet physical activity levels were found to increase in girls at 3 months and sedentary time decreased in boys at 8 months [26]. These results could be explained by the lower activity levels of the girls in the study compared with boys; therefore, girls could achieve greater gains in activity levels [26]. However, it is also possible that there are different degrees of engagement, ranging from brief glances at the device [36], to tracking activity across the day using the monitor, to using specific functions within the accompanying app (eg, trends data), which may mediate the efficacy of the device on activity levels. Research is needed to provide further evidence about how youth engage with these devices and the accompanying apps over time (eg, attrition rates), whether this differs by population subgroups (eg, sex, age), whether the engagement with different features may have differential effects, and reasons for potential changes in use over time. Such research will be critical for identifying how to incorporate wearable activity trackers into interventions, and for informing best practice in future physical activity interventions and health promotion practice.

Feasibility of Wearable Activity Trackers in Youth

This review identified few studies that have examined the feasibility of using wearable activity trackers in children and adolescents, either as part of an intervention or as a stand-alone study. While these technologies offer significant promise for increasing physical activity levels and benefiting health promotion practice [14] or, potentially, clinically relevant outcomes in healthy and clinical populations [27], it is important to establish whether wearable activity trackers are a practical tool for youth to use regardless of the setting. Overall, the results from this review suggest that such devices were viewed positively by youth and their parents [26-29], and that they appreciated the devices' range of functions, which included the tracking of physical activity. Ease of use, comfort, and aesthetics were important to the participants [28,29], which is an interesting point to note given that such devices are unlikely to have been developed with youth in mind [35]. Such factors have been previously identified as important and potential barriers to use in adults [37,38]. Interestingly, while there was some evidence that youth used their wearable activity tracker to

compete with (boys) or support (girls) each other [28], which has also been observed in adults [37], few studies noted concerns over the accuracy of the devices. Schaefer and colleagues found that, while some children did test the accuracy of the monitors [28], this did not appear to influence use. Of potentially greater concern, access to technology was identified as a potential issue in low-socioeconomic-status youth, such as the ability to sync and access data at home [28]. This supports a previous study that found that the use of a mobile phone app that enabled boys from low-socioeconomic areas to track their goals and behavior was moderate, due to their prioritizing data for entertainment over the app [39]. As such, we recommend that researchers examine the feasibility and acceptability of different devices, and we suggest that youth are interested in using wearable activity trackers and that these devices are feasible for use in both clinical and healthy populations. However, it must be noted that there are issues concerning the age required (≥13 years old) to hold an account associated with wearable activity trackers that may preclude their use with children, unless alternative feedback mechanisms are employed (eg, researcher-led feedback [25,27]). Identifying how best to integrate the functions and features of wearable activity trackers into physical activity interventions, in order to maximize their potential within the population of interest, should be a future area of research.

Limitations

There are some limitations in this review that warrant attention. First, this review highlights the surprisingly small number of studies that have used wearable activity trackers in youth despite their widespread prevalence in daily life. This makes it difficult to establish any firm conclusions. Given the pervasiveness of these technologies and widespread appeal, further research is needed to explore these issues in youth in order to inform interventions and public health guidance. In addition, research is needed to establish whether differential effects are observed based on the age of the participants. Second, the quality of the included studies was low, with only 1 study using a randomized controlled design and evaluating intervention effects over a 3-month period. There is a need for more rigorous and robust intervention designs to evaluate the longer-term effectiveness of these devices on youth physical activity levels. Third, while a range of monitors were used in the included studies, some monitors have since been discontinued (eg, PAM, SenseWear) or largely usurped by new models (eg, Fitbit). In addition, the validity and reliability of these devices in youth has not been established.

Conclusions

There is a paucity of literature concerning the effectiveness of wearable activity trackers as a tool for increasing children's and adolescents' physical activity levels. Additionally, little research has documented the feasibility of such technology in youth. While there are some preliminary data to suggest that wearable activity trackers are feasible and may have the potential to increase physical activity levels through self-monitoring and goal setting in youth, there is a clear need for more research to examine these issues in youth using robust studies with longer measurement periods. Given the constant changes in the wearable technology market (eg, newer devices and models are

regularly available), research should primarily focus on the use of the device as a tool for changing behavior in youth in interventions. Focusing on features common to different wearable device brands (eg, self-monitoring displays, accompanying apps, and biofeedback features) may be important for establishing the feasibility of these technologies in youth rather than that of the individual monitors per se. Based on this review, feasibility research should establish how youth engage with this technology, whether adherence and engagement are sustained or change over time, and whether effects vary based on age and sex. Such information will inform the development of future interventions and identify how to maximize the potential contribution that such pervasive technologies could make to physical activity promotion in youth.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Search strategy for each database.

[PDF File (Adobe PDF File), 23KB - mhealth_v4i4e129_app1.pdf]

References

1. Kohl HW, Craig CL, Lambert EV, Inoue S, Alkandari JR, Leetongin G, Lancet Physical Activity Series Working Group. The pandemic of physical inactivity: global action for public health. Lancet 2012 Jul 21;380(9838):294-305. [doi: 10.1016/S0140-6736(12)60898-8] [Medline: 22818941]

2. Andersen LB, Harro M, Sardinha LB, Froberg K, Ekelund U, Brage S, et al. Physical activity and clustered cardiovascular risk in children: a cross-sectional study (The European Youth Heart Study). Lancet 2006 Jul 22;368(9532):299-304. [doi: 10.1016/S0140-6736(06)69075-2] [Medline: 16860699]

3. Janssen I, Leblanc AG. Systematic review of the health benefits of physical activity and fitness in school-aged children and youth. Int J Behav Nutr Phys Act 2010;7:40 [FREE Full text] [doi: 10.1186/1479-5868-7-40] [Medline: 20459784]

4. Hallal PC, Victora CG, Azevedo MR, Wells JCK. Adolescent physical activity and health: a systematic review. Sports Med 2006;36(12):1019-1030. [Medline: 17123326]

5. Department of Health. Does your child get 60 minutes of physical activity every day? Make your move - Sit less! Be active for life! Australia's Physical Activity & Sedentary Behaviour Guidelines for Children (5-12 years). 2014. URL: http://www.health.gov.au/internet/main/publishing.nsf/content/health-pubhlth-strateg-phys-act-guidelines [accessed 2016-11-17] [WebCite Cache ID 6m693PR1O]

6. U.S. Department of Health & Human Services. Physical activity guidelines for Americans. 2008. URL: https://health.gov/paguidelines/ [accessed 2016-11-18] [WebCite Cache ID 6m69Tq6UI]

7. Chief Medical Officer. UK physical activity guidelines: children and young people. Department of Health. 2011. URL: https://www.gov.uk/government/publications/uk-physical-activity-guidelines [accessed 2016-08-16] [WebCite Cache ID 6L0w5lLY5]

8. Schranz N, Olds T, Cliff D, Davern M, Engelen L, Giles-Corti B, et al. Results from Australia's 2014 Report Card on Physical Activity for Children and Youth. J Phys Act Health 2014 May;11 Suppl 1:S21-S25. [doi: 10.1123/jpah.2014-0164] [Medline: 25426909]

9. Dentro KN, Beals K, Crouter SE, Eisenmann JC, McKenzie TL, Pate RR, et al. Results from the United States' 2014 Report Card on Physical Activity for Children and Youth. J Phys Act Health 2014 May;11 Suppl 1:S105-S112. [doi: 10.1123/jpah.2014-0184] [Medline: 25426905]

10. Standage M, Wilkie HJ, Jago R, Foster C, Goad MA, Cumming SP. Results from England's 2014 Report Card on Physical Activity for Children and Youth. J Phys Act Health 2014 May;11 Suppl 1:S45-S50. [doi: 10.1123/jpah.2014-0165] [Medline: 25426913]

11. 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]

12. Corder K, van Sluijs EMF, Goodyer I, Ridgway CL, Steele RM, Bamber D, et al. Physical activity awareness of British adolescents. Arch Pediatr Adolesc Med 2011 Jul;165(7):603-609. [doi: 10.1001/archpediatrics.2011.94] [Medline: 21727272]

13. Lubans DR, Plotnikoff RC, Miller A, Scott JJ, Thompson D, Tudor-Locke C. Using pedometers for measuring and increasing physical activity in children and adolescents: the next step. Am J Lifestyle Med 2014 Jul 10;9(6):418-427. [doi: 10.1177/1559827614537774]

14. Fritz T, Huang E, Murphy G, Zimmerman T. Persuasive technology in the real world: a study of long-term use of activity sensing devices for fitness. Presented at: CHI Conference on Human Factors in Computing Systems; April 26-May 1, 2014; Toronto, ON, Canada. [doi: 10.1145/2556288.2557383]

XSL•FO
RenderX

15. Pureprofile. Wearables work - Australians more active with fitness trackers. 2015. URL: http://businesses.pureprofile.com/news/wearables-work-australians-more-active-with-fitness-trackers [accessed 2016-11-18] [WebCite Cache ID 6m6AFkxZf]

16. Ledger D, McCaffrey D. Jan. Inside wearables: how the science of human behavior change offers the secret to long-term engagement. Endeavour Partners LLC. 2014 Jan. URL: http://endeavourpartners.net/assets/Wearables-and-the-Science-of-Human-Behavior-Change-EP4.pdf [accessed 2016-11-18] [WebCite Cache ID 6m74dmoVY]

17. 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;16(8):e192 [FREE Full text] [doi: 10.2196/jmir.3469] [Medline: 25131661]

18. 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;12(1):159 [FREE Full text] [doi: 10.1186/s12966-015-0314-1] [Medline: 26684758]

19. Lewis ZH, Lyons EJ, Jarvis JM, Baillargeon J. Using an electronic activity monitor system as an intervention modality: a systematic review. BMC Public Health 2015;15:585 [FREE Full text] [doi: 10.1186/s12889-015-1947-3] [Medline: 26104189]

20. Salmon J. Novel strategies to promote children's physical activities and reduce sedentary behavior. J Phys Act Health 2010 Nov;7 Suppl 3:S299-S306. [Medline: 21116014]

21. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA Statement. Open Med 2009;3(3):e123-e130 [FREE Full text] [Medline: 21603045]

22. Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. Cochrane Database Syst Rev 2011 Aug 10(8):MR000026. [doi: 10.1002/14651858.MR000026.pub2] [Medline: 21833989]

23. Hoy MB. Personal activity trackers and the quantified self. Med Ref Serv Q 2016;35(1):94-100. [doi: 10.1080/02763869.2016.1117300] [Medline: 26794199]

24. Viswanathan M, Ansari M, Berkman N, Chang S, Hartling L, McPheeters L, et al. Assessing the risk of bias of individual studies in systematic reviews of health care interventions. Agency for Healthcare Research and Quality Methods Guide for Comparative Effectiveness Reviews 2012 Mar 08 [FREE Full text] [Medline: 22479713]

25. Hayes LB, Van Camp CM. Increasing physical activity of children during school recess. J Appl Behav Anal 2015 Sep;48(3):690-695. [doi: 10.1002/jaba.222] [Medline: 26119136]

26. Slootmaker SM, Chinapaw MJM, Seidell JC, van Mechelen W, Schuit AJ. Accelerometers and Internet for physical activity promotion in youth? Feasibility and effectiveness of a minimal intervention [ISRCTN93896459]. Prev Med 2010 Jul;51(1):31-36. [doi: 10.1016/j.ypmed.2010.03.015] [Medline: 20380847]

27. Hooke MC, Gilchrist L, Tanner L, Hart N, Withycombe JS. Use of a fitness tracker to promote physical activity in children with acute lymphoblastic leukemia. Pediatr Blood Cancer 2016 Apr;63(4):684-689. [doi: 10.1002/pbc.25860] [Medline: 26756736]

28. Schaefer SE, Ching CC, Breen H, German JB. Wearing, thinking, and moving: testing the feasibility of fitness tracking with urban youth. Am J Health Educ 2016 Jan 08;47(1):8-16. [doi: 10.1080/19325037.2015.1111174]

29. Schaefer SE, Van Loan M, German JB. A feasibility study of wearable activity monitors for pre-adolescent school-age children. Prev Chronic Dis 2014;11:E85 [FREE Full text] [doi: 10.5888/pcd11.130262] [Medline: 24854236]

30. Tudor-Locke C, Craig CL, Beets MW, Belton S, Cardon GM, Duncan S, et al. How many steps/day are enough? for children and adolescents. Int J Behav Nutr Phys Act 2011;8:78 [FREE Full text] [doi: 10.1186/1479-5868-8-78] [Medline: 21798014]

31. Nguyen S, Häcker A, Henderson M, Barnett T, Mathieu M, Pagani L, et al. Physical activity programs with post-intervention follow-up in children: a comprehensive review according to categories of intervention. Int J Environ Res Public Health 2016;13(7) [FREE Full text] [doi: 10.3390/ijerph13070664] [Medline: 27376315]

32. Salmon J, Booth ML, Phongsavan P, Murphy N, Timperio A. Promoting physical activity participation among children and adolescents. Epidemiol Rev 2007;29:144-159. [doi: 10.1093/epirev/mxm010] [Medline: 17556765]

33. Williams SL, French DP. What are the most effective intervention techniques for changing physical activity self-efficacy and physical activity behaviour—and are they the same? Health Educ Res 2011 Apr;26(2):308-322 [FREE Full text] [doi: 10.1093/her/cyr005] [Medline: 21321008]

34. Locke EA, Latham GP. Building a practically useful theory of goal setting and task motivation: a 35-year odyssey. Am Psychol 2002 Sep;57(9):705-717. [Medline: 12237980]

35. Clawson J, Pater J, Miller A, Mynatt E, Mamykina L. No longer wearing: investigating the abandonment of personal health-tracking technologies on Craigslist. Presented at: UbComp'15; September 7-11, 2015; Osaka, Japan. [doi: 10.1145/2750858.2807554]

36. Gouveia R, Karapanos E, Hassenzahl M. How do we engage with activity trackers? A longitudinal study of study of Habito. Presented at: UbComp'15; September 7-11, 2015; Osaka, Japan.

37. Harrison D, Marshall P, Bianchi-Berthouze N, Bird J. Activity tracking: garriers, workarounds and customisation. Presented at: UbComp'15; September 7-11, 2015; Osaka, Japan p. 617-621.

38. Ledger D. Inside Wearables - Part 2: A Look at the Uncertain Future of Smart Wearable Devices, and Five Industry Developments That Will Be Necessary for Meaningful Mass Market Adoption and Sustained Engagement. Cambridge, MA: Endeavour Partners LLC; 2014.

XSL·FO
RenderX

39.   Smith JJ, Morgan PJ, Plotnikoff RC, Dally KA, Salmon J, Okely AD, et al. Smart-phone obesity prevention trial for
       adolescent boys in low-income communities: the ATLAS RCT. Pediatrics 2014 Sep;134(3):e723-e731 [FREE Full text]
       [doi: 10.1542/peds.2014-1012] [Medline: 25157000]

Abbreviations

MVPA: moderate- to vigorous-intensity physical activity
PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

XSL•FO
RenderX

Corrigenda and Addenda

Correction of: Sleep Quality Prediction From Wearable Data Using Deep Learning

Aarti Sathyanarayana[1], MSc; Shafiq Joty[1], PhD; Luis Fernandez-Luque[1], PhD; Ferda Ofli[1], PhD; Jaideep Srivastava[1], PhD; Ahmed Elmagarmid[1], PhD; Teresa Arora[2], PhD; Shahrad Taheri[2], MBBS, PhD

[1]Qatar Computing Research Institute, Hamad Bin Khalifa University, Qatar Foundation, Doha, Qatar
[2]Department of Medicine, Weill Cornell Medical College in Qatar, Qatar Foundation, Doha, Qatar

Corresponding Author:
Luis Fernandez-Luque, PhD
Qatar Computing Research Institute
Hamad Bin Khalifa University, Qatar Foundation
HBKU Research Complex
Doha, 5825
Qatar
Phone: 974 50173040
Fax: 974 44540630
Email: lluque@qf.org.qa

Related Article:

Correction of: http://mhealth.jmir.org/2016/4/e125/

The authors of "Sleep Quality Prediction From Wearable Data Using Deep Learning" (JMIR Mhealth Uhealth 2016;4(4):e125) have mistakenly written "linear regression" instead of "logistic regression" in the following 9 instances. Linear regression cannot be used for binary classification, and was not utilized in this methodology. In all the tables and in most of the cases authors wrote correctly "logistic regression," however this typo needs to be fixed to avoid confusion in the following 9 instances:

1. In the results section within the abstract, the sentence "More specifically, the deep learning methods performed better than traditional linear regression" should be changed to "More specifically, the deep learning methods performed better than traditional logistic regression."

2. In the results section within the abstract, the sentence "CNN had the highest specificity and sensitivity, and an overall area under the receiver operating characteristic (ROC) curve (AUC) of 0.9449, which was 46% better as compared with traditional linear regression (0.6463)." should be changed to "CNN had the highest specificity and sensitivity, and an overall area under the receiver operating characteristic (ROC) curve (AUC) of 0.9449, which was 46% better as compared with traditional logistic regression (0.6463)."

3. The first subheading of the results section in the body of the paper, "Comparison Between Deep Learning and Linear Regression" should be changed to "Comparison Between Deep Learning and Logistic Regression."

4. Under the first subheading in the results section of the text, the first sentence, "As shown in Table 1 and Figure 6, the performance of the linear regression in the metrics previously explained performed worse than the models based on deep learning" should be changed to "As shown in Table 1 and Figure 6, the performance of the logistic regression in the metrics previously explained performed worse than the models based on deep learning."

5. Under the first subheading in the results section of the text, the second sentence, "Only the simple RNN performed worse than linear regression in both F1-score (harmonic mean of precision and recall) and accuracy" should be changed to "Only the simple RNN performed worse than logistic regression in both F1-score (harmonic mean of precision and recall) and accuracy."

6. Under the first subheading in the results section of the text, the third sentence, "As shown in Table 1, the AUC of the linear regression model was low. The AUC value for LR was 0.6463, which was close to 0.5 (equivalent to a random prediction)" should be changed to "As shown in Table 1, the AUC of the logistic regression model was low. The AUC value for LR was 0.6463, which was close to 0.5 (equivalent to a random prediction)."

7. Under the subheading Prinicpal Findings in the discussion section, "This was the case of linear regression, which had a high sensitivity but a specificity of 0.3, meaning that in such models many 'poor sleeps' would have been wrongly classified

as good sleep" should be changed to "This was the case of logistic regression, which had a high sensitivity but a specificity of 0.3, meaning that in such models many 'poor sleeps' would have been wrongly classified as good sleep."

8. Under the subheading of relevance of findings in the discussion section, "Furthermore, the good results of deep learning showed that raw accelerometer data had more 'signal' regarding sleep quality, which traditional models such as linear regression are not able to capture right now" should be changed to "Furthermore, the good results of deep learning showed that raw accelerometer data had more 'signal' regarding sleep quality, which traditional models such as logistic regression are not able to capture right now."

9. Under the subheading Limitations in the discussion section, "Other techniques such as linear regression can provide insights on which features contribute to the prediction" should be changed to "Other techniques such as logistic regression can provide insights on which features contribute to the prediction."

All these alterations have been made in the online version of the paper on the JMIR website on November 25, 2016, together with publishing this correction notice. Because these were made after submission to PubMed and other full-text repositories, the correction notice has been submitted to PubMed, and the original paper has been resubmitted to PubMed Central. The corrected metadata have also been updated on PubMed manually and were resubmitted to CrossRef.

XSL•FO

RenderX

XSL•FO

RenderX