JMIR mHealth and uHealth

Impact Factor (2024): 5.4
Volume 7 (2019), Issue 5 ISSN 2291-5222 Editor in Chief: Lorraine Buis, PhD, MSI

Contents

Original Papers

Viewpoint

Original Paper

Smartphones as Sleep Duration Sensors: Validation of the iSenseSleep Algorithm

Matteo Ciman[1], MSc, PhD; Katarzyna Wac[1], BSc, MSc, PhD

Quality of Life Technologies Lab, University of Geneva, Carouge, Switzerland

Corresponding Author:
Katarzyna Wac, BSc, MSc, PhD
Quality of Life Technologies Lab
University of Geneva
Route de Drize 7
Battelle A, UNIGE
Carouge, 1227
Switzerland
Phone: 41 792506308
Email: katarzyna.wac@unige.ch

Abstract

Background: Smartphones are becoming increasingly ubiquitous every day; they are becoming more assimilated into our everyday life, being the last thing used before going to sleep and the first one after waking up. This strong correlation between our lifestyle choices and smartphone interaction patterns enables us to use them as sensors for sleep duration assessment to understand individuals' lifestyle and sleep patterns.

Objectives: The objective of this study was to estimate sleep duration based on the analysis of the users' ON-OFF interaction with their smartphone alone using the iSenseSleep algorithm.

Methods: We used smartwatch sleep assessment data as the ground truth. Results were acquired with 14 different subjects collecting smartwatch and smartphone interaction data for up to 6 months each.

Results: Results showed that based on the smartphone ON-OFF patterns, individual's sleep duration can be estimated with an average error of 7% (24/343) [SD 4% (17/343)] min of the total duration), enabling an estimate of sleep start and wake-up times as well as sleep deprivation patterns.

Conclusions: It is possible to estimate sleep duration patterns using only data related to smartphone screen interaction.

(JMIR Mhealth Uhealth 2019;7(5):e11930) doi:10.2196/11930

KEYWORDS

mobile phone use; mobile health; behavioral research; well being

Introduction

In the last decade, smartphones and mobile or connected devices, in general, are taking on a bigger role in our everyday life. Day by day, the time humans spend interacting with their smartphone is increasing [1], both because these devices can now serve more tasks, thereby helping us along daily life activities (eg, for navigation and communication) and also because they are designed to be engaging. Moreover, smartphones are equipped with several built-in sensors, for example, accelerometer, light sensor, and microphone that can provide valuable data, which can be used to get an insight into an individual's life. For example, smartphones can be instrumented to provide recommendations about lifestyle and to follow a specific exercise regimen [2], understand the user's stress level [3], or just provide overall support for lifestyle choices (eg, for exercise and nutrition) to facilitate better aging [4].

One of the most important aspects of everyday life of individuals that has gained a lot of attention recently is *sleep*. People's feelings and actions throughout the day are strongly correlated with how much they slept (ie, the sleep duration) and how well they slept (ie, the sleep quality). In general, sleep affects personal health. An insufficient amount of sleep can cause fatigue and lack of concentration during the day [5]. Moreover, clinical studies show that poor sleep habits and sleep disorders are related to many serious diseases, including obesity and depression [6-8]. Increasing and converging evidence indicates that much like the majority of other organisms on the planet

[9], biochemistry of the human body varies predictably throughout the day [10], a phenomenon known as the circadian rhythm. It has even been proven that circadian rhythms affect our mood, levels of concentration, digestion, sleep patterns, and much more [11].

The importance of sleep in everyday life has driven researchers to study if, and to which extent, more accessible devices, for example, smartphones or wearable devices, can be used to assess sleep quantity and quality of individuals and, in the long term, help them understand how their sleep-related behaviors could be changed. For these reasons, several developments related to smart, personal, miniaturized, and affordable devices, including smartwatches such as Fitbit [12], Withings [13], and Apple Watch [14] that use an array of embedded accelerometers, temperature, and heart rate sensors to estimate the sleep of the user, will make accurate sleep information, measured at the user's home, available to medical experts.

Moreover, a recent study by the Pew Internet and the American Life Project found that 65% of mobile phone owners (and impressively 90% of teens) sleep with their phone on or near their bed, with many users using their smartphone as an alarm clock [15]. Results presented in this study show how the closeness of the smartphone in everyday life makes this device suitable for understanding the sleep habits of its users. Compared with approaches that require external devices, by only using the smartphone, the cost of entry for a sleep analysis tool is reduced, thus making this information more accessible to everyone.

In this study, we hypothesized that it is possible by using only a smartphone, and in particular, the information related to the users' interactions with the smartphone screen, to understand and estimate their sleep habits. In particular, we compare the sleep duration derived from the smartphone interaction patterns with the sleep duration estimated by a smartwatch worn by the healthy study volunteers during the entire day and night. In our research, we leverage the BASIS Peak Smartwatch (by Intel Corp, Santa Clara, CA, USA) [16]. In this study, we show that it is possible to estimate sleep duration of each user, based solely on the smartphone interaction datasets, collected longitudinally in a minimally obtrusive and lightweight way. On the one hand, this information can be very useful as a contextual background to better diagnose sleep-related disorders once the individual conducts the sleep lab study. On the other hand, according to the literature [17], this information can be very useful to assess (and potentially mitigate) the risk of developing an illness in the long term (eg, cardiovascular disease and diabetes) correlated with unhealthy sleep patterns. The proposed approach is not intended to facilitate real-time sleep detection and real-time intervention but to facilitate longitudinal assessment and behavior change.

Methods

Study Design

The goal of our research is to understand how and to what extent it is possible to use only smartphones, and in particular, the data related to the interactions of the user with the smartphone screen, to evaluate the sleeping patterns of the user and have an insight about his or her routines (eg, sleep deprivation) that could lead to a disease later in the individuals' life.

Sleep Logs

The first point necessary to address is related to the ground truth data, for example, sleep logs of each individual that will later be compared with the results of the smartphone-based computational modeling of the sleep behavior. The first option for annotating data was the usage of a diary where users could annotate sleep moments before going to sleep and when waking up [18,19] or daily reconstruction methods [20] to interview the users with the aim of reconstructing their daily habits, including sleep. These methods are prone to possible subjective errors and imprecision stemming from forgetfulness of the participants to complete the logs, memory bias, or providing information about when they went to bed but not necessarily when they started sleeping. For these reasons, we decided to base our research on more reliable data, which could give a precise insight into their sleeping behavior.

We provided the users with a BASIS Peak Smartwatch [16] to wear in their daily life. The Basis Health Tracker (by Intel Corp, Santa Clara, CA, USA) is a wristwatch with an embedded actigraph. Besides the standard features provided by this smartwatch, for example, step counter, calories burned, and heart rate, the important aspect of this device is that it can automatically detect sleep episodes. Given the evidence from the medical literature, BASIS can be considered as a baseline, as it is the closest to the ground truth of sleep assessment than any other self-reported method. The BASIS sleep duration estimation has been successfully evaluated within diverse studies against gold standard polysomnography, for example, by Patel et al [21], who showed no statistically significantly different results for the total sleep time comparing the 2 methods for 40 participants.

The on-board algorithm of the smartwatch provides information about the different sleep phases (rapid eye movement, light, and deep sleep) based on the user's physiological measures; however, we are interested mainly in the following: the start, the end, and the duration of the sleep of each user. Another feature of the BASIS is the ability to identify sleep interruptions during the night. *Interrupted sleep* occurs when a subject wakes up for few minutes and, for example, goes to a bathroom, and then goes back to sleep in less than 15 min. In this case, the sleep duration for the night is reduced by the duration of the interruptions. There is also an option for *unknown values* being provided by the BASIS, which represent the cases in which the BASIS' on-board algorithm is unable to interpret the physiological measures and to classify the type of sleep (typically such a lack of knowledge is due to insufficient heart rate coverage). In such a case, if the set of *unknown values* is between 2 sleep episodes and is less than 15 min, it is also considered as a sleep period; it is discarded otherwise.

Smartphone Logs

This study aims to understand if it is possible to analyze the users' interaction with the smartphone to infer their sleep patterns. For this reason, we were interested in collecting data about the interaction of the user with his or her smartphone,

and, in particular, when the user turns ON and OFF the smartphone screen—assuming that it is the smartphone owner turning ON or OFF the phone; interacting with his or her smartphone and hence, not sleeping.

To collect these data, we instrumented the smartphone with the *mobile Quality of Life* logger (mQoL-log) [22]. This app, developed by the Quality Of Life Technologies (QoL) Group at the University of Geneva (Switzerland, qol.unige.ch) and currently used in the QoL Living Lab (mqol.unige.ch), can collect and register most of the events that take place in the smartphone. For example, it collects time-stamped data about the apps used (eg, Facebook and email), the screen events such as the screen being turned ON or OFF, the physical activity of the users (walking and running) and used network 3G/4G or Wi-Fi performance information. All the data collection and the task of uploading it to the dedicated QoL lab server are made automatically in the app background, and there is no interference with the daily routine of the smartphone use. In this way, the mQoL-log collects data unobtrusively, without affecting the daily life of the users. Despite all the information available and the vast data being collected from the smartphone users, the algorithm presented in this study uses only information related to the *state* of the screen (screen ON and screen OFF), to minimize the amount of data being used and its potential privacy-obtrusiveness, and hence, maximize the user acceptance for the algorithm. In particular, the mQoL-log logs 3 different events: SCREEN ON, SCREEN OFF, and screen PRESENT. SCREEN ON represents the screen being turned ON, PRESENT when the screen is unlocked and the smartphone is ready for interaction with apps, and SCREEN OFF when the screen is turned OFF. It is important to note that the ON state of screen considered in this study only corresponds to user-interaction events, that is, SCREEN ON is recorded only when the user touches the ON button and not when the screen lights up in response to, for example, notifications and without necessarily being initiated by the user.

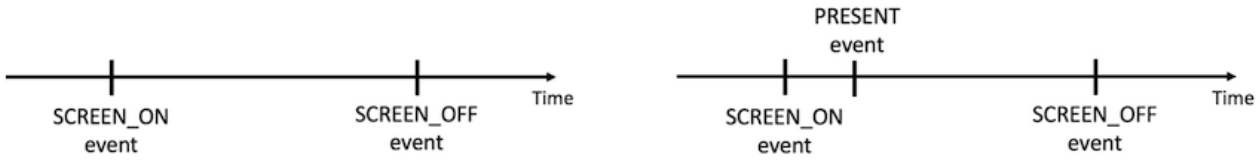
Figure 1 represents the timeline of the events when the screen is turned ON and OFF, or turned ON, unlocked (PRESENT), and turned OFF.

The basic requirement of the algorithm, denoted iSenseSleep, is to identify when the user is sleeping by only analyzing the user's smartphone interaction data, and, in particular, when the screen is turned ON and OFF.

iSenseSleep relies on two consecutive functions: (1) identifying what are all the possible time intervals along 2 consecutive days that could identify as sleep episodes and (2) evaluating these episodes to determine which one is the most probable one to represent the longest (likely overnight) sleep.

The first function works as follows: We denote the *screen event* as a tuple of SCREEN ON and SCREEN OFF events recorded consecutively by the mQoL-log (so, both cases in Figure 1 represent a screen event). Given a list of screen events for 2 consecutive days, the first step of the iSenseSleep algorithm is to identify all tuples, that is, consecutive screen events that are separated by at least 4 hours. Second, the iSenseSleep algorithm evaluates the SCREEN ON events in the morning hours after at least 4 hours since the last SCREEN OFF event. The algorithm reasons if the SCREEN OFF event before the SCREEN ON at the wake-up time is really the event when the user went to sleep (along the evening hours) or an event in the middle of the night when he or she woke up and checked the smartphone. For this reason, the algorithm clusters all possible events around this SCREEN OFF event (if there are other events in the range of 5 min) and picks up the previous SCREEN OFF event (before the one in the middle of the night) that is registered at least 2 hours earlier that night. If this is the case, iSenseSleep assumes the last screen event considered for sleep duration calculation as the last one before the sleep break in the middle of the night.

Figure 1. Different screen events recorded by the mQoL-log (Left: Screen ON and OFF; Right: Screen turned ON, unlocked (PRESENT), and turned OFF). mQoL-log: mobile Quality of Life logger.



The operation of identifying the possible sleep episodes is repeated until all the screen events (SCREEN OFF and SCREEN ON) are analyzed. Once (1) the function of the algorithm provides the list of time intervals relating to sleep episodes, (2) the evaluation function derives 3 scores to provide a likelihood value for all the episodes to understand which one most likely relates to the overnight sleep episode of the user. To do this, the function (2) assigns a score to the different episodes identified by (1) along 2 consecutive days, comparing it with the assumed *time to bed* (eg, how far it is from 22 hours or 10 pm) and the identified time of wake up, for example, how far it is from a *normal* wake-up time (derived based on the entire smartphone log dataset). Function (2) also assigns a score to a sleep episode by evaluating how much time passed between the usual wake-up time during weekends and weekdays, and what is the total duration of the sleep for this episode. On the basis of these 3 values, the function provides an evaluation of all the possible episodes to select the one that identifies the overnight sleep of the user with the highest probability.

User Study and Data Collection

For this research, we recruited both young university students (10) and adults (working mothers, 4), 14 in total, to have 2 different groups of individuals and to understand the behavior and the performance of the algorithm with 2 different daily behaviors and daily usage of the smartphone. In particular, students can be considered as more digitally native [23], hence more attached and closer to their phone, while working mothers

are probably less attached, as they have an entirely different lifestyle from students and are busy with work and caregiving activities. This is also reflected in the fact that we had aimed to recruit 10 working mothers but given the time constraints on the recruitment and the lack of immediate availability of the potential participants, we failed to do so. The study has been approved by the IRB of University of Copenhagen (Denmark) in 2015 under the protocol number 2015-15-0117/519-0019 /15-5000 and took place at the end of 2015-mid 2016 in Copenhagen, Denmark.

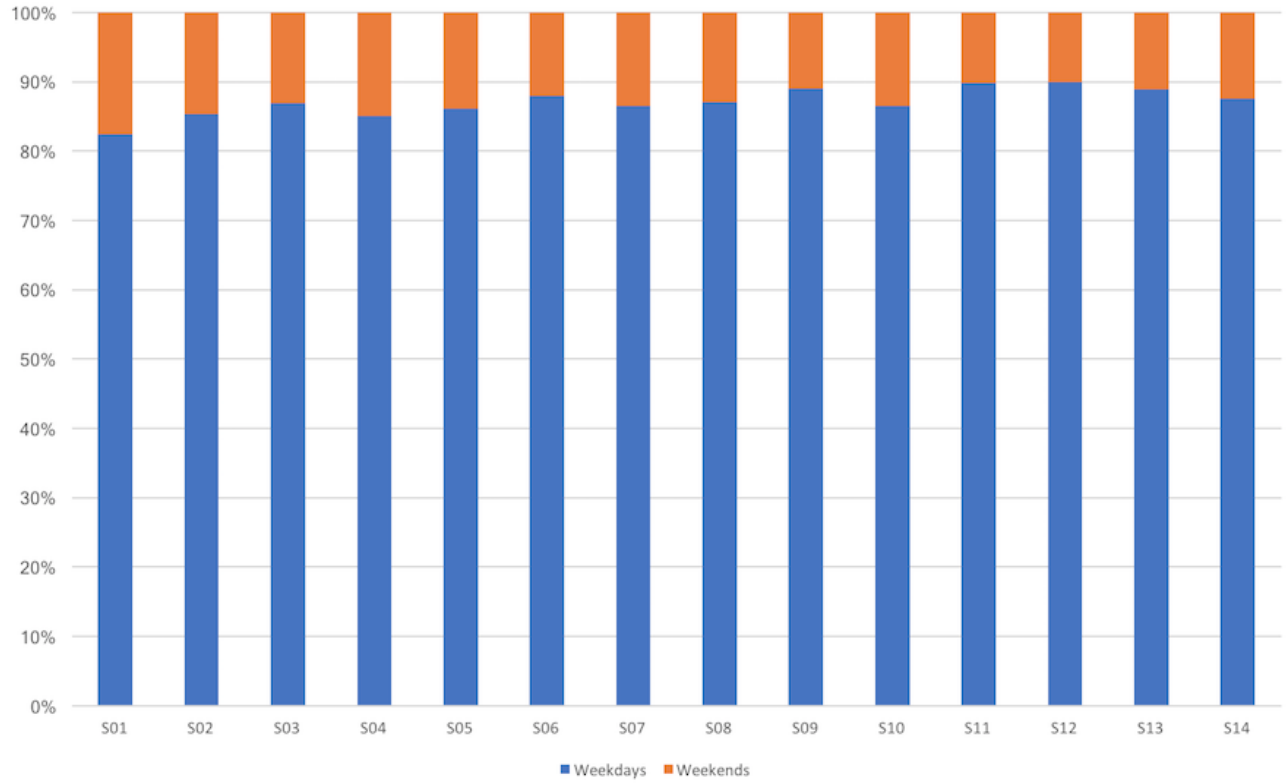
For the women group (S01-S04), we initially recruited 4 working mothers (in their 30s and 40s, caring for at least 1 child), whereas for the student group (S05-S14), we recruited 10 participants (aged between 18 and 30 years, 8 of them are males). For each user, we provided 1 BASIS Peak smartwatch and installed the mQoL-log app on his or her smartphone. The users were encouraged to wear the watch at all times, even for sports and under the shower. The smartphone and smartwatch data were automatically synchronized with the QoL servers without the intervention or interaction of the users.

Participation in the study was free, and participants could leave the study whenever they wanted, in addition to having autonomy of wearing the smartwatch or not. The working mothers group provided an average of a bit more than a month of data—that is, 46 days of data (±23), whereas the students' group—almost 5 months of data—that is, 148 days (±64), with some contributing full 6 months of data. Table 1 provides the total amount of participation days, whereas Figure 2 shows the distribution between weekdays and weekends (Saturday and Sunday).

Table 1. Number of participation days (S01-04: working mothers; S05-S14: students).

Subject ID	Weekdays	Weekend days	Total days
S01	14	3	17
S02	70	12	82
S03	40	6	46
S04	34	6	40
S05	187	30	217
S06	168	23	191
S07	147	23	170
S08	167	15	192
S09	145	18	163
S10	115	18	133
S11	114	13	127
S12	45	5	127
S13	16	2	18
S14	190	27	217

Figure 2. Percentage distribution of week or weekend days for each study subject.



Trial Registration

Study protocol "Understanding Diverse Factors Influencing Individuals' Sleep Quality And Smartphone-Based Ubiquitous Assessment Of Individual Sleep Patterns": Protocol 2015-15-0117/519-0019/15-5000 approved by the Institutional Review Board, Faculty of Science, University of Copenhagen, Denmark; Protocol Director: Professor K Wac, Active since 2015.

Results

Sleep Duration

The sleep duration derived by the iSenseSleep algorithm presented in the previous section has been compared with the ground truth data provided by the BASIS smartwatch. The average sleep duration, as measured by the BASIS and the iSenseSleep algorithm, are provided for both the working mothers' group and the students' group (Table 2).

Recommended Sleep Duration

Figure 3 provides the data calculated using iSenseSleep algorithm and the BASIS data compared with the recommended amount of sleep hours for healthy adults [24], as we will discuss further in this paper.

We additionally compared differences between the beginning and the end of the sleep, to have a closer look in how much time passes between when the person stops using the smartphone (ie, when the last smartphone event (ON or OFF or PRESENT) is recorded) and he or she falls asleep according to the smartwatch. We report the cumulative average difference between the BASIS smartwatch and iSenseSleep-derived time for the beginning and the end of the sleep. Values are calculated considering the absolute value of the difference. Table 3 shows the average error (SD) of the iSenseSleep algorithm for sleep duration, sleep beginning, and sleep end times for working mothers' and the students' group.

As we can see from the results, the error made by the iSenseSleep algorithm is on average approximately 13% of the sleep duration value for the working mothers' group (corresponding to average of 53 min), or approximately 7% for the students' group (corresponding to average of 24 min). If we divide the days between weekdays and weekends, the iSenseSleep error is similar for weekdays (12%±10% for working mothers and 7%±6% for students), whereas it increases to 13% when considering only weekends for working mothers and to 9% for the students' group. Overall, the sleep duration error is persistent for working mothers, while it is lower for weekdays and larger for weekends when considering the student population.

The iSenseSleep error for the start time and the end time of the sleep is slightly higher both for the women and the students' group. In particular, the starting time error for the working mothers' datasets is about 108 min (26% of the entire sleep duration), whereas for the students, this error is about 79 min (18%) on average for all days together. During the weekdays, the error is the same for working mothers, and it increases by an additional 18 min for students—resulting in 97 min (23%). The sleep start time error is the same for the working mothers across the whole dataset, whereas it is higher for the students at weekends.

Table 2. Average sleep duration comparing the BASIS smartwatch and the iSenseSleep algorithm.

Group and subject ID	Sleep duration (min), mean (SD)					
	All days		Weekdays		Weekend days	
	BASIS	iSenseSleep	BASIS	iSenseSleep	BASIS	iSenseSleep
Mothers						
S01	393 (10)	418 (14)	393 (13)	409 (15)	389 (20)	458 (18)
S02	402 (15)	507 (16)	406 (17)	509 (25)	378 (23)	493 (19)
S03	455 (13)	377 (34)	464 (10)	377 (29)	392 (30)	379 (17)
S04	446 (15)	444 (25)	437 (18)	433 (17)	503 (15)	506 (23)
Students						
S05	429 (15)	481 (23)	436 (15)	483 (20)	386 (29)	465 (15)
S06	473 (16)	478 (25)	474 (10)	485 (15	461 (22)	429 (24)
S07	377 (22)	377 (24)	378 (30)	376 (15	374 (23)	384 (37)
S08	450 (13)	454 (35)	449 (13)	450 (23	453 (35)	479 (17)
S09	482 (14)	459 (24)	486 (29)	463 (18	443 (35)	429 (27)
S10	478 (16)	446 (36)	481 (24)	441 (37	453 (27)	476 (19)
S11	409 (19)	378 (45)	417 (17)	381 (23	344 (34)	350 (24)
S12	417 (24)	374 (16)	408 (27)	376 (29	501 (24)	362 (16)
S13	462 (16)	346 (25)	454 (24)	355 (35	519 (12)	272 (23)
S14	452 (29)	426 (39)	450 (23)	424 (34	468 (29)	434 (26)

Figure 3. Average sleep duration by the BASIS smartwatch and iSenseSleep versus the recommended sleep per night (gray line: 7 hours=420 min; yellow line: 9 hours=540 min).

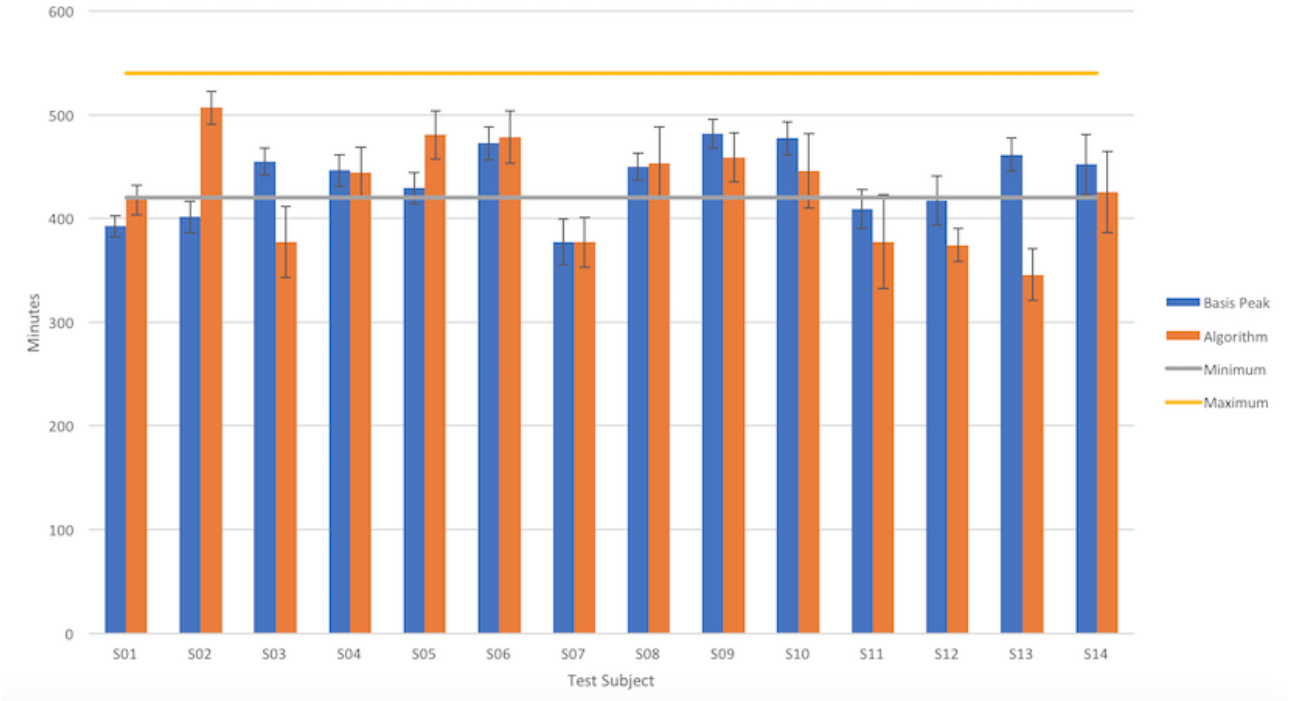


Table 3. Sleep statistics' differences between the BASIS smartwatch and the iSenseSleep algorithm.

Group	All days, mean (SD)		Weekdays, mean (SD)		Weekend days, mean (SD)	
	Min	Percentage	Min	Percentage	Min	Percentage
Working mothers						
Sleep duration	53 (41)	13 (10)	53 (43)	12 (10)	50 (45)	13 (12)
Sleep start time difference	108 (28)	28 (8)	108 (32)	26 (9)	108 (43)	27 (12)
Sleep end time difference	83 (28)	20 (9)	86 (34)	21 (9)	65 (41)	17 (9)
Students						
Sleep duration	24 (17)	7 (4)	32 (27)	7 (6)	41 (40)	9 (8)
Sleep start time difference	79 (16)	18 (3)	78 (17)	18 (3)	97 (18)	23 (6)
Sleep end time difference	68 (22)	17 (5)	72 (25)	16 (5)	84 (47)	20 (13)

The end time error is similar to the start time error but higher than the error for the sleep duration. In particular, the working mothers' group has an average error of 83 min (20%) considering all the days together, or 86 min (21%) for weekdays and 65 min (17%) for weekends. On the other hand, for students, the average error is 68 min (17%) considering all the days together, and 72 min (16%) for weekdays and 84 min (20%) during weekends. The sleep end time error is higher for the weekdays for the working mothers than at the weekends, whereas the opposite is true for the students.

In the following section, we provide a discussion about these results and how they can be interpreted and leveraged when designing technologies helping individuals to develop healthier sleep habits.

Discussion

Principal Findings

Results show that based on the smartphone ON-OFF patterns, an individual's sleep duration can be estimated with an average error of 24 ± 17 min (7% ± 4% of the total duration), enabling estimates of sleep start and wake-up times as well as sleep deprivation patterns.

To evaluate the accuracy of the iSenseSleep algorithm, we calculated the statistical significance between the average duration of the sleep calculated by the algorithm and the ground truth data provided by the BASIS smartwatch. Assuming a normal distribution of the datasets, if the *P* value is larger than .05, it indicates that there is no significant difference between values provided by the two, and therefore, the iSenseSleep algorithm is adequate. We performed a paired, one-tailed *t* test to compare the 2 values of sleep duration. For the all days, for the working mothers' group, *P*=.35, whereas for the students' group *P*=.14. We conclude that the iSenseSleep algorithm that is based only on the smartphone interaction data analytics can be used to estimate the sleep duration within a group of subjects. Moreover, *t* test results conducted for each single study subject show the algorithm's performance within subject. Results are provided in Table 4.

As we can see from Table 4, for less than 50% of participants (6 over 14), the average sleep duration estimated by iSenseSleep is statistically different from the ground truth data; this is mostly because of iSenseSleep underestimating the sleep duration of the individual. Given this result, it is possible to make several conclusions about the performance and validity of our algorithm.

First of all, using only the data related to screen events of the smartphone, for example, SCREEN ON and SCREEN OFF, in this study, we have shown that it is possible to provide an estimate of the smartphone owners' sleep duration with a small probability of the estimation error being at most around 100 min.

In addition, considering the sleep duration recommendations [24] for the adult category, from the data calculated using the algorithm, it is possible to see that only a few of the users sleep on average enough during the night (eg, S03, S06, S09, and S10), whereas all the others lack sleep. Some of them, for example, S01, S02, S07, and S11, may even suffer from sleep deprivation and should increase the amount of sleep time. For smartphone users who are likely to rely on their phones more, sleep deprivation is very likely to be captured by the iSenseSleep (S01, S07, and S11), despite the sleep duration's estimation error. For others, such as S03, S12, or S13, the feedback to the user could be to stay away from interacting with their phone during the night (eg, implementing automatic switching OFF function) because even if they are sleeping long enough (as underestimated by the iSenseSleep algorithm), the night phone *sleep breaks* lasting more than 5 min are not healthy.

Moreover, there are other aspects of the algorithm that are interesting. First of all, the accuracy of the algorithm is higher when calculating the duration of the sleep than when evaluating the individual's sleep start or end time. This result can be explained by the fact that usually when people wake up, before using the phone or turning it ON, they stay in bed without sleeping and just wait for the right moment to get up. Moreover, when going to sleep, even if the smartphone is one of the last things an individual is interacting with, there is an amount of time that passes before the person will effectively fall asleep. These 2 conditions make the identification of the moment of sleep start and wake-up times more challenging than the calculation of the duration of the sleep, which is then shifted in time with respect to sleep start or end time yet adequately estimates the duration.

Table 4. Average sleep duration by the BASIS smartwatch versus iSenseSleep algorithm, and statistical significance tests.

Subject ID	Min, mean (SD)		*P* value[b]	Algorithm estimate (Under or over)	Sleep deprived?	
	BASIS	iSenseSleep			BASIS	Algorithm
S01	393 (10)	418 (14)	.31	__[a]	Yes	No
S02	402 (15)	507 (16)	<.001	Over	Yes	No
S03	455 (13)	377 (34)	<.001	Under	No	Yes
S04	446 (15)	444 (25)	.46	—	No	No
S05	429 (15)	481 (23)	<.001	Over	No	No
S06	473 (16)	478 (25)	.27	—	No	No
S07	377 (22)	377 (24)	.50	—	Yes	Yes
S08	450 (13)	454 (35)	.38	—	No	No
S09	482 (14)	459 (24)	.06	—	No	No
S10	478 (16)	446 (36)	.23	—	No	No
S11	409 (19)	378 (45)	.002	Under	Yes	Yes
S12	417 (24)	374 (16)	.20	—	No	Yes
S13	462 (16)	346 (25)	.003	Under	No	Yes
S14	452 (29)	426 (39)	.002	Under	No	No

[a]The iSenseSleep algorithm estimates sleep duration accurately (does not underestimate or overestimate)

[b]Results were deemed statistically significant at $P<.05$: the algorithms differ.

The second aspect we can see from the results is a comparison between weekdays and weekends. Both for the working mothers' group and the students' group, the accuracy of the algorithm during weekdays is higher with respect to the accuracy during weekends. This may come from 2 different lifestyle choice aspects. The first one is related to the fact that during the weekends, lifestyle of each person is less *standard* and *coherent* with respect to the rest of the days, such as socializing in the evening, eating together outside of the house, etc, thus strongly reducing the amount of time spent with the smartphone. Interestingly enough, mothers have less error for the wake-up time at weekend, than at the weekdays—meaning that they must pick up their phone early at the weekend to, for example, leave the house with the family. In weekdays, they may be occupied with preparing the children alone—before they pick up their phone when leaving the house. In contrast, the students pick up the phone earlier at the weekdays (exhibiting smaller error for the wake-up times)—most likely being late for school or work, whereas leaving it longer behind at the weekends (most likely staying longer at home before leaving for shopping or weekend activities). On the other hand, the number of weekends is much lower than the number of weekdays in our dataset, thus reducing the accuracy of the algorithm because there is fewer data available to compare with the ground truth.

Comparing the 2 user groups, what is clear is that the accuracy of the algorithm is higher when considering the students' group, compared with the working mothers' group. Despite the fact that the number of working mothers is lower than the number of students, what influences the accuracy of the algorithm is probably the users' lifestyle and the closeness to the smartphone. In particular, students can be considered more digitally native than working mothers, thus making their last interaction with the smartphone closer to their sleep time. Working mothers may have family commitments and make behavioral choices resulting in a different relation between their sleep and their smartphone usage.

Limitations

iSenseSleep provides an estimation of the sleep duration patterns of its user; however, it has some study limitations as well as algorithm-specific limitations that can reduce its total accuracy.

Overall, the limitation of the algorithm is that it may not be representative for all populations and be more suitable for *digitally native* populations [23]. Overall, given the small sample of working mothers, the results gathered for this population are rather indicative and cannot be conclusive. We admit that limitation and at the same time indicate that, overall, as mobile users become increasingly more attached to their phones, the algorithm will be able to provide more accurate sleep assessments for a larger variety of populations than initially planned for, for example, for older users relying on smartphone for their daily life tasks, including using their smartphone as an alarm clock [15].

Somehow related to this, a limitation of the algorithm may stem from the choice of baseline method for its comparison. Namely, leaving out a choice of the smartwatch datasets as a baseline, we could have selected other, more smartphone-related baseline methods. In the related work section, we indicate the other smartphone or HCI-analysis methods with an error of 42 min (Chen et al [25]) or 45 min (Abdullah et al [26]) for sleep duration; these methods leveraged microphone and luminosity sensors besides, for example, charging patterns. In contrast, in this work, we leverage only ON or OFF button of the smartphone without privacy-sensitive data sources such as

microphone. In addition, our results are supported by 14 participants engaged in data collection of up to 6 months compared with 8 people 1 week per person study (Chen et al) or a study with 9 persons for 3 months (Abdullah et al). Therefore, even if the choice of baseline may seem to be a limitation, we claim that the choice of validated BASIS smartwatch data is motivated toward this end.

As the first limitation related to the algorithm itself, we recognize that the algorithm is suitable for modeling sleep duration and sleep patterns of users who study or work in conventional daylight hours [27], and it is not suitable for modeling sleep duration and sleep patterns of users who study or work on shifts or along unconventional hours (7% of population of employed adults in Europe and 16% in the United States [27]), who do not sleep during *standard* sleeping times but whenever they can depending on working shifts. This problem is not easy to solve using only the screen events of the smartphone, as the algorithm assumes sleeping episodes occurring during the night (after 10 pm or 22 hours), thus penalizing the possible sleeping episodes that are not within this time range. On the other hand, removing the penalization of the possible sleep episodes that are far from the night-sleep times imposes to calculate the likelihood function using only, for example, the length of the possible sleep duration, hence strongly reducing the accuracy of the algorithm in standard conditions and sleeping patterns. For example, imagine a primary school teacher not touching the phone over consecutive 6 or 8 hours of work time, which would end up being assessed as sleeping along the day, while completely inaccurate. One possible way to overcome this limitation is to introduce a short user profile that may calibrate the iSenseSleep given the individual's self-declared *standard* sleeping times.

The second limitation is related to the usage of the smartphone by the user and is something that is already highlighted by the accuracy of the algorithm during the weekends or with the working mothers' group. In particular, if the individual is not very attached to the smartphone, meaning that the smartphone is not used frequently along the day, especially in the evenings, the algorithm overestimates the time the user goes to bed and underestimates the time when he or she wakes up, as there is a non-negligible delay between the last and the first smartphone interaction versus when the user goes to bed and wakes up. To mitigate this problem, one possible solution is to use other sensors available on the smartphone, such as, for example, the accelerometer and built-in activity recognition, or sense if the phone is charging, to understand if the smartphone, and therefore the user, is potentially still awake and moving around.

Comparison With Previous Work

This paper aims to understand in which way inclusion of a smartphone in our everyday life activities can be leveraged to understand our sleep patterns. For example, Dey et al [28] conducted an empirical experiment estimating how close the smartphone is to its user along daily life activities. With this study, they showed that the smartphone is at the room distance for almost 90% of the time and that it is possible to predict the proximity level of the smartphone with about 80% accuracy with features simple to collect and model on the smartphone itself (eg, Wi-Fi AP name). This and studies by Patel et al [29,30] show how smartphones are getting closer to their users day by day; hence, they can be leveraged to longitudinally and in a minimally obtrusive and lightweight way sense users' lifestyle and infer some aspects (including sleep) without users' intervention or explicit data input.

As sleep is one of the most important aspects of our life, several researchers have focused their attention on how it is possible to understand and infer sleep patterns of smartphone users. For example, Abdullah et al [26] showed how phone usage patterns could be used to detect and predict individual daily variations indicative of temporal preference, sleep duration, and sleep deprivation. They followed 9 participants for 97 days and collected ground truth data using a manual journal where participants annotated their sleep moments during the night. An average sleep duration error among all the participants was about 45 min (10% of the average sleep duration).

In addition, Lane et al [31] assessed the sleep, physical activity, and social interactions leveraging several smartphone sensors; for sleep, they leverage phone charging patterns, accelerometer, as well as microphone data to understand the noise level of the environment the user is in and to classify the activity as sleeping (or not). Lane et al used logistic regression models for sleep duration estimate, and the results were provided only for 5 persons, each collecting 1-week of data via self-reports. The error with respect to the sleep duration was reaching up to 1.5 hours.

Chen et al [25] developed an algorithm for sleep detection using sensors available on the smartphone, in particular, charge events, time, and length of smartphone usage and microphone or luminosity sensor. Combining all the information available, they developed the Best Effort Sleep model, which was tested on a 1-week 8-person study comparing the calculated data with the Zeo headband and the Jawbone wristband, resulting in an average error of sleep duration of about 42 min.

Contrary to the above-mentioned works of Abdullah et al [26], Lane et al [31], and Chen et al [25], iSenseSleep approach results in smaller average error for the sleep duration without using privacy-sensitive datasets from the user's smartphone. Furthermore, we have used wearable datasets as a ground truth, to avoid the use of subjective self-reported data used by Abdullah et al [26].

Min et al [32] developed an algorithm for sleep detection based on smartphone sensors, and in particular, using the accelerometer, the microphone, the ambient light sensor, the screen proximity sensor, the running processes, the battery state, and the display screen state. They used all these sensors to develop an algorithm tested with 27 participants during 1 month. Ground truth data were collected using a sleep diary. The system classified correctly sleep state with a 93% accuracy and overall sleep quality with an 81% accuracy. Sleep duration was neither modeled nor evaluated.

Hao et al [33] developed iSleep, a smartphone app that uses the built-in microphone of the smartphone to detect the events that are related to sleep quality, such as body movements, coughing, and snoring, and infer quantitative measures of sleep quality.

The model is based on a decision-tree algorithm to classify various events and calculate acoustic features. They tested the algorithm with 7 participants during 51 nights of sleep, and the accuracy was above 90% for event classification in different scenarios. Sleep duration was neither modeled nor evaluated. Contrary to Min et al [32] and Hao et al [33], iSenseSleep approach focuses exclusively on the sleep duration without leveraging the privacy-sensitive datasets from the user's smartphone.

Furthermore, the authors of Somnometer (by Shirazi et al [34]) developed their own app (also connected to social network) that acts as an alarm clock, and they have evaluated the sleep duration for an individual based on data collected within the app. The authors compared the sleep duration assessed by the app with the wearable device (HedgeHog) for 20 nights and concluded that their app can be used as a sleep duration sensor (without providing numerical results supporting its accuracy). The work on Somnometer focused then on the social sharing features and their update among the Somnometer users.

Compared with the presented approaches, our approach is slightly different because it aims at understanding sleep duration of each individual considering only the interaction they have with their phone, without using information collected from other sensors, especially privacy-sensitive sensors such as microphone, used by many other authors indicated above. Moreover, ground truth data are not collected using diaries or other methods that are affected by a subjective self-report error but using objective data collected from a smartwatch. In this research, we show how smartphones are becoming accurate proxies of our everyday life and that they can be easily used to provide an estimation of individuals' sleep duration.

Conclusions

Smartphones are getting becoming increasingly ubiquitous every day, getting closer to their owners, being carried around in a pocket, and becoming more integrated into the everyday life of individuals. This proximity of the smartphone with the daily life of users opens the door to many different opportunities for leveraging smartphone use to bring more understanding of the daily activities and routines of the users.

In this study, we presented the possibility of evaluating users' sleep patterns by analyzing their interaction with their smartphone, and in particular, only the smartphone screen interaction data (screen ON and OFF). The approach presented here, denoted iSenseSleep; is lightweight; nonintrusive, as it does not affect individuals' life; privacy preserving because it does not use any privacy-related information (eg, phone microphone or light sensors); and it is low cost because it does not require any other external devices for sleep duration estimation. iSenseSleep has been evaluated against the wearable BASIS peak datasets for 2 different user groups, one consisting of 4 working mothers, and another comprising 10 students. These 2 groups can be considered very different from each other because the second one is significantly more digitally native than the first one; usually the usage of smartphone is higher in adolescents and young adults than in adults. Results show how, on average, the difference between the sleep duration calculated with the algorithm and the ground truth data of the smartwatch is about 53 min for the working mothers' group and 24 min for the students' group, that is about 13% and 7% with respect to their total sleep duration. This error is almost the same for weekdays and slightly higher (13% and 12%) for weekends. Moreover, the results show that the difference in sleep duration evaluated by the iSenseSleep and by the BASIS is not statistically significant. These results support the possibility of using smartphones as a nonintrusive, cheap sleep duration pattern analyzer.

In the future, we plan to increase the accuracy of the algorithm to have more precise data about individuals' sleep behavior (beyond the sleep duration), relying on, for example, historical trends of the user's behavior that could help us understand, which are the sleep episodes, and to analyze if and to what extent, there is a correlation between sleeping patterns of the users and their interaction with the smartphone in terms of usage time, applications used, etc. In addition, more specific efforts will be provided for estimating the wake-up time—as a consistent wake-up time is being recognized by the medical experts as an important contributor to one's wellness and health state in the long term.

Acknowledgments

The authors wish to thank AAL CoME (Caregivers and Me, 7-127) and H2020 WellCo (769765) projects.

Conflicts of Interest

None declared.

References

1. Flurry Analytics Blog. Media, Productivity & Emojis Give Mobile Another Stunning Growth Year URL: http://www.flurrymobile.tumblr.com/post/136677391508 [accessed 2018-08-13] [WebCite Cache ID 71e5V3qZU]
2. Ciman M, Donini M, Gaggi O, Aiolli F. Stairstep recognition and counting in a serious Game for increasing users' physical activity. Pers Ubiquit Comput 2016 Sep 22;20(6):1015-1033. [doi: 10.1007/s00779-016-0968-y]
3. Ciman M, Wac K. Individuals' stress assessment using human-smartphone interaction analysis. IEEE Trans Affective Comput 2018 Jan 1;9(1):51-65. [doi: 10.1109/TAFFC.2016.2592504]
4. Silva PA, Holden K, Jordan P. Towards a list of heuristics to evaluate smartphone apps targeted at older adults: a study with apps that aim at promoting health and well-being, in 2015 48th Hawaii International Conference on System Sciences,

2015. 2015 Jan 1 Presented at: 48th Hawaii International Conference on System Sciences; January 5-8, 2015; Hawaii. [doi: 10.1109/HICSS.2015.390]
5. Miyata S, Noda A, Ozaki N, Hara Y, Minoshima M, Iwamoto K, et al. Insufficient sleep impairs driving performance and cognitive function. Neurosci Lett 2010 Jan 22;469(2):229-233. [doi: 10.1016/j.neulet.2009.12.001] [Medline: 19969042]
6. Foley D, Ancoli-Israel S, Britz P, Walsh J. Sleep disturbances and chronic disease in older adults: results of the 2003 National Sleep Foundation Sleep in America Survey. J Psychosom Res 2004 May;56(5):497-502. [doi: 10.1016/j.jpsychores.2004.02.010] [Medline: 15172205]
7. Winkelman JW. Treatment of nocturnal eating syndrome and sleep-related eating disorder with topiramate. Sleep Med 2003 May;4(3):243-246. [doi: 10.1016/S1389-9457(03)00060-1] [Medline: 14592329]
8. Veldi M, Aluoja A, Vasar V. Sleep quality and more common sleep-related problems in medical students. Sleep Med 2005 May;6(3):269-275. [doi: 10.1016/j.sleep.2004.12.003] [Medline: 15854858]
9. Aschoff J. Circadian rhythms in man. Science 1965 Jun 11;148(3676):1427-1432. [Medline: 14294139]
10. Blatter K, Cajochen C. Circadian rhythms in cognitive performance: methodological constraints, protocols, theoretical underpinnings. Physiol Behav 2007 Feb 28;90(2-3):196-208. [doi: 10.1016/j.physbeh.2006.09.009] [Medline: 17055007]
11. Roenneberg T, Wirz-Justice A, Merrow M. Life between clocks: daily temporal patterns of human chronotypes. J Biol Rhythms 2003 Feb;18(1):80-90. [doi: 10.1177/0748730402239679] [Medline: 12568247]
12. Fitbit. URL: http://www.fitbit.com [accessed 2018-08-13] [WebCite Cache ID 71e4y0ihN]
13. Withings. URL: https://www.withings.com/dk/en/steel-hr [accessed 2019-05-08] [WebCite Cache ID 78EMfUSQq]
14. Apple Inc. Apple Watch URL: https://www.apple.com/watch/ [accessed 2018-08-13] [WebCite Cache ID 71e69YbVx]
15. Pew Research Center. Do you sleep with your cell phone URL: http://www.pewresearch.org/fact-tank/2010/09/13/do-you-sleep-with-your-cell-phone/ [accessed 2018-08-13] [WebCite Cache ID 71e5YN5iZ]
16. Basis Inc. BASIS Watch URL: https://www.mybasis.com [accessed 2018-08-13] [WebCite Cache ID 71e5Zdgz4]
17. Buysse DJ. Sleep health: can we define it? Does it matter? Sleep 2014 Jan 01;37(1):9-17 [FREE Full text] [doi: 10.5665/sleep.3298] [Medline: 24470692]
18. Monk TH, Reynolds CF, Kupfer DJ, Buysse DJ, Coble PA, Hayes AJ, et al. The Pittsburgh Sleep diary. J Sleep Res 1994;3:111-120 [FREE Full text] [Medline: 11537903]
19. Carney CE, Buysse DJ, Ancoli-Israel S, Edinger JD, Krystal AD, Lichstein KL, et al. The consensus sleep diary: standardizing prospective sleep self-monitoring. Sleep 2012 Feb 01;35(2):287-302 [FREE Full text] [doi: 10.5665/sleep.1642] [Medline: 22294820]
20. Kahneman D, Krueger AB, Schkade DA, Schwarz N, Stone AA. A survey method for characterizing daily life experience: the day reconstruction method. Science 2004 Dec 03;306(5702):1776-1780 [FREE Full text] [doi: 10.1126/science.1103572] [Medline: 15576620]
21. Patel S, Ruoff L, Ahmed T, Unadkat T, Lee J. Basis Sci. 2014. Validation of basis science advanced sleep analysis URL: https://nanopdf.com/download/validation-of-basis-science-advanced-sleep-analysis_pdf [accessed 2019-05-09] [WebCite Cache ID 78FNWpEXo]
22. De Masi A, Ciman M, Gustarini M, Wac K. mQoL Smart Lab: Quality of Life Living Lab for Interdisciplinary Experiments. In: UbiComp (Adjunct). NY, USA: ACM; 2016 Presented at: International Workshop on Ubiquitous Mobile Instrumentation (UbiMI); September 7–9, 2016; Turin, Italy. [doi: 10.1145/2968219.2971593]
23. Prensky M. H sapiens digital: from digital immigrants and digital natives to digital wisdom. Innovate (North Miami Beach) 2009;5:3.
24. Hirshkowitz M, Whiton K, Albert SM, Alessi C, Bruni O, DonCarlos L, et al. National Sleep Foundation's updated sleep duration recommendations: final report. Sleep Health 2015 Dec;1(4):233-243. [doi: 10.1016/j.sleh.2015.10.004] [Medline: 29073398]
25. Chen Z, Lin M, Chen F, Lane ND, Cardone G, Wang R, et al. Unobtrusive sleep monitoring using smartphones. In: Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare. 2013 Presented at: 7th International Conference on Pervasive Computing Technologies for Healthcare; May 05-08, 2013; Venice, Italy.
26. Abdullah S, Matthews M, Murnane EL, Gay G, Choudhury T. Penn State. 2014. Towards circadian computing: "Early to bed and early to rise" makes some of us unhealthy and sleep deprived URL: http://pac.cs.cornell.edu/pubs/clockwise-ubicomp-2014.pdf [accessed 2019-05-09] [WebCite Cache ID 78FONFbQv]
27. Hamermesh SD, Stancanelli E. Long Workweeks and Strange Hours. ILR Review, Cornell University, ILR School 2015;68(5):1007-1018. [doi: 10.1177/0019793915592375]
28. Dey AK, Wac K, Ferreira D, Tassini K, Hong J, Ramos J. Getting closer: an empirical investigation of the proximity of user to their smart phones. In: Proceedings of the 13th international conference on Ubiquitous computing. 2011 Presented at: 13th international conference on Ubiquitous computing; September 17-21, 2011; China.
29. Patel SN, Kientz JA, Hayes GR, Bhat S, Abowd GD. Farther than you may think: an empirical investigation of the proximity of users to their mobile phones. 2006 Presented at: International Conference on Ubiquitous Computing; 2006; Pittsburg, USA URL: https://link.springer.com/chapter/10.1007/11853565_8
30. Oulasvirta A, Rattenbury T, Ma L, Raita E. Habits make smartphone use more pervasive. Pers Ubiquitous Comput 2012;16(1):105-114.

31. Lane ND, Lin M, Mohammod M, Yang X, Lu H, Cardone G, et al. BeWell: sensing sleep, physical activities and social interactions to promote Wellbeing. Mobile Netw Appl 2014 Jan 9;19(3):345-359. [doi: 10.1007/s11036-013-0484-5]
32. Min JK, Doryab A, Wiese J, Amini S, Zimmerman J, Hong JI. Toss N' Turn: Smartphone As Sleep And Sleep Quality Detector. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2014 Presented at: ACM CHI; April 26-May 01, 2014; Toronto, Ontario, Canada. [doi: 10.1145/2556288.2557220]
33. Hao T, Xing G, Zhou G. iSleep: unobtrusive sleep quality monitoring using smartphones. In: Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems. 2013 Presented at: 11th ACM Conference on Embedded Networked Sensor Systems; November 11-15, 2013; Roma, Italy. [doi: 10.1145/2517351.2517359]
34. Sahami Shirazi A, Clawson J, Hassanpour Y, Tourian MJ, Schmidt A, Chi EH, et al. Already up? using mobile phones to track & share sleep behavior. Int J Hum Comput Stud 2013 Sep;71(9):878-888. [doi: 10.1016/j.ijhcs.2013.03.001]

Abbreviations

mQoL-log: mobile quality of life logger
QoL: Quality Of Life

Edited by G Eysenbach; submitted 13.08.18; peer-reviewed by V Osmani, JH Hong, M Lang; comments to author 13.10.18; revised version received 01.03.19; accepted 30.03.19; published 21.05.19.

Please cite as:
Ciman M, Wac K
Smartphones as Sleep Duration Sensors: Validation of the iSenseSleep Algorithm
JMIR Mhealth Uhealth 2019;7(5):e11930
URL: http://mhealth.jmir.org/2019/5/e11930/
doi:10.2196/11930
PMID:31115341

Viewpoint

Technical Guidance for Clinicians Interested in Partnering With Engineers in Mobile Health Development and Evaluation

Lochan M Shah[1], BA; William E Yang[1], MD; Ryan C Demo[2], BSc; Matthias A Lee[2], PhD; Daniel Weng[1], BA; Rongzi Shan[1,3], BSc; Shannon Wongvibulsin[1,2], PhD; Erin M Spaulding[4], RN, BSN; Francoise A Marvel[1], MD; Seth S Martin[1,2], MD, MHS

[1]Johns Hopkins University School of Medicine, Baltimore, MD, United States

[2]Johns Hopkins University Whiting School of Engineering, Baltimore, MD, United States

[3]David Geffen School of Medicine at University of California, Los Angeles, Los Angeles, CA, United States

[4]Johns Hopkins University School of Nursing, Baltimore, MD, United States

Corresponding Author:
Lochan M Shah, BA
Johns Hopkins University School of Medicine
733 N Broadway
Baltimore, MD, 21231
United States
Phone: 1 410 550 3350
Email: lochanshah2019@gmail.com

Related Article:

This is a corrected version. See correction statement: https://mhealth.jmir.org/2022/8/e41813

Abstract

The explosion of mobile health (mHealth) interventions has prompted significant investment and exploration that has extended past industry into academia. Although research in this space is emerging, it focuses on the clinical and population level impact across different populations. To realize the full potential of mHealth, an intimate understanding of how mHealth is being used by patients and potential differences in usage between various demographic groups must also be prioritized. In this viewpoint, we use our experiences in building an mHealth intervention that incorporates an iOS app, Bluetooth-enabled blood pressure cuff, and Apple Watch to share knowledge on (1) how user interaction data can be tracked in the context of health care privacy laws, (2) what is required for effective, nuanced communication between clinicians and engineers to design mHealth interventions that are patient-centered and have high clinical impact, and (3) how to handle and set up a process to handle user interaction data efficiently.

(JMIR Mhealth Uhealth 2019;7(5):e14124) doi:10.2196/14124

KEYWORDS

mHealth; cardiology; myocardial infarction; personalized medicine

Introduction

Mobile health (mHealth) interventions promise to improve patient outcomes, increase self-management and health literacy, reduce health disparities, increase access to health services, and lower health care costs in ways previously unachievable [1,2]. However, much work is needed to move mHealth from the ivory tower of academia or the offices of startup companies to clinical practice, where mHealth interventions can be prescribed as adjuncts to therapy. To facilitate clinician and policy-maker decision making, researchers need to create an actionable knowledge base to identify the most effective, safe, and scalable interventions for improving individual and population health.

mHealth studies evaluating clinical and population-level impact have made strides in this direction. For example, BlueStar by WellDoc, a Food and Drug Administration (FDA) cleared mHealth intervention for patients with type 2 diabetes, demonstrated a significantly lower mean decrease in hemoglobin A_{1c} over 1 year compared with the usual care group [3]. Similarly, a 2016 systematic review and meta-analysis found

that across 27 studies, the mHealth group compared with the usual care group had increased adherence to medical therapy and ability to reach blood pressure and exercise goals [1]. However, other studies have found mixed results from mHealth interventions and have further concluded that the evidence for efficacy is still limited [4]. One concern that has been raised is the diversity of patient demographics and how this might lead to differential intervention usage and effectiveness. Indeed, studies have found that sex, age group, ethnicity, and family history significantly affect user engagement and that patients who are more engaged with a given mHealth intervention may be more likely to achieve the intended clinical outcomes [5-7].

Thus, to realize the full potential of mHealth interventions, validation of individual clinical outcomes is not enough. mHealth interventions are typically complex with multiple aims and components, including educational material, self-monitoring tools, and various other aspects of health behavior change constructs. Making meaningful insights and developing recommendations require understanding the causal model of how the intervention will achieve its intended benefit, how key components of the intervention interact with one another, and what combinations will optimize a return on investment [8]. Smartphones, in addition to functioning as a platform for such interventions, send data to a back-end system, where it can be leveraged to understand intervention use patterns in a way that avoids potential limitations of self-reported survey data such as recall bias. This kind of granular app usage analysis, while traditionally not associated with clinical research, is necessary for transitioning mHealth interventions toward real-world implementation and answering the question, "What works and for whom?"

Importantly, an analysis of mHealth back-end data presents important challenges, and few studies have examined clinical outcomes in the context of app usage data [8]. Here, we use our experiences in building the Corrie mHealth intervention as a framework to explore key questions that arise in constructing meaningful insights from back-end data. Corrie guides patients through recovery after acute myocardial infarction and incorporates an iOS app, a Bluetooth-enabled blood pressure cuff, and an Apple Watch. In addition to sharing nuances identified from our research and development process, we aimed to provide practical knowledge to help other clinicians work effectively with an engineering team to design mHealth interventions that are patient-centered and have high clinical impact.

How is User Interaction Data Tracked?

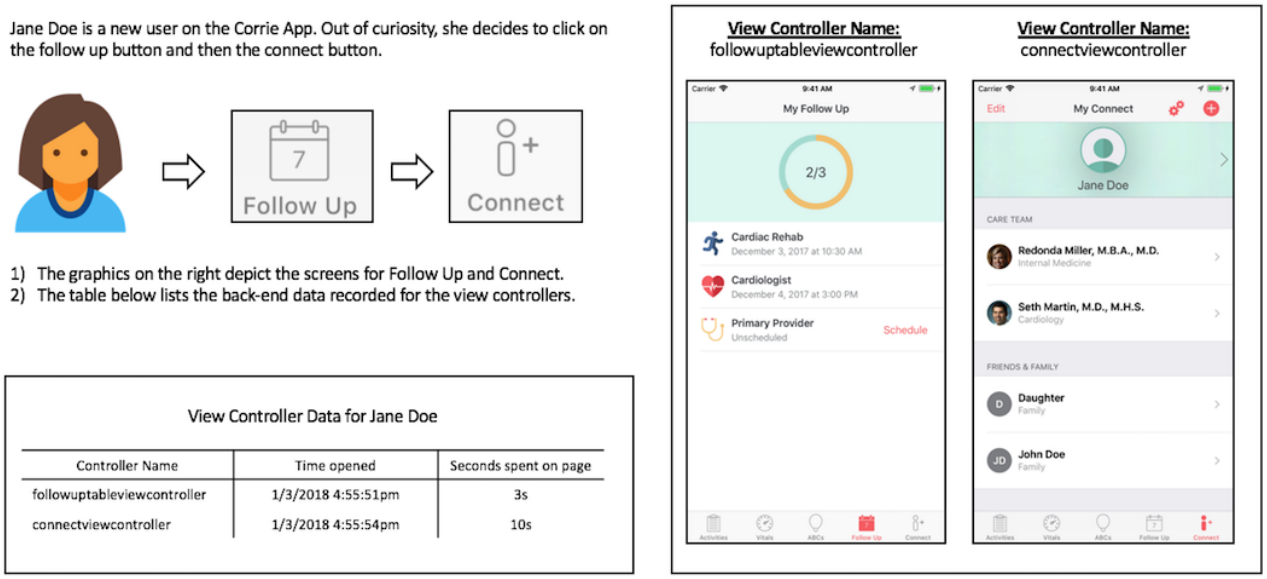
The back-end data analysis allows mHealth teams to identify features that patients use most regularly and take that into account in an iterative design process, thus quantifying and supporting insights collected through patient surveys and focus groups. This enables teams to begin to understand if specific features may be associated with improved clinical outcomes and ultimately helps make interventions more patient-centered.

Although there are a number of different ways to track app usage [9], our intervention tracking is divided into 2 main categories: (1) the tracking of unique *events* and (2) the recording of time elapsed on *view controllers*. Such collection of intrasession measures (within sessions) of user engagement allow for the evaluation of specific user interactions and behaviors, while intersession measures (across sessions) evaluate long-term user engagement and satisfaction [9].

In the Corrie iOS app, unique app *events* are generated any time a patient (user) undertakes a specific action—for example, taking vitals, creating follow-up appointments, and completing learning goals. *View controllers* correspond to individual screens on iPhone. These *view controllers* can then be grouped by the feature they are a part of, allowing for the identification of features that patients spend the most time on (Figure 1). Commonly, nonhealth apps outsource this kind of analytics to third-party software development kits (SDKs) such as Firebase and Google Analytics. However, as many of those third-party software tools are not Health Insurance Portability and Accountability Act (HIPAA) compliant, mHealth teams may need to create their own HIPAA-compliant implementations, especially for detailed insights that can be correlated to patients and their clinical outcomes [10].

To make such meaningful insights, clinical teams must have continuous back-and-forth discussions with their engineering team to understand the context of data being obtained and the nuances of how it is captured. In addition to improving team efficiency, this detailed, continuous communication minimizes misinterpretation of the underlying data that could ultimately yield false conclusions. However, achieving this level of communication can be complicated by a technological language barrier and the inherent intricacies of back-end data collection.

Figure 1. An example of view controllers and their use in user interaction tracking.



View Controller Data for Jane Doe

Controller Name	Time opened	Seconds spent on page
followuptableviewcontroller	1/3/2018 4:55:51pm	3s
connectviewcontroller	1/3/2018 4:55:54pm	10s

What Do Clinicians and Engineers Need to Know About the Other Side?

To address this, we found that it was critical for the clinical and engineering teams to collaborate more closely than is traditionally done. For example, our clinical team obtained the list of *view controllers* from our engineering team and classified them into 4 broad categories for feature usage analytics. We also walked through multiple use case scenarios with our engineering team to understand and fine-tune how data would be recorded in various clinical situations. This was particularly important for the development of a robust system to track unique app *events*. For example, we wanted to capture both time spent watching an educational video as well as completion of the video. If a patient went back to watch a video multiple times, we also wanted to capture that. However, if a patient accidentally pressed a *medication taken* button multiple times, we wanted that recorded as one *event* rather than multiple *events* (Figure 2).

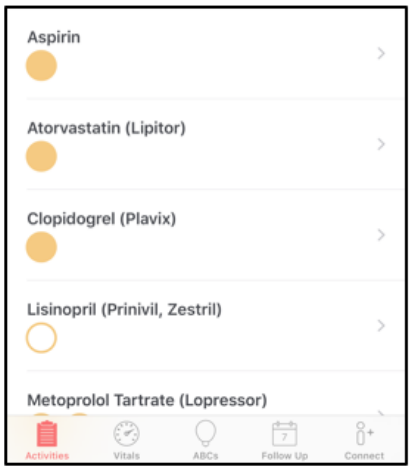
Beyond recording app usage *events*, we also wanted to capture the values of clinical measurements such as heart rate, blood pressure, step count, weight, and mood to create rich datasets for a health outcome–based analysis. For these types of data, we emphasized the importance of querying and displaying the frequency of measurement and timestamps, as well as having well-defined data collection time windows. These details require a significant understanding of variable-capturing capabilities on the back-end but we believe they are key to building accurate and clinically relevant datasets. Ultimately, ensuring high quality and appropriately contextualized data will be critical as interest grows in applying machine learning–based analytical tools to generate actionable insights from large volumes of mHealth data, in the setting of both mHealth validation research as well as clinical scenarios using patient-generated data from commercial devices.

Similarly, when clinicians and engineers team up to solve a problem and work together as equal partners, the engineering team can gain an in-depth understanding of the clinical picture over time. Initially, many mHealth interventions start off with ad hoc combinations of the engineering team and interested clinical professionals partly serving as the data scientists [11]. For our team, this meant that when data were needed, the engineering team provided the relevant information to the clinical team. Close collaboration with the engineering team was required to sort through all the recorded variables and determine which were important to analyze and which were not. To facilitate that decision making, the engineering team needed to have a strong understanding of the clinical context. In addition to taking our engineering team through the discharge process and hosting our engineers on rounds in the hospital, we found it helpful to come together with our engineering team in clinically focused discussions by using real-time team messaging and collaboration platforms such as Slack (Slack Technologies, Inc). We have found this more efficient and effective than email and it provides an ongoing, continuous form of discussion outside the context of our regular face-to-face meetings and calls to facilitate clinical engineering collaboration. Slack discussions enable moment-to-moment understanding of clinical challenges, with the bonus of being able to answer urgent technical questions in real time. Importantly, this streamlined troubleshooting and prioritization of tasks for our engineering team as well, who were able to leverage real-time clinical perspective on potential features, content, and user workflow.

Figure 2. An example of data collection and product considerations when developing a new feature.

Feature: Tracking medications

Each day, patients can toggle whether they have taken a prescribed medication on their medication list.
- Filled circle=taken
- Unfilled=not taken



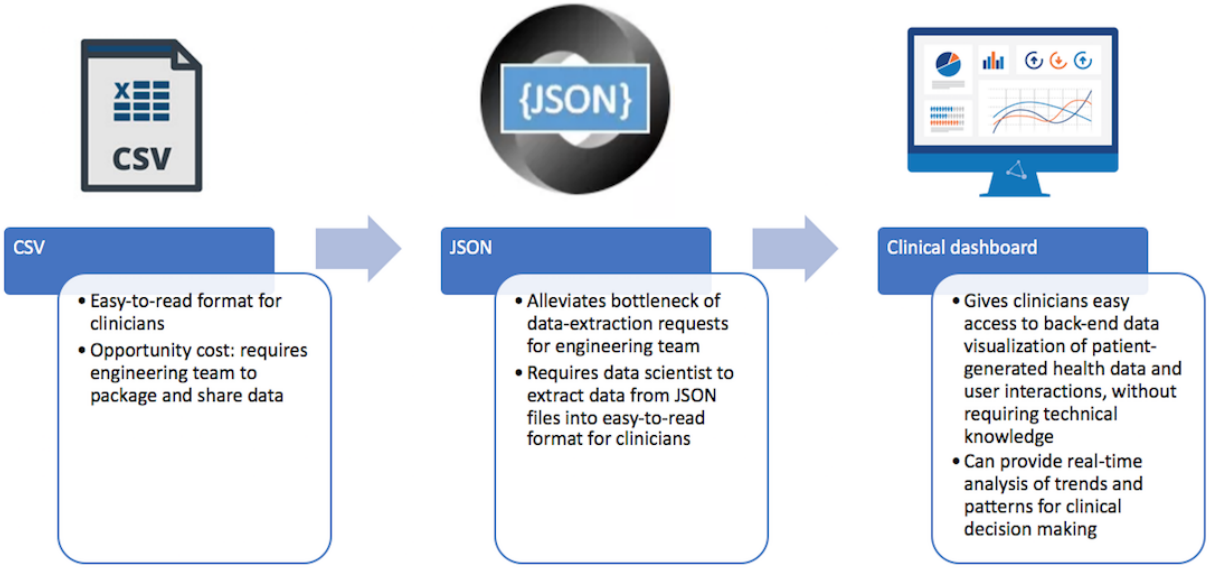
Data collection considerations	Consensus
When are these buttons reset for the next day?	At 12:00am (based on the user's time zone), all buttons are reset to "not taken".
When a medication is prescribed to be taken as needed (PRN), how can users record taking those medications?	Users can add a medication for that specific day, and toggle it as "taken". That specific medication will only be present if the user adds it to their daily list.
How can we capture why a medication was not taken?	Users can swipe right on the medication and select why they were unable to take the medication: refills, adverse effects, etc.
What happens if a user accidentally toggles a button as *taken*? How is that captured in user data?	By pressing the button again, they can reset the medication as *untaken*. The status of the button (taken vs not taken, reason for not taken) at the end of the day (11:59pm) is recorded as the user's true medication adherence status.

Who Handles User Interaction Data and How Is This Affected by Health Care Privacy Laws?

Teams developing mHealth interventions must recognize that having the engineering team manually export back-end data is not a feasible long-term solution. As the intervention scales and the clinical team's need for data extraction increases, the potential for severe bottlenecks warrants an early discussion about data management and access. For example, as our clinical team expanded, it became clear that identifying appropriate variables and packaging data into easy-to-read formats cost the engineering team valuable time that could have been better spent on other endeavors such as further developing the app.

In our case, this led to a stepwise progression of methods to access back-end data in a HIPAA-compliant manner. At first, our engineering team shared data via standard comma separated value (CSV) files over the HIPAA-compliant Johns Hopkins secure virtual desktop. We then transitioned to sharing a JavaScript Object Notation (JSON) file (a data-interchange format) and hiring a data scientist, giving clinicians access to all of the back-end data without the need for extensive teamwide technical training. Although we advocate hiring a dedicated data scientist to bridge the gap between clinicians and engineers early on, as studies scale up and move toward pragmatic trials that evaluate real-world implementation, we believe the end-goal also includes an HIPAA-compliant, queryable dashboard for clinicians, with data visualization and electronic medical record integration (Figure 3) Creating a system that allows clinical team members to intuitively access raw usage data not only decreases the burden on engineers but also empowers team members to better understand their data architecture and more confidently perform analyses based on their individual priorities.

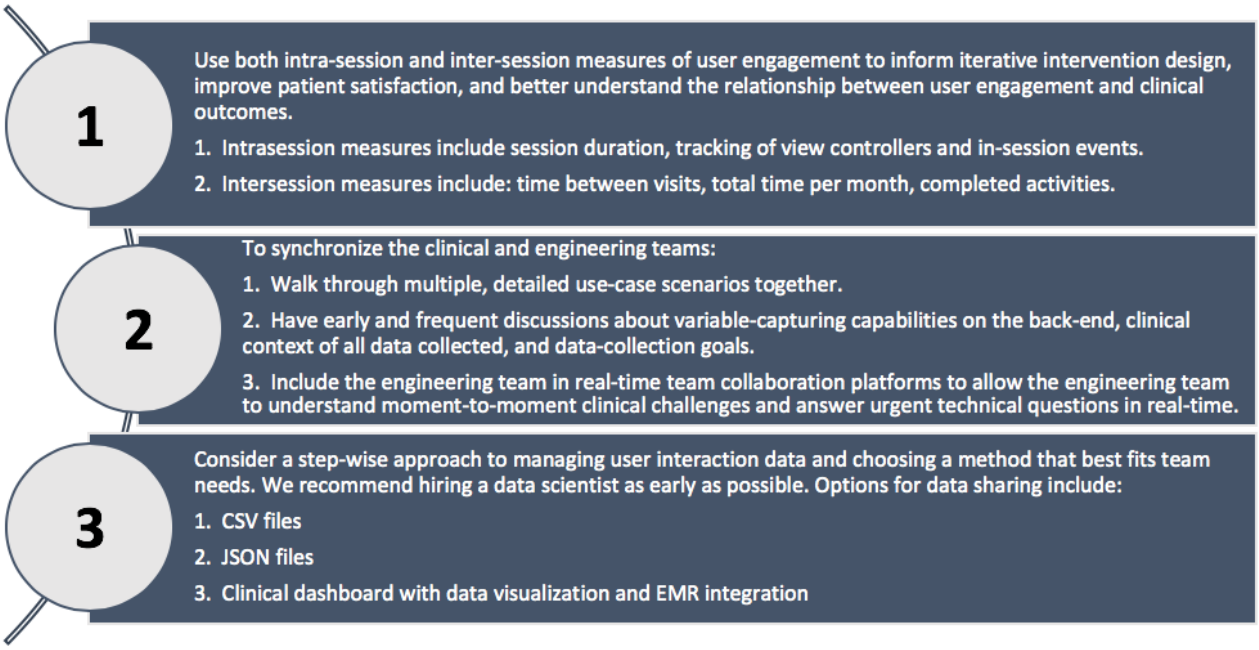
Figure 3. Progression of user-interaction data management and extraction. CSV: comma separated value; JSON: JavaScript Object Notation.



Conclusions

Going forward, it will be essential for research teams to evaluate mHealth interventions in the context of potentially differential app usage patterns and to create environments that support continuous, technically nuanced communication between their clinical and engineering counterparts. Although more business than clinical in nature, these changes to practice are essential to building interventions that patients actively want to use in real-world settings. In our experience with mHealth, being consumer-centered is synonymous with being patient-centered and is the future of health care delivery. We hope our experiences help interdisciplinary teams navigate similar challenges (Figure 4), as the capacity of mHealth interventions to revolutionize health care is directly linked to the quality of interventions developed through clinical-engineering partnership and studied through careful research.

Figure 4. Key takeaways. CSV: comma separated value; JSON: JavaScript Object Notation; EMR: electronic medical record.



Authors' Contributions

The authors LS, WY, ES, RD, ML, DW, RS, FM, and SM provided substantial contributions to the conception and design of the work, contributed significantly to drafts of the work, and critically revised it for key intellectual content. Additionally, final approval of the version to be published was solicited from each, and each agreed to be accountable for all aspects of the work in

ensuring that questions related to the accuracy or integrity of any part of the work were appropriately investigated and resolved. Additionally, the author LS took primary responsibility for communication with the journal during the manuscript submission, peer review, and publication process.

Conflicts of Interest

There was no specific funding for this paper. EMS has received the following financial support for research, authorship, and/or publication: NIH/NINR F31 NR017328, Ruth L Kirschstein National Research Service Award and NIH/NINR T32 NR012704, Pre-Doctoral Fellowship in Interdisciplinary Cardiovascular Health Research. The sponsor had no role in study design, collection, analysis, and interpretation of data, writing the report, and the decision to submit the report for publication. SSM reports receiving grants and research support from the PJ Schafer Cardiovascular Research Fund, the David and June Trone Family Foundation, American Heart Association, Aetna Foundation, Maryland Innovation Initiative, Nokia, iHealth, Google, and Apple, as well as receiving personal fees for serving on scientific advisory boards for Amgen, Sanofi/Regeneron, Quest Diagnostics, Esperion, Novo Nordisk, Akcea Therapeutics. In addition, SSM reports having patent applications pending. FAM and SSM are founders of and hold equity in Corrie Health, which intends to further develop the platform. This arrangement has been reviewed and approved by the Johns Hopkins University in accordance with its conflict of interest policies. Furthermore, the Corrie Health Digital Platform has received material support from Apple and iHealth, and funding from the Maryland Innovation Initiative, Wallace H Coulter Translational Research Partnership, Louis B Thalheimer Fund, and JH Individualized Health Initiative. Corrie Health, as described in this work, was developed by FAM, MAL, and SSM. They are also founders of and hold equity in Corrie Health, which intends to further develop the digital platform. This arrangement has been reviewed and approved by the Johns Hopkins University in accordance with its conflict of interest policies.

References

1. Gandhi S, Chen S, Hong L, Sun K, Gong E, Li C, et al. Effect of mobile health interventions on the secondary prevention of cardiovascular disease: systematic review and meta-analysis. Can J Cardiol 2017 Feb;33(2):219-231. [doi: 10.1016/J.CJCA.2016.08.017]
2. Beatty A, Fukuoka Y, Whooley M. Using mobile technology for cardiac rehabilitation: a review and framework for development and evaluation. J Am Heart Assoc 2013 Nov 01;2(6):e000568 [FREE Full text] [doi: 10.1161/JAHA.113.000568] [Medline: 24185949]
3. Quinn C, Shardell M, Terrin M, Barr E, Ballew S, Gruber-Baldini A. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. Diabetes Care 2011 Sep;34(9):1934-1942 [FREE Full text] [doi: 10.2337/dc11-0366] [Medline: 21788632]
4. Marcolino M, Oliveira J, D'Agostino M, Ribeiro A, Alkmim M, Novillo-Ortiz D. The impact of mHealth interventions: systematic review of systematic reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
5. Bell A, Fonda S, Walker M, Schmidt V, Vigersky R. Mobile phone-based video messages for diabetes self-care support. J Diabetes Sci Technol 2012 Mar 01;6(2):310-319 [FREE Full text] [doi: 10.1177/193229681200600214] [Medline: 22538140]
6. Lin P, Grambow S, Intille S, Gallis J, Lazenka T, Bosworth H, et al. The association between engagement and weight loss through personal coaching and cell phone interventions in young adults: randomized controlled trial. JMIR Mhealth Uhealth 2018 Oct 18;6(10):e10471 [FREE Full text] [doi: 10.2196/10471] [Medline: 30341051]
7. Goyal S, Morita P, Picton P, Seto E, Zbib A, Cafazzo J. Uptake of a consumer-focused mHealth application for the assessment and prevention of heart disease: the <30 Days study. JMIR Mhealth Uhealth 2016 Mar 24;4(1):e32 [FREE Full text] [doi: 10.2196/mhealth.4730] [Medline: 27012937]
8. Heffner J, Vilardaga R, Mercer L, Kientz J, Bricker J. Feature-level analysis of a novel smartphone application for smoking cessation. Am J Drug Alcohol Abuse 2015 Jan;41(1):68-73 [FREE Full text] [doi: 10.3109/00952990.2014.977486] [Medline: 25397860]
9. Lalmas M, O'Brien H, Yom-Tov E. Measuring user engagement: Synthesis Lectures on Information Concepts, Retrieval, and Services. San Rafael, California: Morgan & Claypool Publishers; Nov 2014:1-132.
10. 2019. Google Analytics Help. Best practices to avoid sending Personally Identifiable Information (PII) URL: https://support.google.com/analytics/answer/6366371?hl=en.Archived [WebCite Cache ID 77CENWt8b]
11. Huesch M, Mosher T. NEJM Catalyst. Using It or Losing It? The Case for Data Scientists Inside Health Care URL: https://catalyst.nejm.org/case-data-scientists-inside-health-care/ [accessed 2019-03-27] [WebCite Cache ID 77CFxTtPA]

Abbreviations

CSV: comma separated value
FDA: Food and Drug Administration
HIPAA: Health Insurance Portability and Accountability Act
JSON: JavaScript Object Notation

mHealth: mobile health
SDK: software development kit

Edited by G Eysenbach; submitted 27.03.19; peer-reviewed by S Kakarmath, A Hussain; accepted 23.04.19; published 15.05.19.

Please cite as:
Shah LM, Yang WE, Demo RC, Lee MA, Weng D, Shan R, Wongvibulsin S, Spaulding EM, Marvel FA, Martin SS
Technical Guidance for Clinicians Interested in Partnering With Engineers in Mobile Health Development and Evaluation
JMIR Mhealth Uhealth 2019;7(5):e14124
URL: http://mhealth.jmir.org/2019/5/e14124/
doi:10.2196/14124
PMID:31094337

Original Paper

Usability and Usefulness of a Mobile Health App for Pregnancy-Related Work Advice: Mixed-Methods Approach

Monique van Beukering[1], MD; Adeline Velu[1], MD; Liesbeth van den Berg[2], MSc; Marjolein Kok[1], MD, PhD; Ben Willem Mol[3], MD, PhD; Monique Frings-Dresen[4], PhD; Robert de Leeuw[5], MD; Joris van der Post[1], MD, PhD; Linda Peute[2,6], MSc, PhD

[1]Department of Obstetrics and Gynecology, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands

[2]Department of Medical Informatics, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands

[3]Department of Obstetrics and Gynaecology, Monash Medical Centre, Monash University, Melbourne, Australia

[4]Coronel Institute of Occupational Health, Amsterdam Public Health Research Institute, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands

[5]Department of Obstetrics and Gynecology, Amsterdam UMC, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

[6]Department of Medical Informatics, Center for Human Factors Engineering of Health Information Technology, Amsterdam UMC, University of Amsterdam, Amsterdam, Netherlands

Corresponding Author:
Monique van Beukering, MD
Department of Obstetrics and Gynecology
Amsterdam UMC
University of Amsterdam
Room H4-222
PO Box 22660
Amsterdam, 1100 DD
Netherlands
Phone: 31 657514795
Fax: 31 (0) 206963489
Email: m.d.vanbeukering@amc.uva.nl

Abstract

Background: Pregnant women are often unaware of the potential risks that working conditions can cause to them and their unborn child. A mobile health (mHealth) app, the *Pregnancy and Work* (P and W) app, developed by a multidisciplinary team and based on an evidence-based guideline for occupational physicians, aims to provide advice on work adjustment during pregnancy.

Objective: This study evaluates the usability of the mHealth P and W app and the perceived usefulness of the *work advice*, the main goal of the app, by potential end users.

Methods: A total of 12 working pregnant women participated in think aloud usability sessions and performed 9 tasks. All think aloud sessions were recorded, transcribed, and coanalyzed. The usability problems were rated for their severity in accordance with Nielsen severity scale. The completion rates and time taken for completion of tasks were registered. In addition, participants were questioned on demographics and user characteristics and were asked to evaluate the value of the app by filling in the Intrinsic Motivation Inventory (IMI) score and the System Usability Scale (SUS) questionnaire.

Results: In total, 82 usability problems with a severity ≥1 were identified, of which 40 had severity ≥3. The main usability problems concerned the interpretation of terminology used in the app's questionnaires and difficulties in finding and understanding the work advice. Furthermore, 10 out of 12 participants were able to open the work advice page in the app. Only 7 out of these 10 participants understood and intended to follow the work advice. The overall mean IMI score was relatively high (5 out of 7), indicating that the participants did indeed value the use of the app. This IMI score corresponded to the overall mean SUS score (68 out of 100) and the mean grade given to the P and W app (7 out of 10).

Conclusions: This think aloud usability study showed that the information provided in the P and W app was considered valuable by the end users, working pregnant women, and it meets their needs; however, usability issues severely impacted the perceived usefulness of the work advice given in the app.

(JMIR Mhealth Uhealth 2019;7(5):e11442) doi:10.2196/11442

KEYWORDS

mHealth; eHealth; mobile phone; pregnancy; work; occupation; occupational exposure; qualitative research

Introduction

Background

Many women continue to work during their pregnancy. In the United States, more than 65% of pregnant women work, whereas in the Netherlands, around 80% [1,2]. During pregnancy, exposure to certain working conditions, such as physically demanding work, long working hours, working in night shifts, and stress, are associated with preterm birth, low birth weight, and fetal abnormalities [3-12]. As pregnant women are often not aware of these risks, they do not adjust their working conditions [13].

Mobile health (mHealth) apps can offer a suitable solution to this problem as women of reproductive age who are expecting a child are frequent consumers of Web-based health information [14-17]. mHealth, defined as the use of mobile devices for medical and public health practice [18], could therefore inform pregnant working women about work-related pregnancy risks, to increase their awareness of these risks and their associated need for change in working conditions.

However, evidence on the effectiveness of mHealth apps in general is limited [17,19]. Prior studies provide little information as to how best to design them [20-24]. Adequate consideration of the needs of their intended users is necessary so that they are easy to use and perceived as useful [25,26]. *The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use* is the definition for applied usability, based on the International Standardization Organization [27]. To assess and improve upon the usability of mHealth apps, a wide range of usability evaluation methods (UEMs) is available to detect problems in the user-system interaction. UEMs thus assess human interaction with a system for the purpose of identifying those facets of this interaction which can be improved [28]. Ideally, the design process of any health-related product is conducted in an iterative fashion to better fit with the end user population. Utilizing UEMs in such an iterative design process in the health care domain is especially important as the poor design and usability of medical products can lead to harmful consequences [29,30]. Therefore, the utilization of UEMs during the development and testing process of health apps is widely recommended throughout research [31,32].

In this study, we developed an mHealth solution (the Pregnancy and Work [P and W]) app that aimed to provide information and advice about work-related pregnancy risks [33]. With this advice, pregnant women can adjust their work. The P and W app content is based on the evidence-based guideline for occupational physicians, *pregnancy, postpartum period, and work* [34]. In a prior study, the results of 2 multidisciplinary focus group meetings provided content and design instructions for the development of the P and W app [35].

Objectives

Think aloud (TA), an UEM method, was chosen in this study to assess the usability of the P and W app with potential end users to reveal cognitive processes in the app's user interaction that result in user-interaction problems. The TA method requires participants to talk aloud (ie, verbalize their thoughts) while performing or solving a task to reveal their cognitive processes while interacting with the app, which may result in user-interaction problems [36-38]. In this way, the TA helps to understand how pregnant woman think—or believe they think—the P and W app works (ie, their mental model) [38]. Mismatches in the end users' mental model of an app and the app's design can severely influence its usability and subsequently its use in practice. This study therefore evaluated the usability of the P and W app and also how potential end users experienced the usefulness of the work advice; this was the main goal of the app.

Methods

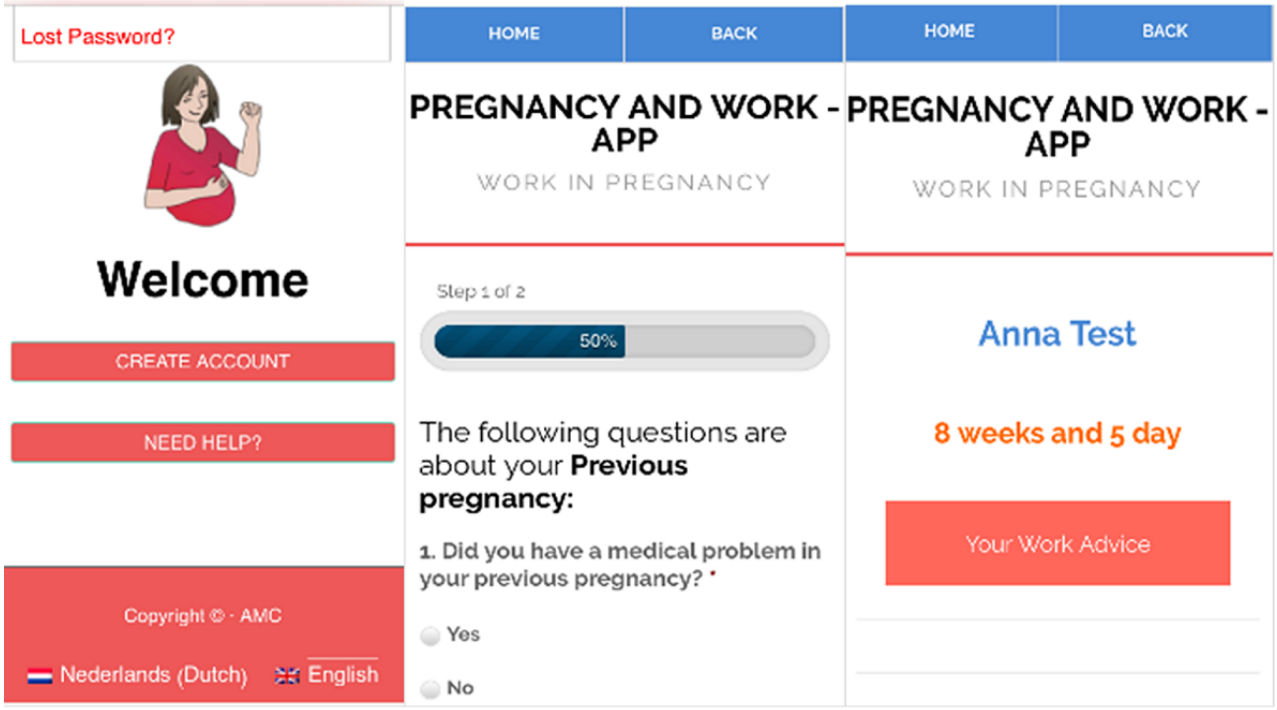
Participant Recruitment

A total of 2 obstetric care facilities, representing a broad variety of patient groups, participated in this study. Posters and flyers were distributed in both locations. The inclusion criteria were drawn up by an obstetrician and occupational physician. If patients met the inclusion criteria, they were invited to participate in the study. The inclusion criteria were Dutch working women, who were less than 20 weeks pregnant. The criterion of being less than 20 weeks pregnant was deliberately stated as the work advice for pregnant women under 20 weeks of pregnancy can be different than that for those after 20 weeks of pregnancy. Eligible participants were recruited in the waiting area of the physician's office. Recruitment of participants continued until a total of 12 female patients agreed to participate in the TA sessions and evaluate the app; this was the first time they used the P and W app. All participants included in the study were offered a gift card worth €15. An app for this research was submitted to the ethical board of the Amsterdam University Medical Center. The board confirmed that the Medical Research Involving Human Subjects Act did not apply to this study. All data from the 12 participants were anonymously processed. Informed consent was obtained from all participants, allowing us to use the data for analysis.

P and W Pregnancy and Work App and Study Flow

The P and W app (Dutch and English) was created as a Web-based app, accessible from every type of mobile browser, with an adaptive design for desktop and mobile phone use.

Figure 1. Examples of screenshots of Pregnancy and Work App: the Welcome page, the Questionnaire page, and the Work Advice page.



The P and W app requires the user to create an account to gain access to its content. After creating an account, a user needs to complete a questionnaire about her pregnancy-related medical and work conditions (Figure 1). When completing this questionnaire, the user will be directed to the home page of the app, from where she can navigate to all other pages. On the home page, users can view monthly pregnancy- and work-related advice messages, which are also sent by email. In addition, the app provides messages about the growth of the unborn baby as the weeks pass. Next to the baby messages, a video with tips and information about pregnancy-related work advice can be viewed on the home page. Participants were given access to a Dutch beta test version of the P and W app.

Phase I: Preparation

Participants were informed about how the TA session would be performed; see Figure 2 for the full study setup. After a 2-week reflection period, a condition for participation in the research, an appointment was made with those women who wanted to participate in the study. The TA session then took place at their next visit (follow-up consultation) to the obstetrics department. After signing an informed consent form, the participant completed a short survey, the validated health literacy (HL) assessment tool— *the Newest Vital Sign*, translated to Dutch—to analyze its potential influence on the TA outcomes (Stage I, Multimedia Appendix 1 [39,40].

Phase II: Think Aloud Usability Testing

Participants started with practice tasks on how to *think aloud* (Multimedia Appendix 2). Each participant was informed that the researcher (LvdB) was solely interested in the app's performance and would only interrupt the participant to provide new tasks and to encourage her to keep talking to break silences longer than 5 seconds [41]. A participant had to complete 9 tasks in total that were centered around the core purpose of the app (Multimedia Appendix 3). Tasks were developed in collaboration with the developer and project supervisors of the P and W project. All TA sessions were recorded via video camera. Voice and screen (of their mobile phone) were also recorded (Figure 3).

Figure 2. Overview of study setup.

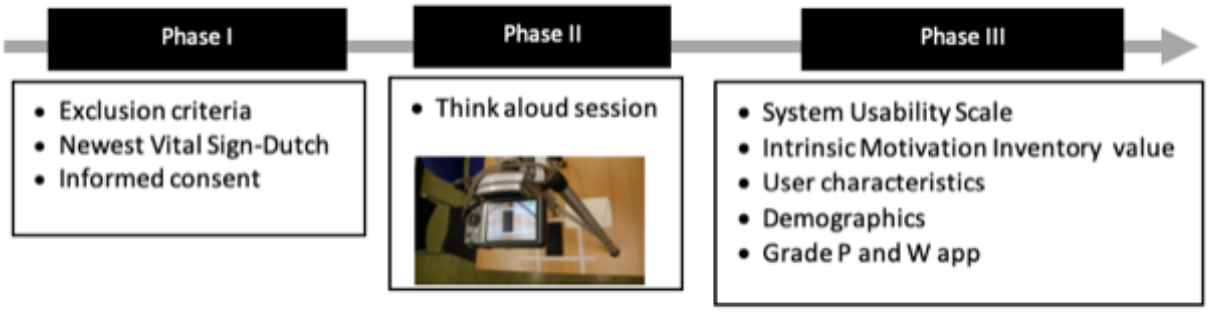
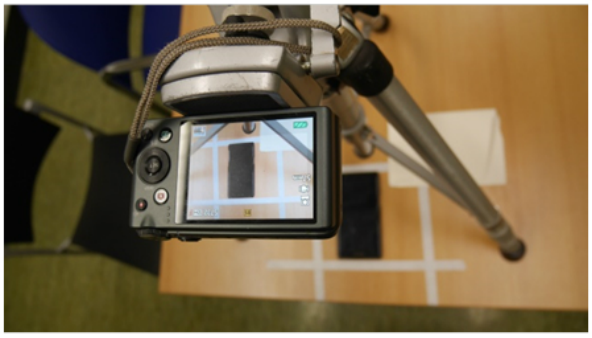


Figure 3. Think aloud session set-up.



Phase III: Usability and Motivation Questionnaires

After the TA test was finished, the SUS survey was given to the participant to assess the perceived usability of the P and W app [42] (Multimedia Appendix 4). The SUS comprises 10 statements which the participant had to rate on a scale from 1 (strongly disagree) to 5 (strongly agree) to indicate the extent to which she agreed. Then, a short survey selection of the Intrinsic Motivation Inventory (IMI) was given to assess a self-reported evaluation on how much the participant valued the P and W app [43] (Multimedia Appendix 4). The IMI value subscale comprised 7 statements which the participant had to rate on a scale from 1 (do not agree) to 7 (strongly agree) to indicate the extent to which she agreed. An additional short survey was developed to gain more insight into participants' demographics, medical history related to pregnancy, prior experience with (pregnancy-related) mobile apps, and working hours (Multimedia Appendix 4). We asked all participants whether they had received and would follow the work advice (Multimedia Appendix 4). Finally, participants were asked to give the P and W app a grade on a scale from 1 to 10, where 1 was the lowest and 10 the highest grade.

Data Collection and Analysis

The TA sessions were videotaped, reviewed multiple times, and transcribed to verbal protocols by 2 researchers (LvdB and LP). To gain insight into the effectiveness and efficiency of the participants in performing tasks, each TA session transcription comprised text spoken by the participant and included task completion time stamps and time taken for task completion. To analyze the usability problems participants encountered in more detail, we performed a thematic analysis for which a coding scheme was developed bottom-up in 3 iterative cycles as described by Jaspers [44]. We analyzed 2 TA sessions in-depth to develop a raw coding scheme (first cycle). Usability issues encountered by participants were then given a specific description. We subsequently discussed the resulting codes and grouped them to determine the main themes in the data (second cycle). The developed coding scheme was then applied to code and analyze all verbalizations, this was performed by LvdB and checked by LP. All new issues were discussed to determine whether they were within the branches of the coding tree or if a new main theme had emerged. Usability problems were rated on severity in accordance with Nielsen severity scale [45]. Nielsen severity scale is a rating scale from 0 to 4 (Textbox 1), that allows for the prioritization of usability problems that need to be revised in the development process. The questionnaires were completed on paper and put in a database for data analysis.

All data filled in by the participants in the P and W app during the TA sessions were specifically transcribed into a different file to test for task efficacy in relation to the IMI-given work advice by the system. Verbalizations of task 6 in the TA sessions (*find the work advice*) were assessed to analyze whether participants would follow the work advice. These results were compared with the results of the IMI on participant level and the questions about the work advice from questionnaire 3 (Multimedia Appendix 4). Finally, the SUS was used to assess the perceived usability of the P and W app.

Textbox 1. Nielsen's severity scale

0–I do not agree that this is a usability problem at all.
1–Cosmetic problem only: need not to be fixed unless extra time is available on project.
2–Minor usability problem: fixing this should be given low priority.
3–Major usability problem: important to fix, so should be given high priority.
4–Usability catastrophe: imperative to fix this before product can be released.

Results

Participant Characteristics

The TA sessions with the participants (N=12) took place between April and June 2017. Most participants scored high (=adequate) HL. All participants had paid jobs and used a mobile phone. The average gestational age of the participants was 15 weeks and 50% (6/12) of the participants were pregnant for the first time (Table 1).

Task Completion

The effectiveness and efficiency of the participants in performing tasks were measured by completion rates and times

and the usability problems. The completion rates and times can be found in Table 2. The average duration of a TA session was 19 min 55 seconds (SD 5 min 25 seconds). Task 1, *create an account*, had a much higher completion time than the other tasks. Tasks 2, 3, 5, and 9 were completed by all participants. Tasks 1, 4, 6, 7, and 8 were not completed by all participants. The first 3 tasks took, on average, the longest time to complete, ranging from 1.5 min to 4 min. Task 9 had the fastest mean completion rate of 4 seconds.

Usability Problems

The TA study identified a total of 101 usability issues, 82 of which were considered *real* usability problems (ie, severity ≥1), whereas 40 usability problems were rated with a severity of 3 (major) or 4 (catastrophic). In addition, the participants encountered 11 unique bugs when using the P and W app. An overview of the most severe usability problems can be found in Table 3. The high completion time with *create an account* (Table 1) seemed to have a connection with the many usability problems in this area (Table 3). None of the participants experienced (severe) usability problems when completing tasks 5, 7, and 9. In the following section, we give an in-depth analysis of the severe usability problems detected regarding *terminology interpretation* and *finding and understanding the work advice* that directly impacted the participants' perceived usefulness of the advice given in the app.

Table 1. Participants' basic demographics and characteristics (N=12).

Characteristics	Statistics
Age (years), mean (SD)	33 (3.8)
Education (secondary school), n	
Higher education	8
Intermediate vocational education	4
Health literacy, n	
High	11
Low	1
Paid job, n	12
Working time (hours per week), mean (SD)	37 (6.15)
Gestational age (weeks), mean (SD)	15 (3)
Previous pregnancy, n	6
Children, n	5
Mobile phone (operating system), n	
Android	7
iPhone	5

Table 2. Completion rates and time taken per task (N=9) by participants.

Task	Completion rate	Time taken for completion (seconds), mean (SD)
1. Create an account	10/12	240 (83)
2. Fill in a questionnaire	12/12	179 (101)
3. Adjust answers to the questionnaire	12/12	96 (74)[a]
4. Find *your rights and tips for consultation* page	11/12	31 (38)
5. Find *baby message(s)*	12/12	16 (10)
6. Find the *your work advice* page	10/12	10 (8)
7. Find the *print/save* button	10/12	9 (9)
8. Find the goal of the Pregnancy and Work P and W app	11/12	32 (18)
9. Log out of the app	12/12	4 (4)

[a]A total of 2 participants initially did not understand this task.

Table 3. Overview of severe usability problems per main problem type.

Usability problem[a]	Frequency	Severity	Source of main problem
Unclear buttons	12	2 to 4	Create account
Functionality with layout	11	4	Create account/home page
Terminology interpretation problems	8	4	Create account/home page
Finding and understanding work advice	8	4	Home page/work advice

[a]Multimedia Appendix 5 shows an overview of all the usability problems.

Qualitative Assessment

Terminology Interpretation Problems

Participants had to complete a questionnaire about their pregnancy-related medical conditions, previous pregnancy (if relevant), and work conditions using the app. Several terminology interpretation problems arose during the TA study, which consequently prevented the participants from receiving accurate personal work advice. For example, when asked whether problems had been experienced during the previous pregnancy, participants were unsure whether *previous pregnancy* implied the immediate previous pregnancy or also the pregnancies before that. One participant who had not experienced problems during her previous pregnancy, but did experience issues during the pregnancy before that, assumed it implied her direct previous pregnancy. Her confusion in answering the question correctly affected the outcome of the work advice, as relevant information was missing:

> *Okay. Um. "Did you have a medical problem in your previous pregnancy?" This is about my last pregnancy, I think, and not the pregnancies before. So, I'm assuming that. And then it's a no.* [Participant 5]

Problems were also prevalent when, in closed-ended questions, the participant did not find the answer that applied to her within the limited selection of possibilities of medical disorders. When given a list of potential problems during a previous pregnancy, participants experienced troubles in selecting the best suited option to describe their problem:

> *...But I do not know if that should be put under "deceased child" or "child born before a gestational age of 37 weeks"? You know what I mean?* [Participant 3]

Another example of a terminology interpretation problem that affected the outcome of the work advice was related to the question of *being exposed to any chemical agents in the work environment*, followed by a list of examples. Several participants did not notice the list of examples and answered *no*. Furthermore, 2 other participants did not know whether an agent that they worked with should be considered chemical, as it was not on the list of examples:

> *...Yes, with hair dye. Is that chemical?* [Participant 9]

> *...I'm having doubts. I work with laughing gas. That's not very chemical, but...I don't know whether I should answer yes or no.* [Participant 11]

Finding and Understanding the Work Advice

Participants also experienced problems in understanding the work advice because of central design problems in the interface. One of the first issues encountered was that the participants expected the app to show them something different than what it actually did. Participants expected the app to show their work advice directly on the homepage, as they perceived this to be the essential goal of the app. They did not expect to have to search for it in the interface or take any other action to find it. For example, participant 6 did not understand that the *your work advice* button was clickable and therefore sought work advice elsewhere or stated that she could not find it (Figure 4):

> *...Oh, let's see if that is somewhere. No idea. [Scrolls down and up] Have a look. Here is my work advice. Uh... [Scrolls up and down, multiple times] No, I have no idea.* [Participant 6]

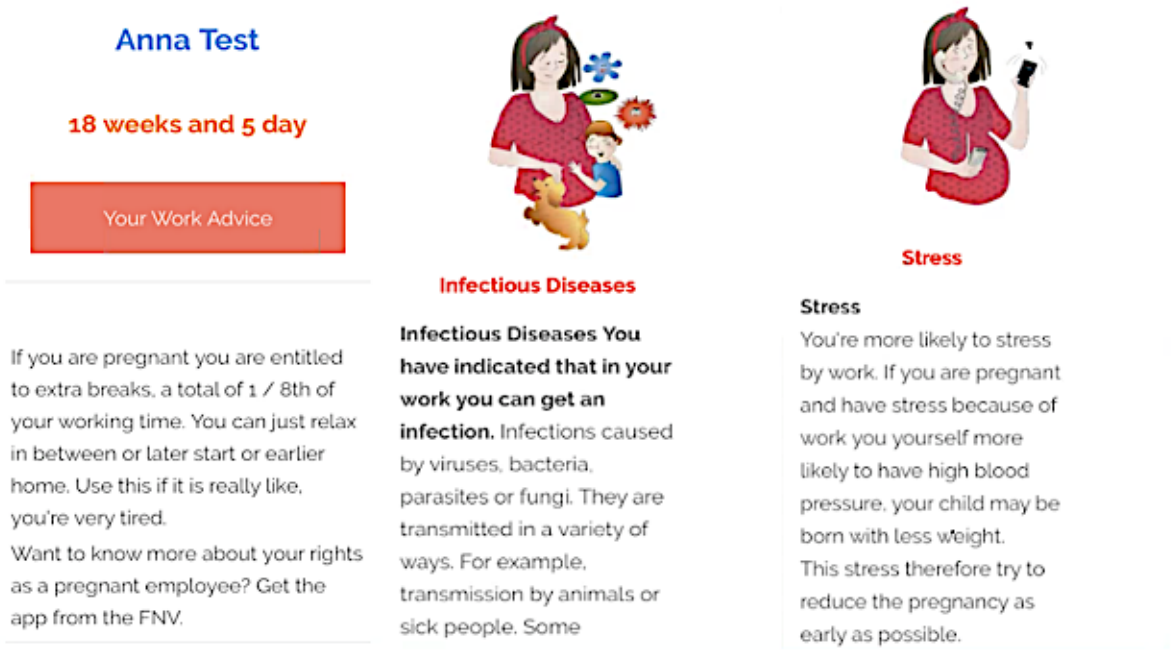
A different example related to the participants stating that they saw their work advice depicted on the home page. However, the home page only provided a small section with tips and information about pregnancy-related work advice, which some clearly interpreted as the entire personal work advice. A total of 2 participants thought this was the case; therefore, both of them missed the actual content of the *your work advice* page:

> *I've just seen my work advice. [Scrolls up and down. Scrolls to top of the page. Taps the back button. Loads page] Yes, your work advice. I have already read it. So, it is here.* [Participant 8]

Another usability issue was related to the fixed structure in which the work advice was presented in the mobile interface. Depending on the answers given in the questionnaire, specific information followed on the work advice page. The resulting advice therefore included some sections without advice and some sections with the advice, spread over the mobile interface. One participant did not get work advice below the *work header*; however, she did receive work advice with regard to issues during her previous pregnancy, but this would only have become visible if she had scrolled the page down. She therefore missed the advice given:

> *None? That's easy. I don't need to make any work adjustments. I don't think so either, because I have an office job.* [Participant 1]

Figure 4. The "Your Work Advice" button on the home page with examples of work advice (infectious diseases and stress) when the button is clicked.



User Evaluation: Intrinsic Motivation Inventory and System Usability Scale

The task efficacy of task 6, *find the "your work advice" page*, was analyzed in relation to the detected usability problems in *finding and understanding the work advice* and combined with the results of the IMI, SUS, and questions about the work advice from questionnaire 3 (Multimedia Appendix 4). Some participants never reached the work advice page on the app (17%) but thought they did, whereas 3 out of 12 participants (25%) were convinced that they had not received this advice (Table 4). However, all participants did actually receive some form of pregnancy-related work advice. Among the 9 participants who stated that they had received work advice, 2 indicated that they would not follow it.

Using the IMI, we assessed the self-reported evaluation of how much the participants valued the P and W app; the overall mean IMI value score was 5 (SD 0.9) out of 7. The perceived usability of the P and W app was stated by the SUS. The overall mean SUS was 68 (SD 11). Finally, the participants were tasked to give the P and W app a grade on a scale from 1 to 10; the mean grade given to the P and W app was a 7 (SD 0.89; Table 4).

Table 4. User evaluation based on the use of work advice, Intrinsic Motivation Inventory (IMI), System Usability Scale (SUS), and grade.

Participant number	Did you receive work advice from the app?[a]	If so, do you intend to do something with this work advice?[b]	IMI[c]	SUS[d]	Grade
1	No[e]	N/A[f]	5.57	85	8
2	No	N/A	4.29	77.5	7
3	Yes	Yes	3.71	55	5
4	Yes	Yes	5.00	77.5	7
5	Yes	Yes	5.14	65	8
6	Yes	No	4.43	77.5	6
7	Yes	Yes	5.57	57.5	7
8	Yes	No	3.00	70	7
9	No	N/A	4.29	75	7
10	Yes	Yes	6.29	55	6
11	Yes	Yes	5.29	50	6
12	Yes	Yes	4.86	72.5	6

[a]Multimedia Appendix 4-III Questionnaire 3, Question 1.

[b]Multimedia Appendix 4-III Questionnaire 3, Question 2.

[c]IMI score; 1=*not at all true* to 7=*very true*.

[d]SUS score; 1=*strongly disagree* to 5=*strongly agree*.

[e]Participants 1, 2, and 9 were convinced that they had not received work advice; however, all participants did receive work advice.

[f]N/A: not applicable as the participant indicated that she did not receive work advice.

Discussion

Principal Findings

The overall effectiveness and efficiency of the 12 participants in performing tasks in the TA sessions are gauged by the completion times and rates and the usability problems. The TA study identified 82 usability problems with a severity ≥1, of which 40 had severity ≥3. The high completion time of the task to create an account seemed to be connected to the many usability problems that participants experienced in this task. As *creating an account* in an mHealth app is not usually part of the core, there is a chance that the design of this first part of the app may be neglected. Design errors in *creating an account*, however, increase the risk of participants dropping out quickly.

We performed an in-depth analysis of the severe usability problems detected regarding *terminology interpretation* and *finding and understanding the work advice* as these issues directly impacted the usefulness of the app. As participants were unable to correctly interpret the terminology in the questionnaire about previous pregnancies, medical disorders, and chemical agents, they did not understand how to complete the questionnaires corresponding to their personal situation. They thus did not receive the correct personal work advice for their circumstances.

Participants also had a different expectation of what the app would show them. Their mental model, the way information is represented in the mind of the end user, affected how they acted in the system in filtering the relevant information. The mental model of the participants did not match how the designer developed the system, as the designer had based it on his own mental model of how future end users would act on the information presented. The mental model of end users, which encompasses values, beliefs, and knowledge, creates perspectives for filtering information and guiding problem solving [46] and has the ability to affect how a person acts [47], differed from that of the designers. The users therefore also experienced problems with understanding the work advice, as their expectations did not match how the designer developed the system (based on his mental model of how future end users should act on information).

Due to the usability problems in its design, 10 out of 12 participants were able to open the work advice page. Only 7 out of these 10 participants understood and intended to follow the work advice given in the app, which was the main goal of the app.

The overall mean IMI score was relatively high (5 out of 7), indicating that the participants did indeed value the use of the app. This corresponded to the overall mean SUS score (68 out of 100) and the mean grade given to the P and W app (7 out of 10).

Comparison With Prior Work

Our main results indicated the effect of the app's navigational structure and screen design on the ability of a specific group of participants—pregnant working women—to find work advice and their intention to follow it thereafter. Other studies in mHealth and electronic health that have applied the TA method have demonstrated that although participants *think* that they have achieved the main goal of using the apps, in reality its intended objective was not reached [48,49]. In one study the researchers observed that the majority of participants, older

cancer patients, were not able to find the requested information although the participants themselves frequently commented during testing that it was easy for them to find it [48]. In a different study, patients with rheumatic diseases were enthusiastic about the possibilities of interactive apps such as peer support forums and online consultations; however, nearly all participants experienced difficulties and were not able to complete all the usability evaluation tasks while interacting with the system [49].

As in our study, other researchers and designers have underlined the importance of an iterative approach in designing mHealth apps to understand the needs of end users as well as improve app usability and feasibility [36,50]. The importance of performing usability studies on mHealth apps to be used in a clinical and patient setting therefore needs serious attention. User testing is an essential part of developing mHealth apps, especially when aiming to effectively change actual patient behavior and/or affect patient outcomes.

Strengths and Limitations

A limitation is that the TA sessions took place in a laboratory setting. In their own home, participants may have taken more time to take a look at the app again. One of the strengths of this study is that the sample size is adequate for obtaining usability problems and that we used a mixed-methods approach— we combined the results of a TA test with the results of questionnaires on demographics, user characteristics, SUS, perceived value (IMI), and evaluation of the app. Another strength of our study is that it was performed by a multidisciplinary team and that the TA study is part of a process in developing an mHealth app, which started with 2 multidisciplinary focus group meetings [29].

Due to a lack of variety in HL levels, we were unable to analyze its potential influence on the TA outcomes. However, the recruitment of only 1 out of 12 participants with limited HL is in line with the estimations of HL prevalence levels in the Netherlands [51]; this certainly applies to a working population.

It is possible that the intention to follow the work advice could change according to the end user's job. However, as a significant proportion of the participants was not able to open the work advice page in the app, and/or understand the work advice or intend to follow it, we think that the influence of profession is limited in this study. For the next study, we would advise asking participants about their job.

To human factor specialists, it is well known that end users should be involved from the beginning when developing an mHealth app. However, those who are well informed about a particular health domain, but less so about medical informatics, should be aware that an iterative multidisciplinary approach with the involvement of the target group from the start by using UEM research in the project is essential and can be very valuable.

The mixed-methods approach provides an insight into the cognitive process of a specific user group—pregnant working women—and their intention to use the P and W app. The TA results, in combination with the questionnaires on the perceived usability and value and the evaluation of the app, showed that incorrect interpretation of terminologies in the system prevented the end users from receiving the correct work advice. They also experienced problems with understanding the work advice because of central design problems in the interface. Despite many usability problems, the participants were relatively positive about the P and W app; the information provided in the app is considered valuable to the end users and meets their needs. The usability findings of this research could then be used to drive recommendations for developers for the next iteration of the P and W app aimed at pregnant working women.

Conclusions

The overall conclusion of this study is that the information provided in the P and W app was considered valuable to the end users, working pregnant women, and meets their needs; however, the usability issues severely impacted the perceived usefulness of the work advice given in the app. The results of this study draw attention to the relation between effective health apps and how their design might hamper their effectiveness in changing patients' behavior. An iterative UEM multidisciplinary approach, with the involvement of the target group from the beginning, is therefore essential for the development of health apps.

The mHealth app will be redesigned and tested in an intervention study, a survey on the effect of the app on actual work adjustment by pregnant women. A future version of the P and W app will be a valuable tool for informing pregnant women about pregnancy-related work risks.

Acknowledgments

The authors would like to thank all the pregnant participants who participated in the TA sessions and the employees of the obstetric care facilities.

This pilot study received funding from ZonMw, the Netherlands Organization for Health Research and Development. This project is part of the Pregnancy and Birth Program.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Questionnaire before think aloud sessions & NVS-D.

[PDF File (Adobe PDF File), 293KB - mhealth_v7i5e11442_app1.pdf]

Multimedia Appendix 2

Think aloud session protocol.

[PDF File (Adobe PDF File), 342KB - mhealth_v7i5e11442_app2.pdf]

Multimedia Appendix 3

Participant tasks during the think aloud sessions: description, achievement and inclusion motivation.

[PDF File (Adobe PDF File), 116KB - mhealth_v7i5e11442_app3.pdf]

Multimedia Appendix 4

Questionnaires after think aloud session: SUS, IMI, and questionnaire 3.

[PDF File (Adobe PDF File), 398KB - mhealth_v7i5e11442_app4.pdf]

Multimedia Appendix 5

Overview of all usability problems and bugs.

[PDF File (Adobe PDF File), 221KB - mhealth_v7i5e11442_app5.pdf]

References

1. Gao G, Livingston G. Pew Research Center. Working while pregnant is much more common than it used to be URL: http://www.pewresearch.org/fact-tank/2015/03/31/working-while-pregnant-is-much-more-common-than-it-used-to-be/ [accessed 2018-04-06] [WebCite Cache ID 6yTgtLQda]
2. Central Statistics Office, the Netherlands. [Labor participation by age and gender] URL: https://www.cbs.nl/nl-nl/achtergrond/2017/07/arbeidsparticipatie-naar-leeftijd-en-geslacht [accessed 2017-11-30] [WebCite Cache ID 6vMHGCxMk]
3. Bonde JP, Jørgensen KT, Bonzini M, Palmer KT. Miscarriage and occupational activity: a systematic review and meta-analysis regarding shift work, working hours, lifting, standing, and physical workload. Scand J Work Environ Health 2013 Jul;39(4):325-334 [FREE Full text] [doi: 10.5271/sjweh.3337] [Medline: 23235838]
4. Vrijkotte TG, van der Wal MF, van Eijsden M, Bonsel GJ. First-trimester working conditions and birthweight: a prospective cohort study. Am J Public Health 2009 Aug;99(8):1409-1416. [doi: 10.2105/AJPH.2008.138412] [Medline: 19542045]
5. van Beukering MD, van Melick MJ, Mol BW, Frings-Dresen MH, Hulshof CT. Physically demanding work and preterm delivery: a systematic review and meta-analysis. Int Arch Occup Environ Health 2014 Nov;87(8):809-834. [doi: 10.1007/s00420-013-0924-3] [Medline: 24390632]
6. Palmer KT, Bonzini M, Harris EC, Linaker C, Bonde JP. Work activities and risk of prematurity, low birth weight and pre-eclampsia: an updated review with meta-analysis. Occup Environ Med 2013 Apr;70(4):213-222 [FREE Full text] [doi: 10.1136/oemed-2012-101032] [Medline: 23343859]
7. Croteau A, Marcoux S, Brisson C. Work activity in pregnancy, preventive measures, and the risk of delivering a small-for-gestational-age infant. Am J Public Health 2006 May;96(5):846-855. [doi: 10.2105/AJPH.2004.058552] [Medline: 16571706]
8. Croteau AA, Marcoux SS, Brisson C. Work activity in pregnancy, preventive measures, and the risk of preterm delivery. Am J Epidemiol 2007 Oct 15;166(8):951-965. [doi: 10.1093/aje/kwm171] [Medline: 17652310]
9. Juhl M, Larsen PS, Andersen PK, Svendsen SW, Bonde JP, Nybo AA, et al. Occupational lifting during pregnancy and child's birth size in a large cohort study. Scand J Work Environ Health 2014 Jul;40(4):411-419 [FREE Full text] [doi: 10.5271/sjweh.3422] [Medline: 24647563]
10. Lee BE, Ha M, Park H, Hong Y, Kim Y, Kim YJ, et al. Psychosocial work stress during pregnancy and birthweight. Paediatr Perinat Epidemiol 2011 May;25(3):246-254. [doi: 10.1111/j.1365-3016.2010.01177.x] [Medline: 21470264]
11. Loomans EM, van Dijk DA, Vrijkotte TG, van Eijsden M, Stronks K, Gemke RJ, et al. Psychosocial stress during pregnancy is related to adverse birth outcomes: results from a large multi-ethnic community-based birth cohort. Eur J Public Health 2013 Jun;23(3):485-491. [doi: 10.1093/eurpub/cks097] [Medline: 22850187]
12. Mocevic E, Svendsen SW, Jørgensen KT, Frost P, Bonde JP. Occupational lifting, fetal death and preterm birth: findings from the Danish National Birth Cohort using a job exposure matrix. PLoS One 2014;9(3):e90550 [FREE Full text] [doi: 10.1371/journal.pone.0090550] [Medline: 24614129]

13. Hooftman WE, van den Bossche SJN. Original Ministry Social Affairs and Employment. 2007. Pregnancy and Work: counseling, measures and sick leave URL: http://www.monitorarbeid.tno.nl/dynamics/modules/SPUB0102/view.php?pub_Id=100111&att_Id=4911 [WebCite Cache ID 6mwmOz1re]
14. 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 May 27;4(2):e53 [FREE Full text] [doi: 10.2196/mhealth.5197] [Medline: 27234926]
15. Abroms LC, Johnson PR, Heminger CL, Van Alstyne J, 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]
16. Evans W, Nielsen PE, Szekely DR, Bihm JW, Murray EA, Snider J, et al. Dose-response effects of the text4baby mobile health program: randomized controlled trial. JMIR Mhealth Uhealth 2015 Jan 28;3(1):e12 [FREE Full text] [doi: 10.2196/mhealth.3909] [Medline: 25630361]
17. Overdijkink SB, Velu AV, Rosman AN, van Beukering MD, Kok M, Steegers-Theunissen RP. The usability and effectiveness of mobile health technology-based lifestyle and medical intervention apps supporting health care during pregnancy: systematic review. JMIR Mhealth Uhealth 2018 Apr 24;6(4):e109 [FREE Full text] [doi: 10.2196/mhealth.8834] [Medline: 29691216]
18. World Health Organization. 2011. mHealth New horizons for health through mobile technologies URL: http://www.who.int/goe/publications/goe_mhealth_web.pdf [WebCite Cache ID 6yTi2OKxp]
19. 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]
20. 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;10(1):e1001363 [FREE Full text] [doi: 10.1371/journal.pmed.1001363] [Medline: 23458994]
21. Tamrat T, Kachnowski S. Special delivery: an analysis of mHealth in maternal and newborn health programs and their outcomes around the world. Matern Child Health J 2012 Jul;16(5):1092-1101. [doi: 10.1007/s10995-011-0836-3] [Medline: 21688111]
22. Majeed-Ariss R, Baildam E, Campbell M, Chieng A, Fallon D, Hall A, et al. Apps and adolescents: a systematic review of adolescents' use of mobile phone and tablet apps that support personal management of their chronic or long-term physical conditions. J Med Internet Res 2015 Dec 23;17(12):e287 [FREE Full text] [doi: 10.2196/jmir.5043] [Medline: 26701961]
23. Tripp N, Hainey K, Liu A, Poulton A, Peek M, Kim J, et al. An emerging model of maternity care: smartphone, midwife, doctor? Women Birth 2014 Mar;27(1):64-67. [doi: 10.1016/j.wombi.2013.11.001] [Medline: 24295598]
24. Badawy SM, Kuhns LM. Texting and mobile phone app interventions for improving adherence to preventive behavior in adolescents: a systematic review. JMIR Mhealth Uhealth 2017 Apr 19;5(4):e50 [FREE Full text] [doi: 10.2196/mhealth.6837] [Medline: 28428157]
25. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
26. Brown W, Yen P, Rojas M, Schnall R. Assessment of the health IT Usability Evaluation Model (Health-ITUEM) for evaluating mobile health (mHealth) technology. J Biomed Inform 2013 Dec;46(6):1080-1087 [FREE Full text] [doi: 10.1016/j.jbi.2013.08.001] [Medline: 23973872]
27. Din E. International Organization for Standardization. 1998. 9241-11 Ergonomic requirements for office work with visual display terminals (VDTs). Part II: Guidance on usability URL: https://www.iso.org/obp/ui/ [accessed 2019-04-27] [WebCite Cache ID 77x2YWeRH]
28. Dumas J. The great leap forward: the birth of the usability profession (1988-1993). J Usability Stud 2007;2(2):54-60 [FREE Full text]
29. Kushniruk AW, Triola MM, Borycki EM, Stein B, Kannry JL. Technology induced error and usability: the relationship between usability problems and prescription errors when using a handheld application. Int J Med Inform 2005 Aug;74(7-8):519-526. [doi: 10.1016/j.ijmedinf.2005.01.003] [Medline: 16043081]
30. Horsky J, Zhang J, Patel VL. To err is not entirely human: complex technology and user cognition. J Biomed Inform 2005 Aug;38(4):264-266 [FREE Full text] [doi: 10.1016/j.jbi.2005.05.002] [Medline: 15967732]
31. Peute LW, Spithoven R, Bakker PJ, Jaspers MW. Usability studies on interactive health information systems; where do we stand? Stud Health Technol Inform 2008;136:327-332. [Medline: 18487752]
32. Castensøe-Seidenfaden P, Reventlov Husted G, Teilmann G, Hommel E, Olsen BS, Kensing F. Designing a self-management app for young people with type 1 diabetes: methodological challenges, experiences, and recommendations. JMIR Mhealth Uhealth 2017 Oct 23;5(10):e124 [FREE Full text] [doi: 10.2196/mhealth.8137] [Medline: 29061552]
33. [Pregnancy and birth An impression of the knowledge network birth care and research projects The Hague]. [Project-The App to work healthy during pregnancy] URL: https://www.zonmw.nl/nl/onderzoek-resultaten/doelmatigheidsonderzoek/programmas/project-detail/zwangerschap-en-geboorte/de-app-gezond-werken-tijdens-de-zwangerschap/ [accessed 2017-11-30] [WebCite Cache ID 6vMIAS1YS]

34. Advice and guidance by the occupational physician. Practice Guideline-Pregnancy, Postpartum Period and Work URL: https://www.nvab-online.nl/sites/default/files/bestanden-webpaginas/Guideline_Pregnancy_Postpartum_Period_and_Work_0.pdf [accessed 2019-04-24] [WebCite Cache ID 77sEBK6rX]
35. Velu AV, van Beukering MD, Schaafsma FG, Frings-Dresen MH, Mol BW, van der Post JA, et al. Barriers and facilitators for the use of a medical mobile app to prevent work-related risks in pregnancy: a qualitative analysis. JMIR Res Protoc 2017 Aug 22;6(8):e163 [FREE Full text] [doi: 10.2196/resprot.7224] [Medline: 28830851]
36. Jaspers MW, Steen T, van den Bos C, Geenen M. The think aloud method: a guide to user interface design. Int J Med Inform 2004 Nov;73(11-12):781-795. [doi: 10.1016/j.ijmedinf.2004.08.003] [Medline: 15491929]
37. Van Engen-Verheul M, Peute L, Kilsdonk E, Peek N, Jaspers M. Usability evaluation of a guideline implementation system for cardiac rehabilitation: think aloud study. Stud Health Technol Inform 2012;180:403-407. [Medline: 22874221]
38. Peute LW, de Keizer NF, Jaspers MW. The value of Retrospective and Concurrent Think Aloud in formative usability testing of a physician data query tool. J Biomed Inform 2015 Jun;55:1-10 [FREE Full text] [doi: 10.1016/j.jbi.2015.02.006] [Medline: 25733166]
39. Weiss BD, Mays MZ, Martz W, Castro KM, DeWalt DA, Pignone MP, et al. Quick assessment of literacy in primary care: the newest vital sign. Ann Fam Med 2005;3(6):514-522 [FREE Full text] [doi: 10.1370/afm.405] [Medline: 16338915]
40. Fransen MP, Leenaars KE, Rowlands G, Weiss BD, Maat HP, Essink-Bot M. International application of health literacy measures: adaptation and validation of the newest vital sign in The Netherlands. Patient Educ Couns 2014 Dec;97(3):403-409. [doi: 10.1016/j.pec.2014.08.017] [Medline: 25224314]
41. Ericsson KA, Simon HA. Verbal reports as data. Psychological Review 1980;87(3):215-251. [doi: 10.1037/0033-295X.87.3.215]
42. Brooke J. SUS-A quick and dirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland IL, editors. Usability evaluation in industry. London: Taylor and Francis; 1996:4-7.
43. Plant RW, Ryan RM. Intrinsic motivation and the effects of self-consciousness, self-awareness, and ego-involvement: an investigation of internally controlling styles. J Personality 1985 Sep;53(3):435-449. [doi: 10.1111/j.1467-6494.1985.tb00375.x]
44. Jaspers MW. A comparison of usability methods for testing interactive health technologies: methodological aspects and empirical evidence. Int J Med Inform 2009 May;78(5):340-353. [doi: 10.1016/j.ijmedinf.2008.10.002] [Medline: 19046928]
45. Nielsen J. Usability Engineering. Fremont: Morgan Kaufmann; 1994.
46. Eckert E, Bell AA. Invisible force: farmers' mental models and how they influence learning and actions. J Ext 2005;43(3):2 [FREE Full text]
47. Rook L. Emerald Insight. Mental models: a robust definition URL: https://www.emeraldinsight.com/doi/abs/10.1108/09696471311288519 [accessed 2019-04-24] [WebCite Cache ID 77sFYSxes]
48. Bolle S, Romijn G, Smets EM, Loos EF, Kunneman M, van Weert JC. Older cancer patients' user experiences with web-based health information tools: a think-aloud study. J Med Internet Res 2016 Dec 25;18(7):e208 [FREE Full text] [doi: 10.2196/jmir.5618] [Medline: 27457709]
49. van der Vaart R, Drossaert CH, de Heus M, Taal E, van de Laar MA. Measuring actual eHealth literacy among patients with rheumatic diseases: a qualitative analysis of problems encountered using Health 1.0 and Health 2.0 applications. J Med Internet Res 2013 Feb 11;15(2):e27 [FREE Full text] [doi: 10.2196/jmir.2428] [Medline: 23399720]
50. White BK, Martin A, White JA, Burns SK, Maycock BR, Giglia RC, et al. Theory-based design and development of a socially connected, gamified mobile app for men about breastfeeding (Milk Man). JMIR Mhealth Uhealth 2016 Jun 27;4(2):e81 [FREE Full text] [doi: 10.2196/mhealth.5652] [Medline: 27349756]
51. Sørensen K, Pelikan J, Röthlin F, Ganahl K, Slonska Z, Doyle G, HLS-EU Consortium. Health literacy in Europe: comparative results of the European health literacy survey (HLS-EU). Eur J Public Health 2015 Dec;25(6):1053-1058 [FREE Full text] [doi: 10.1093/eurpub/ckv043] [Medline: 25843827]

Abbreviations

HL: health literacy
IMI: Intrinsic Motivation Inventory
mHealth: mobile health
P and W app: Pregnancy and Work app
SUS: System Usability Scale
TA: think aloud
UEM: usability evaluation method

Edited by C Dias; submitted 29.06.18; peer-reviewed by C Doarn, S Chokshi, R Marcilly, A Rosman; comments to author 26.07.18; revised version received 14.10.18; accepted 27.12.18; published 09.05.19.

Please cite as:
van Beukering M, Velu A, van den Berg L, Kok M, Mol BW, Frings-Dresen M, de Leeuw R, van der Post J, Peute L
Usability and Usefulness of a Mobile Health App for Pregnancy-Related Work Advice: Mixed-Methods Approach
JMIR Mhealth Uhealth 2019;7(5):e11442
URL: https://mhealth.jmir.org/2019/5/e11442/
doi:10.2196/11442
PMID:31094353

Original Paper

Achieving Sustainability and Scale-Up of Mobile Health Noncommunicable Disease Interventions in Sub-Saharan Africa: Views of Policy Makers in Ghana

Daniel Opoku[1], MScPH; Reinhard Busse[1], MD, MPH; Wilm Quentin[1], PD, MD, MSc HPPF

Faculty VII Economics and Management, Institute of Technology and Management, Department of Health Care Management, Technische Universität Berlin, Berlin, Germany

Corresponding Author:
Daniel Opoku, MScPH
Faculty VII Economics and Management
Institute of Technology and Management
Department of Health Care Management
Technische Universität Berlin
Straße des 17. Juni 135
Administrative office H80
Berlin, 10623
Germany
Phone: 49 3031429222
Fax: 49 30 314 28433
Email: Daniel.Opoku@TU-Berlin.de

Abstract

Background: A growing body of evidence shows that mobile health (mHealth) interventions may improve treatment and care for the rapidly rising number of patients with noncommunicable diseases (NCDs) in sub-Saharan Africa (SSA). A recent realist review developed a framework highlighting the influence of context factors, including predisposing characteristics, needs, and enabling resources (PNE), for the long-term success of mHealth interventions. The views of policy makers will ultimately determine implementation and scale-up of mHealth interventions in SSA. However, their views about necessary conditions for sustainability and scale-up remain unexplored.

Objective: This study aimed to understand the views of policy makers in Ghana with regard to the most important factors for successful implementation, sustainability, and scale-up of mHealth NCD interventions.

Methods: Members of the technical working group responsible for Ghana's national NCD policy were interviewed about their knowledge of and attitude toward mHealth and about the most important factors contributing to long-term intervention success. Using qualitative methods and applying a qualitative content analysis approach, answers were categorized according to the PNE framework.

Results: A total of 19 policy makers were contacted and 13 were interviewed. Interviewees had long-standing work experience of an average of 26 years and were actively involved in health policy making in Ghana. They were well-informed about the potential of mHealth, and they strongly supported mHealth expansion in the country. Guided by the PNE framework's categories, the policy makers ascertained which critical factors would support the successful implementation of mHealth interventions in Ghana. The policy makers mentioned many factors described in the literature as important for mHealth implementation, sustainability, and scale-up, but they focused more on enabling resources than on predisposing characteristics and need. Furthermore, they mentioned several factors that have been rather unexplored in the literature.

Conclusions: The study shows that the PNE framework is useful to guide policy makers toward a more systematic assessment of context factors that support intervention implementation, sustainability, and scale-up. Furthermore, the framework was refined by adding additional factors. Policy makers may benefit from using the PNE framework at the various stages of mHealth implementation. Researchers may (and should) use the framework when investigating reasons for success (or failure) of interventions.

(JMIR Mhealth Uhealth 2019;7(5):e11497) doi:10.2196/11497

KEYWORDS

implementation science; mHealth; eHealth; noncommunicable diseases; disease management; sub-Saharan Africa; qualitative research; health policy

Introduction

Background

With the drastic decline of communicable, maternal, and neonatal diseases as cause of death and burden of disease across the globe [1], and particularly in Africa [1,2], the epidemiologic transition toward noncommunicable diseases (NCDs) is in full swing. By 2030, 42% of all projected deaths in sub-Saharan Africa (SSA) will be caused by NCDs, which will then surpass communicable diseases as the leading cause of death in the subregion [3,4]. In some African countries such as Ghana, statistics show that already today about 42% of the total annual deaths are caused by NCDs, led by cardiovascular diseases [5-7].

At the same time, most countries in SSA have become eager adopters and innovators of the use of mobile and digital technologies. In Ghana, as early as 2013, "[m]ore than four out of every five households (80.3%) in the country own[ed] a mobile phone" [8], thereby expanding the opportunities for the implementation of mobile phone–based health (mHealth) interventions [9,10]. Numerous studies and reviews have reported positive results of mHealth interventions against NCDs [11-15]. The World Health Organization (WHO) promotes the further development and more widespread use of mHealth interventions as part of its Global Action Plan for the prevention and control of NCDs [16]. Nevertheless, most mHealth interventions remain at the stage of pilot projects, and they are almost never scaled-up to entire countries [10,13,17-19].

Efforts at international and European levels have aimed to provide guidance to countries to support scale-up of mHealth interventions and integration into routine care practices [20-22]. For example, WHO and the International Telecommunication Union have produced a detailed toolkit to support the development of national electronic health (eHealth) strategies [21]. The toolkit is focused on the role of enabling legislation and regulation, government and sector buy-in, and planning and funding for implementation and sustainability. More recently, the European Union–funded Momentum project for successful implementation of telemedicine into routine health care has published a list of 18 factors to make telemedicine a success, which also include legislation and sector buy-in, but further recommend consideration of the cultural readiness toward telemedicine and the identification of a compelling need [22].

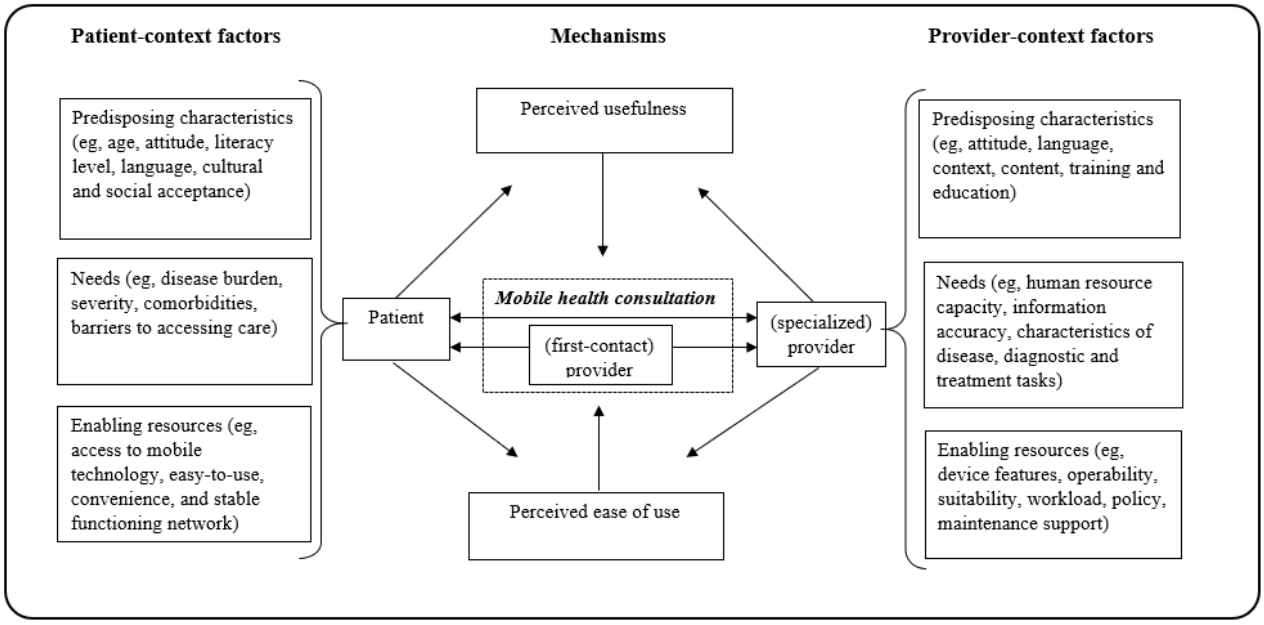
In a recent realist review, Opoku et al [23] developed a theoretical framework that aims to provide guidance to policy makers and other decision makers working on implementing, sustaining, and scaling-up mHealth interventions for NCD management in SSA. The framework hypothesizes that "predisposing characteristics and need of patients and healthcare providers as well as the availability of enabling resources in the community influence the perceptions of patients and providers that mHealth interventions are useful and easy to use—and these perceptions are essential for the successful implementation of an mHealth intervention" [23]. As shown in Figure 1, the framework focuses attention on the influence of context factors, including predisposing characteristics, needs, and enabling resources (PNE), for the long-term success of mHealth interventions. Therefore, we use the term *mHealthPNE framework* for the rest of the paper.

The mHealth PNE framework is grounded in the experiences of patients and health care providers as reported in 20 studies of 18 mHealth interventions for NCDs [14,24-42] conducted in 10 SSA countries. It combines the Andersen behavioral model of health services utilization with the Davis technology acceptance model [43,44]. The framework focuses attention on a large set of—yet to be further refined—contextual factors that can be grouped under PNE. For example, cultural readiness mentioned as one of the 18 factors of the momentum group would fall under *predisposing characteristics*, whereas establishment of an appropriate legal environment would fall under *enabling resources* and identification of patients' needs under *needs*. However, it remains unknown whether the context factors that have so far been identified under the categories of PNE are in line with the views of policy makers and other decision makers about the most important conditions for implementation, sustainability, and scale-up.

Ghana is one of the countries in SSA where efforts to support the development of mHealth interventions have been most pronounced [45-48]. These efforts include the development of the Ghana eHealth Strategy, which aims at supporting the improvement of the overall performance of the health sector [47]. In addition, several mHealth interventions have been implemented, including the Millennium Villages telemedicine project in the Amansie West district [49] and the Mobile Technology for Community Health program in 7 districts [50,51]. As a result, policy makers in Ghana can be expected to have considerable experience with mHealth interventions, and they are likely to have thought about factors that support intervention sustainability and scale-up.

The views of policy makers and other decision makers will ultimately determine the implementation, sustainability, and scale-up of mHealth interventions in SSA. To assure that the mHealth PNE framework is useful as a guide for policy makers, it is important that the framework is sufficiently aligned with their thinking, that is, policy makers should find the categories of the framework useful when considering the most important factors for implementation, sustainability, and scale-up. In addition, the experiences of policy makers may provide additional insights about the important factors contributing to a successful implementation, sustainability, and scale-up of mHealth NCD interventions that might be missing in the existing literature [23].

Figure 1. Mobile heatlh predisposing characteristics, needs, and enabling resources framework.



Objectives

Therefore, the aim of this study was to understand the views of policy makers with regard to the most important factors that should be considered to assure successful implementation, sustainability, and scale-up of mHealth NCD interventions, thus contributing to the improvement of the mHealth PNE framework. More specifically, the study sought to (1) assess policy makers' knowledge of and attitude toward mHealth NCD interventions, (2) identify whether the categories of the framework are useful to structure the thinking of policy makers, and (3) integrate the perspectives of policy makers into the various components of the framework.

Methods

Application for ethical review of the study was submitted to the Committee on Human Research, Publications, and Ethics at the Kwame Nkrumah University of Science and Technology, School of Medical Sciences, and Komfo Anokye Teaching Hospital, Kumasi, Ghana, and final approval was received on February 25, 2016. The study was conducted using qualitative methods (interviews) and by applying a qualitative content analysis (QCA) approach [52]. The paper was drafted following the Consolidated Criteria for Reporting Qualitative Studies [53].

Qualitative Interviews

Informed consent was first sought and participants were given sufficient information, including about the risks and benefits of participating in the study. Guided by a semistructured questionnaire (Multimedia Appendix 1), qualitative interviews were conducted by DO between November 2015 and January 2016 among stakeholders at the health policy direction level, who were actively involved in national health policy decision making and implementation in Ghana. These one-to-one interviews lasted for an average of 45 min and were recorded for transcription and analysis. In addition, field notes were taken for purposes such as capturing off-tape records and explaining why an interview might have been poorly conducted.

Participants

Participants were from diverse backgrounds and generally worked at high levels of hierarchy and responsibility in different institutions. They had a long-standing experience working in various sectors of the Ghana national health system, particularly on NCDs and other related subjects. All participants were involved in drafting and developing the 2012 National Policy for the Prevention and Control of Chronic Non-Communicable Diseases in Ghana. As such, all participants had been involved in defining the technical direction and framework for implementing NCD-related programs in the country [54].

Selection Criteria

A list of the members of the technical working group for Ghana's national NCD policy was retrieved from the document titled *Strategy for the Management, Prevention and Control of Non-Communicable Diseases in Ghana* by the Republic of Ghana, Ministry of Health [54]. The list consisted of a total number of 19 members who were contacted by DO and who received information about the study via emails, telephone calls, and Skype calls. They were medical doctors, including general practitioners and public health consultants, epidemiologists, political scientists, (public health) lecturers, public health researchers, health educators, program managers, disease surveillance officers, international health specialists, program coordinators, (public health) pharmacists, dieticians, public health practitioners, policy advisors or analysts, planning officers, and freelance nutritionists. Participation was voluntary, and participants were assured that information such as names and addresses that could lead to their identification would be avoided to ensure privacy. Those who responded were followed up for the interviews. No restrictions were imposed except that

the participation was based on the availability and willingness to contribute during the period of data collection.

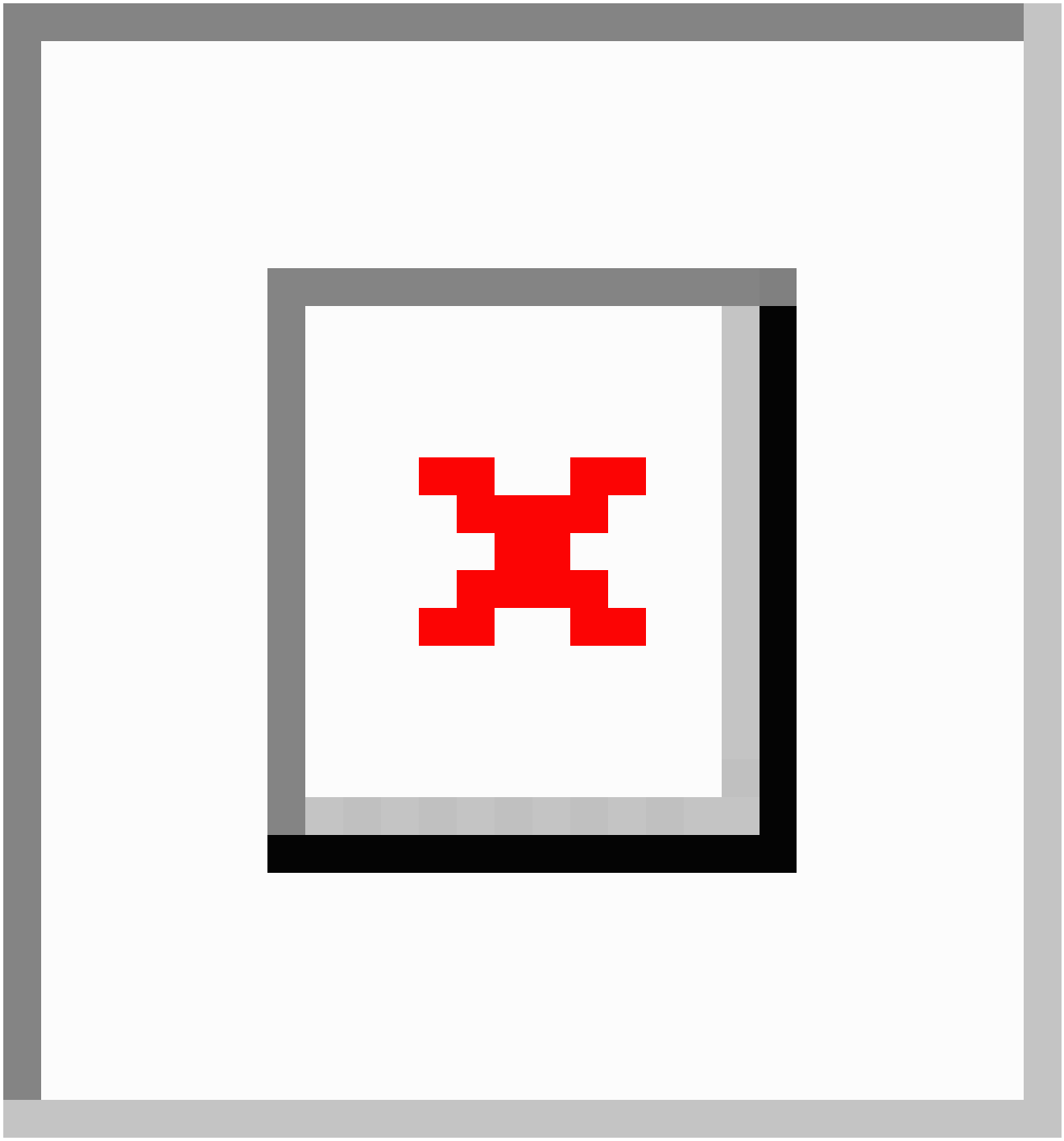
Analysis

Following the QCA approach [52,55], the coding frame in Figure 2 was used for the analysis. It was largely based on the mHealth PNE framework. The framework theorizes that successful implementation of mHealth interventions is determined by context factors— *predisposing characteristics, enabling factors, and needs* —of patients and health care providers, which influence their perceptions on the usefulness and ease of use of the intervention [23]. Thus, for example, whether a mobile phone–based self-monitoring blood glucose intervention designed for diabetes care in Ghana will be successful or not depends on whether both diabetic patients and their health care providers perceive the intervention to be useful and easy to use.

According to the framework, the perceived usefulness and ease of use of an intervention are determined by (1) patients' predisposing characteristics, such as age, attitude, literacy, language, and cultural or social acceptability; (2) their need, such as reducing financial burden of care and avoiding long travel or waiting time; and (3) the factors that will enable them to utilize the intervention well, which may include access to a mobile phone and a stable network [23]. In addition, perceived usefulness and ease of use of providers depend on predisposing characteristics (eg, technology-related training), needs (eg, human resource capacity), and enabling resources (eg, tolerable workload and incentives) [23].

The interview transcripts for the analysis were first coded by DO and subsequently reviewed by both DO and WQ, according to the various components of the framework and grouped into main categories and subcategories. The 2 main categories were *knowledge of and attitudes toward mHealth* and *context factors determining sustainability of mHealth (patient-context factors* and *provider-context factors).* The results were analyzed mainly based on the 3 subcategories of the framework (ie, *predisposing characteristics, needs, and enabling resources*) and then presented thematically under each of the main categories. The analysis also sought to identify other potentially relevant factors missing in this framework.

Figure 2. Coding frame for analysis based on the mobile health predisposing characteristics, needs, and enabling resources framework.



Results

Characteristics of Participants

Out of the 19 policy makers who were contacted, 13 participated in the study. The participant policy makers had long-standing experiences with an average of 26 years in managing various health programs, interventions, and departments and were actively involved in health policy making processes in Ghana. With the exception of 4 participants who had retired at the time the interviews were conducted, all were serving in high-level (national) capacities at the Ministry of Health, Ghana Health Services, academia, and the public and private sectors. Table 1 summarizes the characteristics of the participants.

Table 1. Characteristics of the participant health policy makers and managers in Ghana.

Gender[a]	Age (years)	Working experience with noncommunicable diseases	Experience (years)
Male	>50	Medical practice, program management, policy	21
Male	>50 (retired)	Medical practice, program management, policy	36
Male	>50 (*retired*)	Medical practice, teaching, research, program management, policy	41
Male	>50	Research, health information management, policy	33
Male	>50	Health education, training, research, communication, program management, policy	25
Male	40-44	Health regulations, disease control and prevention, policy	19
Male	40-44	Medical practice, disease control and prevention, policy	13
Female	45-49	Clinical practice, health promotion, policy	20
Male	45-49	Teaching, research, consultancy	16
Female	>50 (retired)	Health promotion, disease prevention, policy	>30
Male	>50	Health sector coordination, program management, policy	30
Male	>50	Teaching, research, consultancy	21
Female	>50 (retired)	Health promotion, advocacy, policy	34

[a]Source: authors' own compilation.

Table 2. Identified beneficial applications of mobile health interventions.

Health promotion and prevention[a]	Education and awareness creation (R[b]: 1, 5, 7, 10), follow-up (R: 1), information centers (R: 11), interactive platform (R: 12)
Health care delivery (maternal and child health care)	Scheduling/adherence/compliance/reminder (appointment and medication) (R: 1, 2, 3, 7, 11, 12), (emergency/specialist) referrals (R: 2, 3, 13), follow-up (R: 1, 11), community-based health care (R: 4, 11), digitalized hospital records (R: 6, 8), record-keeping (vital statistics) (R: 4), creating access to health care (R: 4), health information and follow-up (for pregnant women) (R: 13)
Noncommunicable disease–related management	Education (regenerative health and nutrition) (R: 4, 8, 11), specialized care (for complicated cases) (R: 9, 12), appointments/reminders for testing fasting blood sugar (R: 3), checking/monitoring vital signs (blood pressure and retina check) (R: 4), control and prevention of hypertension (awareness creation, and reminders for drug refill) (R: 7), early detection of complications [R: 9], cancer registry (R: 11), follow-ups (R: 11)

[a]Source: authors' own compilation.

[b]R stands for respondent and the following numbers assigned in this study.

Knowledge of and Attitude Toward Mobile Health Interventions

Interviewed policy makers had considerable knowledge of mHealth interventions, broadly in relation to the general field of health (prevention and health promotion and health care delivery—maternal and child health care) and some specifically relating to NCDs (hypertension prevention and control). Table 2 presents the beneficial applications of mHealth interventions, sorted out according to the highest number of participants who identified them. Some of the policy makers actually had a long-time experience with mHealth and had been involved in the use of mobile phones to support health care delivery, either as providers or as patients themselves.

All interviewees agreed that mHealth interventions can contribute to improved NCD management in Ghana. They identified a range of potentially beneficial applications: awareness creation and (regenerative) health education, early detection of NCD conditions, reduction of waiting time, follow-ups and monitoring, keeping track of the appointments of patients, vital statistics and adherence to medication, emergency alert, creating registries, record keeping, dissemination of evidence, and ensuring sustainable health care. Most importantly, the policy makers highlighted that mHealth could potentially help patients to better manage their NCDs and improve treatment compliance:

> *[it] can keep patients in care, reduce morbidity, reduce mortality. Definitely because it's all a matter of keeping them [ie, patients] in care and ensuring that they learn the good practices and all that. So, I think it would help the outcomes; we will get better outcomes, reduce the disabilities from NCDs, and also reduce the mortalities from NCDs, definitely!* [Respondent: 1]

> *For me the biggest impact is that it will help to manage treatments, it will reduce treatment failures, and it will help people to be more productive so that people can then take better care of themselves and not spend all the time going to the hospitals.* [Respondent: 4]

At the same time, interviewees noted that mHealth interventions provide a solution for only some of the problems of NCD management in Ghana:

> *I always get worried when people try to use technology as a "fix all". Technology is not a fix all, it fixes some problems but not all problems and it's contextual.* [Respondent: 1]

Given the complexity of NCD management, mHealth consultations were considered to be safe and suitable mostly for follow-ups, after an initial contact between patients and providers has been established:

> *...there are huge potentials when it comes to the use of mobile phones but for noncommunicable diseases, the evidence is not very clear for us...from our experience it has to be a "postcontact" intervention. There is always the first contact that has to be made [at the facility] and thenthe intervention kicks in as a follow-up,only after the initial contact. If you don't have initial contact with the hospital, the opportunity to rope in ICT to help you to readjust and to be healthier becomes a bit of a problem.* [Respondent: 4]

However, the policy makers maintained that the use of mobile phones in health care is becoming an important strategy in Ghana, particularly in reducing maternal mortalities and controlling epidemics. In fact, the policy makers were enthusiastic about the potential of mHealth to improve NCD management in Ghana:

> *It is a very good idea, brilliant idea! I mean it is something that we've always been talking about that people should be able to stay in their houses and manage or even call doctors to come or even call for advice from doctors.* [Respondent: 5]

Perspectives of Policy Makers on Context Factors Determining Sustainability of Mobile Health in Ghana

This section presents the interview results categorized along the 3 context factors of the analytical framework, that is, *predisposing characteristics, needs, and enabling resources*. Table 3 provides a summary of the identified factors supporting and/or expanding the framework, arranged in a descending order of the most frequently mentioned factors by the participants.

Predisposing Characteristics

According to the framework, the most important predisposing characteristics supporting the implementation of mHealth interventions are a positive attitude, cultural/social acceptance, and a common language of communication. The interviewed policy makers, however, identified age, literacy, and level of education, as well as providers' continuous training, upgrade, and education as more important factors:

> *The youth are very good at these things and so if you work it out with them it would work. The problem is, are the youth the people who actually go for the services? And so, the majority of the people who would be having noncommunicable diseases are not the youth and they are the people who would not understand this.* [Respondent: 3]

> *My fear is the illiteracy rate. How many people with mobile phones know how to send a text message? How many people can even store or delete numbers or messages? They don't know, they have the phones for receiving calls and making calls, that's all!* [Respondent: 2]

> *We need to do more in terms of training our providers. We call it the two ends, or supply and demand. That is, we supply and the population would demand. So, let us tailor a kind of training for the suppliers of the services in terms of the use of mobile phones and then also teach some of our clients on the use of the mobile phone in terms of getting access to some of these specialists, because I still want to believe that the management of NCDs is a specialized service, which cannot be left in the hands of 'ordinaries'.* [Respondent: 11]

They also stressed that, more generally, trust and confidence among health care providers is a prerequisite for successful implementation of mHealth:

> *The other thing would have to do with the [health] staff attitude. The major question is that, for example, if I [a specialized health care provider] at Korle-Bu teaching hospital [in Accra] give instructions to those [health care providers] in the rural areas, how sure am I that they are doing what I am instructing them to do? And if anything at all should go wrong, who is to be blamed, me or those out there? So, amongst the health staff, there are usually pessimistic views and so some of them will not be interested in these innovations but others might.* [Respondent: 9]

In addition, interviewees highlighted that attitudes of patients related to myths, misconceptions, fear of change, and phobia for technological innovations may negatively impact patients' perceptions about the usefulness of mHealth interventions:

> *We need to look at people's phobia for technology and see how that barrier can be broken.* [Respondent: 6]

> *Maybe we still maintain the old ways of doing things; we don't like change. Africans in general, but Ghanaians especially, we fear change. It is a fact that we fear about what if it doesn't work out well and who takes the fall for it!* [Respondent: 10]

With regard to predisposing characteristics influencing the perceived ease of use, interviewees believed that urban populations are more familiar with mobile technologies. However, in general, the Ghanaian population was thought to be ready to use mobile phones for health care given the high penetration of the technology.

Table 3. Summary of the identified factors supporting the mobile health predisposing characteristics, need, and enabling resources framework.

Mechanism context[a]	Patient		First contact/specialized provider	
	Perceived usefulness	Perceived ease of use	Perceived usefulness	Perceived ease of use
Predisposing characteristics	(Local) language [R[b]:5,6,11,12]; myths, fear/phobia, misconceptions [R:2,5,6,10];*informed, convinced, trust, and confidence (satisfaction)* [R:2,8,11]; *locality (urban/rural)*[c] [R:2,8]; socioculture [R:4,7]; acceptance [R:5,6]; (positive) attitude [R:5,12]; self-motivation [R:3]; age [R:8]; *gender* [R:8]; social class (middle) [R:1]	Literacy and level of education [R:1,2,3,4,5,6,7,11,12]; age (youth ≥10 years, adults) [R:2,3,5,7,10,13]; penetration, and familiarity (urban) [R:1,5,6,13]; training, know-how, confidence [R:3,4,5,12]; basic, simple [R:6,8]; *personalization* [R:8,11]	(Positive) attitude interest, dedication, willingness, and motivation [R:1,8,12]; *good (provider-patient/community)* relationship [R:4,8,11]; language [R:5,9]; trust and confidence [R:11]; ready to support [R:13]	Continuous training, upgrade, and education [R:4,7,9,10,11,12]
Need	Health care access barriers (poverty, transportation, ineffective health facilities, distance, travel and waiting time, cost, urgency and quality of care, stress reduction, and satisfaction) [R:2,3,4,9,10,12,13]; disease condition (severity, upsurge, uncertainties of care) [R:1,2,4,6,9,13]; *need for urgent/special care* [R:7,8,9,13]	*Technology-drivenneed/demand* [R:2,3,4,6,13]	Reduce burden of cases/workload [R:2,6,10,11,12,13]; lack of human resources (limited specialists, unequal distributions of professionals, lack of motivation) [R:9,11,12,13]; *integrated care* [R:3,10,13]; lack of necessary systems and infrastructure (health facility, referral system, transport) [R:9,11]; *continuity of care* [R:1,13]; lack of accurate information [2,11]; reduce morbidity/mortality [R:11,12]; *exchange of expertise* [R:9]; cost-saving [R:9]; enhance emergency care [R:11]	Characteristics of disease, diagnostic and treatment tasks (stage) [R:4,9,11,12]; information need [R:2,10]
Enabling resources	Functioning infrastructure (mobile network/connectivity, transport system, electricity, basic test equipment) [R: 1,4,6,7,8,9,11,12,13]; access to mobile phone [R:1,4,6,7,8,11,12,13]; availability and affordability of (telecommunication) services [R:1,3,5,6,11,12,13]; *partnership and support* [R:2,3,7,9]; awareness creation [R:2,5]; *avoidance of abuse* [R:4,12]; convenience [R:6]; confidentiality and privacy [R:8]; (community) support [R:10]	Portability and easy to use [R:6,13]; *(family) support* [R:8]; maintenance (battery recharge) [R:12]	Legislation and policy (phone usage, liability, funding mechanisms and reimbursement, data security and privacy, staff job description, partners) [R:1,2,4,5,6,7,8,9,13]; *(government, institutional, sectoral, stakeholders') support)* [R:1,4,5,7,9,10,12,13]; infrastructure (functioning network services, equipment) [R:1,5,6,8,10,11]; financial resources and incentives [R:1,6,9,10,11,12]; quality, availability and affordability of services [R:1,7,10,12]; sustainability plan [R:7,10,12,13]; phone access [R:1,4,10]; *documentation and record-keeping* [R:1,2,9]; cost-effectiveness [R:5,8,10]; *evidence-informed (research, expert advice)* [R:5,10,11]; awareness [R:10]; (mobile health) guidelines [R:1]; *abuse/corruption* [R:11]	Simple, safest and easy technologies/intervention (apps and softwares) [R:1,4]; type of (available) technologies [R:1]; maintenance [R:6]; phone features (screen, tailored operability) [R:7]

[a]Source: authors' own compilation based on interview results.

[b]R indicates the reference citations.

[c]Text in italics are the additional patient- and provider-context factors of the mobile health PNE framework identified in this study.

Needs

The framework stipulates that patient and provider needs, such as access barriers for patients (eg, long travel times and costs) and providers' lack of capacity to provide adequate care, influence the utilization of mHealth interventions in SSA. In this study, the interviewed policy makers suggested that patients who face health care access barriers of various forms and nature are more likely to perceive mHealth interventions as useful (see Table 3). They considered patients with severe conditions and/or in need of special/urgent care to benefit most from mHealth interventions, particularly if the interventions contribute to reduced travel times and better access to providers:

> *It would be useful; it would reduce a whole lot of travelling time and reduce some stress levels in getting vehicle/transport. It may even cut down on mortality because it can enhance emergency treatment and emergency care.* [Respondent: 11]

> *It would be very much useful in our settings and circumstances where many people do not even have access to the health facilities because of absence of the health facility, low numbers of health workers; that is, the low patient to health worker ratio. If health professionals can be reached via mobile phones or other ICTs, that would improve the chances of more people getting access and it would even lead to realizing the universal health coverage.* [Respondent: 12]

> *Once we are able to do this, we would save more lives and then again we would have lesser cases developing into complications to demand more attention and more time from the experts.*" [Respondent: 9]

Furthermore, interviewees mentioned several tasks for which mHealth interventions would respond to the needs of health care providers, thus contributing to perceived usefulness and ease of use of mHealth interventions. For example, to reduce the workload on providers and to use mobile phones for regular monitoring of blood sugar levels of diabetic patients:

> *We are aware that the health system in Ghana is stricken by lack of facilities and diagnostics, and the health staffs are not motivated to go and stay in the rural areas...and because we won't have enough doctors and enough experts in the rural areas then we can't run away from telemedicine.* [Respondent: 9]

> *...if we are talking specifically about testing fasting blood sugar, it shouldn't be that the patients wait at the clinic...because we all know that the patients have to fast and for a diabetic, once you haven't taken the blood sample s/he cannot eat. There should be enough health care providers available at all times to attend to them immediately. And so that should be organized well, a mobile phone can help do that easily...* [Respondent: 3]

> *I think it is time we do it, it would even reduce the workload on me [the provider].* [Respondent: 2]

Notably, the framework did not specify which particular needs of patients influence their perceived ease of use of mHealth interventions. Nonetheless, the interviewed policy makers suggested that the general trend to use information technology for other services may create a need to use mHealth in the management of NCDs, while simultaneously making it easier to use the technology:

> *We are in a technology age; whether we like it or not, technology is taking over and the earlier we get ourselves involved the better, because there would be a time where all banking would be done online. So, the fact that one is not computer literate nor mobile phone literate it cannot be assumed that the world should wait for us. So, it has to be done and it is being done.* [Respondent: 3]

Enabling Resources

Enabling resources were the most emphasized considerations of the interviewees in determining the sustainability of mHealth. The framework suggested that the 2 most important enabling resources for the successful implementation of mHealth interventions were access to mobile phones (or devices) and the availability of functioning stable telecommunication networks. Accordingly, the interviewees maintained and also suggested that mHealth interventions could be perceived as useful by both patients and health care providers if access to mobile phones, availability and affordability of the infrastructure for good quality (telecommunication) services, reduced burden of work for providers, the avoidance of system abuse, financial resources, and government and institutional support as well as legislation and policy support are assured (see Table 3).

> *Now we are having a lot of mobile phone services but we do have challenges with them. We need to have stable mobile phone services that are good. The services must be available everywhere.* [Respondent: 3]

> *When you are doing a project and you have somebody funding it like we did for the [mHealth] project, it's cool. But then when the project comes to an end and the realities dawn on us, our governments should give money for some of these things.* [Respondent: 1]

> *It has to be a priority and all these things have to fit in the priorities of the Ministry of Health.* [Respondent: 5]

In addition, policy makers suggested that legislation, policies, and guidelines are needed to guide the activities of (health care) providers. However, they maintained that such policies for the explicit purposes of mHealth interventions should be appropriately informed by the evidence from, for example, pilot projects that first need to be conducted. Furthermore, they highlighted that the availability of financial resources would be an important enabling resource but that financial support and commitment from governments for mHealth interventions still remains low because of resource constraints.

> *[...] Yea, will you buy vaccines or you buy phones. I will rather buy vaccines than buy phones. Those are the realities that we deal with as people at the policy level. [...] So those are the trade-offs that we make at the national level and it's not an easy trade-offs [...] Also we need to really come out clearly what the parameters should be. We developed a mobile device guideline and we did advocate that the mobile phone is a medical device and so the health facilities have to provide them.* [Respondent: 1]

> *I believe in doing pilot projects before developing the policies because the findings of the pilot project should guide the policy. So, the immediate thing is to have a project with the NCD programme... it can be part of the priorities of the Ministry of Health.*" [Respondent: 5]

Interviewees identified several conditions that would enable patients to easily use mHealth interventions, including, for example, family support and availability of maintenance services. In the same vein, they emphasized that attention should be given to the suitability of the technologies for health care

providers including certain specific features, such as the size of the screen.

> *...one mobile phone platform they created for health professionals to monitor those [patients] who are on medications, they secured an Android phone for all of them, I mean something with a bigger screen that they could do so many things on it. I think it has been tailored.* [Respondent: 7]

Discussion

Principal Findings

To our knowledge, this is the first study that has investigated the views of policy makers about factors that support successful implementation and scale-up of mHealth interventions. We found that policy makers in Ghana were well informed about the potential of using mobile phones for health promotion, prevention, and health service delivery—and they strongly supported the further expansion of mHealth in the country. The results of the study also showed that the mHealth PNE framework's categories of *predisposing characteristics, needs, and enabling resources* are a useful guide for policy makers in ascertaining what critical factors would support the successful implementation of mHealth interventions. None of the policy makers stated any view that suggested that the framework has shortcomings. Rather, the responses of interviewed policy makers showed that they are thinking of many of the factors suggested by the mHealth PNE framework but that they tend to focus more on enabling resources than on predisposing characteristics and need. Finally, policy makers added several relevant factors under the categories of the mHealth PNE framework that should be considered when aiming to assure sustainability and scale-up of mHealth interventions.

These findings have several important implications for policy makers and researchers, as well as for the further refinement of the mHealth PNE framework. First, this study shows that policy makers are aware of many of the factors that have been described in the literature as particularly important for assuring successful implementation and scale-up of mHealth interventions for NCDs. For example, in line with previous literature [23], the participating policy makers highlighted that a positive attitude of both patients and providers toward mobile technologies is one of the most important factors influencing the perception of patients and providers that mHealth interventions are useful and easy to use. Similarly, their assessment that patients in need of special/urgent care are likely to benefit most from mHealth is in accordance with previous findings in the literature. This implies that policy makers in Ghana broadly agree with the findings of our systematic review [23] that it is important to consider context factors, that is, PNE, when developing and implementing mHealth interventions for NCDs. In fact, these context factors can be more important than the technical aspects of an intervention in determining its success [23,56-58].

Second, as the thinking of policy makers tends to focus on enabling resources, such as functioning telecommunication infrastructure, sustainable financing, and support from stakeholders, the mHealth PNE framework can be useful to facilitate a more holistic and systematic assessment of other factors supporting successful implementation, sustainability, and scale-up of mHealth interventions for NCDs. For example, future revisions of the Ghana eHealth Strategy [47] may benefit from considering the categories of the PNE framework to assure that new policies are developed, which will be adjusted to the needs of patients and providers, while taking into account their predisposing characteristics. The mHealth PNE framework (see Table 3) provides a long list of predisposing characteristics of patients and providers as well as of their needs, which can be used as a guide by policy makers during implementation, sustainability, and scale-up.

Third, this study has contributed to the refinement of the mHealth PNE framework by identifying additional patient- and provider-context factors that should be considered during implementation, sustainability, and scale-up of mHealth interventions. This includes patients' predisposing characteristics, such as gender, urban/rural location, and personalization of technologies; patients' need, such as their need for urgent/specialized care; and patients' enabling resources, such as avoidance of abuse, partnership, and (family) support. Concerning providers, policy makers identified additional predisposing characteristics, such as good (provider-to-patient/community) relationships; additional need factors, such as the need for exchange of expertise and for continuity of care; and additional enabling resources, such as support from government and other stakeholders (see Table 3). Interestingly, policy makers noted that the increasing utilization of mobile phones by patients for services of other sectors, for example, in the financial/banking sector [59,60], may create a desire (or *need*) to also have mobile phone–based services in the health sector, which, in turn, may contribute to patients finding these technologies easy to use.

Finally, although the study shows that the categories of the framework are useful for policy makers, further (quantitative) research is required to test the validity of the framework and to explore the relative importance of the identified context factors for successful implementation, sustainability, and scale-up of mHealth interventions for NCDs. This may include, for example, studies testing the relevance of the identified context factors during the implementation of the WHO's Package for Essential NCD Interventions (ie, integration of NCDs into primary health care) [9,61] using mobile technologies.

Limitations

This study has several limitations. The recruitment of participants relied on the list of members of the technical working group for Ghana's national NCD policy. This does not constitute a representative sampling of all relevant policy makers, and it has a bias toward the inclusion of policy makers with expertise in the area of NCDs, whereas possibly missing policy makers with expertise in the area of mHealth. However, the selected policy makers demonstrated that they had considerable knowledge in the area of mHealth, in addition to their long-standing experience from working in the health sector in Ghana.

The scope of this study was also limited by the use of qualitative methods. As a result, the contextual factors summarized in Table

3 are rather indicative. It is very likely that there are further predisposing characteristics, enabling resources, and need that are relevant for the implementation and scale-up of mHealth interventions for NCDs beyond those identified by the interviewed policy makers or by our systematic review [23]. In addition, the relative importance of the identified factors remains unknown. Therefore, more research is needed to confirm the mHealth PNE framework and to operationalize some of its categories. For example, concerning the interplay of predisposing characteristics and perceived usefulness (see Table 3), quantitative research is needed to confirm that a positive attitude toward mHealth is a predictor of perceived usefulness. This requires an operationalization for measuring a positive attitude and for quantifying its impact on the sustained use of mHealth for NCDs. Ideally, the mHealth PNE framework would be tested using a large dataset from a multicountry mHealth trial, allowing sufficient variation in the context factors that are hypothesized to influence long-term success of interventions.

Conclusions

There is great potential for mHealth interventions to improve treatment and care for patients with NCDs in SSA. However, the views of policy makers about factors that support the successful implementation, sustainability, and scale-up of these interventions used to be unexplored. Our qualitative study found that policy makers in Ghana are aware of many of the factors that have been described in the literature as particularly important for assuring successful implementation, sustainability, and scale-up of mHealth interventions for NCDs. In addition, the study showed that the mHealth PNE framework is useful to guide policy makers toward a more holistic and systematic assessment of context factors that support intervention implementation, sustainability, and scale-up, such as predisposing characteristics of patients and providers, as well as their need. Furthermore, the study allowed to refine the mHealth PNE framework by identifying additional context factors under the categories of PNE that support implementation, sustainability, and scale-up of mHealth interventions for NCDs.

The implication of these findings is that policy makers may benefit from using the mHealth PNE framework at various stages of implementation and scale-up of mHealth interventions for NCDs. However, it is important to be aware that the framework is still in its early stages of development. Researchers may (and should) use the framework when investigating reasons for success (or failure) of interventions. Over the years, such an emerging body of evidence will contribute to confirming and/or refining the factors proposed by the mHealth PNE framework, and it may ultimately allow quantifying the relative importance of these factors.

Acknowledgments

This study was conducted as part of a PhD research with scholarship from Paul and Maria Kremer Stiftung in Germany. The doctoral and postdoctoral services at the Technische Universität Berlin supported the corresponding author to participate in a special colloquium in Ghana, which opportunity enabled the data collection for this study. The authors thank all the study participants for their time and Victor Stephani for supporting the study. In addition, the authors acknowledge support by the German Research Foundation and the Open Access Publication Funds of Technische Universität Berlin.

Conflicts of Interest

None declared

Multimedia Appendix 1

Semistructured questionnaire.

[PDF File (Adobe PDF File), 107KB - mhealth_v7i5e11497_app1.pdf]

References

1. GBD 2016 DALYsHALE Collaborators. Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet 2017 Sep 16;390(10100):1260-1344 [FREE Full text] [doi: 10.1016/S0140-6736(17)32130-X] [Medline: 28919118]
2. Lawn JE, Blencowe H, Oza S, You D, Lee AC, Waiswa P, Lancet Every Newborn Study Group. Every newborn: progress, priorities, and potential beyond survival. Lancet 2014 Jul 12;384(9938):189-205. [doi: 10.1016/S0140-6736(14)60496-7] [Medline: 24853593]
3. The Economist Intelligence Unit Limited. 2014. Sub-Saharan African healthcare: the user experience: a focus on non-communicable diseases URL: https://africacapacityalliance.org/sub-saharan-african-healthcare/ [accessed 2019-03-28] [WebCite Cache ID 77DAiUdNR]
4. Dalal S, Beunza JJ, Volmink J, Adebamowo C, Bajunirwe F, Njelekela M, et al. Non-communicable diseases in sub-Saharan Africa: what we know now. Int J Epidemiol 2011 Aug;40(4):885-901. [doi: 10.1093/ije/dyr050] [Medline: 21527446]
5. World Health Organization. In: Noncommunicable Diseases (NCD) Country Profiles 2014. 20 Avenue Appia, 1211 Geneva 27, Switzerland: WHO Library Cataloguing-in-Publication Data; 2014.

6. de-Graft Aikins A, Addo J, Ofei F, Bosu W, Agyemang C. Ghana's burden of chronic non-communicable diseases: future directions in research, practice and policy. Ghana Med J 2012 Jun;46(2 Suppl):1-3 [FREE Full text] [Medline: 23661810]
7. Agyei-Mensah S, de-Graft AA. Epidemiological transition and the double burden of disease in Accra, Ghana. J Urban Health 2010 Sep;87(5):879-897 [FREE Full text] [doi: 10.1007/s11524-010-9492-y] [Medline: 20803094]
8. Ghana Statistical Services. 2014 Aug. Ghana living standards survey, Round 6 (GLSS 6), Main report, 2014 URL: http://www.statsghana.gov.gh/docfiles/glss6/GLSS6_Main Report.pdf [accessed 2019-03-29] [WebCite Cache ID 77ESpbTb9]
9. World Health Organization. 20 Avenue Appia, 1211 Geneva 27, Switzerland: WHO Press, World Health Organization; 2010. Package of Essential Noncommunicable (PEN) Disease Interventions for Primary Health Care in Low-Resource Settings URL: https://www.who.int/nmh/publications/essential_ncd_interventions_lr_settings.pdf [accessed 2019-03-29] [WebCite Cache ID 77ESyMM5P]
10. Lee S, Cho Y, Kim S. Mapping mHealth (mobile health) and mobile penetrations in sub-Saharan Africa for strategic regional collaboration in mHealth scale-up: an application of exploratory spatial data analysis. Global Health 2017 Aug 22;13(1):63 [FREE Full text] [doi: 10.1186/s12992-017-0286-9] [Medline: 28830540]
11. Beratarrechea A, Lee AG, Willner JM, Jahangir E, Ciapponi A, Rubinstein A. The impact of mobile health interventions on chronic disease outcomes in developing countries: a systematic review. Telemed J E Health 2014 Jan;20(1):75-82 [FREE Full text] [doi: 10.1089/tmj.2012.0328] [Medline: 24205809]
12. Beratarrechea A, Moyano D, Irazola V, Rubinstein A. mHealth interventions to counter noncommunicable diseases in developing countries: still an uncertain promise. Cardiol Clin 2017 Feb;35(1):13-30. [doi: 10.1016/j.ccl.2016.08.009] [Medline: 27886783]
13. Bloomfield GS, Vedanthan R, Vasudevan L, Kithei A, Were M, Velazquez EJ. Mobile health for non-communicable diseases in Sub-Saharan Africa: a systematic review of the literature and strategic framework for research. Global Health 2014 Jun 13;10:49 [FREE Full text] [doi: 10.1186/1744-8603-10-49] [Medline: 24927745]
14. Opoku D, Scott P, Quentin W. Healthcare professionals' perceptions of the benefits and challenges of a teleconsultation service in the Amansie-West district of Ghana. Telemed J E Health 2015 Sep;21(9):748-755. [doi: 10.1089/tmj.2014.0210] [Medline: 25942579]
15. Stephani V, Opoku D, Quentin W. A systematic review of randomized controlled trials of mHealth interventions against non-communicable diseases in developing countries. BMC Public Health 2016 Jul 15;16:572 [FREE Full text] [doi: 10.1186/s12889-016-3226-3] [Medline: 27417513]
16. Global Action Plan for the Prevention and Control of Noncommunicable diseases 2013-2020. In: World Health Organization. Geneva: WHO Press; 2013:1-55.
17. Eckman M, Gorski I, Mehta K. Leveraging design thinking to build sustainable mobile health systems. J Med Eng Technol 2016 Aug;40(7-8):422-430. [doi: 10.1080/03091902.2016.1218560] [Medline: 27535325]
18. Obasola OI, Mabawonku I, Lagunju I. A review of e-Health interventions for maternal and child health in Sub-Sahara Africa. Matern Child Health J 2015 Feb 5;19(8):1813-1824. [doi: 10.1007/s10995-015-1695-0] [Medline: 25652059]
19. Christine ZQ, Masatake Y, Vicky HA. World Bank. Mobile Applications for the Health Sector URL: http://siteresources.worldbank.org/INFORMATIONANDCOMMUNICATIONANDTECHNOLOGIES/Resources/mHealth_report.pdf [accessed 2019-03-28] [WebCite Cache ID 77DBPLpLu]
20. Myers K, Nelson E, Rabinowitz T, Hilty D, Baker D, Barnwell SS, et al. American telemedicine association practice guidelines for telemental health with children and adolescents. Telemed J E Health 2017 Dec;23(10):779-804. [doi: 10.1089/tmj.2017.0177] [Medline: 28930496]
21. World Health Organization. 2012. National eHealth Strategy Toolkit URL: https://www.itu.int/dms_pub/itu-d/opb/str/D-STR-E_HEALTH.05-2012-PDF-E.pdf [accessed 2019-03-28] [WebCite Cache ID 77DCVKUrW]
22. European Momentum for Mainstreaming Telemedicine Deployment in Daily Practice. 2014. Eighteen critical success factors for deploying telemedicine URL: http://www.telemedicine-momentum.eu/18-factors/ [accessed 2019-03-28] [WebCite Cache ID 77DCX0lOm]
23. Opoku D, Stephani V, Quentin W. A realist review of mobile phone-based health interventions for non-communicable disease management in sub-Saharan Africa. BMC Med 2017 Feb 06;15(1):24 [FREE Full text] [doi: 10.1186/s12916-017-0782-z] [Medline: 28162090]
24. Temmingh H, Claassen A, van Zyl S, Carrara H, Dayakalashe N, Myer L, et al. The evaluation of a telephonic wellness coaching intervention for weight reduction and wellness improvement in a community-based cohort of persons with serious mental illness. The Journal of Nervous and Mental Disease 2013;201(11):977-986. [doi: 10.1097/nmd.0000000000000036]
25. Rotheram-Borus MJ, Tomlinson M, Gwegwe M, Comulada W, Kaufman N, Keim M. Diabetes buddies: peer support through a mobile phone buddy system. Diabetes Educ 2012;38(3):357-365 [FREE Full text] [doi: 10.1177/0145721712444617] [Medline: 22546740]
26. Qin R, Dzombak R, Amin R, Mehta K. Reliability of a telemedicine system designed for rural Kenya. J Prim Care Community Health 2013 Jul 01;4(3):177-181. [doi: 10.1177/2150131912461797] [Medline: 23799704]
27. Pastakia SD, Karwa R, Kahn CB, Nyabundi JS. The evolution of diabetes care in the rural, resource-constrained setting of western Kenya. Ann Pharmacother 2011 Jun;45(6):721-726. [doi: 10.1345/aph.1P779] [Medline: 21558485]

28. Osei-tutu A, Shih T, Rosen A, Amanquah N, Chowdhury M, Nijhawan RI, et al. Mobile teledermatology in Ghana: sending and answering consults via mobile platform. J Am Acad Dermatol 2013 Aug;69(2):e90-e91. [doi: 10.1016/j.jaad.2012.08.008] [Medline: 23866892]
29. Odigie VI, Yusufu LM, Dawotola DA, Ejagwulu F, Abur P, Mai A, et al. The mobile phone as a tool in improving cancer care in Nigeria. Psychooncology 2012 Mar;21(3):332-335. [doi: 10.1002/pon.1894] [Medline: 22383275]
30. Ndlovu K, Littman-Quinn R, Park E, Dikai Z, Kovarik CL. Scaling up a mobile telemedicine solution in Botswana: keys to sustainability. Front Public Health 2014 Dec 11;2:275. [doi: 10.3389/fpubh.2014.00275] [Medline: 25566520]
31. Littman-Quinn R, Mibenge C, Antwi C, Chandra A, Kovarik C. Implementation of m-health applications in Botswana: telemedicine and education on mobile devices in a low resource setting. J Telemed Telecare 2013 Feb;19(2):120-125. [doi: 10.1177/1357633x12474746] [Medline: 23454821]
32. Kivuti-Bitok LW, McDonnell G, Pokhariyal G, Roudsari A. Self-reported use of internet by cervical cancer clients in two National Referral Hospitals in Kenya. BMC Res Notes 2012 Oct 09;5:559 [FREE Full text] [doi: 10.1186/1756-0500-5-559] [Medline: 23046538]
33. Kiser M, Beijer G, Mjuweni S, Muyco A, Cairns B, Charles A. Photographic assessment of burn wounds: a simple strategy in a resource-poor setting. Burns 2013 Feb;39(1):155-161. [doi: 10.1016/j.burns.2012.04.003] [Medline: 22647494]
34. Kingue S, Angandji P, Menanga A, Ashuntantang G, Sobngwi E, Dossou-Yovo R, et al. Efficiency of an intervention package for arterial hypertension comprising telemanagement in a Cameroonian rural setting: The TELEMED-CAM study. Pan Afr Med J 2013;15:153 [FREE Full text] [doi: 10.11604/pamj.2013.15.153.2655] [Medline: 24396559]
35. Holeman I, Evans J, Kane D, Grant L, Pagliari C, Weller D. Mobile health for cancer in low to middle income countries: priorities for research and development. Eur J Cancer Care (Engl) 2014 Oct 17;23(6):750-756. [doi: 10.1111/ecc.12250] [Medline: 25324023]
36. Greisman L, Nguyen TM, Mann RE, Baganizi M, Jacobson M, Paccione GA, et al. Feasibility and cost of a medical student proxy-based mobile teledermatology consult service with Kisoro, Uganda, and Lake Atitlán, Guatemala. Int J Dermatol 2014 Dec 29;54(6):685-692. [doi: 10.1111/ijd.12708]
37. Frühauf J, Hofman-Wellenhof R, Kovarik C, Mulyowa G, Alitwala C, Soyer H, et al. Mobile teledermatology in sub-Saharan Africa: a useful tool in supporting health workers in low-resource centres. Acta Derm Venereol 2013 Jan;93(1):122-123 [FREE Full text] [doi: 10.2340/00015555-1404] [Medline: 22735496]
38. Fiander A, Ndahani C, Mmuya K, Vanneste T. Results from 2011 for the transportMYpatient program for overcoming transport costs among women seeking treatment for obstetric fistula in Tanzania. Int J Gynaecol Obstet 2013 Mar;120(3):292-295. [doi: 10.1016/j.ijgo.2012.09.026] [Medline: 23265834]
39. Stewart A, Noakes T, Eales C, Shepard K, Becker P, Veriawa Y. Adherence to cardiovascular risk factor modification in patients with hypertension. Cardiovasc J S Afr 2005;16(2):102-107 [FREE Full text] [Medline: 15915277]
40. Krüger C, Niemi M. A telemedicine network to support paediatric care in small hospitals in rural Tanzania. J Telemed Telecare 2012 Jan;18(1):59-62. [doi: 10.1258/jtt.2011.110312] [Medline: 21968000]
41. Chindo L. Mobile phone use in Cameroon: an increasingly useful tool for the follow-up of children with Burkitt lymphoma. Pediatr Blood Cancer 2013 Mar;60(3):524. [doi: 10.1002/pbc.24226] [Medline: 23065945]
42. Azfar RS, Lee RA, Castelo-Soccio L, Greenberg MS, Bilker WB, Gelfand JM, et al. Reliability and validity of mobile teledermatology in human immunodeficiency virus-positive patients in Botswana: a pilot study. JAMA Dermatol 2014 Jun;150(6):601-607 [FREE Full text] [doi: 10.1001/jamadermatol.2013.7321] [Medline: 24622778]
43. Andersen RM. Revisiting the behavioral model and access to medical care: does it matter? J Health Soc Behav 1995 Mar;36(1):1-10. [Medline: 7738325]
44. Davis F. Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly 1989 Sep;13(3):319. [doi: 10.2307/249008]
45. Piot P, Caldwell A, Lamptey P, Nyrirenda M, Mehra S, Cahill K, et al. Addressing the growing burden of non-communicable disease by leveraging lessons from infectious disease management. J Glob Health 2016 Jun;6(1):010304 [FREE Full text] [doi: 10.7189/jogh.06.010304] [Medline: 26955469]
46. Afagbedzi S, Obuobi H, Aryeetey R, Bosomprah S. A review of Ghana's E-health strategy. J Health Inform Africa 2013 Sep 18;1(1):2013. [doi: 10.12856/JHIA-2013-v1-i1-52]
47. Ministry of Health Ghana. National E-Health Strategy URL: https://www.isfteh.org/files/media/ghana_national_ehealth_strategy.pdf [accessed 2019-03-28] [WebCite Cache ID 77DBwt6Ei]
48. Afarikumah E. Electronic health in Ghana: current status and future prospects. Online J Public Health Inform 2014 Feb 05;5(3):230 [FREE Full text] [doi: 10.5210/ojphi.v5i3.4943] [Medline: 24678382]
49. Novartis Foundation for Sustainable Development. Telemedicine Project in Ghana URL: http://http//www.novartisfoundation.org [accessed 2019-03-28] [WebCite Cache ID 77DBvO9yU]
50. LeFevre AE, Mohan D, Hutchful D, Jennings L, Mehl G, Labrique A, et al. Mobile technology for community health in Ghana: what happens when technical functionality threatens the effectiveness of digital health programs? BMC Med Inform Decis Mak 2017 Dec 14;17(1):27 [FREE Full text] [doi: 10.1186/s12911-017-0421-9] [Medline: 28292288]
51. Macleod B, Phillips J, Stone A, Walji A, Awoonor-Williams JK. The architecture of a software system for supporting community-based primary health care with mobile technology: the Mobile Technology for Community Health (MoTeCH)

initiative in Ghana. Online J Public Health Inform 2012;4(1):- [FREE Full text] [doi: 10.5210/ojphi.v4i1.3910] [Medline: 23569631]
52. Schreier M. Qualitative Content Analysis in Practice. London: SAGE Publications Ltd; 2012.
53. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Int J Qual Health Care 2007 Dec;19(6):349-357 [FREE Full text] [doi: 10.1093/intqhc/mzm042] [Medline: 17872937]
54. Republic of Ghana Ministry of Health. 2012. Strategy for the Management, Prevention and Control of Chronic Non-Communicable Diseases in Ghana URL: http://www.nationalplanningcycles.org/sites/default/files/country_docs/Ghana/ghana.pdf [accessed 2019-03-28] [WebCite Cache ID 77DBn8Njc]
55. Uwe F. The Sage Handbook Of Qualitative Data Analysis. London: Sage Publications; 2019.
56. Broens TH, Huis IV, Vollenbroek-Hutten MM, Hermens HJ, van Halteren AT, Nieuwenhuis LJ. Determinants of successful telemedicine implementations: a literature study. J Telemed Telecare 2007 Jun;13(6):303-309. [doi: 10.1258/135763307781644951] [Medline: 17785027]
57. L'Engle K, Plourde KF, Zan T. Evidence-based adaptation and scale-up of a mobile phone health information service. Mhealth 2017 Mar;3:11 [FREE Full text] [doi: 10.21037/mhealth.2017.02.06] [Medline: 28567408]
58. Slater H, Campbell JM, Stinson JN, Burley MM, Briggs AM. End user and implementer experiences of mHealth technologies for noncommunicable chronic disease management in young adults: systematic review. J Med Internet Res 2017 Dec 12;19(12):e406 [FREE Full text] [doi: 10.2196/jmir.8888] [Medline: 29233804]
59. Hossain M, Ahmed F. University of Dhaka. 2014. Evaluating the Impact of Mobile Banking Deployment for Microfinance Institutions URL: http://www.dumarketing.ac.bd/html5boilerplatemkt/uploads/2013/11/08Evaluating-the-Impact-of-Mobile-Banking-Deployment.pdf [accessed 2019-03-29] [WebCite Cache ID 77EXWPCPa]
60. Dzokoto V, Appiah E. Virginia Commonwealth University. Making Sense of Mobile Money in Urban Ghana: Personal, Business, Social and Financial Inclusion Prospects URL: https://www.imtfi.uci.edu/files/docs/2014/dzkoto_final_report_feb2014.pdf [accessed 2019-03-28] [WebCite Cache ID 77DBXUFPK]
61. Nyarko KM, Ameme DK, Ocansey D, Commeh E, Markwei MT, Ohene S. Capacity assessment of selected health care facilities for the pilot implementation of Package for Essential Non-communicable Diseases (PEN) intervention in Ghana. Pan Afr Med J 2016;25(Suppl 1):16 [FREE Full text] [doi: 10.11604/pamj.supp.2016.25.1.6252] [Medline: 28149441]

Abbreviations

eHealth: electronic health
mHealth: mobile health
NCD: noncommunicable disease
PNE: predisposing characteristics, need, and enabling resources
QCA: qualitative content analysis
SSA: sub-Saharan Africa
WHO: World Health Organization

Edited by G Eysenbach; submitted 05.07.18; peer-reviewed by M Mars, V Stephani; comments to author 09.08.18; revised version received 04.10.18; accepted 25.01.19; published 03.05.19.

Please cite as:
Opoku D, Busse R, Quentin W
Achieving Sustainability and Scale-Up of Mobile Health Noncommunicable Disease Interventions in Sub-Saharan Africa: Views of Policy Makers in Ghana
JMIR Mhealth Uhealth 2019;7(5):e11497
URL: https://mhealth.jmir.org/2019/5/e11497/
doi:10.2196/11497
PMID:31066706

Original Paper

A Smartphone Game to Prevent HIV Among Young Africans (Tumaini): Assessing Intervention and Study Acceptability Among Adolescents and Their Parents in a Randomized Controlled Trial

Gaëlle Sabben[1], MPH; Victor Mudhune[2], MBA, MPH; Ken Ondeng'e[2], MPH; Isdorah Odero[2], BSc; Richard Ndivo[2], BSc; Victor Akelo[2], MBChB, MPH; Kate Winskell[1], PhD

[1]Hubert Department of Global Health, Rollins School of Public Health, Emory University, Atlanta, GA, United States

[2]Centre for Global Health Research, HIV Research Branch, Kenya Medical Research Institute, Kisumu, Kenya

Corresponding Author:
Gaëlle Sabben, MPH
Hubert Department of Global Health
Rollins School of Public Health
Emory University
1518 Clifton Road
Atlanta, GA, 30322
United States
Phone: 1 4046833103
Email: gaelle.sabben@emory.edu

Abstract

Background: Young people aged 15 to 24 years account for one-third of new adult HIV infections. Controlling the HIV epidemic requires effective interventions targeted toward young people and their needs. Smartphone games offer a promising avenue for reaching this population with evidence-based HIV prevention interventions. It is crucial to the effectiveness of these interventions that they be acceptable and intrinsically motivating to adolescents as well as acceptable to their parents.

Objective: *Tumaini* is a narrative-based smartphone game designed to help prevent HIV among young Africans aged 11 to 14 years by delaying first sex and increasing condom use at first sex. Following a 16-day feasibility study of *Tumaini*, we assessed the acceptability (1) of the intervention, where acceptability was operationalized as appeal, relevance, value, usability, and understandability, and (2) of this study and a planned future randomized controlled efficacy trial.

Methods: During the randomized feasibility study (n=60) of *Tumaini* in western Kenya in spring 2017, 30 participants used the intervention on a study-provided smartphone. The app automatically logged participant interaction with the game in time-stamped log files. All 30 participants completed an Audio Computer-Assisted Self-Interview–based game experience survey, and 27 took part in 4 focus group discussions (FGDs) about the game's appeal, relevance, value, usability, and understandability. Their parents (n=22) also participated in 4 FGDs about the acceptability of the intervention, of this study, and of a planned efficacy trial. Survey data were analyzed using SAS software (SAS Institute Inc); FGD transcripts were coded and analyzed in MAXQDA 12 (Verbi GmbH); and gameplay log files were analyzed using Microsoft Excel.

Results: Adolescent participants' survey responses indicated that *Tumaini* scored well with players on all indicators of acceptability (appeal, relevance, value, usability, and understandability). Focus group analyses aligned with these findings and emphasized a high degree of player engagement with the game, which was supported by log file analysis. Adolescent participants were eager for additional content, and parents were receptive to a longer study involving biomarkers, based on their positive experiences with this study. There is scope to improve communication with parents about their role in the intervention. As the game was tested in beta version, there is also scope to fine-tune some of the game mechanics to increase usability.

Conclusions: This study shows the strong acceptability of an interactive smartphone-based game both to adolescents and their parents in western Kenya and that of the study methods used to pilot-test the intervention. It also suggests that longitudinal efficacy studies of this type of intervention, including those using biomarkers, have the potential to be acceptable among parents.

Trial Registration: ClinicalTrials.gov NCT03054051; https://clinicaltrials.gov/ct2/show/NCT03054051 (Archived by WebCite at http://www.webcitation.org/70U2gCNtW)

(JMIR Mhealth Uhealth 2019;7(5):e13049) doi:10.2196/13049

KEYWORDS

HIV; youth; sub-Saharan Africa; Kenya; serious game; narrative; smartphone; pilot study; randomized controlled trial; acceptability

Introduction

Background

Young people aged between 15 and 24 years, particularly young women and those from sub-Saharan Africa, account for a large proportion of new HIV infections among adults [1]. Due to demographic change, a youth bulge, or a particularly large cohort of young people aged 15 to 24 years, makes up a large segment of the population in many sub-Saharan African countries, including those affected by HIV [2]. Without a significant decrease in HIV incidence, this youth bulge will result in large numbers of new HIV infections among young people each year. Any effort to reduce the incidence and prevalence of HIV must address this age group's specific needs through targeted prevention efforts. Reaching them with prerisk interventions can help them to delay sex and to develop safer sexual behaviors from the onset of sexual activity [3,4], including by establishing patterns of consistent condom use [5].

Mobile technologies, including smartphones, are rapidly becoming more affordable and accessible, including in sub-Saharan Africa [6,7]. Their increasing ubiquity offers exciting opportunities for delivering targeted, culturally relevant prevention messaging and skills-building interventions at scale and remotely [8]. This is especially appealing for settings where local capacity for implementation and facilitation of group-based interventions may be limited: remotely delivered interventions would not rely on trained facilitators to ensure intervention fidelity or to regularly provide updated content [9].

Smartphones provide a platform for interactive and immersive interventions, including games, which may be particularly well suited for young people [10] and have the potential to be leveraged in ways that cannot be achieved with traditional intervention models [11]. Serious games, that is, those designed with a primary purpose that is not entertainment [12], adequately grounded in behavioral, instructional, and communication theory [13-15], can provide an engaging and safe environment to build and rehearse health protective knowledge and skills. Studies of existing serious games have shown their effects on clinical outcomes among adolescents [14,16], and an emphasis is being placed on assessing the efficacy of newly developed technology-based interventions on both health and behavior [11].

Establishing intervention efficacy is crucial; however, it is no less important to assess factors likely to affect adoption, preferably early in the development process [9]. These factors include the acceptability of the intervention to its intended users, their intrinsic motivation to engage with it, and any potential barriers to its uptake [17]. As a key component of implementation research, appropriateness of an intervention can be defined as its "perceived fit, relevance, or compatibility" for a particular purpose and consumer [17]. For an intervention delivered using a technological platform, acceptability of the delivery mode itself, including its perceived ease of use and relevance, must be considered as influencing users' eventual uptake [18]. In this paper, these aspects of user acceptance and perception of the intervention are combined under the umbrella term *acceptability*.

An intervention aimed at minors, especially one intended to be delivered remotely without a facilitator, must be acceptable to the young users as well as to their parents, who are likely to control access. Assessing intervention acceptability is particularly important when using new modes of delivery because little prior research exists to indicate what children and their parents are likely to find most palatable. In the case of interventions intended for low-resource settings, evaluation and dissemination of technologically advanced interventions and systems have received even less attention.

The success of an efficacy trial for an intervention aimed at minors also lies in the acceptability of the trial's design to parents because their consent is necessary for children's participation and thus the feasibility of the study. Parents can justifiably force study procedures to change or even a trial to end if the methods are unacceptable to them. Therefore, their acceptance of the methods is critical to studies' success, in particular, that of studies focusing on adolescent sexual health, including HIV prevention research [19]. As such, parents must be involved early in the process of developing study methods to identify potential barriers to a study's acceptability and smooth the way for successful efficacy trial implementation.

Intervention

Tumaini is a narrative-based smartphone game designed to help prevent HIV among young Africans. It was tested in a randomized controlled feasibility trial in Kisumu, western Kenya, in collaboration with the Kenya Medical Research Institute (KEMRI) in spring 2017. Pilot outcome data from this study were promising, showing significant effects on behavioral mediators of sexual debut and condom use [20]. The objectives of the pilot study included assessing the acceptability (1) of the game-based intervention, where acceptability is defined as appeal, usability, understandability, relevance, and value and (2) of the study itself and of a potential future randomized controlled trial to establish the efficacy of the game to delay first sex and increase condom use at first sex. This planned study would test participants for HIV and herpes simplex virus type 2 (HSV-2) as proxy for cumulative sexual activity, a measure motivated by the recognized challenges with inconsistent self-reporting of sexual behavior (as a risk factor for HIV) among adolescents [21]. This paper thus shares findings on the acceptability to adolescent participants and their parents of the *Tumaini* intervention and the study itself with a view to informing the development of other related interventions and studies.

Methods

Data Collection

The participants allocated to the intervention arm of the study (n=30) were provided with low-cost Android smartphones loaded with the game *Tumaini* and asked to play for at least 1 hour per day for the duration of the 16-day study. All other phone functions, with the exception of the alarm clock (for reminders), were disabled. The app automatically collected time-stamped gameplay data. Phones were provided to participants for the duration of the study to ensure consistency of technology and to avoid any socioeconomic bias resulting from using smartphone ownership as a criterion for inclusion in the study.

Both adolescents and their parents were told that *Tumaini* was the adolescents' "space to learn" and that adolescents should feel free to navigate the game as they wished. It was explained that adolescents could seek help or advice from and speak to others about the game and its contents if they so desired, and that parents should feel comfortable speaking to their children about the game and what he or she was learning. Parents were also asked to remind their children to play daily and to contact the study staff with any concerns, problems, or questions.

Immediately postintervention, all 30 participants completed a survey about their experience playing the game, delivered via Audio Computer-Assisted Self-Interview. Participants and their parents were also invited to take part in focus group discussions (FGDs) to provide additional insight into their experiences during the study. Ethical approval was granted by the Emory University and KEMRI institutional review boards, and the study was registered with ClinicalTrials.gov. Further details of the study protocol are provided elsewhere [22].

Intervention

Tumaini aims to help prevent HIV among young Africans by delaying first sex and increasing condom use at first sex. It does this by increasing knowledge about sexual health and HIV; building risk-avoidance and risk-reduction skills and related self-efficacy; challenging HIV stigma and harmful gender norms and attitudes; fostering future orientation, goal setting, and planning; and promoting dialogue with adult mentors. *Tumaini* was developed in collaboration with a US commercial game developer, Realtime Associates, and US-based and Kenyan specialists in adolescent sexual health and HIV prevention and with input from Kenyan adolescents and their parents. The game is in English with an audio track featuring Kenyan voice talent.

Tumaini is grounded in narrative and narrative-based applied communication and social behavioral theory and existing evidence-based HIV prevention interventions to promote problem solving, cognitive and behavioral rehearsal, observational learning, and immersion (manuscript under review), and it draws on extensive research on HIV-themed narratives written by young Africans [23-25].

The choose-your-own-adventure interactive narrative follows 6 characters, 3 male and 3 female, as they progress through adolescence and face real-life challenges that players are likely to face in their own lives. These challenges include peer pressure, puberty, violence, and decisions about smoking, alcohol, drugs, and sex. Players make decisions for the characters that affect their narrative trajectories and reflect possible real-world consequences of these decisions. The content explored in the narrative is bolstered by 2 additional components: minigames that tie into the themes of the narrative chapters and strengthen players' knowledge and skills and *My Story* in which players create a personal avatar and respond to questions that connect the contents of the role-playing narrative to their lives. *Tumaini* includes approximately 12 hours of discrete gameplay; it is designed to be replayed so that players can observe the outcomes of different decisions.

Participants

All 30 intervention arm participants played the game and responded to the survey. Subsequently, 27 participated in 4 FGDs, stratified by participant gender and age (11-12 and 13-14 years). The 4 parent FGDs were stratified by gender and age of the enrolled child and included 22 parents, mostly female. The demographics of the players, adolescent FGD participants, and parents are presented in Table 1.

Table 1. Participant demographics.

Characteristics	Intervention adolescents (n=30)	FGD[a] adolescents (n=27)	FGD parents (n=22)
Gender, n (%)			
Female	14 (47)	12 (44)	19 (83)
Male	16 (53)	15 (56)	3 (17)
Age (years), mean (SD)	12.8 (1.0)	12.9 (0.9)	—[b]
Religion, n (%)			
Catholic	14 (47)	13 (48)	—
Protestant/Anglican	8 (27)	6 (22)	—
Muslim	2 (7)	2 (7)	—
Seventh-day Adventist	4 (13)	4 (15)	—
Other	2 (7)	2 (7)	—
Living with both parents, n (%)	22 (73)	19 (70)	—
Housing type, n (%)			
Permanent	8 (27)	8 (30)	—
Semipermanent	11 (37)	9 (33)	—
Temporary	9 (30)	8 (30)	—
Iron sheets	2 (7)	2 (7)	—
Smartphone ownership (check all that apply), n (%)			
Parent	21 (70)	19 (70)	—
Self	2 (7)	2 (7)	—
Sibling	11 (37)	10 (37)	—
Other adult	4 (13)	4 (15)	—
No one	3 (10)	3 (11)	—
Have used a smartphone before baseline, n (%)	22 (73)	20 (74)	—

[a]FGD: focus group discussion.

[b]Demographic data not collected directly from parents but expected to reflect adolescents' self-report.

Measures and Analysis

The Technology Acceptance Model's (TAM) framework for the acceptability of technology-based systems separates acceptability or *user acceptance* into (1) a system's usefulness and (2) its ease of use. These influence a user's *affective response* and, as a result, likelihood of using the system [18]. In line with this framework, the survey included questions on the game's appeal, relevance and value (usefulness), and usability and understandability (ease of use). The appeal of specific aspects of the game, including its graphics and characters, was also assessed, and as an indirect measure of likability, participants were asked whether they would recommend the game to a friend. Ease-of-use measures focused on how easy it was to understand how to play the game (usability) and to understand the English used in the game (understandability). The game's usefulness questions focused on what players had learned, how useful they found that content (value), and how suitable they found the game to be for different player age groups and genders (relevance).

The FGD guide for adolescents included questions about players' impressions of the game, parts they liked and did not like, and sections that were hard to play; individual game components; what they learned; and recommendations for additional changes or other comments. The discussion guide for parents focused on their reactions to the game and study, their suggestions for improvements, and their recommendations for a future larger longitudinal study likely to involve blood tests. The focus groups were conducted in a mixture of English, Dholuo, and Kiswahili by moderators fluent in all 3 languages. The participants were encouraged to use any combination of languages they were most comfortable using.

Survey data were imported into and analyzed using SAS software (SAS Institute Inc), version 9.4. Descriptive statistics were calculated and stratified by participant sex and age group. Transcripts were translated into English and uploaded to MAXQDA 12 software (Verbi GmbH) for thematic coding using inductive (eg, *age appropriateness*) and deductive (eg, *misunderstandings*) codes. The data were analyzed thematically and compared across demographics. Once the phones were collected at the end of the intervention, the *Tumaini* app gameplay log files were exported to Microsoft Excel and analyzed for mean exposure time.

Results

Acceptability of the Game

Appeal

Survey responses indicated that most players played every day as instructed (87%, 26/30), in most cases for at least an hour each time (77%, 23/30; Table 2). Almost all players reported that playing was "very fun" (90%, 27/30), with the 3 others, all older male participants, responding that playing was “fun.” Most also responded that given the opportunity, they would like to play the game much more (93%, 28/30) and that they would tell friends to play (97%, 29/30), a measure of participants’ engagement with the intervention. Log file analyses indicate that players used the game for close to 27 hours on average, and all but 1 participant completed the game at least once.

Participants’ enthusiasm was apparent during both participant and parental focus groups. The adolescents explicitly and repeatedly described the game as "fun", with 1 young girl saying she “felt like laughing” when she played (FGD for 11- to 12-year-old females) and all groups saying they would tell their friends to play also. One older male participant noted that playing a variety of characters was very enjoyable:

> *I felt good playing them because even when I was watching TV and I found it boring, I would just go to my bed and look at the game* [FGD for 13- to 14-year-old males]

Participants identified the skills they learned and the future goal planning practice as especially motivating and useful. They suggested that the emphases on refusal skills, condom use, HIV prevention, and future planning would also be appealing to their friends. Parents’ comments about their children’s engagement with the game provided further indication that it had high motivational appeal. Children appear to have played “willingly,” not needing or relying on the alarm to remind them to play, and there were several spontaneous reports of children being reluctant to return the phones at the end of the study as they wanted to play more.

Also speaking to the game’s appeal to players, participants enthusiastically suggested adding further content to the game and expanding prizes in the game’s reward system. The older male participants were particularly enthusiastic about this, asking for “30 chapters,” “or even 55,” as well as more characters of both genders, growing up over a longer period.

When asked which aspect of the game they liked most (story, minigames, *My Story*, rewards, or look of the game), 14 of the 29 players who indicated a preference (48%) chose the story, although 29 (97%) players also reported liking the design and minigame components “a lot.”

Usefulness

Players were also enthusiastic about *Tumaini* ’s value to their lives and its relevance for their age group. All players found *Tumaini* to be equally suitable for male and female players (Table 3). When asked what age group the game was "best for", half of the players (n=15/30) selected the 13 to 14 years age range, and 12 others (40%) identified the 11 to 12 years age range. Two-thirds of the participants indicated that the game was best suited to their own age group (11-12 or 13-14 years).

Relevance

A total of 13 participants (43%, 13/30) further indicated that *Tumaini* would be suitable for adolescents over 14 years and 6 participants (20%, 6/30) for children under 11 years. However, during FGDs, 5 younger children expressed discomfort around the age-appropriateness of certain content: 2 boys and 1 girl said that because of their age they were "confused" about or "uncomfortable" with the reproductive system and information about condom use, and 2 girls reported being "frightened" by scenarios of older men pressuring young women (FGDs for 11- to 12-year-old males and females). Older participants, in contrast, appreciated playing older characters in preparation for encountering those situations in their own lives. As one 13- to 14-year-old male explained:

> *...it felt good because as I was playing the game I was imagining these things happening in real life.* [FGD for 13- to 14-year-old males]

Concerns about content age-appropriateness were expressed in one of the focus groups by 2 parents of younger children (FGD2 for parents of 11- to 12-year-old children). The mother of an 11-year-old boy felt that he should not yet be exposed to information about condoms. When asked if the subject should be removed, she demurred:

> *...it’s hard for a parent to teach them because it’s of no use to them […] the 11-12 year olds are still young and don’t even know how to use condoms*

Table 2. Frequency and percentage of respondents responding most positively to questions on appeal by gender.

Variables	Age 11-12 years, n (%)		Age 13-14 years, n (%)		All (n=30), n (%)
	Males (n=7)	Females (n=5)	Males (n=9)	Females (n=9)	
Played every day	7 (100)	4 (80)	7 (78)	8 (89)	26 (87)
Played 1 hour or more each time	7 (100)	2 (40)	7 (78)	7 (78)	23 (77)
In general playing was very fun	7 (100)	5 (100)	6 (67)	9 (100)	27 (90)
Would like to play much more	6 (86)	4 (80)	9 (100)	9 (100)	28 (93)
Would tell friends to play	6 (86)	5 (100)	9 (100)	9 (100)	29 (97)

Table 3. Frequency and percentage of participant responses to questions on relevance of game by gender.

Variables	Age 11-12 years, n (%)		Age 13-14 years, n (%)		All (n=30), n (%)
	Males (n=7)	Females (n=5)	Males (n=9)	Females (n=9)	
Age groups suitable for? (years; check all that apply)					
Younger than 9	0 (0)	2 (40)	1 (11)	0 (0)	3 (10)
9-10	0 (0)	1 (20)	1 (11)	2 (22)	4 (13)
11-12	4 (57)	5 (100)	7 (78)	3 (33)	19 (63)
13-14	4 (57)	5 (100)	8 (89)	7 (78)	24 (80)
15-16	1 (14)	2 (40)	5 (56)	3 (33)	11 (37)
Older than 16	0 (0)	0 (0)	3 (33)	3 (33)	6 (20)
Age group best for? (years)					
Younger than 9	0 (0)	1 (20)	0 (0)	0 (0)	1 (3)
9-10	0 (0)	0 (0)	1 (11)	0 (0)	1 (3)
11-12	5 (71)	3 (60)	2 (22)	2 (22)	12 (40)
13-14	2 (29)	1 (20)	6 (67)	6 (67)	15 (50)
15-16	0 (0)	0 (0)	0 (0)	1 (11)	1 (3)
Older than 16	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Best for both genders	7 (100)	5 (100)	9 (100)	9 (100)	30 (100)

Most parents, however, indicated that although they had some initial fears about the appropriateness of the content, these were soon allayed:

> *I was initially worried that there would be harm, that the game would contain sexual issues or even pornography that would influence her attitude...after playing the game for like a week, I stopped being concerned* [FGD2 for parents of 13-14 year olds]

One parent (FGD2 for parents of 13-14 year olds) soon realized that the game, rather than being a bad influence, actually "discouraged her from such bad thoughts about sex."

Value

The game also scored high on the measures of its value to participants. All 30 participants indicated that they had learned "a lot" by playing *Tumaini* and that the information they had acquired would be "very useful for the future" (Table 4). All but 1 participant (97%, 29/30), a young male, also indicated that they had found the game's contents to be immediately useful. Participants also felt that the game had increased their preparedness and self-efficacy for managing risky situations: 28 (93%, 28/30) participants responded that they felt more prepared to handle difficult situations, and 29 (97%, 29/30) responded that they were more sure that they could say no firmly in situations of pressure.

During FGDs, adolescents explained that an important aspect of the game's intrinsic appeal lay in the value of the game and its content to their lives both in the immediate and in preparation for situations they might face in the future. Some reported feeling encouraged by characters' positive developments and, in some cases, applying what they had learned in their own lives or sharing their new knowledge to help others. When discussing why they would recommend the game to others, 1 young boy summarized:

> *I would like to tell them that the game of Tumaini is a very beautiful game that you can play and understand things that may help you in life.* [FGD for 11- to 12-year-old males]

Table 4. Frequency and percentage of participants responding most positively to questions on value of game by gender.

Variables	Age 11-12 years, n (%)		Age 13-14 years, n (%)		All (n=30), n (%)
	Males (n=7)	Females (n=5)	Males (n=9)	Females (n=9)	
Learned a lot	7 (100)	5 (100)	9 (100)	9 (100)	30 (100)
Information very useful now	6 (86)	5 (100)	9 (100)	9 (100)	29 (97)
Information very useful for future	7 (100)	5 (100)	9 (100)	9 (100)	30 (100)
Since playing Tumaini, I feel more prepared for difficult situations I might face in the future	7 (100)	3 (60)	9 (100)	9 (100)	28 (93)
Since playing Tumaini, I feel more sure I can say no firmly when people are trying to pressure me	7 (100)	5 (100)	9 (100)	8 (89)	29 (97)

Parents also expressed satisfaction with the ways in which the content of the game related to the real-life challenges young people face. They reported that their children had shown increased interest in topics such as HIV after playing the game and that the lessons their children had shared with them were beneficial. These benefits ranged from recognizing and distancing themselves from peers' negative influences to increasing empathy for others and effectively planning for the future. Part of the appeal of the game for parents was its resonance with their children's lives:

> *...the game touched many things; school life, life at home, diseases like HIV...to me it was all okay because it touches all the parts of life.* [FGD1 for parents of 13-14 year olds)]

Ease of Use

The game's usability and understandability scored reasonably well on the game experience survey, with 67% (20/30) of adolescents rating *Tumaini*"very easy to understand how to play" and a further 10% (3/30) rating it "a little easy" to understand (Table 5). Usability rated higher among older participants.

Usability

During FGDs, adolescents also gave a positive assessment of the game's usability, with participants in all groups reporting that overall the game was "easy", both to play and understand. Most parents confirmed that their children had not needed much help to navigate *Tumaini*. Minor usability concerns centered on certain of the minigame mechanics being too complex, particularly among younger players, though players were able to overcome those hurdles by seeking advice from parents, siblings, or other trusted adults. Similarly, although adolescent participants in every FGD noted challenges that arose from content that they had not encountered previously, they succeeded by seeking help or, as a 13- to 14-year-old girl noted, through perseverance:

> *...some parts like drawing the condom steps were difficult but I tried and managed.* [FGD for 13- to 14-year-old females]

Parents in 2 FGDs mentioned having contacted the study staff for troubleshooting help. Although a few participants mentioned minor bugs that remained in the version of the game used for the trial, these do not seem to have affected players' engagement with the game or its usability.

Understandability

The game's level of English did not prove a challenge to players, with 83% (25/30) saying the English was "easy to understand" and 1 parent (FGD2 for parents of 11-12 year-olds) reporting that her child had told her that "the English in the game was simple, they are able to understand." One parent (FGD2 for parents of 11-12 year olds) even reported that her son had asked the rest of the household to speak English as a result of playing the game. Indeed, in the survey, 25 participants (83%, 25/30) also said that English was the language in which they would most like to play, rather than Dholuo or Kiswahili. Of the others, 2 participants expressed a preference for Dholuo and 3 for Kiswahili.

The intervention showed a high level of acceptability to adolescents, judging by their positive reviews in terms of its appeal, value, and relevance. Participants found the game a highly appealing way to learn important information, would recommend the intervention to others, and, given the chance, opt to play longer. Adolescents also confirmed that, overall, *Tumaini* had been appropriately pitched to their age group, although they indicated that older adolescents might also enjoy and benefit from playing. The content of the game also spoke to the adolescents, who reported gaining knowledge, skills, and self-efficacy with both immediate and future relevance.

Acceptability of the Study

This Study

Parents' experiences with the pilot study appear to have been very positive. They thanked the team for engaging them in the study, and several spontaneously asked to be included in future studies. Most comments about the acceptability of the study related to the acceptability of the intervention and its positive influence on their children, either through its content or by keeping the child busy during the school holidays. Delivering the intervention via smartphone was also acceptable to parents, who noted the family and adolescents' joy at having a phone to play with as part of the study. They appreciated that the phones' other functions were blocked so that it could not be used for communication by their children. They had no concerns about their children's safety while in possession of the phone, in part because the functions were blocked.

A small number of parents reported not having felt free to interact with their children around *Tumaini* and its content. Although many parents engaged with their children on a range of topics and with positive results, 3 parents noted that the study team's instructions had discouraged them from playing the game or interacting with their children as they would have liked, with 1 parental FGD agreeing on this point when it was raised. These parents had understood the study team's instructions about allowing the child to play without interference as restricting parents from responding to their children's questions or helping their children navigate difficult parts of the game.

Table 5. Frequency and percentage of respondents responding most positively to questions on usability and understandability by gender.

Variables	Age 11-12 years, n (%)		Age 13-14 years, n (%)		All (n=30), n (%)
	Males (n=7)	Females (n=5)	Males (n=9)	Females (n=9)	
Very easy to understand how to play	4 (57)	3 (60)	7 (78)	6 (67)	20 (67)
English was easy to understand	4 (57)	4 (80)	8 (89)	9 (100)	25 (83)
Want to play game in English	6 (86)	4 (80)	7 (78)	8 (89)	25 (83)

In one focus group, 2 parents agreed that:

> *...there were some parts that when he reached he could have wanted to ask me, but from the way they were told to own the phone he could not ask*

and that the study team:

> *...didn't set us free to discuss with this kids when they were playing this game […] maybe if you left us free that maybe we can interact with them at the time they are playing that game it would have been fair.* [FGD1 for parents of 11-12 year olds]

Despite this misunderstanding, there were many reports of parent-child communication across a broad range of topics, which are detailed in a separate publication [26].

Future Studies

Parents discussed the possibility of future studies involving *Tumaini*, focusing on the logistics of a longer study, as well as the possibility of such a study including blood draws and HSV-2 testing. Parents cited their positive experience with this trial in support of a longer study and their expectation that any gains their children made during a short intervention period would be magnified if they were exposed to the game for longer. Several, in fact, volunteered their family's participation in such a study.

In general, there was no opposition to blood collection or to HSV-2 testing. Several parents of younger children wanted to be present for blood draws and emphasized that communication from the study team about the purpose of the blood draw would be important to ensure acceptability both among potential participants and in their communities. One mother explained:

> *It's important you educate so that we be open and free. If you don't educate us it can be a flop because people have been doing monkey business with people's blood, which has made people develop some diseases that they cannot explain.* [FGD1 for parents of 11-12 year olds]

Some also highlighted the importance of preparing the child for the test and for the possibility of positive results, including appropriate explanations, support, and counseling. They also expressed a desire to know the results of any test that was conducted.

Parents questioned when the study intervention periods and study activities would take place, whether during school or holiday periods. Most voiced a preference for the study to only take place during holidays so that it would not coincide with school responsibilities, and children would not be distracted from schoolwork by the game.

Discussion

Principal Findings

Both participants and their parents expressed a high level of acceptability for the game *Tumaini*. Adolescents and parents indicated that *Tumaini* held a great deal of appeal, was highly valuable and relevant to adolescents' lives, and was overall easy to use and understand. Parents responded positively to a planned multiyear study involving biological biomarkers. These findings, in combination with the data on the intervention's effects on secondary (behavioral) outcomes, presented in a separate publication [20], suggest that a large-scale efficacy trial of *Tumaini* is warranted and would be acceptable to parents of the target adolescent population.

Acceptability of the Intervention

Adolescents' positive assessments of the interventions suggest that these types of serious games and smartphone-based intervention are acceptable and engaging ways to reach young people with important sexual health content, including about HIV. Whether the participants were motivated to use the intervention because of its entertaining nature or its educational promise, it is clear that acceptability was high. In the context of a self-regulated and remotely delivered intervention, such as this one, it is crucial that users be motivated to access and use the intervention repeatedly. This group of adolescent participants reported regular extended use of the game, requested additional time with the intervention, and suggested that additional content be made available, reflecting high intrinsic motivation to use the app. Parents' reports that their children had played willingly and had not needed the alarm reminder to do so imply that *Tumaini's* intrinsic motivational appeal to adolescents may be high enough to result in unprompted use by the target audience in the context of a longer efficacy trial and beyond.

Although acceptability is understood to be a key aspect of early intervention development [9], definitions of the term *acceptability* with regard to an intervention are inconsistent; reliable measures to assess this construct are few [27]. *Tumaini's* nature as a technology-based intervention resulted in this study's definition of acceptability drawing on Davis's TAM framework [18], in addition to Proctor's definitions of acceptability [17], to account for the game's perceived usefulness to and ease of use for participants, both of which affect its overall likeability. Participants linking the appeal of *Tumaini* to its perceived usefulness and ease of use suggests that the TAM is a useful framework for assessing acceptability of this type of intervention. At the time of study design and measures development, there were, to the best of our knowledge, no validated acceptability scale or measures. However, the acceptability and appropriateness dimensions of Weiner et al's [28] proposed implementation outcomes instrument, developed since, do align well with the items used in this study. As the field of mobile health (mHealth) continues to grow, instruments specifically designed to assess acceptability for this type of intervention will facilitate comparison of acceptability across individual interventions.

It is unlikely that all adolescents who fall within the intervention's target age range will have their own smartphones (only 3 of the 60 adolescent participants in this study did at the time of the study). This intervention, if proven efficacious and made available for download, would likely be used, at least by some young adolescents, on a parent's or other family member's smartphone. It is very encouraging that during this study, parents appeared to find *Tumaini* highly appealing, valuable, and relevant to their children's lives. A small number of parents' minor concerns about the age-appropriateness of the content

presented were outweighed by the majority's endorsement of the positive changes they saw in their children and attributed to the intervention as well as the role they saw the intervention playing in preparing their children for challenges they would face.

In Kenya, primary school education has been free and compulsory since 2003, and all subjects are taught in English from standard year 4 (age 10+ years) onward, so this study's eligibility criterion of English proficiency equivalent to grade 3 to 4 did not limit enrollment. The game's English level, supported by an audio track, appears to have successfully matched adolescents' proficiency, making it accessible to study participants.

The study had been scheduled to take place during the August 2017 school holiday period. However, the timeline was advanced by several months to avoid coinciding with the August 2017 Kenyan national elections. As a result, there was not enough time before the study to prepare a fully beta-tested version of *Tumaini*. This resulted in users noticing some bugs and some of the minigame mechanics proving more challenging than intended. These minor usability issues did not appear to affect the overall acceptability of the game or its appeal and, in fact, resulted in additional communication with older family members about the game as they sought help [26]. These issues will be addressed and a full bug test conducted before further game testing as originally planned.

Acceptability of Study

Parents had no negative experiences with or concerns about their children's involvement in the study and indicated that they approved of the use of study-provided functionally limited phones as an intervention delivery mechanism. In addition, there was no loss to follow-up of adolescent participants, who continue to engage with the study team on additional postintervention qualitative research to this day. The only aspect of the study procedures that raised any issues were the instructions given to parents about their interactions with their children around the intervention. Although the instructions to let children play as they wished were intended to maximize adolescents' ownership of their game experience without undue parental interference, the message was misinterpreted by a small number of parents as barring them from discussing the game with their children. In isolated cases, this resulted in less conversation and freedom of communication than anticipated. With parent-child communication, such an important protective factor for young people's healthy decision making around sexual health [29], inadvertently minimizing discussions between adolescents and their parents about the contents of this type of intervention through misleading study instructions has the potential to limit the effects of the intervention. Before any future studies of *Tumaini*, instructions relating to parent-child communication and parents' role in the intervention will be pretested to ensure that they are clearly understood as intended.

Based on their positive experiences with this study, parents had few reservations about a study spanning multiple years or one including the collection of blood for biomarkers. Any anticipated longitudinal efficacy trial of *Tumaini* would be planned such that study visits take place during Kenyan school holidays in line with parents' recommendations. HIV and HSV-2 testing would follow Kenyan testing guidelines [30], which provide for pre- and posttesting counseling as well as parental presence for young adolescents, addressing parents' concerns regarding testing procedures. In light of this, the planned efficacy study is expected to prove acceptable to parents of other adolescents in this age group. Other research from Botswana involving HSV-2 testing as a biomarker for sexual activity identified some potential barriers to acceptability from parents, which aligned with parents' concerns in this study [31].

Gooding et al [32] note that defining trial acceptability is challenging, that it may change over the course of a study's duration, and that it is likely to differ across participants. They encourage an early definition of trial acceptability specific to the study and its aims rather than an all-purpose definition applied to studies across the board. They additionally warn against using loss to follow-up as an indicator of continued acceptability as withdrawal, or continued participation, may result from factors external to study methods. This study was relatively short and parents' comments on the acceptability of the planned efficacy trial may not reflect how acceptable the trial is with participants once the trial is underway. However, with these caveats, it is encouraging that participants engaged in this study do not foresee barriers to the acceptability of a longer study or one involving blood collection. Additional data collection should be carried out before and alongside the planned trial to identify, minimize, and account for emerging obstacles to continued acceptability, and not merely tolerance, of the study to all participants [33]. This study adds to the current body of literature on the development of digital interactive sexual and reproductive health interventions aimed at adolescents and young adults. It provides detailed insights into different facets of acceptability of this intervention intended to inform future trials of *Tumaini*. In addition, it discusses the importance of assessing the acceptability of an intervention and of study methods before undertaking large-scale efficacy trials and the value of developing standards to allow for the comparison of interventions in terms of acceptability.

Limitations

This study included 30 adolescents in the intervention group and 27 adolescents and 22 parents in the FGDs, and the intervention period was just over 2 weeks. As most of the parents were female, it is possible that fathers may have had a different perception of the game and study. Adolescents were not asked about the acceptability of blood tests or of a longitudinal study, and acceptability of a future study was based on a hypothetical description of a study rather than a detailed study plan. This study was conducted in Kisumu, western Kenya, and the findings may not be generalizable throughout Kenya or to other parts of East Africa.

Conclusions

As researchers develop innovative mHealth prevention interventions and strengthen efficacy testing, they must consider the acceptability of the intervention and trial as crucial to success. For interventions aimed at adolescents, acceptability must include acceptability to parents to minimize barriers to trial retention and intervention uptake. Our study findings show

a high level of acceptability for a smartphone-based HIV prevention serious game intervention among adolescents and their parents as well as the methods used to conduct the study. The acceptability constructs used in our study align with new frameworks proposed by Weiner [28] and Gooding [32], and we hope they can help inform similar intervention studies.

Acknowledgments

Research reported in this publication was supported by the National Institute of Mental Health of the US National Institutes of Health under Award Number 5R34MH106368 (principal investigator: KW). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. This research was also supported by the Emory Center for AIDS Research (P30 AI050409) and the Andrew W Mellon Foundation. The sponsors played no role in the review and approval of the manuscript for publication. We are grateful to Rob Breiman, Eugen Condur, Jeff Cook, Meghan Duffy, Amy Eisenbach, Gilian Njika, Chris Obong'o, David Schnabel, and David Warhol. Special thanks to Kim Miller for her guidance and mentorship. This paper is published with approval from the Director KEMRI.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Sample Tumaini Graphics.

[DOCX File, 740KB - mhealth_v7i5e13049_app1.docx]

Multimedia Appendix 2

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 8MB - mhealth_v7i5e13049_app2.pdf]

References

1. UNAIDS. Geneva; 2018. UNAIDS Data 2018 URL: http://www.unaids.org/en/resources/documents/2018/unaids-data-2018 [accessed 2018-09-16] [WebCite Cache ID 786Vz2DhO]
2. UNAIDS. Geneva; 2018. The Youth Bulge and HIV URL: http://www.unaids.org/en/resources/documents/2018/the-youth-bulge-and-hiv [accessed 2018-06-11] [WebCite Cache ID 786XQmv7C]
3. Poobalan A, Pitchforth E, Imamura M, Tucker J, Philip K, Spratt J, et al. Characteristics of effective interventions in improving young people's sexual health: a review of reviews. Sex Education 2009 Aug;9(3):319-336. [doi: 10.1080/14681810903059185]
4. Michielsen K, Chersich MF, Luchters S, De Koker P, Van Rossem R, Temmerman M. Effectiveness of HIV prevention for youth in sub-Saharan Africa: systematic review and meta-analysis of randomized and nonrandomized trials. AIDS 2010 May 15;24(8):1193-1202. [doi: 10.1097/QAD.0b013e3283384791] [Medline: 20375876]
5. Pedlow CT, Carey MP. Developmentally appropriate sexual risk reduction interventions for adolescents: rationale, review of interventions, and recommendations for research and practice. Ann Behav Med 2004 Jun;27(3):172-184 [FREE Full text] [doi: 10.1207/s15324796abm2703_5] [Medline: 15184093]
6. Groupe Speciale Mobile Association. London: Groupe Speciale Mobile Association; 2018. The Mobile Economy, Sub-Saharan Africa URL: https://www.gsma.com/mobileeconomy/sub-saharan-africa/ [accessed 2018-11-02] [WebCite Cache ID 786YDBqKp]
7. Groupe Speciale Mobile Association. London: Groupe Speciale Mobile Association; 2017. Accelerating affordable smartphone ownership in emerging markets URL: https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2018/08/Accelerating-affordable-smartphone-ownership-in-emerging-markets-2017_we.pdf [accessed 2018-11-01] [WebCite Cache ID 786YOCeXM]
8. Pettifor A, Bekker L, Hosek S, DiClemente R, Rosenberg M, Bull SS, et al. Preventing HIV among young people: research priorities for the future. J Acquir Immune Defic Syndr 2013 Jul;63(Suppl 2):S155-S160 [FREE Full text] [doi: 10.1097/QAI.0b013e31829871fb] [Medline: 23764629]
9. Noar SM. Computer technology-based interventions in HIV prevention: state of the evidence and future directions for research. AIDS Care 2011 May;23(5):525-533 [FREE Full text] [doi: 10.1080/09540121.2010.516349] [Medline: 21287420]
10. Hightow-Weidman LB, Muessig KE, Bauermeister JA, LeGrand S, Fiellin LE. The future of digital games for HIV prevention and care. Curr Opin HIV AIDS 2017 Sep;12(5):501-507. [doi: 10.1097/COH.0000000000000399] [Medline: 28692490]

11. Allison S, Bauermeister JA, Bull S, Lightfoot M, Mustanski B, Shegog R, et al. The intersection of youth, technology, and new media with sexual health: moving the research agenda forward. J Adolesc Health 2012 Sep;51(3):207-212 [FREE Full text] [doi: 10.1016/j.jadohealth.2012.06.012] [Medline: 22921129]
12. Abt C. Serious Games. New York: University Press of America; 1987.
13. Shegog R. Application of behavioral theory in computer game design for health behavior change. In: Cannon- Bowers J, Bowers C, editors. Serious Game Design and Development: Technologies for Training and Learning. Hershey, PA: Information Science Reference; 2010:196-232.
14. Lieberman DA. Management of chronic pediatric diseases with interactive health games: theory and research findings. J Ambul Care Manage 2001 Jan;24(1):26-38. [Medline: 11189794]
15. Singhal A, Cody M, Roger E, Sabido M, editors. Entertainment-Education and Social Change: History, Research, and Practice. Mahwah, NJ: Lawrence Erlbaum Associates; 2004.
16. Kato PM, Cole SW, Bradlyn AS, Pollock BH. A video game improves behavioral outcomes in adolescents and young adults with cancer: a randomized trial. Pediatrics 2008 Aug;122(2):e305-e317 [FREE Full text] [doi: 10.1542/peds.2007-3134] [Medline: 18676516]
17. Proctor E, Silmere H, Raghavan R, Hovmand P, Aarons G, Bunger A, et al. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. Adm Policy Ment Health 2011 Mar;38(2):65-76 [FREE Full text] [doi: 10.1007/s10488-010-0319-7] [Medline: 20957426]
18. Davis F. User acceptance of information technology: system characteristics, user perceptions and behavioral impacts. Int J Man Mach Stud 1993 Mar;38(3):475-487. [doi: 10.1006/imms.1993.1022]
19. Vig J, Miller K, Chirwa-Motswere C, Winskell K, Stallcup E. Involving parents from the start: formative evaluation for a large randomised controlled trial with Botswana Junior Secondary School students. Afr J AIDS Res 2016 Mar 25;15(1):9-15. [Medline: 27002354]
20. Winskell K, Sabben G, Akelo V, Ondeng'e K, Obong'o C, Stephenson R, et al. A smartphone game-based intervention (Tumaini) to prevent HIV among young Africans: pilot randomized controlled trial. JMIR Mhealth Uhealth 2018 Aug 01;6(8):e10482 [FREE Full text] [doi: 10.2196/10482] [Medline: 30068501]
21. Plummer M, Ross D, Wight D, Changalucha J, Mshana G, Wamoyi J, et al. "A bit more truthful": the validity of adolescent sexual behaviour data collected in rural northern Tanzania using five methods. Sex Transm Infect 2004 Dec;80 Suppl 2:ii49-ii56 [FREE Full text] [doi: 10.1136/sti.2004.011924] [Medline: 15572640]
22. Sabben G, Akelo V, Mudhune V, Ondeng'e K, Ndivo R, Stephenson R, et al. A smartphone game to prevent HIV among young Africans: protocol for a randomized pilot study of a mobile intervention. JMIR Res Protoc 2019 Mar 27;8(3):e11209 [FREE Full text] [doi: 10.2196/11209] [Medline: 30916661]
23. Beres LK, Winskell K, Neri EM, Mbakwem B, Obyerodhyambo O. Making sense of HIV testing: social representations in young Africans' HIV-related narratives from six countries. Glob Public Health 2013;8(8):890-903 [FREE Full text] [doi: 10.1080/17441692.2013.827734] [Medline: 24004339]
24. Winskell K, Beres LK, Hill E, Mbakwem BC, Obyerodhyambo O. Making sense of abstinence: social representations in young Africans' HIV-related narratives from six countries. Cult Health Sex 2011 Sep;13(8):945-959 [FREE Full text] [doi: 10.1080/13691058.2011.591431] [Medline: 21787256]
25. Winskell K, Obyerodhyambo O, Stephenson R. Making sense of condoms: social representations in young people's HIV-related narratives from six African countries. Soc Sci Med 2011 Mar;72(6):953-961 [FREE Full text] [doi: 10.1016/j.socscimed.2011.01.014] [Medline: 21388731]
26. Winskell K, Sabben G, Ondeng'e K, Odero I, Akelo V, Mudhune V. A smartphone game to prevent HIV among young Kenyans: household dynamics of gameplay in a feasibility study. Health Educ J 2019 Feb 28:001789691983234. [doi: 10.1177/0017896919832344]
27. Lewis CC, Fischer S, Weiner BJ, Stanick C, Kim M, Martinez RG. Outcomes for implementation science: an enhanced systematic review of instruments using evidence-based rating criteria. Implement Sci 2015 Nov 04;10:155 [FREE Full text] [doi: 10.1186/s13012-015-0342-x] [Medline: 26537706]
28. Weiner BJ, Lewis CC, Stanick C, Powell BJ, Dorsey CN, Clary AS, et al. Psychometric assessment of three newly developed implementation outcome measures. Implement Sci 2017 Dec 29;12(1):108 [FREE Full text] [doi: 10.1186/s13012-017-0635-3] [Medline: 28851459]
29. Miller KS, Levin ML, Whitaker DJ, Xu X. Patterns of condom use among adolescents: the impact of mother-adolescent communication. Am J Public Health 1998 Oct;88(10):1542-1544. [Medline: 9772860]
30. National AIDS and STI Control Programme. Ministry of Health Kenya. Nairobi: Ministry of Health Kenya; 2015. The Kenya HIV Testing Services Guidelines URL: https://aidsfree.usaid.gov/sites/default/files/hts_policy_kenya_2015.pdf [accessed 2017-01-14] [WebCite Cache ID 786ZKjhSm]
31. Cham H, Lasswell S, Miller K. Parents' reactions to testing for herpes simplex virus type 2 as a biomarker of sexual activity in Botswana junior secondary school students. Sex Health 2016 Jan;13(2):148-154 [FREE Full text] [doi: 10.1071/SH15092] [Medline: 26886026]

32. Gooding K, Phiri M, Peterson I, Parker M, Desmond N. Six dimensions of research trial acceptability: how much, what, when, in what circumstances, to whom and why? Soc Sci Med 2018 Jan;213:190-198. [doi: 10.1016/j.socscimed.2018.07.040] [Medline: 30142500]
33. King K, Kolopack P, Merritt M, Lavery J. Community engagement and the human infrastructure of global health research. BMC Medical Ethics 2014 Dec 13;15(1):84. [Medline: 25495054]

Abbreviations

FGD: focus group discussion
HSV-2: herpes simplex virus type 2
KEMRI: Kenya Medical Research Institute
mHealth: mobile health
TAM: Technology Acceptance Model

Edited by G Eysenbach; submitted 21.12.18; peer-reviewed by A Thomas, K Tufts; comments to author 03.04.19; revised version received 09.04.19; accepted 09.04.19; published 21.05.19.

Please cite as:
Sabben G, Mudhune V, Ondeng'e K, Odero I, Ndivo R, Akelo V, Winskell K
A Smartphone Game to Prevent HIV Among Young Africans (Tumaini): Assessing Intervention and Study Acceptability Among Adolescents and Their Parents in a Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(5):e13049
URL: http://mhealth.jmir.org/2019/5/e13049/
doi:10.2196/13049
PMID:31115348

Original Paper

Using Mobile Health to Support Clinical Decision-Making to Improve Maternal and Neonatal Health Outcomes in Ghana: Insights of Frontline Health Worker Information Needs

Hannah Brown Amoakoh[1,2], MPH, MD; Kerstin Klipstein-Grobusch[2,3], MSc, PhD; Diederick E Grobbee[2], MD, PhD; Mary Amoakoh-Coleman[2,4], MSc, MPH, MD, PhD; Ebenezer Oduro-Mensah[5], MD; Charity Sarpong[6], MBA, MPH, MD; Edith Frimpong[7], MSc; Gbenga A Kayode[2], MSc, MD, PhD; Irene Akua Agyepong[8], MD, DrPH; Evelyn K Ansah[9], MD, PhD

[1]Department of Epidemiology, School of Public Health, University of Ghana, Accra, Ghana

[2]Julius Center for Health Sciences and Primary Care, University Medical Center, Utrecht University, Utrecht, Netherlands

[3]Division of Epidemiology and Biostatistics, School of Public Health, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

[4]Department of Epidemiology, Noguchi Memorial Institute for Medical Research, University of Ghana, Accra, Ghana

[5]La General Hospital, Ghana Health Service, Accra, Ghana

[6]Regional Health Directorate, Ghana Health Services, Koforidua, Ghana

[7]Dodowa Research Centre, Ghana Health Service, Accra, Ghana

[8]Research and Development Division, Ghana Health Service, Accra, Ghana

[9]Centre for Malaria Research, University of Health and Allied Sciences, Ho, Ghana

Corresponding Author:
Hannah Brown Amoakoh, MPH, MD
Julius Center for Health Sciences and Primary Care
University Medical Center
Utrecht University
PO Box 85500
Utrecht, 3508 GA
Netherlands
Phone: 233 264545762
Email: ansomaame@hotmail.com

Abstract

Background: Developing and maintaining resilient health systems in low-resource settings like Ghana requires innovative approaches that adapt technology to context to improve health outcomes. One such innovation was a mobile health (mHealth) clinical decision-making support system (mCDMSS) that utilized text messaging (short message service, SMS) of standard emergency maternal and neonatal protocols via an unstructured supplementary service data (USSD) on request of the health care providers. This mCDMSS was implemented in a cluster randomized controlled trial (CRCT) in the Eastern Region of Ghana.

Objective: This study aimed to analyze the pattern of requests made to the USSD by health workers (HWs). We assessed the relationship between requests made to the USSD and types of maternal and neonatal morbidities reported in health facilities (HFs).

Methods: For clusters in the intervention arm of the CRCT, all requests to the USSD during the 18-month intervention period were extracted from a remote server, and maternal and neonatal health outcomes of interest were obtained from the District Health Information System of Ghana. Chi-square and Fisher exact tests were used to compare the proportion and type of requests made to the USSD by cluster, facility type, and location; whether phones accessing the intervention were shared facility phones or individual-use phones (*type-of-phone*); or whether protocols were accessed during the day or at night (*time-of-day*). Trends in requests made were analyzed over 3 6-month periods. The relationship between requests made and the number of cases reported in HFs was assessed using Spearman correlation.

Results: In total, 5329 requests from 72 (97%) participating HFs were made to the intervention. The average number of requests made per cluster was 667. Requests declined from the first to the third 6-month period (44.96% [2396/5329], 39.82% [2122/5329],

and 15.22% [811/5329], respectively). Maternal conditions accounted for the majority of requests made (66.35% [3536/5329]). The most frequently accessed maternal conditions were postpartum hemorrhage (25.23% [892/3536]), *other conditions* (17.82% [630/3536]), and hypertension (16.49% [583/3536]), whereas the most frequently accessed neonatal conditions were prematurity (20.08% [360/1793]), sepsis (15.45% [277/1793]), and resuscitation (13.78% [247/1793]). Requests made to the mCDMSS varied significantly by cluster, type of request (maternal or neonatal), facility type and its location, *type-of-phone*, and *time-of-day* at 6-month interval (P<.001 for each variable). Trends in maternal and neonatal requests showed varying significance over each 6-month interval. Only asphyxia and sepsis cases showed significant correlations with the number of requests made (r=0.44 and r=0.79; P<.001 and P=.03, respectively).

Conclusions: There were variations in the pattern of requests made to the mCDMSS over time. Detailed information regarding the use of the mCDMSS provides insight into the information needs of HWs for decision-making and an opportunity to focus support for HW training and ultimately improved maternal and neonatal health.

(JMIR Mhealth Uhealth 2019;7(5):e12879) doi:10.2196/12879

KEYWORDS

mHealth; maternal health; neonatal health; health care systems; developing countries; decision-making; information retrieval systems

Introduction

Background

Weak health systems are a major barrier to achieving improved health outcomes in low- and middle-income countries [1]. It is therefore not surprising that many countries that could not attain the Millennium Development Goals (MDGs) 3 and 4 which targeted improvements in maternal, neonatal, and child health (MNCH), were from the parts of the globe with poorly developed health systems such as sub-Saharan Africa and Southern Asia [2]. As global efforts to improve MNCH intensifies through the Sustainable Development Goals (SDGs) 3.1 and 3.2 [3], health system strengthening has become imperative to attain these SDGs.

Among the many interventions currently being implemented to address MNCH challenges, mobile health (mHealth) interventions have been widely used in low- and middle-income countries [4] as a potential solution to maximize health worker (HW) impact, efficiency, and health outcomes [5,6] and improve service utilization [7]. Common areas of application of mHealth tools include point-of-care decision-making support, provider-to-provider communication, and data collection [4,8,9]. Though mHealth interventions are well received by HWs and the community [9-12], data about their effectiveness with regards to patient health outcomes, improved efficiency of health systems, or their use by HWs are limited [5,7-9,13,14].

Ghana, a sub-Saharan African country, reports unacceptably high maternal and neonatal deaths that fell short of the MDGs targets [2]. Ghana's maternal mortality is presently estimated at 319 per 100,000 live births [15] and its neonatal mortality rate is 25 deaths per 1000 live births, with higher mortality rates being reported in rural areas of the country [16-19]. Though numerous training programs and maternal audits are performed in Ghana to improve the quality of MNCH services [20], health system constraints still remain. Health system constraints contributing to persistently high maternal and neonatal mortality in Ghana include cost, distance, availability of health facilities (HFs), attitude of nurses toward pregnant women [21,22], and nonadherence of HWs to clinical guidelines [23,24]. To address the constraint of poor adherence to clinical guidelines by HWs, we designed an mHealth intervention—a clinical decision-making support system (CDMSS) to facilitate easy access to maternal and neonatal guidelines for routine and emergency obstetric, antenatal, and neonatal care for frontline providers of maternal and neonatal care in Ghana [25].

Description of the Intervention

This mHealth clinical decision-making support system (mCDMSS) consisted of 4 components: (1) Phone calls (to facilitate verbal communication between frontline health workers, FHWs), (2) SMS text messaging (short message service, SMS; to facilitate communication between FHWs during periods of nonsustained network connectivity), (3) Access to an unstructured supplementary service data (USSD) for standard emergency obstetric and neonatal protocols via SMS text messaging (to provide quick and easy access to the standard guidelines to maternal and neonatal health protocols in Ghana), and (4) Access to the internet (to facilitate access to health information that may not be found in the USSD protocols). All these components were embedded in a composite intervention on a project nonsmart mobile phone. The multifaceted nature of the mCDMSS was aimed to assure access to clinical decision-making support for HWs at all times following suggestions from FHWs for clinical decision-making support in a formative study [26]. Access to the USSD was considered to be the main intervention component. Health workers were expected to use the phones primarily to access neonatal and maternal health emergency protocols via the USSD and obtain additional support from colleagues and the internet via the other intervention components. The messages on the USSD were created by a team of FHWs, family physicians, obstetricians, and pediatricians in the Greater Accra Region, drawing on the Ghana's Safe Motherhood protocols [27]. The development of the intervention was done using an iterative process that piloted and tested the USSD messages among FHWs in the Greater Accra Region to assure comprehension and appropriateness of the USSD messages. The USSD was designed such that new protocol requests needed to be initiated if a request session was terminated prematurely. FHWs, mainly midwives were provided with 312 dedicated nonsmart mobile

phones to access the intervention. These phones were classified by the research team as shared facility phones if dedicated for shared-use by all providers of maternal and neonatal health care services in a HF or, as individual-use phones if dedicated to personal use of midwives. Each midwife at post in each HF during baseline assessment was provided with 1 mobile phone (individual-use phone) as they work closely with maternal and neonatal patients. FHWs were assumed to be familiar with the basic functioning of a mobile phone (making calls, texting, and accessing the internet) as documented in previous studies [28,29], so the training concerning the use of the mCDMSS focused on how to use the USSD. Navigation through the USSD has been demonstrated in Multimedia Appendix 1.

We tested the intervention in a cluster randomized controlled trial (CRCT) in the Eastern Region of Ghana. The CRCT has been described in detail elsewhere [25]. Vodafone Ghana, a telecommunication company, provided technical support for the mCDMSS and collected routine data regarding how the intervention was used throughout the intervention period.

Study Objectives

The USSD component of the intervention explicitly and objectively provides insight into the information needs of FHWs. As details of protocols accessed from the USSD by FHWs are not known, we aimed to, first, describe the pattern of USSD protocol requests made by frontline providers of maternal and neonatal health services in district level HFs in the Eastern Region of Ghana and second, to examine the relationship between the patterns of requests made and the incidence of maternal and neonatal morbidity in HFs accessing the intervention.

Methods

Study Design and Sampling

This study was conducted within the context of the aforementioned CRCT, which aimed to assess the impact of the mCDMSS on institutional neonatal mortality in the Eastern Region of Ghana and comprised 16 districts randomized into 8 intervention and 8 control clusters. In a given cluster, all public and private HFs that work with the Ghana Health Service participated in the CRCT. We extracted all requests made to the USSD during the 18 months of intervention implementation (August 1, 2015 to January 31, 2017) from the USSD server of Vodafone Ghana; and all morbidity cases for the aforementioned timeframe for which requests were made, from the District Health Information Management System (DHIMS2) in Ghana. The DHIMS2 is a data recording, collection, collation, and analysis tool that hosts the entire national institutional health data of Ghana mainly from the public sector and a few private facilities [25].

This study was approved by the Ghana Health Service Ethics Review Committee before its commencement; study approval number GHS-ERC: 04/09/16.

Data Collection

Before data extraction, phone numbers assigned to various users was collated such that each intervention user, the HF as well as the district (cluster) the user worked in, was documented and coded in Vodafone Ghana's database. This ensured that requests made to the USSD could be traced back to the clusters, HFs, and FHWs using the phone. A total of 5 of the individual-use phones could not be traced back to the FHWs who received them as they were not signed for, and efforts to reach these numbers were futile. These 5 phone numbers were thus, not included in analysis. The USSD data were extracted monthly. Due to technical challenges at Vodafone Ghana, 22 days of data were lost during the first 6 months of the intervention. From the DHIMS2 database, maternal cases of postpartum hemorrhage (PPH), antepartum hemorrhage (APH), hypertensive disorders in pregnancy (HDP), and neonatal cases of prematurity, asphyxia, jaundice, cord sepsis, and sepsis occurring in the intervention period were extracted. In the DHIMS2, data captured regarding the aforementioned maternal cases cover hospital in-patients only. In the case of neonatal morbidity, the DHIMS2 captures data regarding neonatal cases of sepsis and prematurity at only hospital level, whereas neonatal cases of asphyxia, jaundice, and cord sepsis are captured as aggregate data for all types of HFs, that is, hospitals, health centers (HCs), and Community-based Health Planning and Services (CHPS) working with or within the Ghana Health Service. Due to challenges with the DHIMS2, some hospitals entered data concerning morbidities of interest that were not captured or could not be extracted from the DHIMS2 onto Excel spreadsheets that were given to the project team for analysis. The data entry in such situations was done by the hospital health information officers responsible for entering those data into the DHIMS2, and the data were validated by the head of the health information unit in these hospitals.

Statistical Analysis

The data were checked for errors and exported from Microsoft Excel (Microsoft Corporation) to Stata version 13 (StataCorp LLC) for cleaning and analysis. We classified HFs into 2 groups of remote and nonremote areas based on access. Remote facilities were either located more than 30-min' walk or more that 15-min motorbike ride from the main district township and had poor road access (uneven and untarred roads overcrowded with weeds and shrubs) leading to them. Nonremote HFs were either located within 30-min' walk or 15-min motorbike ride from the main district township and had good road access leading to them. Due to the similarities in organizational structure, personnel and health services provided by CHPS, and maternity homes, requests from these 2 facility types were combined for analysis. Time of accessing the USSD was coded as day if requests were made from 6 am to 6 pm; all other time periods were coded night. Maternal morbidities— gestational hypertension, chronic hypertension, eclampsia, pre-eclampsia, and hypertensive encephalopathy were all classified as HDP. Placenta praevia and abruption were considered as APH, and retained placenta was considered as PPH as patients are usually hospitalized because of bleeding from these conditions. Unspecified cause of bleeding and vomiting were excluded during analysis. The Vodafone data were not corrected for the 22 days of missing data in the first 6 months of intervention implementation as the data were considered missing completely at random [30].

Descriptive analysis of requests made to the USSD server from clusters, HFs, *type-of-phone* (individual-use or shared-use), HF location, and *time-of-day* (explanatory variables) was done and expressed in numbers and percentages, first, as a combined 18-month data and then at 6-month intervals. Trends in maternal and neonatal requests were assessed. Chi-square and Fisher exact tests were applied to these analyses to assess the significance of the observed pattern of USSD requests. Morbidity from aforementioned cases of interest were estimated from the DHIMS2. The relationship between USSD requests and morbidity from cases for which requests were made was also estimated using Spearman correlation. All analyses were performed using Stata 13 statistical software and using 2-tailed tests at alpha=.05.

Results

User Statistics

A total of 74 HFs in all 8 intervention clusters were recruited into this study (Table 1). Each cluster included at least 1 district hospital but a varying mix of HCs and CHPS. In all, data from 307 mobile phones were analyzed; 74 were shared-use phones, whereas the rest were individual-use phones. At the end of the intervention period, a total of 5329 requests were made to the USSD. Of these requests, 2396 (44.96% [2396/5329]) were made during the first 6 months, 2122 (39.82% [2122/5329]) in the second 6 months, and 811 (15.22% [811/5329]) in the last 6 months. Throughout the intervention period, maternal requests (66.35% [3536/5329]) were made more frequently compared with neonatal requests (33.65% [1793/5329]). Requests per cluster ranged from 1167 (representing 21.90% [1167/5329] of total requests) to 403 (representing 7.56% [403/5329] of requests); the average request made per cluster was 667. All clusters made a request to the USSD. Of the 74 HFs (combined from all clusters), 72 accessed the USSD at least once during the intervention period. The 2 HFs that did not access the intervention included a privately owned maternity home that had no midwife at post throughout intervention implementation and a CHPS compound whose midwife shared during a routine supervisory visit by the research team that she trusted her competence in midwifery practice and so did not see the need to consult the USSD protocols. Among HFs, requests from hospitals declined from the first to the last 6 months, whereas requests from HCs and CHPS increased. Close to hundred percent (98.44% [2904/2950]) of all requests made from hospitals were with individual-use phones compared with the proportion of requests made with individual-use phones in HCs (52.87% [654/1237]) and CHPS (30.74% [351/1142]; *P*<.001).

At night, the proportion of requests made from HCs (27.81% [344/1237]) and CHPS (27.67% [316/1142]) was lower than the proportion of requests from hospitals (34.17% [1008/2950]; *P*<.001). There were similarities in the observed proportion of maternal protocols assessed by individual-use phones (65.49% [2560/3909]) and shared-use phones (68.73% [976/1420]; *P*=.03); and in the proportion of requests made at night by both phone types (27.62% [450/1630] for shared-use phones and 31.52% [1371/4350] for individual-use phones; *P*=.003). Shared-use phones were used more often in remote areas (78.24% [1111/1420]) compared with individual-use phones (11.49% [499/3909]) in accessing the intervention (*P*<.001). The frequency of shared-use phones accessing the intervention increased over time, whereas the frequency of individual-use phone decreased.

Table 1. Background characteristics of clusters.

Cluster[a]	Number of health facilities				Demographic location of health facilities, n (%)		Number of deliveries per midwife[b]	Proportion of shared phones received, n (%)
	Hospital[c]	HCs[d]	CHPS[e]	Maternity home[f]	Remote	Nonremote		
A	1	1	7	0	7 (78)	2 (22)	80.0	9[g] (43)
B	1	7	3	0	6 (55)	5 (45)	130.4	11 (38)
C	3	2	9	1	6 (40)	9 (60)	99.0	15 (26)
D	2	3	3	0	5 (63)	3 (37)	94.8	8 (20)
E	1	3	2	0	3 (50)	3 (50)	107.6	6 (24)
F	1	8	1	1	4 (36)	7 (64)	101.6	11 (22)
G	3	2	0	1	1 (17)	5 (83)	75.0	6 (11)
H	1	3	3	1	4 (50)	4 (50)	96.4	8[g] (26)

[a]Clusters have been named A-H for anonymity.

[b]Reference year is 2014.

[c]Includes both private and public hospitals.

[d]HCs: Health Centers.

[e]CHPS: Community-based Health Planning and Services.

[f]Includes only private maternity homes.

[g]This may differ slightly from the sum of the number of midwives in the cluster and the number of health facilities as 2 individual-use phones from these clusters could not be traced.

The proportion of maternal requests from remote (69.81% [1089/1560]) and nonremote areas (64.92% [2447/3769] as well as the proportion of requests made at night from remote (31.09% [485/1560]) and nonremote areas (31.39% [1183/3769]) were similar (*P*=.001 for request type, *P*=.046 for time of day requests were made). The frequency of remote areas accessing the intervention increased over time, whereas the frequency of nonremote areas decreased. Requests by clusters, HFs and their location, type of request (maternal or neonatal), *type-of-phone*, and *time-of-day* varied significantly at 6-month intervals during the intervention period (Table 2).

Trends in Maternal Requests

Detailed analysis of maternal requests show that PPH protocols were accessed the most (27.22% [450/1653]) in the first 6 months, followed by *other conditions* protocols (16.76% [277/1653]) and HDP protocols (16.21% [268/1653]). This trend in requests was repeated in the second 6 months (PPH: 22.69% [300/1322], *other conditions*: 20.57% [272/1322], and HDP: 16.34% [216/1322]). In the last 6 months, HDP (17.7% [99/561]) was the second most accessed protocol after PPH (25.3 [142/561]), whereas APH and *other conditions* contributed 14.4% (81/561) each to requests made. Across clusters, this trend in maternal requests was significant at each 6-month interval (*P*<.001 for each timeframe). Across HFs, the trend of maternal requests aforementioned differed significantly only in the first and second timeframe (*P*=.04, *P*=.03, and *P*=.15, respectively); by *type-of-phone*, this trend varied at all 3 time points (*P*=.05, *P*=.01, and *P*<.001, respectively); and across HFs, maternal request trends differed in the third 6th month alone (*P*=.57, *P*=.42, and *P*=.001, respectively, for each timeframe).

Table 2. Distribution of unstructured supplementary service data requests at 6-monthly intervals.

Variable	First 6 months, frequency (%)	Second 6 months, frequency (%)	Third 6 months, frequency (%)	Total, frequency (%)	*P* value for χ^2 test
Cluster[a]					
A	244 (10.18)	216 (10.18)	198 (24.41)	658 (100.00)	<.001
B	262 (10.93)	184 (8.67)	42 (5.18)	488 (100.00)	<.001
C	406 (16.94)	311 (14.66)	97 (11.96)	814 (100.00)	<.001
D	174 (7.26)	220 (10.37)	98 (12.08)	492 (100.00)	<.001
E	173 (7.22)	153 (7.21)	77 (9.49)	403 (100.00)	<.001
F	552 (23.04)	438 (20.64)	177 (21.82)	1167 (100.00)	<.001
G	261 (10.89)	468 (22.05)	48 (5.92)	777 (100.00)	<.001
H	324 (13.52)	132 (6.22)	74 (9.12)	530 (100.00)	<.001
Type of request					
Maternal care	1653 (68.99)	1322 (62.30)	561 (69.17)	3536 (100.00)	<.001
Neonatal care	743 (31.01)	800 (37.70)	250 (30.83)	1793 (100.00)	<.001
Type of facility					
Hospitals	1563 (65.23)	1069 (50.38)	318 (39.21)	2950 (100.00)	<.001
HCs[b]	418 (17.45)	587 (27.66)	232 (28.16)	1237 (100.000	<.001
CHPS[c] and maternity homes	415 (17.32)	466 (21.96)	261 (32.18)	1142 (100.00)	<.001
Type of phone					
Individual-use	1921(80.18)	1531 (72.15)	457 (56.35)	3903 (100.00)	<.001
Shared-use	475 (19.82)	591 (27.85)	354(43.65)	1420 (100.00)	<.001
Demographic location					
Nonremote	1906 (79.55)	1435 (67.62)	457 (56.35)	3769 (100.00)	<.001
Remote	490 (20.45)	687 (32.38)	354 (43.65)	1560 (100.00)	<.001
Time of day					
Day	1573 (65.65)	1526 (71.91)	562 (69.30)	3661 (100.00)	<.001
Night	823 (34.35)	596 (28.09)	249 (30.70)	1668 (100.00)	<.001

[a]Clusters have been named A-H for anonymity.

[b]HCs: Health Centers.

[c]CHPS: Community-based Health Planning and Services.

There was no variation in maternal requests trends by *time-of-day* requests were made at 6-month intervals (*P*=.16, *P*=.58, and *P*=.93, respectively). Detailed analysis of maternal requests pertaining to *other conditions* shows that hyperemesis was the most frequently requested protocol accounting for 26.3% (47/179) and 37.4% (70/187) of requests in the first and second 6 months, respectively. This was followed by fetal distress, which accounted for 18.4% (33/179) and 13.9% (26/187) of requests and premature rupture of membranes for gestation <37 weeks, which accounted for 17.3% (31/179) and 13.9% (26/187) of requests for *other conditions* for the same timeframe. In the third 6 months, cord prolapse, hyperemesis, and premature rupture of membranes for gestation <37 weeks accounted for 28% (15/54), 28% (15/54), and 19% (10/54) of *other conditions* request, respectively. Figure 1 describes in detail the pattern of maternal requests made by the clusters, HFs and their location, *type-of-phone*, and *time-of-day* for each 6-month period. Overall, there was a 20.02% (331/1653) and a 57.56% (761/1322) decline respectively, in the number of maternal requests made from the first to the second 6 months and from the second to the third 6 months of intervention implementation.

Trends in Neonatal Requests

Trends in neonatal requests show that prematurity protocols were accessed the most (22.6% [168/743]) in the first 6 months, followed by *abnormal breathing* protocols (15.8% [117/743]) and neonatal sepsis protocols (16.2% [113/743]). In the second 6 months, prematurity was most requested (16.9% [135/800]), followed by neonatal sepsis (16.38% [131/800]) and then *neonatal more* (16.3% [130/800]). In the last 6 months, frequently requested protocols were prematurity, resuscitation, and asphyxia in descending order of 22.8% (57/250), 14.8% (37/250), and 14.0% (35/250), respectively. Across clusters, this trend in neonatal requests was significantly different during the first and second 6 months of intervention implementation (*P*<.001 in each interval and *P*=.12 in the third 6 months). Across HFs, this trend of neonatal requests was again significantly different during the first 6 months (*P*=.001, *P*=.07, and *P*=.15, respectively, per interval); by *type-of-phone*, the observed trend aforementioned varied significantly during the first 6 months (*P*<.001, *P*=.38, and *P*=.07, respectively, per timeframe); by HF location and *time-of-day*, requests varied during the second 6 months only (*P* values for location type=.31, *P*=.001, and *P*=.13, respectively; *P* values for *time-of-day* analysis=.20, <.001, and .78, respectively). Detailed analysis of *neonatal more* requests show that 55.4% (72/130) of requests concerned neonatal seizures and the rest concerned birth trauma. Figure 2 describes in detail the pattern of neonatal requests made by the clusters, HFs and their location, *type-of-phone*, and *time-of-day* for each 6-month period. Overall, there was a 7.7% (57/743) increase and then a 68.8% (550/800) decline, respectively, in the number of neonatal requests made from the first to the second 6 months and from the second to the third 6 months of intervention implementation.

Figure 1. Maternal requests at 6-month interval by cluster, health facility type and location, phone type, and time of day requests were made. P values are chi-square tests or Fisher exact test comparing requests within each subcategory of each explanatory variable at 6-month intervals. Clusters have been labeled A-H for anonymity; CHPS: Community-based Health Planning and Services.

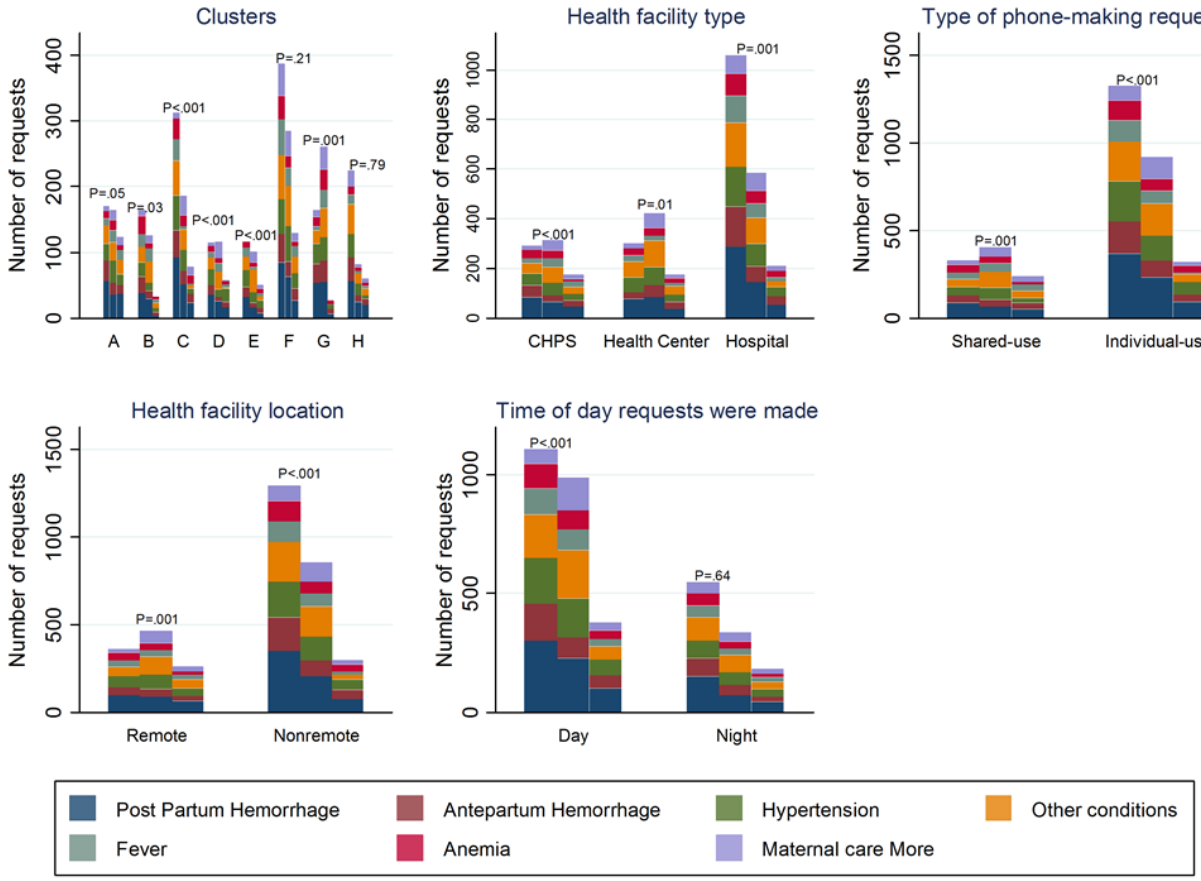
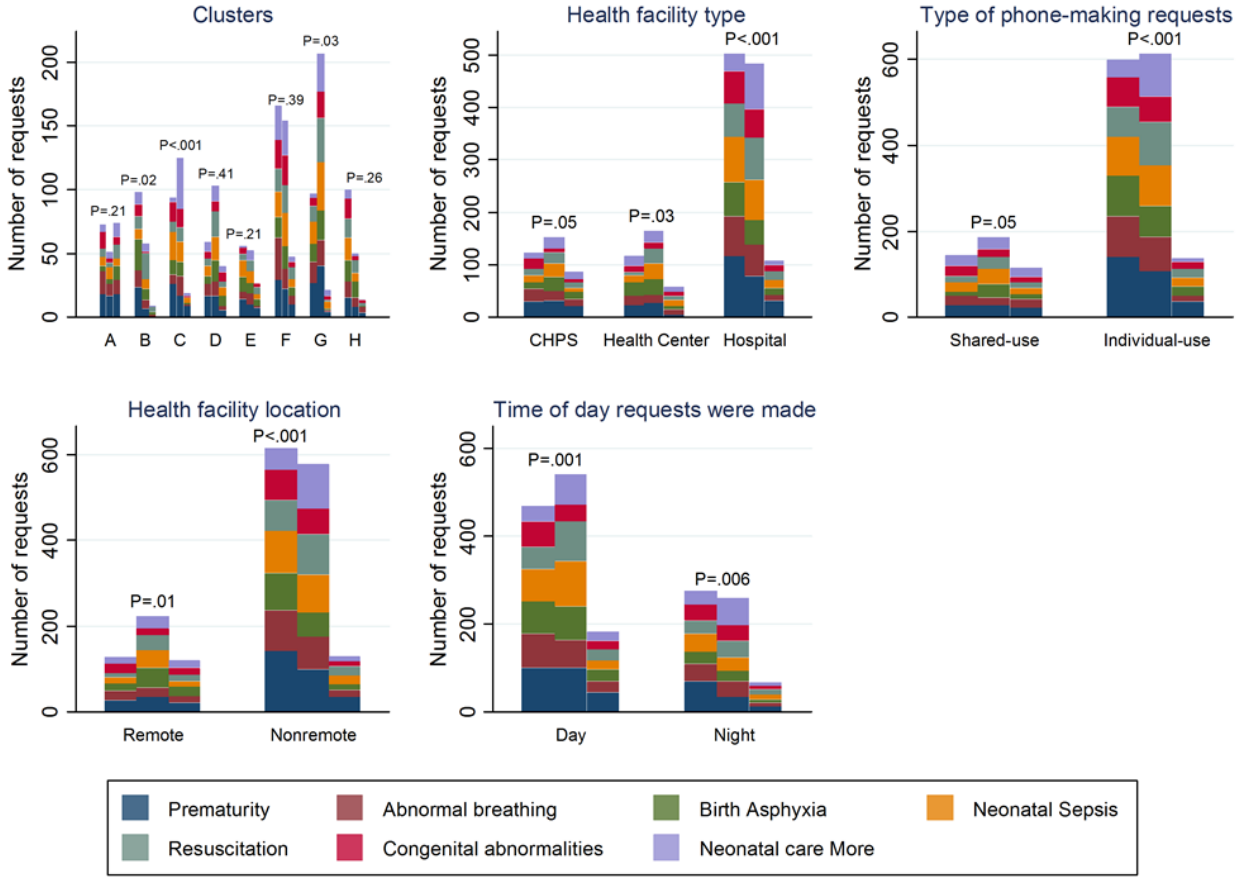


Figure 2. Neonatal requests at 6-month interval by cluster, health facility type and location, phone type, and time of day requests were made. P values are chi-square tests or Fisher exact test comparing requests within each subcategory of each explanatory variable at 6-month intervals. Clusters have been labeled A-H for anonymity; CHPS: Community-based Health Planning and Services.



Correlation Between Requests Made and Incidence of Cases

Generally, the number of maternal and neonatal cases exceeded the number of requests made except in the case of PPH. The correlation between requests made and actual number of cases recorded in HFs ranged from weak to strong positive and negative correlations. Spearman correlation was, however, significant for only asphyxia (Spearman rho=.44; *P*<.001) and sepsis cases (Spearman rho=.79; *P*=.03). Table 3 details the correlation coefficients for all outcomes of interest.

Table 3. Correlation between requests made to the intervention and actual number of cases recorded in health facilities.

Type of case	Number of requests[a] (%)	Number of cases (%)	Spearman rho	*P* value
Maternal[b]				
Antepartum hemorrhage	231 (51.7)	242 (100.0)	.05	.90
Postpartum hemorrhage	438 (49.1)	298 (100.0)	–.03	.93
Hypertension	267 (45.8)	1339 (100.0)	–.32	.41
Neonatal				
Asphyxia	320 (100.0)	2004 (100.0)	.44	<.001
Jaundice	15 (100.0)	158 (100.0)	.18	.12
Cord sepsis	6 (100.0)	63 (100.0)	–.07	.57
Sepsis[b]	124 (40.4)	185 (100.0)	.79	.03
Prematurity[b]	208 (57.9)	831 (100.0)	–.22	.58

[a]Represents the proportion of requests from only health facilities included in analysis.

[b]Represents hospital level data excluding 3 of the 12 hospitals in the intervention arm. Two of these 3 hospitals are in the same district and are privately owned; data from hospitals excluded were unavailable to researchers as of time of data analysis (August 2018).

Discussion

Principal Findings

This study describes the pattern of requests made to a SMS text messaging–based mobile CDMSS by FHWs providing maternal and neonatal health services in Ghana. We assessed the relationship between protocol requests made and types of maternal and neonatal morbidities for which requests were made. All clusters accessed the intervention, which is consistent with known findings of general acceptability of mhealth interventions among HWs and communities [9-12]. Maternal protocols were requested for more often than neonatal protocols, suggesting differences in information needs among FHWs with regards to maternal and neonatal care. Such differences in information needs was previously documented among community-level HWs in Nepal who seemed to be more knowledgeable in neonatal than maternal care matters [31]. This observation could also be a reaction of FHWs to the extensive maternal death audits conducted in Ghana [20]. Neonatal deaths, on the other hand, have not received such attention.

The high number of requests for protocols of PPH, HDP, prematurity, and sepsis in our study reflects the global and local trend in maternal and neonatal morbidity where these morbidities top the list [2,32-34]. This observation emphasizes the persistence of these morbidities in low-resource settings and the consequent need for health system strengthening in this regard and focus on these areas during HW training.

Within clusters, PPH and prematurity protocols were most commonly requested, suggesting that FHWs in the different clusters have a common information gap regarding these 2 morbidities that was bridged by this intervention. Differences in requests by category of HFs appears to reflect the differences in information needs of FHWs at the different levels of health care [35]. Surprisingly, the majority of requests emanated from hospitals where one would assume resource availability to be higher. The higher number of requests from hospital FHWs could indicate an unmet need for clinical decision-making support that is action-oriented even at higher levels of the health system. D'Adamo and his colleagues made similar findings of a near or complete lack of access to current useful information for district- and community-level HWs in their study [35]. It is striking that the trend in both maternal and neonatal requests did not differ significantly by HF location at all 3 time points in our study. A plausible explanation for this observation is similarities in competencies of FHWs in both remote and nonremote settings. This may be particularly true as in the Ghana Health Service, HWs may be freely transferred from 1 HF type to another. Similarities in FHW competencies across HFs may explain the absence of differences in request trends by the *time-of-day* requests were made. Detailed information about the FHWs who utilized the intervention could have provided more information regarding this analysis but was not collected in this study. The higher proportion of requests made by nonremote areas compared with remote areas is most likely because of the higher proportion of individual-use phones in nonremotes areas where there are generally higher numbers of midwives. Hence, a collinear relationship between requests made by *type-of-phone* and HF location can be observed. However, it is remarkable that nearly all requests from hospitals were made with individual-use phones implying a near absolute redundancy of shared-use phones in hospitals. This observation suggests that FHWs who were given the project phones are probably the same and only people who used the intervention in hospitals. Lack of knowledge transfer concerning the availability and use of the intervention with other FHWs who missed the project team's training sessions, the practice of keeping project phones under lock and key in senior colleagues offices, and the use of project phones as though they were individual-use phones by HWs who received these phones on behalf of the HFs, as documented in a study to understand how and why the intervention was used [36], could explain the low number of USSD requests by hospital shared-use phones. These observations are common health system challenges in low-resource settings that need to be addressed as not all HWs may attend the various training programs constantly organized for staff, and scarce resources have to be shared.

The setup of our intervention database is unique and allowed for in-depth analysis of requests made to the USSD by individual users unlike previous work [11]. Our study shows varying pattern of requests for emergency protocols across and within clusters, HFs and their location, *type-of-phone*, and *time-of-day* and type of request made at all 3 time points considered in this study. This reflects the dynamic nature of information needs of FHWs. Such dynamism has been reported [10,37,38] and is important to take into account in the design and maintenance of CDMSS [10,12,37,38] as well as training for FHWs.

The intervention phones were predominantly used for voice calls (64%), followed by data (28%), SMS text messaging (5%), and USSD to access protocols (2%), respectively [36]. Differential baseline technological literacy among FHWs may have impacted the use of the different intervention components [36]. The declined usage of the USSD over time can be explained by the so-called *novelty effect* associated with mhealth interventions [37,39-42]. Novelty effect is the tendency for performance to initially improve when new technology is instituted, not because of any actual improvement in learning or achievement, but in response to increased interest in the new technology [43]. However, the *novelty-effect* alone cannot be considered as the reason for the much lower number of requests made in the last 6 months of the study. Another possible explanation for this phenomenon may be *testing effect in learning* (that long-term memory is often increased when some of the learning period is devoted to retrieving the to-be-remembered information [44]). The much lower usage of the USSD in the last 6 months is most likely because of conversion of USSD protocols into tacit knowledge of FHWs [36]. Availability of specialist obstetricians, doctors, and senior midwives in hospitals and the need to conform to instructions from superior colleagues (eg, doctors) can also explain this finding, particularly in hospitals where the observed decline in requests was highest [36]. Conflict from overreliance on CDMSS [37] by users of the CDMSS and provider knowledge and experience from nonusers of the CDMSS may lead to abandonment of the CDMSS in resolving such conflicts with the mindset that critical thinking of the human mind must not

be taken over by a CDMSS [37]. Where there is a disconnect between protocols and the reality on the ground (such as lack of equipment), HW may also decide not to access electronic resources [10]. Technical and supervisory support to motivate users may also play a role in the decline in requests observed [12,36], and thus, this observation warrants further probing.

The moderate to strong correlation between the number of sepsis and asphyxia requests suggest that FHWs actually encountered these cases and used these protocols in their decision-making for these morbidities. The converse may be true where weak and negative associations are observed; exploration of the USSD protocols to satisfy FHW curiosity and mobile network problems [11,45,46] necessitating that FHWs send multiple requests may explain these weak and negative associations.

Limitations

Though our study highlights important patterns of use of a SMS text messaging–based CDMSS, the use of information accessed in the care of patients or clients is undetermined in this study. Though this limitation is inherent in the design of this study, this study provides much needed insights as to how an mHealth SMS text messaging–based CDMSS functioned in a low-resource setting and quantifies the information needs of FHWs providing maternal and neonatal health care in this type of setting. Insight into the information needs of FHWs can inform the design of interventions mHealth (or otherwise) to meet these needs.

Conclusions

This study demonstrates that health care providers of maternal and neonatal health services in Ghana readily use a mobile SMS text messaging–based CDMSS in their clinical decision-making. These FHWs used the mHealth tool to request emergency protocols depending on their information needs, which varied across and within clusters, HFs and their location, and with time. Thus, the information needs of HWs is not static but continues to change over time requiring health system strengthening strategies that take cognizance of this dynamism. Mechanisms to sustain utilization of similar mHealth CDMSS interventions must be designed to suit relevant context if such interventions will be up-scaled as health system strengthening strategies in future.

Acknowledgments

The authors would like to thank the Ghana Health Service Research and Development Division, the Eastern Region Health Directorate, members of the district health management teams in the Eastern Region, and the School of Public Health, University of Ghana for their support in the conduct of this study. They also thank Richard Kwofie, Theophilus Botchway, and Ebenezer Amankwah of Vodafone Ghana for their support in extracting the USSD data. They also thank the Netherlands Foundation for Scientific Research, Global Health Policy and Health Systems Research Program (Grant Number: 07.45.102.00), and the Julius Center, University Medical Center, Utrecht, Netherlands for financial support. HBA has been supported by a Global Health Scholarship Grant from the University Medical Center Utrecht.

Authors' Contributions

HBA, KK-G, and EKA designed and performed the study. HBA drafted the manuscript, and KK-G and EKA reviewed the manuscript. IAA, DEG, MA-C, CS, EF, GAK, and EO-M provided critical comments on the review of the manuscript. All authors have read and approved the final manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Steps to request for emergency maternal or neonatal protocols.

[PDF File (Adobe PDF File), 414KB - mhealth_v7i5e12879_app1.pdf]

References

1. Haines A, Victora C. Evidence-based action needed on health systems. Lancet 2004 Oct;364(9441):1204. [doi: 10.1016/S0140-6736(04)17160-2]
2. United Nations. New York, USA: United Nations; 2015. The Millennium Development Goals Report 2015 URL: http://www.un.org/millenniumgoals/2015_MDG_Report/pdf/MDG%202015%20rev%20%28July%201%29.pdf [accessed 2018-09-01] [WebCite Cache ID 725fATTvg]
3. Division for Sustainable Development UN-DESA. New York, USA: United Nations; 2016. Sustainable Development Knowledge Platform URL: https://sustainabledevelopment.un.org/sdg3 [accessed 2018-09-03] [WebCite Cache ID 729mV36k2]
4. Labrique AB, Vasudevan L, Kochi E, Fabricant R, Mehl G. mHealth innovations as health system strengthening tools: 12 common applications and a visual framework. Glob Health Sci Pract 2013 Aug;1(2):160-171 [FREE Full text] [doi: 10.9745/GHSP-D-13-00031] [Medline: 25276529]

5. Betjeman TJ, Soghoian SE, Foran MP. mHealth in Sub-Saharan Africa. Int J Telemed Appl 2013;2013:482324 [FREE Full text] [doi: 10.1155/2013/482324] [Medline: 24369460]
6. Amoakoh-Coleman M, Borgstein AB, Sondaal SF, Grobbee DE, Miltenburg AS, Verwijs M, et al. Effectiveness of mHealth interventions targeting health care workers to improve pregnancy outcomes in low- and middle-income countries: a systematic review. J Med Internet Res 2016 Aug 19;18(8):e226 [FREE Full text] [doi: 10.2196/jmir.5533] [Medline: 27543152]
7. Sondaal SFV, Browne JL, Amoakoh-Coleman M, Borgstein A, Miltenburg AS, Verwijs M, et al. Assessing the effect of mHealth interventions in improving maternal and neonatal care in low-and middle-income countries: a systematic review. PLoS One 2016;11(5):e0154664 [FREE Full text] [doi: 10.1371/journal.pone.0154664] [Medline: 27144393]
8. Hall CS, Fottrell E, Wilkinson S, Byass P. Assessing the impact of mHealth interventions in low- and middle-income countries--what has been shown to work? Glob Health Action 2014;7:25606 [FREE Full text] [Medline: 25361730]
9. Agarwal S, Perry HB, Long L, Labrique AB. Evidence on feasibility and effective use of mHealth strategies by frontline health workers in developing countries: systematic review. Trop Med Int Health 2015 Aug;20(8):1003-1014 [FREE Full text] [doi: 10.1111/tmi.12525] [Medline: 25881735]
10. Park E, Masupe T, Joseph J, Ho-Foster A, Chavez A, Jammalamadugu S, et al. Information needs of Botswana health care workers and perceptions of wikipedia. Int J Med Inform 2016 Dec;95:8-16. [doi: 10.1016/j.ijmedinf.2016.07.013] [Medline: 27697235]
11. Lemay NV, Sullivan T, Jumbe B, Perry CP. Reaching remote health workers in Malawi: baseline assessment of a pilot mHealth intervention. J Health Commun 2012;17 Suppl 1:105-117. [doi: 10.1080/10810730.2011.649106] [Medline: 22548604]
12. Adepoju IO, Albersen BJA, De Brouwere V, van Roosmalen J, Zweekhorst M. mHealth for clinical decision-making in sub-Saharan Africa: a scoping review. JMIR Mhealth Uhealth 2017 Mar 23;5(3):e38 [FREE Full text] [doi: 10.2196/mhealth.7185] [Medline: 28336504]
13. Oyeyemi SO, Wynn R. The use of cell phones and radio communication systems to reduce delays in getting help for pregnant women in low- and middle-income countries: a scoping review. Glob Health Action 2015 Sep 10;8(1):28887 [FREE Full text] [doi: 10.3402/gha.v8.28887] [Medline: 26362421]
14. Mahmud N, Rodriguez J, Nesbit J. A text message-based intervention to bridge the healthcare communication gap in the rural developing world. Technol Health Care 2010;18(2):137-144. [doi: 10.3233/THC-2010-0576] [Medline: 20495253]
15. The World Bank. Geneva, Switzerland: The World Bank Group; 2016. Maternal mortality ratio (modeled estimate, per 100,000 live births) | Data | Table URL: https://data.worldbank.org/indicator/SH.STA.MMRT?locations=GH [accessed 2018-09-03] [WebCite Cache ID 72A3Z3fmG]
16. Kayode GA, Ansah E, Agyepong IA, Amoakoh-Coleman M, Grobbee DE, Klipstein-Grobusch K. Individual and community determinants of neonatal mortality in Ghana: a multilevel analysis. BMC Pregnancy Childbirth 2014 May 12;14:165 [FREE Full text] [doi: 10.1186/1471-2393-14-165] [Medline: 24884759]
17. The DHS Program. Maryland, USA; 2015. Ghana Demographic and Health Survey 2014 URL: https://dhsprogram.com/pubs/pdf/fr307/fr307.pdf [accessed 2018-09-03] [WebCite Cache ID 72A4IArhu]
18. The DHS Program. Maryland, USA; 2009. Ghana Demographic and Health Survey 2008 URL: https://www.dhsprogram.com/pubs/pdf/FR221/FR221%5B13Aug2012%5D.pdf [accessed 2018-09-03] [WebCite Cache ID 72A4u6Pfj]
19. Ghana Statistical Service, Ghana Health Service, ICF. Ghana Statistical Services. 2018. Ghana Maternal Health Survey 2017 Key Indicators URL: http://www.statsghana.gov.gh/docfiles/PR95.pdf [accessed 2018-09-03] [WebCite Cache ID 72A5Lr6oO]
20. Owusu-Sarpong A, Boamah KA, Baiden F. African Journal of Reproductive Health. Benin City, Nigeria: Women's Health and Action Research Centre; 2017 Jun. Associated factors and quality of care received among maternal deaths at a Regional Hospital in Ghana : maternal death audit review URL: https://journals.co.za/content/journal/10520/EJC-90103bdb9 [accessed 2018-09-05] [WebCite Cache ID 72Cfj3uGs]
21. Amoakoh-Coleman M, Ansah EK, Agyepong IA, Grobbee DE, Kayode GA, Klipstein-Grobusch K. Predictors of skilled attendance at delivery among antenatal clinic attendants in Ghana: a cross-sectional study of population data. BMJ Open 2015 May 19;5(5):e007810 [FREE Full text] [doi: 10.1136/bmjopen-2015-007810] [Medline: 25991459]
22. Mills S, Bertrand JT. Use of health professionals for obstetric care in northern Ghana. Stud Fam Plann 2005 Mar;36(1):45-56. [Medline: 15828524]
23. Issah K, Nang-Beifubah A, Opoku CF. Maternal and neonatal survival and mortality in the Upper West Region of Ghana. Int J Gynaecol Obstet 2011 Jun;113(3):208-210. [doi: 10.1016/j.ijgo.2011.01.007] [Medline: 21457982]
24. Howe LD, Manu A, Tawiah-Agyemang C, Kirkwood BR, Hill Z. Developing a community-based neonatal care intervention: a health facility assessment to inform intervention design. Paediatr Perinat Epidemiol 2011 Mar;25(2):192-200. [doi: 10.1111/j.1365-3016.2010.01178.x] [Medline: 21281331]
25. Amoakoh HB, Klipstein-Grobusch K, Amoakoh-Coleman M, Agyepong IA, Kayode GA, Sarpong C, et al. The effect of a clinical decision-making mHealth support system on maternal and neonatal mortality and morbidity in Ghana: study protocol for a cluster randomized controlled trial. Trials 2017 Dec 4;18(1):157 [FREE Full text] [doi: 10.1186/s13063-017-1897-4] [Medline: 28372580]

26. Oduro-Mensah E, Kwamie A, Antwi E, Amissah BS, Bainson HM, Marfo B, et al. Care decision making of frontline providers of maternal and newborn health services in the greater Accra region of Ghana. PLoS One 2013;8(2):e55610 [FREE Full text] [doi: 10.1371/journal.pone.0055610] [Medline: 23418446]
27. Family Health Division, Ghana Health Service. Accra, Ghana: Ghana Health Service; 2016. 2015 Annual Report URL: https://www.ghanahealthservice.org/downloads/2015_FAMILY_HEALTH_DIVISION_ANNUAL_REPORT.pdf [accessed 2019-04-04] [WebCite Cache ID 77Nh5vGi6]
28. Little A, Medhanyie A, Yebyo H, Spigt M, Dinant G, Blanco R. Meeting community health worker needs for maternal health care service delivery using appropriate mobile technologies in Ethiopia. PLoS One 2013;8(10):e77563 [FREE Full text] [doi: 10.1371/journal.pone.0077563] [Medline: 24204872]
29. DeSouza SI, Rashmi MR, Vasanthi AP, Joseph SM, Rodrigues R. Mobile phones: the next step towards healthcare delivery in rural India? PLoS One 2014;9(8):e104895 [FREE Full text] [doi: 10.1371/journal.pone.0104895] [Medline: 25133610]
30. Janssen KJ, Donders AR, Harrell FE, Vergouwe Y, Chen Q, Grobbee DE, et al. Missing covariate data in medical research: to impute is better than to ignore. J Clin Epidemiol 2010 Jul;63(7):721-727. [doi: 10.1016/j.jclinepi.2009.12.008] [Medline: 20338724]
31. Acharya D, Paudel R, Gautam K, Gautam S, Upadhyaya T. Knowledge of maternal and newborn care among primary level health workers in Kapilvastu district of Nepal. Ann Med Health Sci Res 2016;6(1):27-32 [FREE Full text] [doi: 10.4103/2141-9248.180266] [Medline: 27144073]
32. Say L, Chou D, Gemmill A, Tunçalp Ö, Moller A, Daniels J, et al. Global causes of maternal death: a WHO systematic analysis. Lancet Glob Health 2014 Jun;2(6):e323-e333 [FREE Full text] [doi: 10.1016/S2214-109X(14)70227-X] [Medline: 25103301]
33. Bauserman M, Lokangaka A, Thorsten V, Tshefu A, Goudar SS, Esamai F, et al. Risk factors for maternal death and trends in maternal mortality in low- and middle-income countries: a prospective longitudinal cohort analysis. Reprod Health 2015;12 Suppl 2:S5 [FREE Full text] [doi: 10.1186/1742-4755-12-S2-S5] [Medline: 26062992]
34. Der EM, Moyer C, Gyasi RK, Akosa AB, Tettey Y, Akakpo PK, et al. Pregnancy related causes of deaths in Ghana: a 5-year retrospective study. Ghana Med J 2013 Dec;47(4):158-163 [FREE Full text] [Medline: 24669020]
35. D'Adamo M, Short Fabic M, Ohkubo S. Meeting the health information needs of health workers: what have we learned? J Health Commun 2012;17 Suppl 2:23-29. [doi: 10.1080/10810730.2012.666626] [Medline: 22724669]
36. Amoakoh HB, Klipstein-Grobusch K, Ansah EK, Grobbee DE, Yveoo L, Agyepong I. How and why front-line health workers (did not) use a multifaceted mHealth intervention to support maternal and neonatal healthcare decision-making in Ghana. BMJ Glob Health 2019 Mar 8;4(2):e001153 [FREE Full text] [doi: 10.1136/bmjgh-2018-001153] [Medline: 30997162]
37. Castillo RS, Kelemen A. Considerations for a successful clinical decision support system. Comput Inform Nurs 2013 Jul;31(7):319-26; quiz 327. [doi: 10.1097/NXN.0b013e3182997a9c] [Medline: 23774450]
38. Pakenham-Walsh N, Bukachi F. Information needs of health care workers in developing countries: a literature review with a focus on Africa. Hum Resour Health 2009 Apr 8;7:30 [FREE Full text] [doi: 10.1186/1478-4491-7-30] [Medline: 19356239]
39. Taki S, Lymer S, Russell CG, Campbell K, Laws R, Ong K, et al. Assessing user engagement of an mHealth intervention: development and implementation of the growing healthy app engagement index. JMIR Mhealth Uhealth 2017 Jun 29;5(6):e89 [FREE Full text] [doi: 10.2196/mhealth.7236] [Medline: 28663164]
40. Guertler D, Vandelanotte C, Kirwan M, Duncan MJ. Engagement and nonusage attrition with a free physical activity promotion program: the case of 10,000 steps Australia. J Med Internet Res 2015;17(7):e176 [FREE Full text] [doi: 10.2196/jmir.4339] [Medline: 26180040]
41. 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]
42. Neve MJ, Collins CE, Morgan PJ. Dropout, nonusage attrition, and pretreatment predictors of nonusage attrition in a commercial Web-based weight loss program. J Med Internet Res 2010;12(4):e69 [FREE Full text] [doi: 10.2196/jmir.1640] [Medline: 21156470]
43. Wikipedia. San Francisco, USA: Wikipedia; 2019. Novelth effect URL: https://en.wikipedia.org/wiki/Novelty_effect [accessed 2019-04-01] [WebCite Cache ID 77JdlcPxN]
44. Wikipedia. San Francisco, USA: Wikipedia; 2019. Testing effect URL: https://en.wikipedia.org/wiki/Testing_effect [accessed 2019-04-02] [WebCite Cache ID 77KRLa67C]
45. Jain N, Singh H, Koolwal GD, Kumar S, Gupta A. Opportunities and barriers in service delivery through mobile phones (mHealth) for severe mental illnesses in Rajasthan, India: a multi-site study. Asian J Psychiatr 2015 Apr;14:31-35. [doi: 10.1016/j.ajp.2015.01.008] [Medline: 25701069]
46. Vedanthan R, Blank E, Tuikong N, Kamano J, Misoi L, Tulienge D, et al. Usability and feasibility of a tablet-based Decision-Support and Integrated Record-keeping (DESIRE) tool in the nurse management of hypertension in rural western Kenya. Int J Med Inform 2015 Mar;84(3):207-219 [FREE Full text] [doi: 10.1016/j.ijmedinf.2014.12.005] [Medline: 25612791]

Abbreviations

APH: antepartum hemorrhage
CDMSS: clinical decision-making support system
CHPS: Community-based Health Planning and Services
CRCT: cluster randomized controlled trial
DHIMS2: District Health Information Management System 2
FHWs: frontline health workers
HCs: health centers
HDP: hypertensive disorders in pregnancy
HFs: health facilities
HWs: health workers
mCDMSS: mHealth clinical decision-making support system
MDGs: Millennium Development Goals
mHealth: mobile health
MNCH: maternal, neonatal, and child health
PPH: postpartum hemorrhage
SDGs: Sustainable Development Goals
SMS: short message service
USSD: unstructured supplementary service data

Edited by G Eysenbach; submitted 21.11.18; peer-reviewed by K Ndlovu, O Ilozumba, E Green; comments to author 31.03.19; revised version received 04.04.19; accepted 04.04.19; published 24.05.19.

Please cite as:
Amoakoh HB, Klipstein-Grobusch K, Grobbee DE, Amoakoh-Coleman M, Oduro-Mensah E, Sarpong C, Frimpong E, Kayode GA, Agyepong IA, Ansah EK
Using Mobile Health to Support Clinical Decision-Making to Improve Maternal and Neonatal Health Outcomes in Ghana: Insights of Frontline Health Worker Information Needs
JMIR Mhealth Uhealth 2019;7(5):e12879
URL: http://mhealth.jmir.org/2019/5/e12879/
doi:10.2196/12879
PMID:31127719

Original Paper

Clinicians' Role in the Adoption of an Oncology Decision Support App in Europe and Its Implications for Organizational Practices: Qualitative Case Study

Christine Jacob[1,2], BSc, MSc; Antonio Sanchez-Vazquez[3], PhD; Chris Ivory[3], PhD

[1]Anglia Ruskin University, Cambridge, United Kingdom

[2]University of Applied Sciences Northwestern Switzerland, Brugg, Switzerland

[3]Innovation and Management Practice Research Centre, Anglia Ruskin University, Cambridge, United Kingdom

Corresponding Author:
Christine Jacob, BSc, MSc
Anglia Ruskin University
Cambridge Campus, East Rd, Cambridge CB1 1PT
Cambridge,
United Kingdom
Phone: 41 798769376
Email: christine.k.jacob@gmail.com

Abstract

Background: Despite the existence of adequate technological infrastructure and clearer policies, there are situations where users, mainly physicians, resist mobile health (mHealth) solutions. This is of particular concern, bearing in mind that several studies, both in developed and developing countries, showed that clinicians' adoption is the most influential factor in such solutions' success.

Objective: The aim of this study was to focus on understanding clinicians' roles in the adoption of an oncology decision support app, the factors impacting this adoption, and its implications for organizational and social practices.

Methods: A qualitative case study of a decision support app in oncology, called ONCOassist, was conducted. The data were collected through 17 in-depth interviews with clinicians and nurses in the United Kingdom, Ireland, France, Italy, Spain, and Portugal.

Results: This case demonstrates the affordances and constraints of mHealth technology at the workplace, its implications for the organization of work, and clinicians' role in its constant development and adoption. The research findings confirmed that factors such as app operation and stability, ease of use, usefulness, cost, and portability play a major role in the adoption decision; however, other social factors such as endorsement, neutrality of the content, attitude toward technology, existing workload, and internal organizational politics are also reported as key determinants of clinicians' adoption. Interoperability and cultural views of mobile usage at work are the key workflow disadvantages, whereas higher efficiency and performance, sharpened practice, and location flexibility are the main workflow advantages.

Conclusions: Several organizational implications emerged, suggesting the need for some actions such as fostering a work culture that embraces new technologies and the creation of new digital roles for clinicians both on the hospitals or clinics and on the development sides but also more collaboration between health care organizations and digital health providers to enable electronic medical record integration and solving of any interoperability issues. From a theoretical perspective, we also suggest the addition of a fourth step to Leonardi's methodological guidance that accounts for user engagement; embedding the users in the continuous design and development processes ensures the understanding of user-specific affordances that can then be made more obvious to other users and increase the potential of such tools to go beyond their technological features and have a higher impact on workflow and the organizing process.

(JMIR Mhealth Uhealth 2019;7(5):e13555) doi:10.2196/13555

KEYWORDS

telemedicine; smartphone; cell phone; oncologists; electronic health record; workflow; workload; workplace; public health practice; technology; perception; health education; mHealth; mobile health; telehealth; eHealth

Introduction

Background

Globally, health care sectors are facing persistent challenges including increasing costs, inconsistent patient care, and a growing burden of chronic disease. One way of overcoming these challenges that have become increasingly the focus of policy change is the transformation of care through a patient-centric service design and taking a more proactive and preventive approach that focuses on quality of life and not only on treating the disease [1].

Smartphones can and do play a strong role here. The latest developments in mobile technology have allowed smartphones to achieve increasingly sophisticated tasks [2], so much so that a new area of health, mobile health (mHealth), has emerged [3].

The global observatory of electronic health in the World Health Organization defines mHealth as "medical and public health practice supported by mobile devices, such as mobile phones, patient monitoring devices, personal digital assistants (PDAs), and other wireless devices" [4]. mHealth solutions differ from other information and communication technology applications in the sense that they are typically user-driven, accessible, and affordable [5]; therefore, a good understanding of the factors impacting user adoption and the roles they play in such adoption are key to the success of such solutions.

However, despite the existence of adequate technological infrastructure and clearer policies, there are situations where users, mainly physicians, resist mHealth solutions [6]. This is particularly relevant, bearing in mind that several studies, both in developed and developing countries, showed that clinicians' adoption is the most influential factor in such solutions' success [7-10]. Remarkably, studies show that social aspects are the major factor behind the nonadoption of new technologies, mainly owing to their complexity and the fact that users tend to prefer existing systems over newly introduced ones [11,12]. Furthermore, clinicians' attitude toward risk is a crucial element in the successful implementation of new health care technologies [13].

Objectives

Therefore, this research focuses on understanding clinicians' roles in the adoption of mHealth solutions and its implications for clinicians' practice. The topic is investigated through the following subquestions:

- What are clinicians' roles in the adoption of mHealth solutions?
- What are the factors that constrain or afford clinicians' adoption of mHealth?
- Does the use of mHealth impact social practices in this context? How?

These questions are addressed through a case study of an oncology decision support mHealth solution called ONCOassist and its adoption by clinicians across Europe.

Created in 2012, ONCOassist, shown in (Multimedia Appendix 1), is a free decision support app for oncology clinicians, with over 11,000 active users worldwide in 2019 at the time of writing this paper. The app gives oncologists and oncology nurses easy access to a range of features that aim to help them save time and improve the quality of patient care. The app's key features are explained in Multimedia Appendix 2.

User research, as a mode of inquiry, has shifted from the assumption that there is a stable material thing (a technology) and separate and identifiable *user preferences* in relation to it toward an understanding in which use and technology are seen as bound together within one another and are constitutive of one another [14-16]. From this ontological perspective, it is challenging to explain *user choices* within existing modes of inquiry [17-20]. It follows that new and more balanced user-research methodologies are needed if questions about user choices and preferences, with regard to new technology, are to be asked.

Contemporary sociomaterial scholars explain that, from their perspective, the social and material combine over time (imbricate) to yield stable sociomaterial constructions—such as stable patterns of use and new organizational structures [21]. Technology partly constrains these constructions, without wholly determining them, as users interact with certain preferred features of technology to create particular *technology affordances* [17,18,22,23]; that is to say that technology is defined not just by its functions and its materiality but by how precisely it is constructed through use. At the same time, use cannot be seen as distinct from those functions or the materiality of the technology in question [22]. Technology *affordance*, in other words, lies somewhere between the user and technology.

Understanding the affordances resulting from this sociomaterial interaction is also key in accounting for how technology impacts on organizations and the process of organizing, as people's interactions with objects influences the way they organize their microlevel relations [24] and also impacts on the definition of people's roles [25,26]. Moreover, as organizations become increasingly digital, scholars have portrayed how the technologies used at the workplace are central to the enablement of new forms of organizing and practice and how, in turn, new practices are essential to the enablement of new technologies and new organizational forms [27,28].

Therefore, to better understand the role of clinicians (the users) in the adoption of mHealth (new technologies) and the implications for health care organizations (organizing), the research themes and questions were developed in line with Leonardi's *Methodological Guidelines for the Study of Materiality and Affordances* to crystallize the focus of the data collected in the interviews [22].

This study used Leonardi's guidance to build a solid analysis of the role of users in technology adoption and its impact on organizations following these 3 main steps:

- Understanding and documenting the material aspects of technology and their limitations by identifying utility and limitations of the app as perceived by the participating clinicians.

- Linking the material aspects of technology to the tasks that they enable and facilitate by recognizing the real constraints upon opportunities faced by clinicians when using the app.
- Recognizing the processes resulting from these affordances and determining the consequential interactions taking place in the organization by understanding the workflow advantages and disadvantages related to the app's usage.

The following section explains the research method and how the interview questions and subsequently the analysis stemmed from these 3 steps.

Methods

A qualitative paradigm was adopted because it gives precedence to *the voices of participants* and the individual and unique *reflexivity of the researcher* [29] and because of the rich insights it offers, which can help draw out and understand clinicians' experiences and perceptions in different ways that quantitative methods cannot [30,31].

Data Collection

The data were collected via in-depth semistructured interviews. Given that participants were in many different locations across Europe, not all interviews were held face-to-face, some of them were conducted via Skype, Google Hangouts, or telephone conferencing. Furthermore, physical artifacts such as screenshots of the app, the devices it can be used on, and example written feedback to the developers (eg, app reviews on the app store) were collected to develop a broader perspective about the solution subject of the study [32]. The data collection took place from February 2018 to January 2019, and a total of 17 interviews were conducted with 13 participants (4 were follow-up interviews with the same participants to clarify some details or to ask more questions that were necessary for the analysis). The interviews lasted between 23 min and 110 min and were all conducted and recorded by the first author (CJ) in English. Data collection stopped when an acceptable level of saturation was reached.

Sampling Techniques and Participants' Profiles

Purposive sampling was used, where potential participants were selected based on their ability to provide rich and in-depth information about the research topic [29,30]. Key informants in the participating company were contacted, and a snowballing sampling was used to identify suitable participants in their solutions' user base. The main selection criteria were that participants must be active clinicians or former users of the app being studied. To avoid the possible selection bias that might result from the key informants selectively choosing users with a positive inclination toward the studied solution, it was agreed that participants would be asked if they can in turn refer to other colleagues who were using the solution and were willing to participate. Some participants did refer other colleagues, who were using the app, and some agreed to participate in the study. However, unfortunately, none of the users who had discontinued use of the app agreed to take part in the research when they were asked.

The participants were working in hospitals and clinics in the United Kingdom, Ireland, France, Italy, Spain, and Portugal; and the sample was composed of 8 oncologists (one of them was also an ONCOassist cofounder), 3 nurses, and 2 other members from the ONCOassist team as shown in Table 1.

Data Analysis and Ethical Considerations

Thematic analysis was used to identify and extract themes that addressed the research questions and explained what each theme could mean as well as the links between themes [33]. QSR's NVivo was used for coding and then excerpts were selected to create an account that tells the narrative of each theme in a way that helps the reader to understand the analysis. The first author (CJ) conducted the interviews and did the initial analysis and coding; she is a digital strategist with more than 17 years of experience and has contributed to the creation and realization of several digital solutions in health care. Then, the second author (ASV) reviewed the coding; any cases of disagreement about coding were discussed in conjunction with the last author (CI) and mutually agreed. The thematic map is represented in Multimedia Appendix 3, and the phases of the thematic analysis are clarified in Multimedia Appendix 4.

Table 1. Sample demographics and characteristics (N=13).

Sample Characteristics	Composition
Function (%)	
8 oncologists (one of them is also an ONCOassist cofounder)	61.5
3 nurses	23
2 other ONCOassist cofounders	15.5
Gender (%)	
5 females	38.5
8 males	61.5
Tech awareness (Participants were asked to define their tech awareness on a scale of 1 to 10), mean	7.25
Health care experience (years), mean	14.5
Mobile health experience (years), mean	7
Location	United Kingdom, Ireland, France, Italy, Spain, and Portugal

The infographic creator, Venngage, was used to visualize some of the findings and not to quantify the data but to emphasize the frequency by showing visually which themes were brought up by more participants than others. The frequency reflected in the infographics counts the theme only once per participant and does not accumulate if the same participant brought up the same theme several times. Such visualization mainly aims to improve the comprehension of the article especially when contrasting 2 elements such as workflow advantages and disadvantages [34] and can provide a clear and simple illustration of the dominant themes and ideas for lay readers [35].

Ethical approval was obtained from the Faculty Research Ethics Panel under the terms of Anglia Ruskin University's Research Ethics Policy, and all participants were briefed about the research context and signed a study consent form agreeing to participate.

Interview Themes and Questions

The interview questions addressing each of the 3 following themes are reflected in the interview topic guides available in Multimedia Appendix 5. The first interview guide was used for clinicians and the second was used for providers.

Theme 1: Accounting for Materials

This theme evolves around understanding the limitations of technology and the types of uses that it enables. Understanding the material aspects of mHealth solutions is crucial because it allows us to identify the different ways it can be used as well as things that cannot be done with it owing to material limitations.

According to Leonardi, technological features "can have various degrees of utility based on the forms into which they are cast" [22]. He explains that a good understanding of the technological features of the solution, recognizing what it can do versus what it cannot do, should help avoid the misconception that users can achieve unlimited tasks with the technologies that they use in their daily work. Therefore, understanding not only the opportunities but also the limitations of mHealth solutions is crucial as a first step of the analysis.

The way a technological tool is built matters because based on that some uses might be very difficult or impossible to achieve, the same way that materials often resist scientists' efforts to control them, implying that materials have agency [36]. Nevertheless, the user's ability to rearrange or reshape the materials impacts the way they are used [37].

Therefore, researchers need to understand the materials that form a specific technological tool, how they are organized into specific features, and what such features do or not do. Answering these questions is key to understanding any potential limitations to the use of a specific technology at work; therefore, recognizing the real constraints upon opportunities is the first step toward the rationalization of the role of materiality [22].

Theme 2: Accounting for Materiality

The focus of this theme is to understand users' perceptions of technology and how they intend to use it. Leonardi clarifies that people's views of technology can influence the way they utilize it in their everyday practice; he explains that claiming that a certain technology has materiality means that its materials are being *entangled* or *imbricated* with users' experiences and culture in forms that make it hard to define the technological tool separately from its context of usage [22]. This understanding acknowledges that users' intentions and the goals that they want to achieve when using a specific technology have an impact on its affordances.

Several scholars believe that materiality is mainly created through observations of affordance and constraint [18,38,39]; such observations can explain how materiality arises at the junction between technology and its users [40,41]. Some scholars suggest that affordances happen when the existing material properties of a specific technology are given a meaning based on its users' behavior and that people will not interact with a technology unless they already recognize its utility [42]; this would entail that one technology can result in different outcomes based on the understanding that materiality can offer various affordances [22]. Conversely, other scholars suggest that affordances are created by design and *waiting to be perceived*, are not altered across the diverse contexts of use, and are not created by the users but rather by the designers, and it is up to the users to discover them [43,44].

Hutchby takes the middle ground explaining that affordances go beyond users and technology's properties and are established based on the kind of relationship formed between people and the technology that they use; accordingly, affordances of a specific solution can change from one context to another even when its material features remain unchanged [45]. This is because users usually have different objectives when approaching materiality, so they distinguish different uses that a specific technology can afford [36]. This relational nature of affordances also means that technology can have various uses, which can result in various ways in how work is organized [46,47].

Therefore, researchers need to understand what social organizations shape the users' objectives, how these objectives impact users' understandings of what a specific technology can or cannot do, and what makes users perceive different constraints or affordances based on the options obtainable by the material aspects of the technology itself [22].

Theme 3: Accounting for Materialization

Leonardi explains that once the researcher understands technology limitations and users' intentions for its use and how this impacts the affordances, it is important to expand the analysis to comprehend the influence of technology on organizing. This evolves around analyzing and realizing the instances when specific affordances influence and transform the actions, hierarchies, and relationships that constitute the organizing process [22].

It is important to remember that not every technology will clearly impact the organizing process, it is the creation and enablement of specific affordances that helps materials and materiality to materialize in a way that actually impacts organizations and how work is being organized [22]. This idea is emphasized in Leonardi's example of how a group of

technicians who were given a new Information Technology (IT) system started by using it for its initially defined feature of assigning jobs to others, and as they realized its affordances, their use developed to use it for documenting completed jobs, which further evolved to use this past documentation to define who is the best technician to be assigned a new job based on previous experiences. Only then did the new system start to materialize as a crucial element in organizing and optimizing work in that department [48].

Therefore, researchers need to understand how the current patterns of organizing rely on the materiality of specific technologies, why some organizing processes create a social context in which technology can materialize in actions and interactions' flows, and how have the affordances enabled by technology supported, changed, or transformed the way that people work or interact in a specific organization [22].

Results

Accounting for the Materials: Utility and Limitations

Participants were first asked to name the features that they use most in the app to better understand the technological artifacts that they find most useful. They named the features in Table 2.

They were then asked to explain how the app helped them and their patients on a daily basis to better understand ONCOassist's utility from their perspective; 3 key themes emerged as visualized in Figure 1. The size of the circles reflects the frequency of the themes; blue represents utility elements and gray, limitations.

All participants mentioned that the app helps them to save time and be more efficient in their daily work. The app also allows a more seamless clinic experience for patients because it enables clinicians to make critical decisions at the point of care, also enhancing patient safety and quality of care. The compact overview of numerous tools and the possibility to switch between the Web application and the mobile app were also highly valued. Sample participants' quotes are reflected in Multimedia Appendix 6.

To complete the picture, the participants were also asked about any limitations they faced when using mHealth apps. And as most participants did not recall many limitations in the ONCOassist app, they were also asked to reflect on limitations associated with other mHealth solutions they had experienced. The limitations are visualized in Figure 2 and explained below.

Factors such as information incompleteness or incorrectness, for example, outdated information, can be perceived as limitations. Apps where the design is too *cluttered* and users cannot find what they are looking for are usually abandoned. This was not the case with ONCOassist, but some participants mentioned it as the reason why they stopped using other apps.

Furthermore, interoperability and electronic medical record (EMR) integration are the key system-related limitations that most mHealth apps are currently facing. Shortage of resources can be an issue, financing is a concern, especially in Europe, as most hospitals are publicly funded. Again, this was not the case with ONCOassist, which is free, but some participants mentioned it as a general barrier with other paid mHealth solutions. Sample participants' quotes are also reflected in Multimedia Appendix 6.

To have a deeper understanding of perceived limitations in the current features, the participants were also asked to suggest new features that they would like to add to the app. They requested a link to guidelines, adding drug interactions, immunotherapy toxicity, access to new research on different protocols and regimes, adding a prostate-specific antigen (PSA) doubling calculator, a feature for patient monitoring and management, the possibility to use big data for predictive models, breast cancer protocols, clinical trial matching, enabling a community aspect for clinician discussions, local drug prices and reimbursement information, geriatric tool improvement, interpersonalization based on subspecialty, Systemic Anti-Cancer Treatment (SACT) datasets, and product characteristics.

Table 2. Most used features.

Feature	Sample quote
Toxicity criteria	"I find it very easy for doses of toxicity… it's very practical because it's all in one" (P3)
Calculators	"It's really useful for when I have to calculate carboplatin dosage using the right formula. It allows me to do it all in the app as opposed to the really, really long calculation that you normally have to make" (P4)
Staging tools	"The main thing that I used the most was the AJCC because it has the super summarized staging" (P7)
Adjuvant tools	"I use ONCOassist for prognostic values in choosing adjuvant chemotherapies" (P6)
Product characteristics	"It's useful … especially with the new treatments. They're evolving all the time" (P4)
Customer service	"And the customer service specifically with ONCOassist is very good. I can directly ask, even if I've got any issues and I need help with that" (P8)
Drug interaction	"And now the drug interaction checker, which is very, very, very important for us" (P6)
Offline functionality	"A lot of hospitals, the way they are, for some reason the signal is never good. And therefore, that also has a bearing on how these apps work" (P8)

Figure 1. Workflow Improvements and Disadvantages.

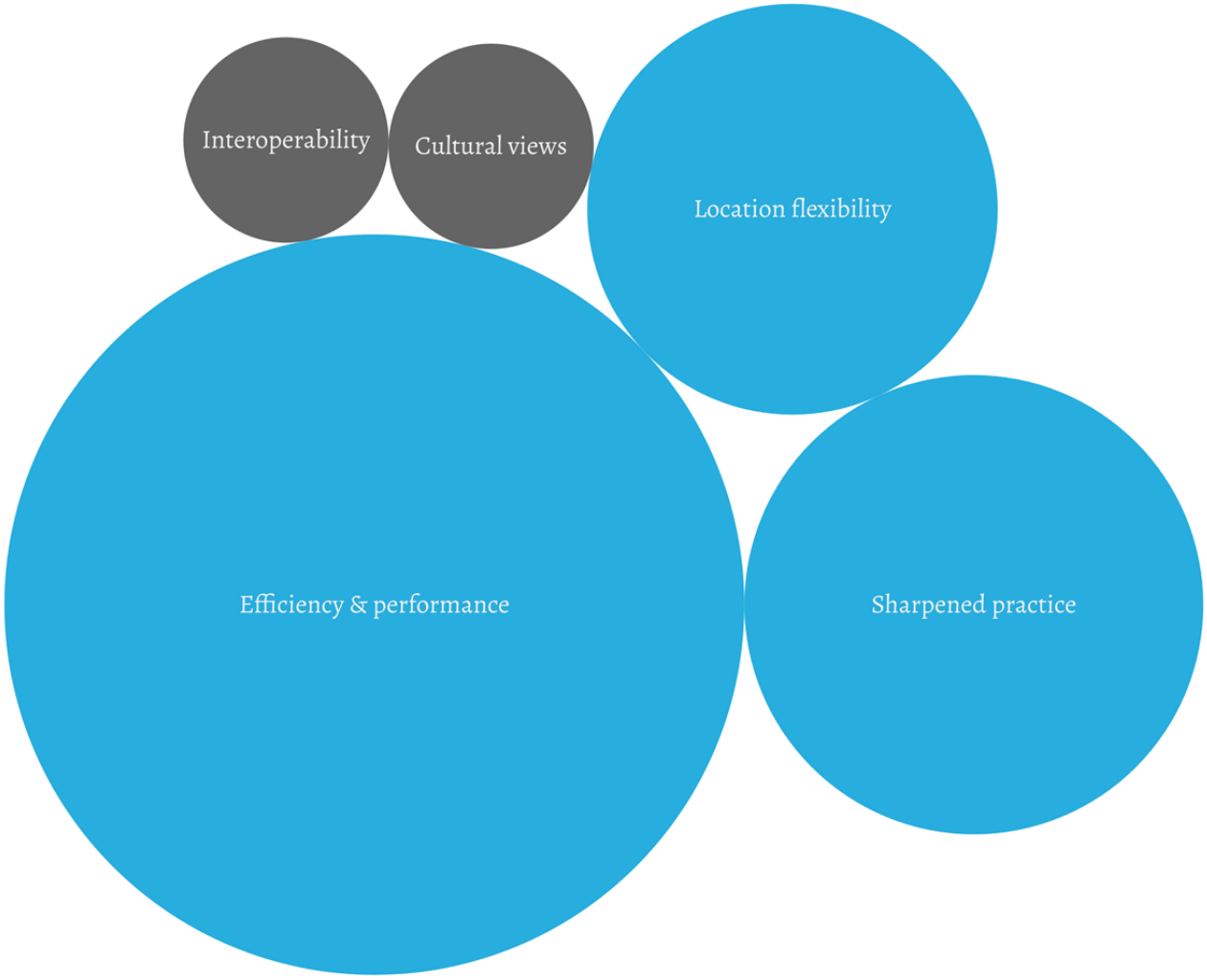
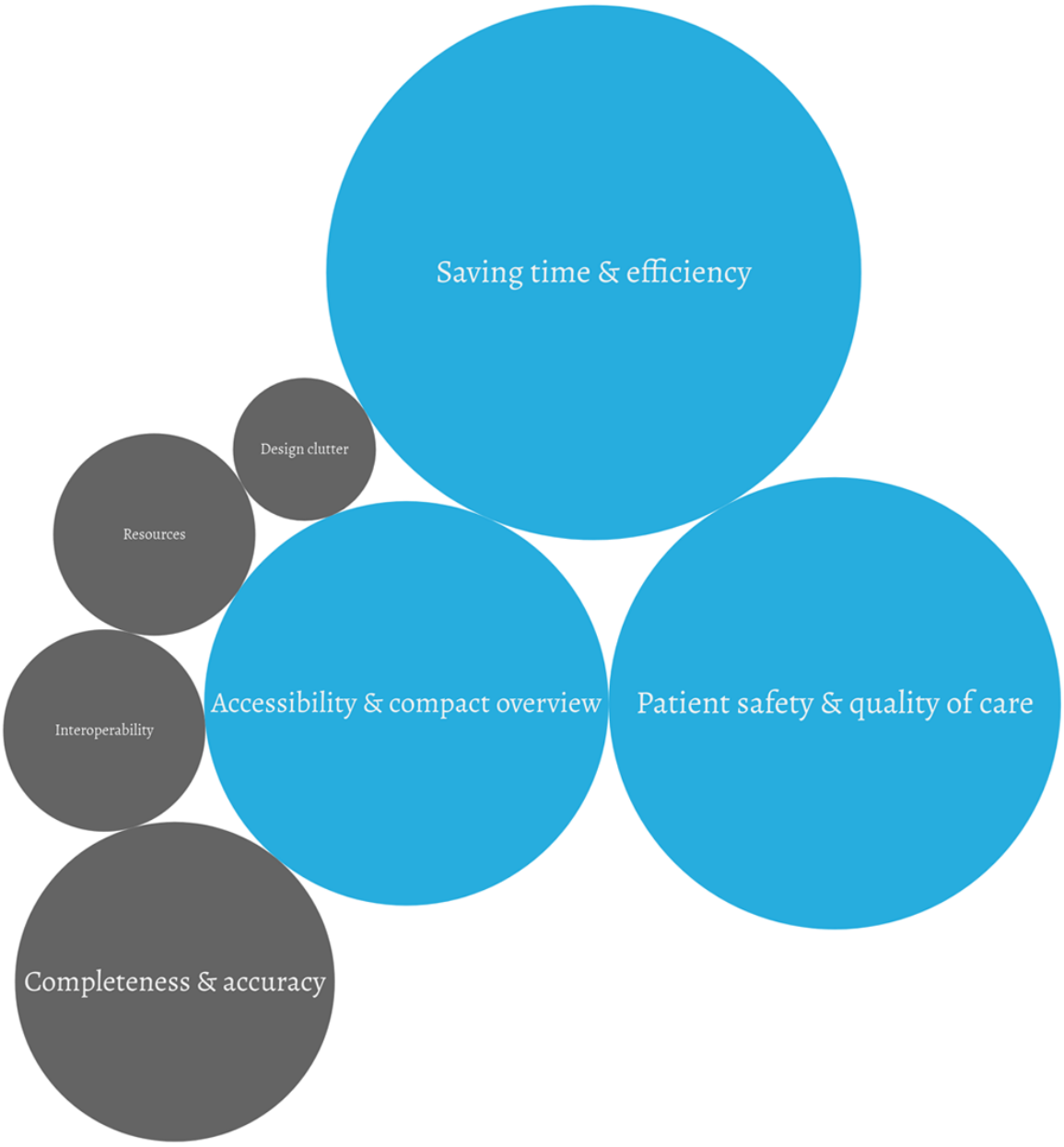


Figure 2. Utility and Limitations.



In a follow-up interview, the ONCOassist team explained that features such as the product characteristics and new research on different protocols and regimes already exist in the app but probably the participants who mentioned them were not aware of it. Furthermore, the features drug interactions and PSA doubling calculator were about to be launched at the time of writing of this paper just in time to meet the participants' requests. The patient monitoring and management feature is also being piloted at the time of writing this paper.

The team explained that other requested features such as drug prices and reimbursement information can be quite challenging to implement because they can vary drastically from country to country and they are subject to frequent change, which requires a large team that is constantly working on updating this information to guarantee it is correct.

Accounting for Materiality: Constraints and Affordances

To understand affordances, it is important to first start by understanding the intention for use. Therefore, the participants were asked to explain why they decided to use the app in the first place and what they intended to achieve with it.

Most participants decided to use the app because it could help them save time and be more efficient as the information that they need on a daily basis is aggregated reliably in one place; it also increases efficiency by offering a standardized and reliable tool. Recommendations from another oncologist, a medical society, a conference, or even social media channels such as LinkedIn motivated them to try the app; furthermore, usefulness and the constant development and update make the information on the app more accurate than other sources and can improve the quality of care by aiding clinical decisions.

For a more comprehensive analysis, it is also important to consider the reasons for nonadoption; however, it was not possible to recruit nonusers, presumably because they were not interested in the topic. To address this gap, the participants were asked about what they thought could be the reasons why other colleagues might not adopt mHealth solutions. The main reasons given were due to social and cultural factors such as attitude toward technology, age, attitude toward change, previous negative experiences, and perceptions of the use of a personal device at work; furthermore, the nature of the role can also be a factor; depending on how complex the role is, the need for apps may differ. Also, there is a greater awareness of apps that are advertised more and, accordingly, use is more widespread.

The interview questions evolved to address the participants' more holistic views of the factors impacting the adoption of mHealth from their perspective, going beyond ONCOassist to have a general understanding of barriers and opportunities for mHealth adoption. The main themes are reflected in Multimedia Appendix 7 with sample participants' quotes and a visual to emphasize their frequency.

App operation and stability including factors such as app size, offline functionality and reliability on connectivity, login issues, and the type of device used may be perceived as constraints or affordances to the use of such apps. Almost all participants mentioned cost as a key factor that could hinder the adoption of such apps, especially for younger clinicians. Most participants mentioned ease of use as a decisive factor for the adoption.

Usefulness is also very important, that the app addresses a need for clinicians to adopt it, purely informative solutions are not perceived as value adding in the form of an app. Endorsement and recommendations from people or institutions that they can trust are an important factor for adoption. Reliability and neutrality of the content, as well as things such as regulatory approval, are decisive factors for the adoption.

Social and cultural factors that impact people's attitudes toward technology, such as digital savviness or perceptions of the use of a personal device at work, can also play a considerable role in the adoption. Work overload and workplace politics can also play a role in the adoption, depending on what each clinician is focusing on and prioritizing. Finally, the portability and location flexibility of the app as a pocket solution is a significant affordance for adoption.

Accounting for Materialization: Social and Organizational Implications

The participants acknowledged several social and organizational impacts of the app usage as visualized in Multimedia Appendix 8 to emphasize their frequency, along with participants' sample quotes.

The app helps users feel more empowered and autonomous. The progress of Health Tech in general is fostering the creation of new digital roles for clinicians and their involvement in digital training and raising awareness about mHealth among their colleagues. The decision to use one mHealth solution or the other lies with the individual clinicians themselves and not centralized through the administration.

Owing to the need for speed, location flexibility, and the limited number of computers available at the workplace, clinicians prefer to reach out to their personal mobile for quick access rather than using the computers at work. Participants also noted that new health technologies, including mHealth, are replacing traditional tools in health care institutions. Also, if systematically embraced by the organization, such an app could be used as a tool for people evaluation and reward.

When asked about workflow improvements, the participants brought up 3 key themes as visualized in Figure 1 and detailed below. The size of the circles reflects the frequency of the themes; the blue color represents workflow advantages, and the gray color represents workflow disadvantages. Sample participants' quotes are reflected in Multimedia Appendix 9.

The app made their work easier, helped them be more efficient, looked more professional, and meant less interruption. It also sharpened their practice by giving them a backup and a safety net. It offered them location flexibility by helping them overcome computer shortage as well as decision making at the point of care or in meetings; it can also be an alternative in places where a computer is not allowed.

Participants also mentioned 2 workflow disadvantages that should be considered, as shown in Figure 1.

There is still an interoperability issue, as the majority of mHealth apps are not integrated into EMRs and therefore could easily be dismissed in favor of bigger vendors who offer such an integration. Cultural views on the usage of mobile phones at work and the perception that it is *wasting time* can impact mHealth users negatively in the workplace.

Future Vision and Clinicians' Role in Adoption

Participants agreed that the future of mHealth is promising and will have a considerable impact on health care. In total, 3 main themes emerged in the discussion about their views of the future of mHealth. They expect a rise in the adoption of mHealth solutions, especially with the standardization of EMRs, and such solutions can then be integrated or plugged into the system. They also expect to be able to benefit from artificial intelligence and all the big data that will result from the implementation of EMRs more widely; such data can be fed through mHealth, for example, to create predictive models and proactive treatment. This could also enable a more patient-centric approach to health care.

They all also agreed that clinicians would play a central role not only in the adoption of mHealth solutions but also their development as shown in Table 3.

Table 3. Clinicians' roles in the adoption of mHealth.

Theme	Sample quote
New digital roles for clinicians, both in digital health start-ups and in health care institutions	"By having CCIOs, clinical information officers, to get involved with the IT side of service provision. So it's not IT companies just serving up a ready-made product, but they're actually engaging clinicians at an earlier stage to make them more useful" (P5); "And so we will have more and more, some doctors involved in the development. Maybe if these companies will hire doctors" (P9); "I think of having an experienced nurse that her daily job won't be to see the patients but rather to see the parameters on her computer … it's like a control centre that will take care of the patients" (P3)
More involvement in the development and design, not only in the early stages but also through the whole lifecycle of the app.	"Well, hopefully, we'll be engaged in the design of these processors, which we are being, a little bit more than we used to be", (P11); "I think the physicians who know a lot can have good ideas and more things that can better these apps" (P5)
Education and awareness by acting as ambassadors, raising awareness, and taking the lead for digital health education in health care institutions	"So it's organising forums that even that they can be discussed, and educating other staff about them" (P12); "Education. Remove the fear. Stigmas about use of mobile phones in workplace, in hospitals need to stop" (P4)
With the efficiencies achieved with mHealth apps, clinicians will be able to focus more on the human side and invest more time with the patients themselves as opposed to spending all the time getting the task done.	"And I don't think there is going to be any machine able to sit with a patient and explain to them and look at them in the eye, all those things that robots don't do but we as humans do" (P7)

How Do These Insights Make Their Way Back to the Design?

Given the interesting affordances and workflow impacts that emerged from the insights, it was important to understand whether the providers have processes that help them identify these affordances, understand any limitation, and embed these insights in the constant development of the tool. In a follow-up interview with the team to better understand how they engage their users, they explained that their key to success is to follow Steve Bank's advice by getting out of the building and talking to their users [49]; one of the company's founders explains, "And part of our CE approval is that we're always engaging with users, always asking how can we improve."

Therefore, they established a process that enables them to systematically engage with their users depending on how engaged they are as visualized in Multimedia Appendix 10.

All active users receive weekly learning emails for the first few weeks about various features on the app where they are also asked to give feedback. Every user that exceeds 25 active sessions automatically gets an email asking them to fill out a survey; then after 50 sessions, they get another email prompting them to leave app store feedback; they also get emails at various stages throughout their life cycle, asking them what they think, "Is there anything we need to improve?"

Inactive users, defined by the providers as every user that has not used the app for 60 days, automatically get an email asking them to fill out a survey to help the team better understand why these users stopped using the app.

Engaged users, also known as power users, are the users that give regular feedback, they are usually invited to a 30- to 60-min interview during one of the major oncology conferences such as the European Society for Medical Oncology to have a more in-depth discussion with the team and give more detailed feedback. The team explains the importance of these interviews to them:

> *I think people are more likely to tell you like—when you meet them in person, they're going to tell you the small little things that annoy them, more so when they're responding to surveys or emails. The slight little things like, "I have to click here to go back here to find this." Those kind of things, which I think is really useful.*

The team started to identify users who were willing to be more involved to be invited to become official advisors as part of the extended team. One of the founders clarifies, "We need more medical input, so we're trying to kind of get this panel of clinicians giving us input and paying them for that, in order to kind of help us go to the next level of tools." This effort has just started, and the team is looking at how they can expand it.

After collecting the feedback, the team discusses together to start screening the users' input and decide on which features to add accordingly; one of the founders explains:

> *We are lucky in that the majority of users request similar tools and content, this makes our jobs easier. I think as ONCOassist develops it's obvious what is missing, and most of our engaged users notice this and point it out. It may be because certain tools and content are becoming more relevant, because of new products on the market. Sometimes we get feedback that may be a little different, which is very useful as well, but 60-70% of the time the requests are similar. Other times we get requests for tools that are difficult for us to build in the short to medium term. We take note of this to see if trends emerge.*

The team also ensured that there was a process to allow them to inform the users when the feature they requested was implemented, as they clarify, "We usually take note of the emails, and then when it goes out, we let them know. And there's also—like we send them a short email, a personalised message, saying, 'Hey, just so you know, this is now live'. But they'll get a group email as well," this helps the team build a personal relationship with these engaged users.

Some of the requests that they receive are also about existing features, as they explain:

> *Other times—all of what our biggest problem at the moment is around discovery. We have content in there that people don't know is available, so that's something we're working on as well. It's like making sure that people know that we have what they need already… It can be an educational process.*

The ongoing challenge that the company is facing, similar to most mHealth start-ups, is monetization, as the team explains:

> *I feel like we've figured out the user interaction, and getting that feedback, and building based on that feedback. Now, for us, the big challenge is figuring out how to kind of have a scalable model with companies that results in us generating good revenues and kind of having on-going relationships in helping and showing improved outcomes.*

Discussion

The results show that technology adoption in our case study is influenced by factors that go beyond the technical and material features of the app; there are other social and organizational factors that play a crucial role in the adoption and success of such new technologies as detailed below.

Understanding the Possibilities and Limitations of Technological Artifacts

Overall, the participants found the app useful, but not all features were seen as equally so; some were emphasized more than others were and some more relevant to specific cases. Features relevant to everyday tasks at the point of care such us toxicity criteria, calculators, staging tools, and adjuvant tools topped the list. Whereas other features, such as product characteristics, were perceived as important in the case of new drugs because clinicians may not be familiar with new medications and this tool can allow them to quickly check drug-related information, especially given that new drugs are continuously being produced.

The main utility of the app is time saving and efficiency as it helps busy clinicians in their daily work in different ways. From one side, it serves as a memory aid as they cannot memorize all the formulas and checking toxicity calculations, for example, takes time, but with the app, they can get that in a couple of seconds while on the go, this makes it faster and easier to extract information. At the same time, the app helps them ensure higher quality of care and patient safety by enabling them to make critical decisions at the point of care, such as recalculating a dose based on body weight fluctuation, contributing to a seamless patient experience at the clinic or hospital. Clinicians created further affordances by using it as a support tool in their discussions with the patients to better explain to them why some decisions were made (eg, why should they continue with chemotherapy even after the tumor is out) extending the use of the app from a clinician tool to a support tool in patient discussions.

Even though the participants explained that there was nothing in the app that they cannot find elsewhere (eg, medical websites or paper-based calculators), it was the compact overview and the presence of all relevant tools in a single app that is accessible anywhere anytime that made it particularly useful for them. Here, it is clear that the material features of the mobile app, its portability, and accessibility played an important role in the way it is being used by the users and why they perceived it as more convenient and easier to use compared with possible alternatives.

However, the app is not without limitations; the lack of information completeness or correctness is perceived as the main limitation, for example, not all adjuvant tools are on the app; hence, if clinicians need to use one that is missing from the app, they will need to use a different tool to do so. Design-related limitations such as clutter also made it difficult for the users to find what they were looking for; this was not the case for ONCOassist, but the participants mentioned that they did face this issue with other medical apps. Hence, reaffirming Kallinikos' idea that the way a material feature is shaped or arranged has an impact on the way it is actually used [37]. The lack of EMR integration was also seen as a limitation, such an integration would make it very easy to adjust calculations and toxicity criteria to specific patient profiles and electronic records, taking the app's utility to a whole new level of convenience and efficiency. This technological interoperability hurdle confirms that often it is the software rather than the feature themselves that can be limiting to some technology uses [22]. Furthermore, there are also social and organizational aspects such as shortage of time and financial resources that can cause limitations to such solutions' adoption.

When talking about limitations and features that they would like to add to the app, it was interesting to see that some existing features were on the requested list, showing that participants were not always aware of everything that the app could do, which underlined the crucial role of awareness and training. Oudshoorn and Pinch emphasized this point when they explained that the success and adoption of new technologies usually depends on the users' knowledge of its features and how to use them, stating that "It has long been recognized that the most sophisticated and complex computer hardware and software will come to naught if users don't know how to use them" [50].

Understanding Constraints and Affordances as Defined by the Users

When exploring the intention of use and the initial motives that encouraged initial use of the app, participants highlighted several features of the app and what it enabled them to achieve, that is, saving them time, increasing work efficiency, and improving quality of patient care. As Gibson proposes, users might refrain from interacting with an object or in this case a technological solution, without learning what it is good for [42]. However, participants also mentioned some aspects related to what technology in general allows, such as the constant development and update of digital sources, which makes them more accurate than media—such as medical books or paper-based tools. However, social aspects, such as endorsement, also impacted

the intention of use, which is a reason for adoption given by nearly all participants. Hence, clinicians mostly trust tools that are endorsed by other clinicians or reliable medical associates they know, showing how social and material aspects start to become entangled, while the decision for adoption does not only rely solely on the tool's features and capabilities but also on social aspects such as trusted endorsement. This implies that without social aspects such as endorsement some exceptional mHealth solutions might not get used.

Social and cultural factors are perceived as the main source of constraints when it comes to mHealth adoption; general attitude toward technology and change, age, previous negative experiences, and perceptions of the use of a personal device at work can be decisive factors in the nonadoption of mHealth according to participants. For instance, an important example of entanglement of the social and the material is when the use of the smartphone at work is routinely perceived as a nonprofessional activity. Even though professional apps that can run on smartphones are no longer a novelty, the stigma related to the use of personal devices at work persists. Users being concerned with negative perceptions resulting from cultural views of smartphone use at work validates previous scholars' argument that "culture tells us what something affords," implying that constraints and affordances are the joint result of the entanglement between the material aspects of technology and cultural practices forming users' perceptions [22,51].

A further look into the participants' views of what elements would afford or constrain mHealth helps us understand materiality at the intersection between social and material aspects as Faraj and Azad and Majchrzak et al explain [40,41]. The material aspects mentioned by clinicians mostly evolved around app stability and operation, cost, ease of use, usefulness, and portability; these findings agree with previous research that also reported similar factors [3,52].

However, these aspects intertwined with other important social aspects namely recommendations from people or institutions they can trust, reliability and neutrality of the content (the content provider must be neutral to be trusted, whereas examples given highlighted concerns of bias where content was provided by Pharma companies), social and cultural factors impacting people's attitudes toward technology, and some organizational factors such as workload and internal politics. These findings complement previous research that shed light on nontechnical factors that play a crucial role in mHealth adoption [53-56].

It's also worth noting, however, how the material properties of different technologies can also afford different use cases depending on the context they are being used in [22]; this was clear with participants' experiences around how they switch between the mobile app and the Web app on the desktop when they are using ONCOassist. Several examples showed that the mobile app's portability, accessibility, and compact overview are the key added values of the solution; however, in specific social contexts, that is, if the clinician is showing a graph to the patients using the app they might prefer to use the Web application to avoid a personal message showing on the phone being seen by the patient, which, in turn, might appear as unprofessional. This highlights the importance of both mobile apps and Web applications as their use can differ from one social context to another.

Understanding How Technology Materialized in the Organizing Process

Various social and organizational impacts of the app usage were mentioned, namely people empowerment, as the app helps users be more autonomous and gives them more power through knowledge and education; it also empowers young doctors that find themselves in an environment that does not encourage questions as they can find answers to any queries via the app without having to ask a more senior colleague. Furthermore, at this early stage of adoption, the decision making is still individualized, meaning that the decision to use one mHealth solution or another lies with the individual clinicians themselves; however, this may change as the rise of the adoption of such solutions is also driving the creation of new roles for clinicians evolving around digital solutions. We now start to see *clinical information officers* in some hospitals and clinics to deal not only with the medical aspects but also the IT side of such services including topics such as interoperability and software integration, which could lead to more centralized decision making in the future.

Clinicians also prefer to use their personal devices rather than the work devices due to the need for speed, location flexibility, and the limited number of computers available at the workplace. In addition, these apps are rapidly replacing some traditional tools such as textbooks, paper-based calculators, and websites.

Another example that can have clear organizational implications is the use of the app for people evaluation and rewards through, for example, streamlining the calculations and SACT datasets and attaching them to Commissioning for Quality and Innovation, based on which staff members get monetary rewards. However, this is only feasible if it is systematically embraced by the organization and not decided on an individual basis. This reinforces the argument that the impact of the materials and materiality on the organizing process can only be seen if something enables their materialization [22]; in our example, this would not happen unless the management enforces the use of the app for people evaluation and rewards.

As for workflow advantages, the app enhanced users' efficiency and performance by making their work easier, quicker, more professional, and with fewer interruptions. It also led to a sharpened practice because it provided a backup solution and a safety net to double check drug dosages and calculations then adjust in case of fluctuations in patient's weight, for example. The personal phone usage also enabled location flexibility; by helping them overcome computer shortages or instant decision making at the point of care or in meetings, it can also provide a solution in places where traditional devices are not permitted such as the clinical room where computers are not allowed because of infection control.

Nevertheless, interoperability and lack of integration with EMRs are seen as key workflow disadvantages because the app as it stands today cannot be directly linked to a specific patient health record, meaning that clinicians have to enter the individual

patient values manually for each calculation. Some participants suggest that apps that do not offer EMR integration could easily be dismissed in favor of bigger vendors who offer such an integration. On the contrary, cultural aspects such as assumptions made regarding mobile phone usage in the workplace and the perception that it is a waste of time could impact mHealth users negatively in the workplace unless tackled and awareness is raised.

These insights complement the existing literature that pinpointed the importance of organizational and workflow factors that can considerably hinder mHealth adoption if not addressed properly [6,54,57]

Accounting for User Engagement: Affordances Enablement

The 3 steps in Leonardi's methodological guidelines smoothly lead us into understanding the materials, followed by materiality, then materialization, and how the materiality impacts the organizing process. What it does not address though is how can the affordances be enabled through user engagement; in other words, how is user feedback considered and how does this feedback make its way into the design process. It is important to find a way to identify mechanisms that would enable the designers to understand and maximize the potential of the affordances created by the users.

ONCOassist offers a best practice example of embedding the users in the development process and how it made for a constant interaction between the social and the material, allowing the material to develop in a way that best serves both organizational and social practices. They explained how they capture feedback from every user category depending on their level of engagement and that they keep adding new features based on this feedback. They also recently established an advisory board that includes some of the apps' power users. This enables them to always be on top of newly emerging medical needs and stay relevant in a world that is constantly changing.

From their side, clinicians showed clear interest in being a part of the development process and foresee a rise in digital roles for clinicians, both in digital health start-ups and in health care institutions. This would entail more involvement in the development and design of such solutions, not only in the early stages but also through the whole lifecycle of the app, and they would act as ambassadors raising awareness and taking the lead for digital health education in health care institutions. Some clinicians are clear advocates of such solutions because they believe that it would help them focus more on the human side through the efficiencies that they create.

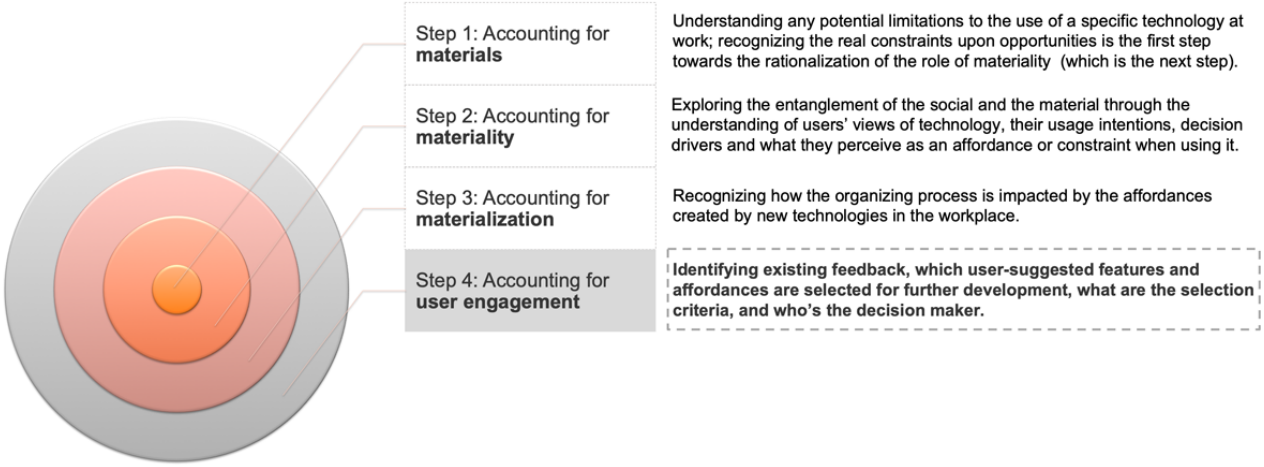
This paper proposes the addition of a fourth step to Leonardi's guidelines to account for user engagement as visualized in Figure 3 [22], by identifying existing feedback loops, understanding which user-suggested features and affordances are selected for further development, how they are selected and based on which criteria, and who the decision maker is. These questions would help to capture any processes that activate the role of the users and the implementation of affordances that might arise.

Including the users in the constant development and testing of innovative tools is not without challenges, especially considering the very competitive environment that mHealth start-ups are operating in; hence, they are protective of their ideas, which results in most of the design and testing being done inside the company without involving anyone external [58]. This, however, needs to change to enable the user-created affordances to make their way to the design and to ensure that these new tools will stay relevant to the users in a world that changes continuously. A balance between user engagement and confidentiality can be achieved through advisory contracts, nondisclosure agreements, or the like.

Practical and Theoretical Implications

This case demonstrates the affordances and constraints of mHealth technology at the workplace, their implications for the organization of work, and clinicians' role in their constant development and adoption. The research findings confirmed that factors such as app operation and stability, ease of use, usefulness, cost, and portability play a major role in the adoption. However, other social factors such as endorsement, neutrality of the content, attitude toward technology, workload, and internal politics are also perceived as key determinants of adoption.

Figure 3. Accounting for User Engagement.



Several organizational implications also emerged, suggesting the need for some actions such as fostering a work culture that embraces new technologies and the creation of new digital roles for clinicians both on the hospitals or clinics and on the development sides but also more collaboration between health care organizations and digital health providers to enable EMR integration and solve any interoperability issues.

From a theoretical perspective, this study suggests the addition of a fourth step to Leonardi's methodological guidance to account for user engagement; embedding the users in the continuous design and development processes ensures the understanding of user-specific affordances that can then be made more obvious to other users and increase the potential of such tools to go beyond their technological features and have a higher impact on workflow and the organizing process.

Limitations and Recommendations for Future Research

This case study is limited to a specific mHealth solution and a specific geography during a specific timeframe, and therefore its findings may not be generalizable to other contexts where, for example, the health care system and its regulations could be considerably different. Furthermore, the sample size is relatively small and excluded nonusers because their recruitment proved to be very challenging. Moreover, given the constantly evolving nature of mHealth, the context of the research might change very quickly necessitating new research to update the findings.

To address some of these limitations, future research should cover other mHealth solutions in other geographies, timeframes, and contexts. It would also be very relevant to include some nonusers to the participants mix to cover their views as well.

Acknowledgments

The authors would like to thank the ONCOassist team, as well as each of the busy clinicians, who agreed to give their time and opinions for the sake of this research; this study would not have been possible without their valued input.

Conflicts of Interest

None declared.

Multimedia Appendix 1

ONCOassist, a decision support app for oncologists.

[PNG File, 302KB - mhealth_v7i5e13555_app1.png]

Multimedia Appendix 2

ONCOassist's key features at the time of writing this paper.

[PDF File (Adobe PDF File), 41KB - mhealth_v7i5e13555_app2.pdf]

Multimedia Appendix 3

NVivo code scheme.

[PDF File (Adobe PDF File), 51KB - mhealth_v7i5e13555_app3.pdf]

Multimedia Appendix 4

Phases of thematic analysis after Braun and Clarke.

[PDF File (Adobe PDF File), 48KB - mhealth_v7i5e13555_app4.pdf]

Multimedia Appendix 5

Interview guides for users and providers.

[PDF File (Adobe PDF File), 53KB - mhealth_v7i5e13555_app5.pdf]

Multimedia Appendix 6

Utility and limitations.

[PNG File, 640KB - mhealth_v7i5e13555_app6.png]

Multimedia Appendix 7

Barriers and opportunities.

[PDF File (Adobe PDF File), 331KB - mhealth_v7i5e13555_app7.pdf]

Multimedia Appendix 8

Social and organizational impacts.

[PDF File (Adobe PDF File), 310KB - mhealth_v7i5e13555_app8.pdf]

Multimedia Appendix 9

Workflow improvements and disadvantages.

[PNG File, 715KB - mhealth_v7i5e13555_app9.png]

Multimedia Appendix 10

ONCOassist's user engagement strategy.

[PNG File, 1MB - mhealth_v7i5e13555_app10.png]

References

1. Pavel M, Jimison HB, Wactlar HD, Hayes TL, Barkis W, Skapik J, et al. The role of technology and engineering models in transforming healthcare. IEEE Rev Biomed Eng 2013;6:156-177. [doi: 10.1109/RBME.2012.2222636] [Medline: 23549108]
2. Putzer GJ, Park Y. The effects of innovation factors on smartphone adoption among nurses in community hospitals. Perspect Health Inf Manag 2010;7:1b [FREE Full text] [Medline: 20697467]
3. Gagnon M, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. J Am Med Inform Assoc 2016 Jan;23(1):212-220. [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
4. Kay M, Santos J, Takane M. mHealth: New Horizons for Health through Mobile Technologies. Geneva, Switzerland: World Health Organisation; 2011.
5. Akter S, Ray P. mHealth - an ultimate platform to serve the unserved. Yearb Med Inform 2010:94-100. [Medline: 20938579]
6. Choi WS, Park J, Choi JY, Yang J. Stakeholders' resistance to telemedicine with focus on physicians: utilizing the Delphi technique. J Telemed Telecare 2018 Jan 01:1357633X18775853 Epub ahead of print(forthcoming). [doi: 10.1177/1357633X18775853] [Medline: 29792080]
7. Hussein R, Khalifa A. Telemedicine in Egypt?: SWOT analysis and future trends. GMS Medizinische Inform Biometrie und Epidemiol 2012;8(1):1-16. [doi: 10.3205/mibe000125]
8. Moffatt J, Eley D. Barriers to the up-take of telemedicine in Australia--a view from providers. Rural Remote Health 2011;11(2):1581 [FREE Full text] [Medline: 21385004]
9. Walter Z, Lopez MS. Physician acceptance of information technologies: role of perceived threat to professional autonomy. Decis Support Syst 2008 Dec;46(1):206-215. [doi: 10.1016/j.dss.2008.06.004]
10. Xue Y, Liang H, Mbarika V, Hauser R, Schwager P, Kassa GM. Investigating the resistance to telemedicine in Ethiopia. Int J Med Inform 2015 May 09;84(8):537-547. [Medline: 25991059]
11. Lewin K. Frontiers in group dynamics: concept, method and reality in social science. Hum Relat 1947;1(1):5-41. [doi: 10.1177/001872674700100103]
12. Markus ML. Power, politics, and MIS implementation. Commun ACM 1983;26(6):430-444. [doi: 10.1145/358141.358148]
13. Hu PJ, Chau PY, Sheng OR. Adoption of telemedicine technology by health care organizations: an exploratory study. J Organ Comput Electron Commer 2009 Nov 18;12(3):197-221. [doi: 10.1207/S15327744JOCE1203_01]
14. Orlikowski WJ. Sociomaterial practices: exploring technology at work. Organ Stud 2016 Jul;28(9):1435-1448. [doi: 10.1177/0170840607081138]
15. Orlikowski WJ. The sociomateriality of organisational life: considering technology in management research. Cambridge J Econ 2009 Nov 05;34(1):125-141. [doi: 10.1093/cje/bep058]
16. Orlikowski WJ, Scott SV. Exploring material-discursive practices. J Manag Stud 2014 Dec 28;52(5):697-705. [doi: 10.1111/joms.12114]
17. Mutch A. Sociomateriality — taking the wrong turning? Inf Organ 2013 Jan;23(1):28-40. [doi: 10.1016/j.infoandorg.2013.02.001]
18. Leonardi PM. When does technology use enable network change in organizations? A comparative study of feature use and shared affordances. MIS Q 2013 Mar 3;37(3):749-775. [doi: 10.25300/MISQ/2013/37.3.04]

19. McPhee R, Canary H. Distanciation and Organizational Constitution. 2013 Presented at: European Group for Organizational Studies Conference; 2013; Canada.
20. Tunçalp D. Questioning the ontology of sociomateriality: a critical realist perspective. Manag Decis 2016 Jun 20;54(5):1073-1087. [doi: 10.1108/MD-07-2014-0476]
21. Leonardi P, Rodriguez-Lluesma C. Sociomateriality as a lens for design: imbrication and the constitution of technology and organization. Scand J Inf Syst 2013;24(2).
22. Leonardi P. Methodological Guidelines for the Study of Materiality and Affordances. New York: Routledge; 2018.
23. Carlile P, Nicolini D, Langley A, Tsoukas H. How matter matters: Objects, artifacts, and materiality in organization studies: Introducting the third volume of Perspectives on Organization Studies. Oxford: Oxford Scholarship Online; 2013.
24. Garfinkel H. Studies in ethnomethodology. Englewoods Cliffs, NJ: Prentice Hall; 1967.
25. Goffman E. The presentation of self in everyday life. Garden City, New York: Doubleday Anchor; 1956.
26. Trevor P. The invisible technologies of Goffman's sociology from the merry-go-round to the internet. Technol Cult 2010;51(2):409-424. [doi: 10.1353/tech.0.0456]
27. Bélanger F, Watson-Manheim MB. Virtual teams and multiple media: structuring media use to attain strategic goals. Group Decis Negot 2006 Jul 11;15(4):299-321. [doi: 10.1007/s10726-006-9044-8]
28. Chudoba K, Wynn E, Lu M, Watson-Manheim M. How virtual are we? Measuring virtuality and understanding its impact in a global organization. Inf Syst J 2005;15(4):279-306. [doi: 10.1111/j.1365-2575.2005.00200.x]
29. Creswell J, Poth C. Five Qualitative Approaches to Inquiry: choosing among five approaches. California: SAGE; 2018.
30. Braun V, Clarke V. Successful Qualitative Research: A Practical Guide For Beginners. California: Sage; 2013.
31. Braun V, Clarke V. What can “thematic analysis” offer health and wellbeing researchers? Int J Qual Stud Health Well-being 2014 Oct 16;9. [doi: 10.3402/qhw.v9.26152]
32. Yin R. Case Study Research And Applications: Design And Methods. New York: Sage Publications, Inc; 2019.
33. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
34. Verdinelli S, Scagnoli NI. Data display in qualitative research. Int J Qual Methods 2013 Feb;12(1):359-381. [doi: 10.1177/160940691301200117]
35. Henderson S, Segal EH. Visualizing qualitative data in evaluation research. New Dir Eval 2013 Sep;2013(139):53-71. [doi: 10.1002/ev.20067]
36. Pickering A. The mangle of practice: agency and emergence in the sociology of science. Am J Sociol 1993 Nov;99(3):559-589. [doi: 10.1086/230316]
37. Kallinikos J. Form, Function, and Matter: Crossing the Border of Materiality. In: Materiality and Organizing: Social Interaction in a Technological World. Oxford: Oxford Scholarship Online; 2012.
38. Leonardi PM. When flexible routines meet flexible technologies: affordance, constraint, and the imbrication of human and material agencies. MIS Q 2011;35(1):147. [doi: 10.2307/23043493]
39. Leonardi P. Materiality, Sociomateriality, and Socio-Technical Systems: What Do These Terms Mean? How Are They Different? Do We Need Them?. Oxford: Oxford University Press; 2012:25-48.
40. Faraj S, Azad B. The Materiality of Technology: An Affordance Perspective. In: Materiality and Organizing: Social Interaction in a Technological World. Oxford: Oxford University Press; 2012.
41. Majchrzak A, Faraj S, Kane GC, Azad B. The contradictory influence of social media affordances on online communal knowledge sharing. J Comput-Mediat Comm 2013 Oct 18;19(1):38-55. [doi: 10.1111/jcc4.12030]
42. Gibson J. The Ecological Approach To Visual Perception (psychology Press & Routledge Classic Editions). London: Routledge; 2014.
43. Norman D. The Design Of Everyday Things: Revised And Expanded Edition. New York: Basic Books; 2019.
44. Norman D. Affordance, convention, and design. Interactions 1999;6(3):38-43. [doi: 10.1145/301153.301168]
45. Hutchby I. Technologies, texts and affordances. Sociology 2001;35(2):441-456. [doi: 10.1017/S0038038501000219]
46. Fayard A, Weeks J. Photocopiers and water-coolers: the affordances of informal interaction. Organ Stud 2007;28(5):605-634. [doi: 10.1177/0170840606068310]
47. Zammuto RF, Griffith TL, Majchrzak A, Dougherty DJ, Faraj S. Information technology and the changing fabric of organization. Organ Sci 2007 Oct;18(5):749-762. [doi: 10.1287/orsc.1070.0307]
48. Leonardi PM. Activating the informational capabilities of information technology for organizational change. Organ Sci 2007 Oct;18(5):813-831. [doi: 10.1287/orsc.1070.0284]
49. Bank S, Dorf B. The Startup Owner's Manual: The Step-By-Step Guide for Building a Great Company. California: K&S Ranch; 2012.
50. Oudshoorn N, Pinch T. How Users Matter: The Co-construction Of Users And Technology (inside Technology). Cambridge: MIT Press; 2005.
51. Gerardhi S. Sociomateriality in posthuman practice theory. In: The Nexus Of Practices: Connections, Constellations, Practitioners. London: Routledge; 2016.
52. Tahamtan I, Pajouhanfar S, Sedghi S, Azad M, Roudbari M. Factors affecting smartphone adoption for accessing information in medical settings. Health Info Libr J 2017 Jun;34(2):134-145. [doi: 10.1111/hir.12174] [Medline: 28406547]

53. Sezgin E, Özkan-Yildirim S, Yildirim S. Understanding the perception towards using mHealth applications in practice. Inf Dev 2016 Dec 19;34(2):182-200. [doi: 10.1177/0266666916684180]
54. Ayatollahi H, Mirani N, Nazari F, Razavi N. Iranian healthcare professionals' perspectives about factors influencing the use of telemedicine in diabetes management. World J Diabetes 2018 Jun 15;9(6):92-98 [FREE Full text] [doi: 10.4239/wjd.v9.i6.92] [Medline: 29988886]
55. Sezgin E, Yildirim S, Yildirim S. Intention vs. Perception: Understanding the Differences in Physicians' Attitudes Toward Mobile Health Applications. In: Current and Emerging mHealth Technologies. Switzerland: Springer; 2018.
56. Wernhart A, Gahbauer S, Haluza D. eHealth and telemedicine: practices and beliefs among healthcare professionals and medical students at a medical university. PLoS One 2019;14(2):e0213067 [FREE Full text] [doi: 10.1371/journal.pone.0213067] [Medline: 30818348]
57. Koivunen M, Saranto K. Nursing professionals' experiences of the facilitators and barriers to the use of telehealth applications: a systematic review of qualitative studies. Scand J Caring Sci 2017 Aug 03;32(1):24-44. [doi: 10.1111/scs.12445]
58. Oudshoorn N, Rommes E, Stienstra M. Configuring the user as everybody: gender and design cultures in information and communication technologies. Sci technol human values 2004 Jan 01;29(1):30-63. [doi: 10.1177/0162243903259190]

Abbreviations

EMR: Electronic Medical Record
IT: Information Technology
mHealth: mobile health
PSA: prostate-specific antigen
SACT: Systemic Anti-Cancer Treatment

Edited by M Focsa; submitted 05.02.19; peer-reviewed by MP Gagnon, P Ware, J Jagarapu, F López Seguí; comments to author 28.02.19; revised version received 13.03.19; accepted 03.04.19; published 03.05.19.

Please cite as:
Jacob C, Sanchez-Vazquez A, Ivory C
Clinicians' Role in the Adoption of an Oncology Decision Support App in Europe and Its Implications for Organizational Practices: Qualitative Case Study
JMIR Mhealth Uhealth 2019;7(5):e13555
URL: https://mhealth.jmir.org/2019/5/e13555/
doi:10.2196/13555
PMID:31066710

Original Paper

Development of Comprehensive Personal Health Records Integrating Patient-Generated Health Data Directly From Samsung S-Health and Apple Health Apps: Retrospective Cross-Sectional Observational Study

Se Young Jung[1,2*], MPH, MD; Jeong-Whun Kim[3*], MD, PhD; Hee Hwang[1,4], MD, PhD; Keehyuck Lee[1,2], MD, MBA; Rong-Min Baek[5], MD, PhD; Ho-Young Lee[1,6], MD, PhD; Sooyoung Yoo[1], PhD; Wongeun Song[1], MS; Jong Soo Han[2], MD, MBA

[1]Office of eHealth Research and Business, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea

[2]Department of Family Medicine, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea

[3]Department of Otorhinolaryngology Head and Neck Surgery, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea

[4]Department of Pediatrics, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea

[5]Department of Plastic Surgery, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea

[6]Department of Nuclear Medicine, Seoul National University Bundang Hospital, Seongnam-si, Republic of Korea

[*]these authors contributed equally

Corresponding Author:
Hee Hwang, MD, PhD
Office of eHealth Research and Business
Seoul National University Bundang Hospital
172 Dolma-ro, Bundang-gu
Seongnam-si, 13620
Republic of Korea
Phone: 82 31 787 0114
Fax: 82 82 31 787 4054
Email: epilepsyguy@gmail.com

Abstract

Background: Patient-generated health data (PGHD), especially lifelog data, are important for managing chronic diseases. Additionally, personal health records (PHRs) have been considered an effective tool to engage patients more actively in the management of their chronic diseases. However, no PHRs currently integrate PGHD directly from Samsung S-Health and Apple Health apps.

Objective: The purposes of this study were (1) to demonstrate the development of an electronic medical record (EMR)–tethered PHR system (Health4U) that integrates lifelog data from Samsung S-Health and Apple Health apps and (2) to explore the factors associated with the use rate of the functions.

Methods: To upgrade conventional EMR-tethered PHRs, a task-force team (TFT) defined the functions necessary for users. After implementing a new system, we enrolled adults aged 19 years and older with prior experience of accessing Health4U in the 7-month period after November 2017, when the service was upgraded.

Results: Of the 17,624 users, 215 (1.22%) integrated daily steps data, 175 (0.99%) integrated weight data, 51 (0.29%) integrated blood sugar data, and 90 (0.51%) integrated blood pressure data. Overall, 61.95% (10,919/17,624) had one or more chronic diseases. For integration of daily steps data, 48.3% (104/215) of patients used the Apple Health app, 43.3% (93/215) used the S-Health app, and 8.4% (18/215) entered data manually. To retrieve medical documentation, 324 (1.84%) users downloaded PDF files and 31 (0.18%) users integrated their medical records into the Samsung S-Health app via the Consolidated-Clinical Document Architecture download function. We found a consistent increase in the odds ratios for PDF downloads among patients with a higher number of chronic diseases. The age groups of ≥60 years and ≥80 years tended to use the download function less frequently than the others.

Conclusions: This is the first study to examine the factors related to integration of lifelog data from Samsung S-Health and Apple Health apps into EMR-tethered PHRs and factors related to the retrieval of medical documents from PHRs. Our findings on the lifelog data integration can be used to design PHRs as a platform to integrate lifelog data in the future.

(JMIR Mhealth Uhealth 2019;7(5):e12691) doi:10.2196/12691

KEYWORDS

personal health record; patient generated health data; lifelog; electronic medical record; mobile phone

Introduction

The increase in the prevalence of preventable chronic diseases has become an enormous problem worldwide in recent decades [1,2]. Chronic diseases impose an immense societal and financial burden [3,4]. Three of every five people with type 2 diabetes develop complications such as stroke, heart disease, and eyesight problems. The rate of avoidable hospitalization for hypertension, which is a major risk factor for heart disease, has increased more than 90% over the last 10 years [5]. The main reason for this increase is that patients tend to neglect chronic conditions, because most of these diseases are indolent during subclinical stages. Therefore, it is important to manage chronic diseases before they transition from the subclinical to the clinical phase [6,7].

Personal health records (PHRs) have been considered an effective tool to engage patients more actively in the management of their chronic diseases [8-10]. However, most provider-based PHRs have been unsuccessful in engaging patients in their health care because most patients and health care professionals have failed to obtain benefits from PHRs [11,12]. The health-related information in PHRs is inconsistent and inaccurate because most administrative functions in PHRs are solely dependent on PHR users [13,14]. If users are not interested in inputting information into PHRs, PHRs become useless. Therefore, patient-generated health data (PGHD) are crucial determinants of the success of PHRs, as doctors can recommend or prescribe treatments based on PGHD in PHRs [15,16]. PGHD are health-related data recorded, created, or gathered by or from patients, caregivers, or family members to monitor and address a health concern [15]. PGHD generally include health history, treatment history, biometric data, symptoms, and lifestyle choices. From these data, biometric and lifelog data can be collected via objective methods such as wearables, smartphones, or smart devices.

Seoul National University Bundang Hospital (SNUBH) has had its own onsite electronic medical record (EMR) solution since 2003. It also started an EMR-tethered PHR system called Health4U in June 2013, integrating PGHD such as daily steps, sleep patterns, weight, and blood pressure directly from Samsung S-Health and Apple Health apps for the first time beginning in November 2017, and was the first hospital in the world to do so [14,17,18]. In particular, daily steps are recorded automatically by the mere use of a smartphone in everyday life, and the information can be streamed into Health4U without much effort. Doctors and nurses can check patients' daily activities by merely monitoring those records. Health4U also adopted medical document download functions for health information exchange. This study explored the features of EMR-tethered PHRs that can integrate S-Health and Apple Health data and investigated factors related to the use pattern of the lifelog integration and medical document download functions in Health4U.

Methods

Development Process

A task-force team (TFT) of 5 attending physicians, 3 nurses, and 4 engineers was formed to upgrade Health4U. In June 2017, the TFT began conducting an analysis of a previous version of Health4U, and for 6 months, they defined functions that should be upgraded based on this analysis (Table 1).

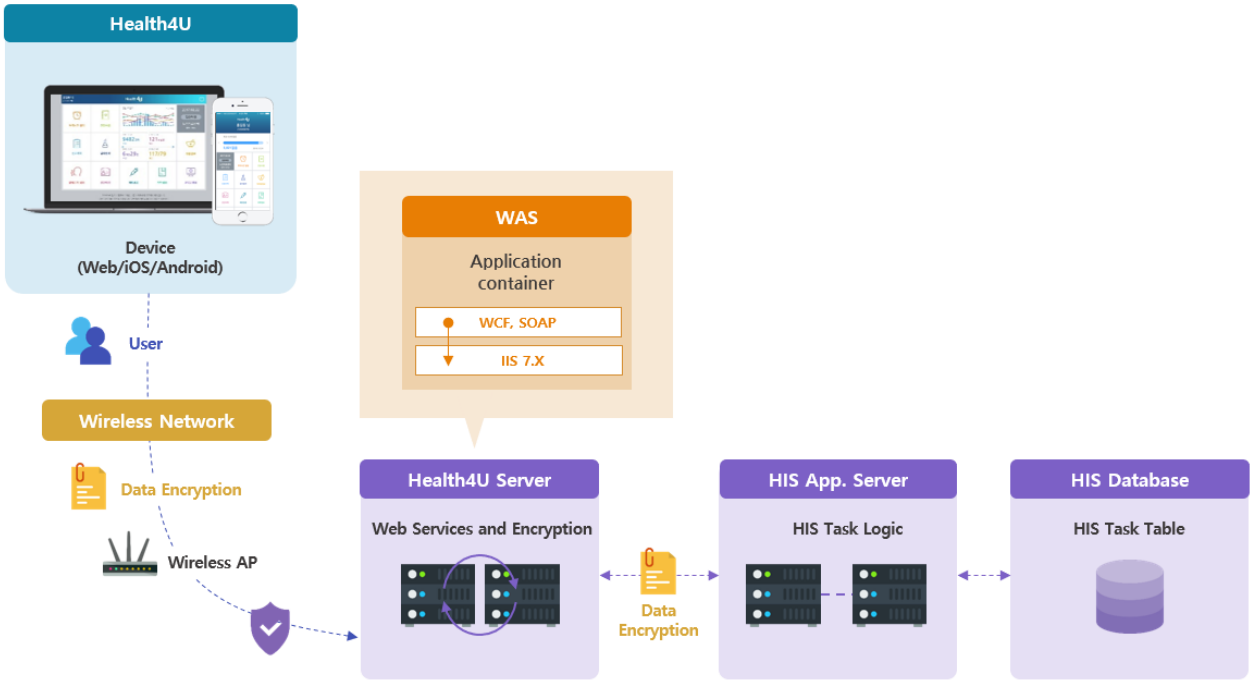
Health4U is installed on a personal computer and mobile. The overall system architecture is described in Figure 1. When a user logs into Health4U, the server communicates through the hospital information system (HIS) server attached to the HIS database. All data streamed between the Health4U server and the HIS app server are encrypted.

Table 1. Important functionalities identified by the task-force team.

Domain	Functions
Medical record	View the medical history
Lab results	Change the date standard from the reporting date to the order-issued date
Prescription history	Provide medication information (eg, drug image, usage, effect, and side effects)
Diagnosis history	View function for diagnosis history
Alarm setting for medication	Push notification for taking medicine/note whether to take a dose
Vaccination	View vaccination history (eg, for children and adults)
Health notes	• Lifestyle data management for activity, weight, blood pressure, blood sugar test, sleep, and stress • Link to Samsung S-Health and Apple Health App data
Medical info	Link to check the encyclopedia (medical information) of the famous web portal called NAVER
Educational material	Provide education animation related to the patient's diagnosis
Blue button	• Download and view the PDF of medical history • Transfer to Samsung S-Health in the form of C-CDA[a] file
View health data	View lifestyle data (eg, activity, weight, blood pressure, and sleep)/push the numerical criteria for recommendation
Login	Fingerprint login
Link to external apps	Link to app for hospital webpage and BESTguide app

[a]C-CDA: Consolidated Clinical Document Architecture.

Figure 1. System architecture. HIS: hospital information system. WAS: web application server.



Upgrade of the Previous Version of Health4U

A previous version of Health4U was developed as an EMR-tethered PHR system based on the needs of users and was composed of five parts: visit history, drug notification, prescription history, laboratory results, and management of a self-administered component (called the Health Notes; Figure 2) [14].

The TFT concluded that the integration of PHRs into EMR was not enough to engage patients to manage their diseases, and PGHD integration into both PHRs and EMR can be a key feature that attracts users' interest. Therefore, we added five new functions to the previous version: tracking of diagnosis history; push notifications for taking medicine; integration of lifelog data including activity, blood pressure, blood sugar, sleep, and stress; fingerprint login; and links to external apps such as an

indoor navigation app called BESTguide and a cardiovascular disease management app called Heart4U (Figure 3) [19].

In particular, the TFT designed a key function that integrated lifelog data directly from Samsung S-Health and Apple Health apps. This is the first instance worldwide of use these two platforms as portals to collect lifelog data for EMR-tethered PHRs (Figure 4). The TFT believed that patients no longer needed to complete structured questionnaires such as the International Physical Activity Questionnaire for tracking daily activities, and doctors could check patients' daily steps on the dashboard without much effort if we were successful in implementing the daily steps integration function. Additionally, information in PHRs can be downloaded in PDF format or directly integrated into the Samsung S-Health app via the Consolidated Clinical Document Architecture (C-CDA) file (Figure 5). Furthermore, we developed a dashboard for doctors to see patients' lifelog records to provide recommendations based on their results (Figure 6).

Figure 2. The Health4U app for personal computer and mobile phone.

Figure 3. Screenshots of Lifelog management.

Figure 4. Direct integration of mobile and wearable data.

Figure 5. Function for the download of medical records. CDA: clinical document architecture.

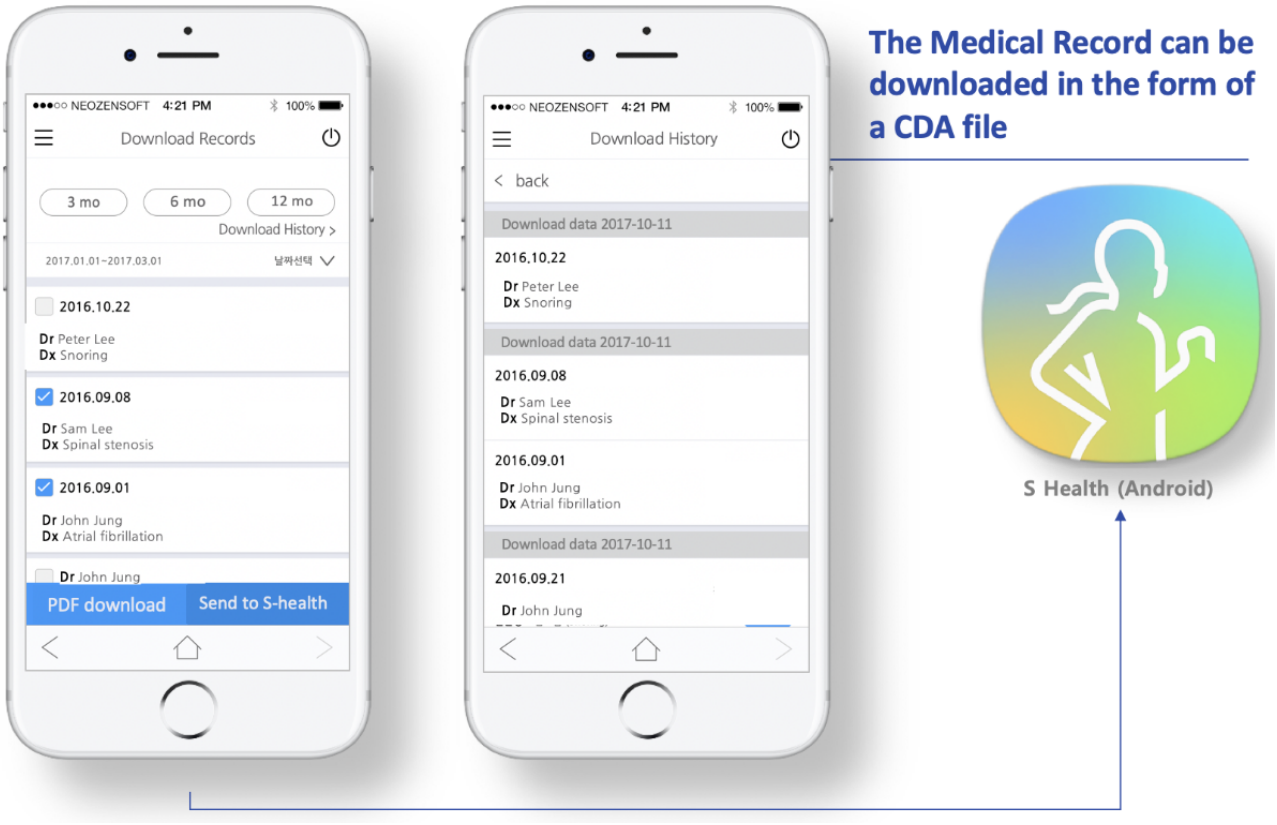


Figure 6. Dashboard for doctors.

Study Population

This study used cross-sectional data extracted from a clinical data warehouse of SNUBH. The enrollees were selected from a sample of adults aged 19 years and older, with prior experience in accessing Health4U in the 7-month period after November 2017, when the service was upgraded. The study participants were selected regardless of prior hospitalization to SNUBH and did not receive any rewards from the research when they downloaded and used the app. A total of 17,624 patients were included in this study.

Activity tracking is important for managing chronic diseases. Therefore, we planned to analyze data on daily steps integration as a surrogate marker of easily obtainable activity data, because steps data are recorded whenever users carry their smartphones or put on their wearables such as smart watches. If patients use a Bluetooth-based scale, sphygmomanometer, or blood sugar checker, the data generated by those devices can be integrated into Health4U through Samsung S-Health and Apple Health apps. Additionally, patients manually input those data when they see the results on their devices. Therefore, we planned to analyze weight, blood pressure, and blood sugar data to explore how patients integrate them into their daily lives. Finally, we planned to evaluate factors associated with the use patterns of the PDF and C-CDA download functions because they can be used to exchange health information via social media, and we wanted to gauge patient interest in those functions.

All analyses were performed using Stata 14.0 (Stata Corp, College Station, TX). A *P* value<.05 was considered significant. The requirement of informed consent was waived because we used anonymized the data. This research was approved by the Institutional Review Board of Human Research of SNUBH.

Results

Principal Results

Among the 17,624 users of Health4U from November 2017 to May 2018, 215 (1.22%) users integrated daily steps data, 175 (0.99%) users integrated weight data, 51 (0.29%) users integrated blood sugar data, and 90 (0.51%) users integrated blood pressure data. Overall, 61.95% (10,919/17,624) had one or more chronic diseases. For the function of retrieving medical documentation, 324 (1.84%) users downloaded PDF files and 31 (0.18%) patients integrated their medical records into Samsung S-Health via the C-CDA download function (Table 2).

Table 2. Baseline characteristics (N=17,624).

Variable	n (%)
Age (years)	
19-40	3848 (21.83)
41-60	7043 (39.96)
60-80	5710 (32.40)
≥80	1023 (5.80)
Gender	
Male	9142 (51.87)
Female	8482 (48.13)
Having a spouse	
Yes	11,713 (79.00)
No	3113 (21.00)
Educational level	
Elementary school and lower	1497 (10.22)
Middle school degree	1249 (8.53)
High school degree	4364 (29.79)
College degree and higher	7539 (51.46)
Occupation	
Blue collar	1853 (12.49)
White collar	4522 (30.48)
Student	405 (2.73)
Other jobs	4605 (31.04)
No job	3449 (23.25)
Patient type	
New	1173 (6.66)
Returning	16,451 (93.34)
Number of chronic diseases[a]	
0	6705 (38.04)
1	4112 (23.33)
2	2534 (14.38)
3	1496 (8.49)
4	989 (5.61)
≥5	1788 (10.15)
Hypertension	
Yes	2562 (14.54)
No	15,062 (85.46)
Diabetes mellitus	
Yes	2267 (12.86)
No	15,357 (87.14)
Dyslipidemia	
Yes	1843 (10.46)
No	15,781 (89.54)

Variable	n (%)
Cancer	
Yes	5948 (33.75)
No	11,676 (66.25)
Prior hospitalization to SNUBH[b]	
Yes	14,834 (84.17)
No	2790 (15.83)
Departments	
Cancer center	2066 (11.72)
Internal medicine or family medicine	4074 (23.12)
Neurovascular center	1566 (8.89)
Surgery	2491 (14.13)
Other departments	7427 (42.14)
Daily steps integration	
Yes	215 (1.22)
No	17,409 (98.78)
Weight integration	
Yes	175 (0.99)
No	17,449 (99.01)
Blood sugar integration	
Yes	51 (0.29)
No	17,573 (99.71)
Blood pressure integration	
Yes	90 (0.51)
No	17,534 (99.49)
PDF download	
Users	324 (1.84)
Nonusers	17,300 (98.16)
C-CDA[c] download	
Users	31 (0.18)
Nonusers	17,593 (99.82)

[a]Chronic diseases include diabetes mellitus, dyslipidemia, hypertension, and any type of cancer.

[b]SNUBH: Seoul National University Bundang Hospital.

[c]C-CDA: Consolidated Clinical Document Architecture

For platforms integrating daily steps, 104 patients used the Apple Health app (HealthKit), 93 patients used the Samsung S-Health app (S-Health), and 18 patients entered data by themselves. For integrating weight, blood sugar, and blood pressure data, participants more often input data manually than in other ways (Table 3).

Table 4 shows the trends of patients' recordings in each quarter, who integrated lifelog data into Health4U. After the upgraded version of Health4U was launched in the fourth quarter of 2017, daily steps data were integrated into Health4U from Samsung S-Health and Apple Health apps.

We hypothesized that the rate of lifelog integration into PHRs and the use of the PDF or C-CDA download functions depend on a user's age, gender, educational level, occupation, and number of chronic diseases. These variables were included as covariates for multivariable logistic regression, and additionally, two other variables (having a spouse and department visited) were included to adjust for other influential factors.

Table 3. Platforms for integrating daily steps, weight, blood sugar, and blood pressure data.

Variable	n (%)
Daily steps integration (n=215)	
Samsung S-Health	93 (43.3)
Apple Health	104 (48.3)
Manual input	18 (8.4)
Weight integration (n=175)	
Samsung S-Health	23 (13.1)
Apple Health	4 (2.3)
Manual input	148 (84.6)
Blood sugar integration (n=51)	
Samsung S-Health	4 (8)
Apple Health	4 (8)
Manual input	43 (84)
Blood pressure integration (n=90)	
Samsung S-Health	10 (11)
Apple Health	18 (20)
Manual input	62 (69)

Table 4. Number of patient recordings from the third quarter of 2017 to the second quarter of 2018. The upgraded version of Health4U was launched in the fourth quarter of 2017.

Variable	n (%)
Weight integration (n=457)	
Third quarter of 2017	110 (24.0)
Fourth quarter of 2017	91 (19.9)
First quarter of 2018	121 (26.5)
Second quarter of 2018	135 (29.6)
Blood pressure integration (n=230)	
Third quarter of 2017	58 (25.2)
Fourth quarter of 2017	58 (25.2)
First quarter of 2018	46 (20.0)
Second quarter of 2018	68 (29.6)
Blood sugar integration (n=164)	
Third quarter of 2017	30 (18.3)
Fourth quarter of 2017	44 (26.8)
First quarter of 2018	40 (24.4)
Second quarter of 2018	50 (30.5)
Daily steps integration (n=370)	
Third quarter of 2017	0 (0)
Fourth quarter of 2017	69 (18.6)
First quarter of 2018	135 (36.5)
Second quarter of 2018	166 (44.9)

Analysis of the Lifelog Data Integration Traits

The age group of ≥60 years used the daily steps sync function less often than the reference group. Women used the function 70% less often than men, and the difference was statistically significant (P<.01). Educational level, occupation, patient type, number of chronic diseases, and department did not influence the use rate (Table 5).

Analysis of the Traits Related to Medical Information Download

We found a consistent increase in the odds ratios for the number of chronic diseases related to the PDF download function. The age groups of ≥60 years and ≥80 years had a lower tendency to use this function. Patients without a spouse used the function 61% more often than those with a spouse. White-collar workers tended to use the function 60% more often than blue-collar workers. Users with a college degree and higher education used the function 2.3 times more often than users who finished elementary school only (Table 6).

In contrast, we could not identify any factors that were significantly associated with the use of the C-CDA download function (Table 7).

We also analyzed the association of daily steps sync function with download functions (PDF and C-CDA) and lifelog integration functions (blood pressure, blood sugar, and weight). We found that daily steps sync was strongly related with the other functions (Table 8). Users of the PDF download, C-CDA integration, blood sugar integration, blood pressure integration, and weight integration functions synced their daily steps 7.8 times, 18.76 times, 18.66 times, 4.68 times, and 41.31 times more, respectively, than users who did not use these functions.

Table 5. Multivariable analysis of the factors associated with daily steps synced with Health4U.

Variable	Adjusted odds ratio	*P* value
Age (years)		
19-40	1	
41-60	0.75	.30
60-80	0.39	.01
≥80	N/A[a]	N/A
Gender		
Male	1	
Female	0.30	<.001
Having a spouse		
Yes	1	
No	0.92	.78
Educational level		
Elementary school and lower	1	
Middle school degree	N/A	N/A
High school degree	4.05	.18
College degree and higher	5.40	.10
Occupation		
Blue collar	1	
White collar	2.19	.03
Student	2.57	.15
Other jobs	1.91	.14
No job	1.42	.43
Patient type		
New	1	
Returning	1.32	.59
Number of chronic diseases[b]		
0	1	
1	1.39	.23
2	1.37	.36
3	2.16	.03
4	1.88	.15
≥5	1.73	.19
Departments		
Other departments	1	
Cancer center	0.76	.46
Internal medicine or family medicine	0.88	.61
Neurovascular center	0.85	.67
Surgery	1.40	.24

[a]N/A: not applicable.

[b]Chronic diseases include diabetes mellitus, dyslipidemia, hypertension, and any type of cancer.

Table 6. Multivariable analysis of the factors associated with the use of the PDF download function in the Health4U app.

Variable	Adjusted odds ratio	*P* value
Age (years)		
19-40	1	
41-60	0.75	.15
60-80	0.41	.001
≥80	0.21	.006
Gender		
Male	1	
Female	0.8	.17
Having a spouse		
Yes	1	
No	1.61	.009
Educational level		
Elementary school and lower	1	
Middle school degree	0.67	.49
High school degree	1.46	.37
College degree and higher	2.31	.046
Occupation		
Blue collar	1	
White collar	1.60	.049
Student	2.25	.06
Other jobs	1.07	.80
No job	0.71	.26
Patient type		
New	1	
Returning	0.79	.41
Number of chronic diseases[a]		
0	1	
1	3.53	<.001
2	4.18	<.001
3	4.26	<.001
4	4.45	<.001
≥5	6.25	<.001
Departments		
Other departments	1	
Cancer center	0.77	.29
Internal medicine or family medicine	0.78	.17
Neurovascular center	0.79	.40
Surgery	0.75	.24

[a]Chronic diseases include diabetes mellitus, dyslipidemia, hypertension, and any type of cancer.

Table 7. Multivariable analysis of the factors associated with the use of the Consolidated Clinical Document Architecture integration function in the Samsung S-Health app.

Variable	Adjusted odds ratio	*P* value
Age (years)		
19-40	1	
41-60	3.00	.17
60-80	1.89	.50
≥80	N/A[a]	N/A
Gender		
Male	1	
Female	0.59	.33
Having a spouse		
Yes	1	
No	2.25	.14
Educational level		
Elementary school and lower	1	
Middle school degree	0.47	.48
High school degree	0.78	.45
College degree and higher	N/A	N/A
Occupation		
Blue collar	1	
White collar	0.89	.85
Student	2.48	.50
Other jobs	0.91	.90
No job	0.40	.30
Patient type		
New	1	
Returning	1.12	.91
Number of chronic diseases[b]		
0	1	
1	1.09	.14
2	1.04	.05
3	1.23	.25
4	1.82	.70
≥5	1.96	.88
Departments		
Other departments	1	
Cancer center	0.69	.63
Internal medicine or family medicine	1.26	.64
Neurovascular center	0.44	.44
Surgery	0.31	.27

[a]N/A: not applicable.

[b]Chronic diseases include diabetes mellitus, dyslipidemia, hypertension, and any type of cancer.

Table 8. Multivariable analysis of the association of daily step sync function with other functions. In this analysis, the outcome variable was daily steps synced.

Variable	Adjusted odds ratio	*P* value
Age (years)		
19-40	1	
41-60	0.62	.12
60-80	0.16	<.001
≥80	N/A[a]	N/A
Gender		
Male	1	
Female	0.36	.001
Having a spouse		
Yes	1	
No	0.67	.27
Educational level		
Elementary school and lower	1	
Middle school degree	N/A	N/A
High school degree	4.79	.22
College degree and higher	7.45	.12
Occupation		
Blue collar	1	
White collar	2.07	.097
Student	2.52	.24
Other jobs	1.93	.20
No job	1.58	.43
Patient type		
New	1	
Returning	1.41	.59
Departments		
Other departments	1	
Cancer center	0.68	.37
Internal medicine or family medicine	0.64	.18
Neurovascular center	0.51	.18
Surgery	0.87	.68
PDF download		
No	1	
Yes	7.80	<.001
C-CDA[b] download		
No	1	
Yes	18.76	<.001
Blood pressure integration		
No	1	
Yes	18.66	<.001

Variable	Adjusted odds ratio	*P* value
Blood sugar integration		
No	1	
Yes	4.68	.02
Weigh integration		
No	1	
Yes	41.31	<.001

[a]N/A: not applicable.

[b]C-CDA: Consolidated Clinical Document Architecture

Discussion

Overview

To the best of our knowledge, this is be the first study to link mobile phone-based health care platforms such as Samsung S-Health and Apple Health to EMR-tethered PHRs to collect lifelog data. We hypothesized that there can be large hurdles despite our efforts to integrate lifelog data directly into PHRs from the Samsung S-Health and Apple Health apps. We also hypothesized that the adoption rate of new functions, such as lifelog integration and medical document downloads, depends on a user's age, gender, occupation, educational status, and number of chronic conditions. We analyzed the results of daily steps integration as a representative marker of lifelog data because it can be generated easily from wearables and smartphones and integrated effortlessly into PHRs.

We found that the use rate of lifelog integration is very low, given the total number of registered users during the study period: 1.22% for daily steps, 0.99% for weight, 0.29% for blood sugar, and 0.51% for blood pressure (Table 2). However, we already presented a hypothesis based on the results of previous research that the use rate of the self-administration function in EMR-tethered PHRs can be lowered if patients have difficulties inputting data into PHRs [14]. A previous study also revealed that the successful adoption of PHRs depends on the patient-clinician relationship and the promotion of the technology to physicians [8]. If we promote the use of self-administered functions that can be integrated automatically from smartphones among patients, the results might be different. Although the use rate for lifelog integration was low, we found interesting results. Patients who integrated daily steps used the Samsung S-Health and Apple Health apps more than the self-administration function. Compared to daily steps integration, weight, blood sugar, and blood pressure data were integrated by patients who input their data manually. If users are registered to access their PHRs and launch the Health4U app, data on their daily steps are automatically transferred from their smartphones to Health4U, thus making it unnecessary for users to remember to integrate their activity data into Health4U. We concluded that the difference between the integration of daily steps and that of other lifelog data is attributable to the different methods by which patient data are collected in PHRs. We proved in a previous study that health outcomes can be improved if we integrate lifelog data into EMR-tethered PHRs in order to enable doctors to provide recommendations based on shared lifelog records [16]. However, we also found that patients usually do not want to put on wearables, such as smart watches, for long periods of time. Therefore, collecting activity data from only smartphones and adjusting these data with possible real-world activity data can be a good option for patients and health care professionals to obtain continuous activity data.

We also found that daily steps integration was associated with gender and the age group of ≥60 years. Women and people aged ≥60 years were less likely to integrate their daily steps data. These results are similar to those of previous studies. Jung et al [14] found that women are less likely than men to use self-administered functions in EMR-tethered PHRs [14]. Studies found that men were more likely to consider computer use enjoyable, be confident about using the internet and PHRs, and be engaged users of PHRs as compared to women [20,21]. We were unable to find consistent results for age groups with statistical significance, but the odds ratios were consistently decreased when patient groups were older: The group of adults aged ≥60 years was 61% less likely to integrate daily steps data into PHRs. To integrate daily steps into the Health4U app from the Samsung S-Health or Apple Health app, users must launch the S-Health or Apple Health app at least once. It is possible that the elderly are not interested in health-related functions installed on their smartphones. As shown above, daily steps can be integrated more easily from the S-Health and Apple Health apps than from self-administration; the age barrier for lifelog integration can be addressed by encouraging patients to use health-related preinstalled apps on their smartphones.

Finally, we found that the more chronic diseases a patient has, the more frequently the patient uses the PDF download function. To our knowledge, this is the first study to show that the PDF download function for medical information in EMR-tethered PHRs is associated with the number of chronic conditions. This finding suggests that patients with multiple chronic diseases want to maintain their medical records independently. In 2015, the Office of the National Coordinator for Health IT issued a paper [22] wherein they proposed that future PHRs should be based on HIE because a PHR tethered to a single institution is not sufficient for managing diseases, as patients usually visit multiple hospitals and institutions. Another study showed that willingness to enroll in PHRs was associated with the presence of doctor-diagnosed chronic diseases [23]. One other study found that HIE can decrease the number of future encounters and future readmissions among patients with chronic diseases

[24]. Our findings, in conjunction with the results of these previous studies, suggest that patients with multiple chronic diseases are already aware of the importance of the HIE and PHR functions for multiple chronic conditions, and these functions may help improve health outcomes, if implemented appropriately. For example, patients may become interested in systemically summarized notes about their chronic health conditions.

Participants in the age groups of ≥60 years and ≥80 years had a lower tendency to use the PDF download function. The result of the lower use rate among these older age groups is similar to that of the lower use rate of daily steps integration, indicating that older age can be a hurdle to using advanced PHR functions. A previous study revealed that usability concerns are barriers to PHR adoption and use [25]. Other research proved that older people are likely to have lower literacy for new technology compared to younger people [26]. Therefore, to increase the use rate of lifelog integration functions, we must consider usability and technology literacy issues. White-collar workers tend to use the function 60% more often than blue-collar workers. White-collar workers today primarily perform their jobs on computers and portable devices [27], which may account for the difference in these use rates. Users with college degrees and higher education used the function 2.3 times more often than users who finished only elementary school. Use rates may also be related to technology literacy issues. To promote the use of PHR functions for HIE, we must consider the use of advertisements targeted toward groups with different educational levels and occupational backgrounds.

Limitations

This research is a single-center retrospective cross-sectional study, which makes it difficult to identify causal relationships, and the study lacks accurate information on improvement in the health outcomes of PHR users who integrated lifelog data. However, this study found characteristics of users related to lifelog integration and medication documentation downloads. Based on these findings, future studies can be performed effectively.

There are limitations to generalizing the results of this study because it involved only one tertiary hospital and the participation rate was low. However, because this study focused on the use of novel functions implemented in EMR-tethered PHRs at a large hospital, the results will serve as an important background for medical institutions intending to develop similar features or for national agencies planning to develop HIE-based PHRs.

Conclusions

This is the first study to identify factors related to the integration of daily steps from Samsung S-Health and Apple Health apps into EMR-tethered PHRs and the factors related to the retrieval function of medical documents from PHRs. The finding that patients with more chronic diseases tend to download their medical information more frequently can serve as the basis for enhancement of the features of an EMR-tethered PHR system for HIE. Additionally, findings on lifelog data integration can be used to design PHRs as a platform to integrate lifelog data in the future.

Acknowledgments

This research is supported by the SNUBH Research Fund (Grant #18-2018-007). This work was also supported by the Creative Industrial Technology Development Program (Grant #10053249, the Development of Personalized Healthcare System Exploiting User Life-Log and Open Government Data for Business Service Model Proof on Whole Life Cycle Care), funded by the Ministry of Trade, Industry & Energy (MOTIE, Korea). The data used in this study cannot be shared due to a policy of the Institutional Review Board of SNUBH.

Authors' Contributions

SYJ and J-WK analyzed the data and drafted the manuscript as first authors. WS contributed to the data analysis. R-MB, KL, SY, and H-YL contributed to the discussion of data. HH supervised the entire process as a corresponding author. We give special thanks to Hyun-ok Bae and Mi-young Kim, who contributed to data acquisition and are not included in the list of authors. We are also very grateful to John Chaesoo Oh, who dedicated time to the development of Health4U.

Conflicts of Interest

None declared.

References

1. Centers for Disease Control. National Center for Chronic Disease Prevention and Health Promotion URL: https://www.cdc.gov/chronicdisease/index.htm [accessed 2018-11-05] [WebCite Cache ID 73h8Ha8jk]
2. Pereira M, Lunet N, Azevedo A, Barros H. Differences in prevalence, awareness, treatment and control of hypertension between developing and developed countries. J Hypertens 2009 May;27(5):963-975. [Medline: 19402221]
3. Hardy GE. The burden of chronic disease: the future is prevention. Introduction to Dr. James Marks' presentation, "The Burden of Chronic Disease and the Future of Public Health". Prev Chronic Dis 2004 Apr;1(2):A04 [FREE Full text] [Medline: 15663880]

4. Gaziano TA, Bitton A, Anand S, Weinstein MC, International Society of Hypertension. The global cost of nonoptimal blood pressure. J Hypertens 2009 Jul;27(7):1472-1477. [doi: 10.1097/HJH.0b013e32832a9ba3] [Medline: 19474763]
5. The Boston Foundation. Healthy People/Healthy Economy coalition seeks to make Massachusetts a national leader in health and wellness URL: https://tinyurl.com/y2lhstau [accessed 2019-05-16] [WebCite Cache ID 78PFW164a]
6. Tattersall R. The expert patient: a new approach to chronic disease management for the twenty-first century. Clin Med (Lond) 2002;2(3):227-229 [FREE Full text] [Medline: 12108472]
7. Mills KT, Bundy J, Kelly T, Reed JE, Kearney PM, Reynolds K, et al. Global Disparities of Hypertension Prevalence and Control: A Systematic Analysis of Population-Based Studies From 90 Countries. Circulation 2016 Aug 09;134(6):441-450 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.018912] [Medline: 27502908]
8. Kerns JW, Krist AH, Longo DR, Kuzel AJ, Woolf SH. How patients want to engage with their personal health record: a qualitative study. BMJ Open 2013 Jul 30;3(7) [FREE Full text] [doi: 10.1136/bmjopen-2013-002931] [Medline: 23901027]
9. 2020health. Personal Health Records - Putting patients in control? URL: http://www.2020health.org/2020health/Publications/publications-2012/Public-Health-Records.html [accessed 2019-05-16] [WebCite Cache ID 78PFvLzzg]
10. Rudd P, Frei T. How personal is the personal health record?: comment on "the digital divide in adoption and use of a personal health record". Arch Intern Med 2011 Mar 28;171(6):575-576. [doi: 10.1001/archinternmed.2011.35] [Medline: 21444848]
11. Brittany V, Brent T, Jena S. Marshall University Marshall Digital Scholar. 2015 Mar 01. Benefits and Barriers for Adoption of Personal Health Records URL: https://mds.marshall.edu/cgi/viewcontent.cgi?article=1135&context=mgmt_faculty
12. Google: Official Blog. 2011. An update on Google Helath and Google PowerMeter URL: https://googleblog.blogspot.com/2011/06/update-on-google-health-and-google.html [accessed 2018-10-25] [WebCite Cache ID 73h8P8XsE]
13. Woods SS, Schwartz E, Tuepker A, Press NA, Nazi KM, Turvey CL, et al. Patient experiences with full electronic access to health records and clinical notes through the My HealtheVet Personal Health Record Pilot: qualitative study. J Med Internet Res 2013 Mar 27;15(3):e65 [FREE Full text] [doi: 10.2196/jmir.2356] [Medline: 23535584]
14. Jung SY, Lee K, Hwang H, Yoo S, Baek HY, Kim J. Support for Sustainable Use of Personal Health Records: Understanding the Needs of Users as a First Step Towards Patient-Driven Mobile Health. JMIR Mhealth Uhealth 2017 Feb 23;5(2):e19 [FREE Full text] [doi: 10.2196/mhealth.6021] [Medline: 28232300]
15. 15 H. HealthIT.gov. 2018. What are patient-generated health data? URL: https://www.healthit.gov/topic/otherhot-topics/what-are-patient-generated-health-data [accessed 2018-11-05] [WebCite Cache ID 73h8TZNu4]
16. Ryu B, Kim N, Heo E, Yoo S, Lee K, Hwang H, et al. Impact of an Electronic Health Record-Integrated Personal Health Record on Patient Participation in Health Care: Development and Randomized Controlled Trial of MyHealthKeeper. J Med Internet Res 2017 Dec 07;19(12):e401 [FREE Full text] [doi: 10.2196/jmir.8867] [Medline: 29217503]
17. Ro HJ, Jung S, Lee K, Hwang H, Yoo S, Baek H, et al. Establishing a Personal Health Record System in an Academic Hospital: One Year's Experience. Korean J Fam Med 2015 May;36(3):121-127 [FREE Full text] [doi: 10.4082/kjfm.2015.36.3.121] [Medline: 26019761]
18. Yoo S, Lee KH, Lee HJ, Ha K, Lim C, Chin HJ, et al. Seoul National University Bundang Hospital's Electronic System for Total Care. Healthc Inform Res 2012 Jun;18(2):145-152 [FREE Full text] [doi: 10.4258/hir.2012.18.2.145] [Medline: 22844650]
19. Yoo S, Jung SY, Kim S, Kim E, Lee K, Chung E, et al. A personalized mobile patient guide system for a patient-centered smart hospital: Lessons learned from a usability test and satisfaction survey in a tertiary university hospital. Int J Med Inform 2016 Jul;91:20-30. [doi: 10.1016/j.ijmedinf.2016.04.003] [Medline: 27185506]
20. Luque AE, van Keken A, Winters P, Keefer MC, Sanders M, Fiscella K. Barriers and Facilitators of Online Patient Portals to Personal Health Records Among Persons Living With HIV: Formative Research. JMIR Res Protoc 2013 Jan 22;2(1):e8 [FREE Full text] [doi: 10.2196/resprot.2302] [Medline: 23612564]
21. Chrischilles EA, Hourcade J, Doucette W, Eichmann D, Gryzlak B, Lorentzen R, et al. Personal health records: a randomized trial of effects on elder medication safety. J Am Med Inform Assoc 2014;21(4):679-686 [FREE Full text] [doi: 10.1136/amiajnl-2013-002284] [Medline: 24326536]
22. Venesco LLC. HealthIT.gov. 2015 Sep 28. Final Report HIEs and Personal Health Records Community of Practice: Key Considerations for HIE-based Personal Health Records URL: https://www.healthit.gov/sites/default/files/phrkeyconsiderations.pdf [accessed 2019-05-16] [WebCite Cache ID 78PGxYjP4]
23. Zhang P, Zhang L, Wang F, Cheng Y, Liang Y. Societal and individual determinants in the enrollment of personal health records: A preliminary investigation from China. Int J Health Plann Manage 2019 Jan;34(1):e752-e762. [doi: 10.1002/hpm.2688] [Medline: 30350379]
24. Atasoy H, Demirezen E, Chen P. The Value of Health Information Exchange Use for Chronic Disease Management. SSRN Journal 2018 Jan 21 [FREE Full text] [doi: 10.2139/ssrn.3191566]
25. Showell C. Barriers to the use of personal health records by patients: a structured review. PeerJ 2017;5:e3268 [FREE Full text] [doi: 10.7717/peerj.3268] [Medline: 28462058]
26. Parker RM, Ratzan S, Lurie N. Health literacy: a policy challenge for advancing high-quality health care. Health Aff (Millwood) 2003;22(4):147-153. [Medline: 12889762]

27. Wikipedia. 2018. Digital literacy URL: https://en.wikipedia.org/wiki/Digital_literacy [accessed 2018-11-05] [WebCite Cache ID 73h8XQBdl]

Abbreviations

C-CDA: Consolidated Clinical Document Architecture
EMR: electronic medical record
HIS: hospital information system
PGHD: Patient-generated health data
PHR: personal health record
SNUBH: Seoul National University Bundang Hospital
TFT: task-force team

Edited by G Eysenbach; submitted 05.11.18; peer-reviewed by S Li, FC Lin; comments to author 12.12.18; revised version received 06.05.19; accepted 12.05.19; published 28.05.19.

Please cite as:
Jung SY, Kim JW, Hwang H, Lee K, Baek RM, Lee HY, Yoo S, Song W, Han JS
Development of Comprehensive Personal Health Records Integrating Patient-Generated Health Data Directly From Samsung S-Health and Apple Health Apps: Retrospective Cross-Sectional Observational Study
JMIR Mhealth Uhealth 2019;7(5):e12691
URL: http://mhealth.jmir.org/2019/5/e12691/
doi:10.2196/12691
PMID:31140446

Original Paper

A Mobile Lifestyle Management Program (GlycoLeap) for People With Type 2 Diabetes: Single-Arm Feasibility Study

David Koot[1], MBBS, MMed; Paul Soo Chye Goh[1], MBBS, MMed; Robyn Su May Lim[2], PhD; Yubing Tian[2], MSc; Teng Yan Yau[3], MBBS, MMed; Ngiap Chuan Tan[1], MBBS, MMed; Eric Andrew Finkelstein[2], MHA, PhD

[1]SingHealth Polyclinics, Singapore, Singapore

[2]Health Services and Systems Research, Duke-NUS Medical School, Singapore, Singapore

[3]KKT Technology Pte Ltd, Singapore, Singapore

Corresponding Author:
Eric Andrew Finkelstein, MHA, PhD
Health Services and Systems Research
Duke-NUS Medical School
8 College Road
Singapore,
Singapore
Phone: 65 6516 2338
Fax: 65 6534 8632
Email: eric.finkelstein@duke-nus.edu.sg

Abstract

Background: Singapore's current prevalence of diabetes exceeds 13.6%. Although lifestyle modification can be effective for reducing the risks for complications of type 2 diabetes mellitus (T2DM), traditional lifestyle interventions are often difficult to administer in the primary care setting due to limited resources. Mobile health apps can address these limitations by offering low-cost, adaptable, and accessible platforms for disseminating lifestyle management interventions.

Objective: Using the RE-AIM evaluation framework, this study assessed the potential effectiveness and feasibility of GlycoLeap, a mobile lifestyle management program for people with T2DM, as an add-on to standard care.

Methods: This single-arm feasibility study recruited 100 patients with T2DM and glycated hemoglobin (HbA_{1c}) levels of ≥7.5% from a single community health care facility in Singapore. All participants were given access to a 6-month mobile lifestyle management program, GlycoLeap, comprising online lessons and the Glyco mobile phone app with a health coaching feature. The GlycoLeap program was evaluated using 4 relevant dimensions of the RE-AIM framework: (1) reach (percentage who consented to participate out of all patients approached), (2) effectiveness (percentage point change in HbA_{1c} [primary outcome] and weight loss [secondary outcome]), (3) implementation (program engagement as assessed by various participatory metrics), and (4) maintenance (postintervention user satisfaction surveys to predict the sustainability of GlycoLeap). Participants were assessed at baseline and at follow-up (≥12 weeks after starting the intervention).

Results: A total of 785 patients were approached of whom 104 consented to participate, placing the reach at 13.2%. Four were excluded after eligibility screening, and 100 patients were recruited. Program engagement (implementation) started out high but decreased with time for all evaluated components. Self-reported survey data suggest that participants monitored their blood glucose on more days in the past week at follow-up compared to baseline (P<.001) and reported positive changes to their diet due to app engagement (P<.001) (implementation). Primary outcome data were available for 83 participants. Statistically significant improvements were observed for HbA_{1c} (–1.3 percentage points, P<.001) with greater improvements for those who logged their weight more often (P=.007) (effectiveness). Participants also had a 2.3% reduction in baseline weight (P<.001) (effectiveness). User satisfaction was high with 74% (59/80) and 79% (63/80) of participants rating the app good or very good and claiming that they would probably or definitely recommend the app to others, respectively (maintenance).

Conclusions: Although measures of program engagement decreased with time, clinically significant improvements in HbA_{1c} were achieved with the potential for broader implementation. However, we cannot rule out that these improvements were due to factors unrelated to GlycoLeap. Therefore, we would recommend evaluating the effectiveness and cost effectiveness of GlycoLeap using a randomized controlled trial of at least 12 months.

Trial Registration: ClinicalTrials.gov NCT03091517; https://clinicaltrials.gov/ct2/show/NCT03091517 (Archived by WebCite at http://www.webcitation.org/77rNqhwRn)

(JMIR Mhealth Uhealth 2019;7(5):e12965) doi:10.2196/12965

KEYWORDS

type 2 diabetes mellitus; self-management; mobile health; mHealth; mobile phone app; mobile apps; health coaching; blood glucose; single-arm feasibility study; RE-AIM

Introduction

Background

The worldwide prevalence of diabetes has almost doubled over the past three decades and continues to rise [1]. Singapore's current prevalence of diabetes exceeds 13.6% [2], and forecasts predict that without successful interventions, the lifetime risk of developing type 2 diabetes mellitus (T2DM) in Singapore will be 1 in 2 by 2050 [3].

Lifestyle modification can be highly effective at reducing the risks for complications of T2DM [4], yet traditional lifestyle interventions are often difficult to administer in the primary care setting due to limited resources and infrequent patient interaction with health care personnel. Mobile health (mHealth) apps can address these limitations by offering low-cost, highly adaptable, and easily accessible platforms for disseminating lifestyle management interventions. Effective T2DM apps stress skill building; self-efficacy; and frequent monitoring of blood glucose levels, weight, dietary intake, and physical activity [5-9]. Several diabetes management mHealth interventions also provide participants with some level of personalized communication with health professionals, including real-time feedback [10-16].

Given the rising health and cost implications of diabetes in Singapore, a comprehensive program that incorporates the most effective mHealth strategies and personalized health coaching offers a potentially scalable model to address Singapore's diabetes epidemic. One potential program is GlycoLeap (Holmusk), a proprietary lifestyle management program for adults with T2DM. The GlycoLeap program, which was originally developed for use in Singapore, comprises two components: a comprehensive T2DM educational curriculum delivered through online lessons and the Glyco mobile phone app with a health coaching feature. The Glyco app enables users to log and monitor their blood glucose levels, weight, meals, and physical activity, which is captured via the mobile phone's built-in pedometer. The app also serves as a vehicle for accredited dietitians, known as health coaches, to provide personalized feedback to participants on their progress and to present opportunities for improvement.

Goal of the Study

The goal of this study was to assess the potential effectiveness and feasibility of the GlycoLeap program as an add-on to standard care using the RE-AIM evaluation framework [17,18]. If results are promising, a 2-arm randomized controlled trial aimed to test effectiveness and cost effectiveness will be recommended.

Methods

Research Design

This effort consisted of a 6-month (24-week), single-arm, preintervention (baseline) and follow-up evaluation. One hundred participants were recruited by a research coordinator from a single community health care facility, SingHealth Polyclinics–Tampines, in Singapore. All participants signed an informed consent to participate in a baseline and follow-up assessment that was conducted at least 12 weeks after starting the intervention and to allow for medical records abstraction. At the baseline assessment, participants completed a brief survey (Multimedia Appendix 1) that captured demographic information, diabetes status, and self-care activity information (from the Summary of Diabetes Self-Care Activities scale [19]). At the follow-up, participants completed a survey (Multimedia Appendix 2) that captured self-reported changes in diabetes self-care [19], dietary consumption, physical activity, program engagement, and user satisfaction. Glycated hemoglobin (hemoglobin A_{1c}, or HbA_{1c}) levels, medications, and weight were measured or obtained from medical records at baseline and follow-up.

Participants were eligible to participate if they (1) were aged 21 to 70 years, (2) had been medically diagnosed with T2DM as listed in electronic health records, (3) had an HbA_{1c} result of ≥7.5% within the past 2 months, (4) had a body mass index (BMI) of >23 kg/m^2, (5) were not on insulin, and (6) owned and were able to use an iPhone or Android mobile phone. Participants were excluded if they (1) had cancer requiring treatment in the past 5 years, (2) had cardiovascular diseases (heart attack or cardiac procedure within the past 3 months), (3) had stroke or history of treatment for transient ischemic attacks in the past 3 months, (4) had chronic renal failure or were on dialysis, (5) had any amputation of lower limbs, (6) were using medication for weight loss, (7) had chronic treatment with systemic corticosteroids, (8) had bariatric surgery or extensive bowel resection, (9) were unable to converse in or read and write English, or (10) did not have a valid HbA_{1c} blood test within the 2 months prior to the date of recruitment. Each participant was compensated SGD20 (approximately US $15) in vouchers at study completion. All study procedures were approved by the SingHealth Centralized Institutional Review Board (CIRB Ref: 2017/2013) and the study is registered at ClinicalTrials.gov [NCT03091517].

Intervention Program

All eligible participants received 24 weeks of free, unlimited access to the GlycoLeap program. Participants downloaded the Glyco app onto their mobile phones upon recruitment. An

Accu-Chek Performa (F. Hoffmann–La Roche Ltd) glucometer kit with lancets and test strips, a BodyTrace (BodyTrace Inc) wireless weighing scale, and a resistance band for strength training were provided at no cost to participants. Although the recommendation to switch glucometers may have been an inconvenience, this was not raised as a concern by any of the participants. Participants were also given printed instruction guides on operating the Glyco app and devices and two guidebooks educating them on how to make healthier food choices and achieve weight loss. Table 1 describes key components of the GlycoLeap program and the recommended frequencies of engagement. Screenshots illustrating the Glyco app user interface are presented in Multimedia Appendices 3-7.

Table 1. Description of GlycoLeap program components and recommendations for engagement.

Component	Description	Recommended frequency
Online health lessons on diabetes and self-management	A total of 24 educational lessons on diabetes and self-management were delivered online. This curriculum was adapted for the local population and covers topics that take reference from the 7 healthy self-care behaviors as described by the American Association of Diabetes Educators. Quizzes tested knowledge on diabetes, obtained information about participants' lifestyle habits, and were designed to keep participants engaged throughout each lesson.	Complete one lesson (lasting about 15 minutes) per week
Blood glucose monitoring	Blood glucose measurements obtained using the Accu-Chek Performa glucometer kit were input manually by participants into their Glyco app accounts.	At least 4 blood glucose logs per week (preferably paired pre- and postmeal readings)
Weight monitoring	Wireless weighing scale readings were automatically synced to participants' accounts via cellular connectivity (3G).	At least one weight log per week
Meal logging	Meal photos taken by participants were uploaded onto the app for health coach evaluation. Health coaches rate meals using a 1 to 5 linear scale. Meal scores are awarded based on the balance of nutrients, food quality, and nutritional content. The meal scores take reference from the Singapore Health Promotion Board's national dietary guidelines.	No recommendation was provided. Participants were encouraged to log as often as they wanted to.
Physical activity tracking	The Glyco app tracks the number of daily steps taken using the participants' built-in phone pedometers.	70,000 steps per week
Health coach	Health coaches rate and respond to all meal logs and regularly send messages to participants to provide recommendations, encouragement, and personalized feedback on progress and answer participants' questions (Multimedia Appendices 6 and 7). Correspondence is 2-way and participants are free to initiate or respond to messages from their health coach. All correspondence is conducted in-app, and all participants receive health coaching regardless of whether they send any messages to their health coach. Coaches capitalize on modalities that have been shown to lead to effective behavior change [20,21], including several theoretical frameworks [22-25], and are trained in the nutrition care process [26].	No recommendation was provided. Participants were encouraged to engage as often as they wanted to.

Program Evaluation

The GlycoLeap program was evaluated using relevant dimensions of the RE-AIM framework [17,18].

Reach

A proxy for Reach was used and defined as the percentage of those who gave informed consent to participate out of all patients approached.

Effectiveness

As this was a feasibility study without a control group, measures of potential effectiveness were assessed as changes in HbA_{1c} levels (primary outcome) and weight (secondary outcome) between baseline and the follow-up. The analysis was conducted separately on the total sample with baseline values carried forward for those with missing data at follow-up (intention-to-treat [ITT] analysis) and on those who completed the study and did not initiate insulin (per-protocol analysis). All HbA_{1c} tests were conducted using the polyclinic's protocols and approved laboratories, and weight was measured using validated weighing scales at the polyclinic. The window for eligible tests was defined as within 2 months before the scheduled baseline assessment and from 12 weeks to a maximum of 38 weeks after the intervention start date for the follow-up. In addition to the above indicators of potential effectiveness, the association between changes in health outcomes and measures of program engagement (as defined in the Implementation section) were evaluated using the per-protocol approach with the expectation that participants with greater levels of program engagement will show greater improvements in health outcomes.

Adoption

Adoption tends to focus more on system level factors and was not captured as part of this feasibility study.

Implementation

Implementation was assessed on the total sample by exploring engagement with key components of the GlycoLeap program as shown in Table 1. This includes the following process measures: (1) number of online health lessons completed, (2)

number of blood glucose measurements logged per week, (3) number of weight measurements logged per week, (4) number of meal logs per week, and (5) number of messages sent by participants to health coaches per week. We did not evaluate results from the physical activity component because it would be an inaccurate reflection of activity given that participants were not expected to carry their mobile phones with them throughout the day and when exercising. We present the percentage of participants that engaged with the components over the intervention period. Self-reported changes in diabetes self-care activities and changes in lifestyle-related behaviors such as measures of dietary consumption and physical activity were assessed from survey responses. See Multimedia Appendices 1 and 2 for the baseline and follow-up surveys, respectively, which include the full list of self-reported measures.

Maintenance

Within-trial maintenance is included in the Implementation domain. Here, we assess sustainability at the setting level by exploring user satisfaction at follow-up (Multimedia Appendix 2) including ease of use, perceived value of each program component, an overall rating of the program, and whether or not participants would recommend the app to others. The idea is that the program has little chance of broader sustainability if it performs poorly in these measures.

Sample Size and Statistical Analysis

Power calculations were not performed because this was a feasibility study. The target sample size of 100 was selected taking into consideration the aims of the study (testing acceptability and potential for effectiveness rather than efficacy) and practical feasibility. Studies on diabetes self-management mobile app interventions often involved a sample size of less than 100, including controlled trials [8,27].

Paired or one-sample *t* tests were conducted to assess for differences in means in HbA_{1c} and weight outcomes from baseline to follow-up. Associations between changes in health outcomes and measures of program engagement were assessed using linear regression, controlling for baseline health outcomes and for age, gender, and ethnicity. For all program engagement explanatory variables, participants were categorized into more engaged users if their total number of logs, messages, or lessons completed were above the median. Conversely, participants were categorized into less engaged users if their total number of logs, messages, or lessons completed were equal to or below the median. Analyses were performed on Stata/MP version 14.0 (StataCorp LLC) and R version 3.5.1 (The R Foundation).

Results

Reach and Sample Statistics

Between June and November 2017, a total of 785 SingHealth Polyclinics–Tampines patients were approached by the research coordinator. Of these patients, 681 declined to participate or were deemed ineligible based on a prescreen assessment, placing the Reach at 13.2% (Figure 1). Although we did not record the reasons for prescreen failures, anecdotal information from the research coordinator suggests that the majority of patients were willing to participate but were assessed to be ineligible during prescreening due to insulin treatment, English illiteracy, or did not own a mobile phone. A total of 100 participants were recruited and all were given access to the 24-week GlycoLeap program.

The average age of participants was 54 years old, and 50 were male. Of the 100 participants, 45 were Chinese, 29 were Malay, 18 were Indian, and 8 were of other ethnicity. Sixty-one participants had high school-equivalent or lower education, and 69 were employed. At baseline, the mean HbA_{1c} and weight were 8.8% and 79.7 kg, respectively. On average, participants were diagnosed with T2DM 9.3 years ago. Other baseline characteristics can be found in Table 2.

Thirteen of the 100 participants withdrew from the study, either due to insulin initiation or free will (Figure 1). Of the 87 that reached follow-up, 4 were excluded from the per-protocol analysis as they either did not take a follow-up HbA_{1c} test or took their test outside the window (Figure 1). No statistically significant differences were found in baseline characteristics between the total sample (n=100, included in ITT analysis) and the completes (n=83, included in per-protocol analysis) (Table 2). Eighty of the 100 participants completed the follow-up survey.

Figure 1. Participant recruitment and retention Consolidated Standards of Reporting Trials flow diagram.

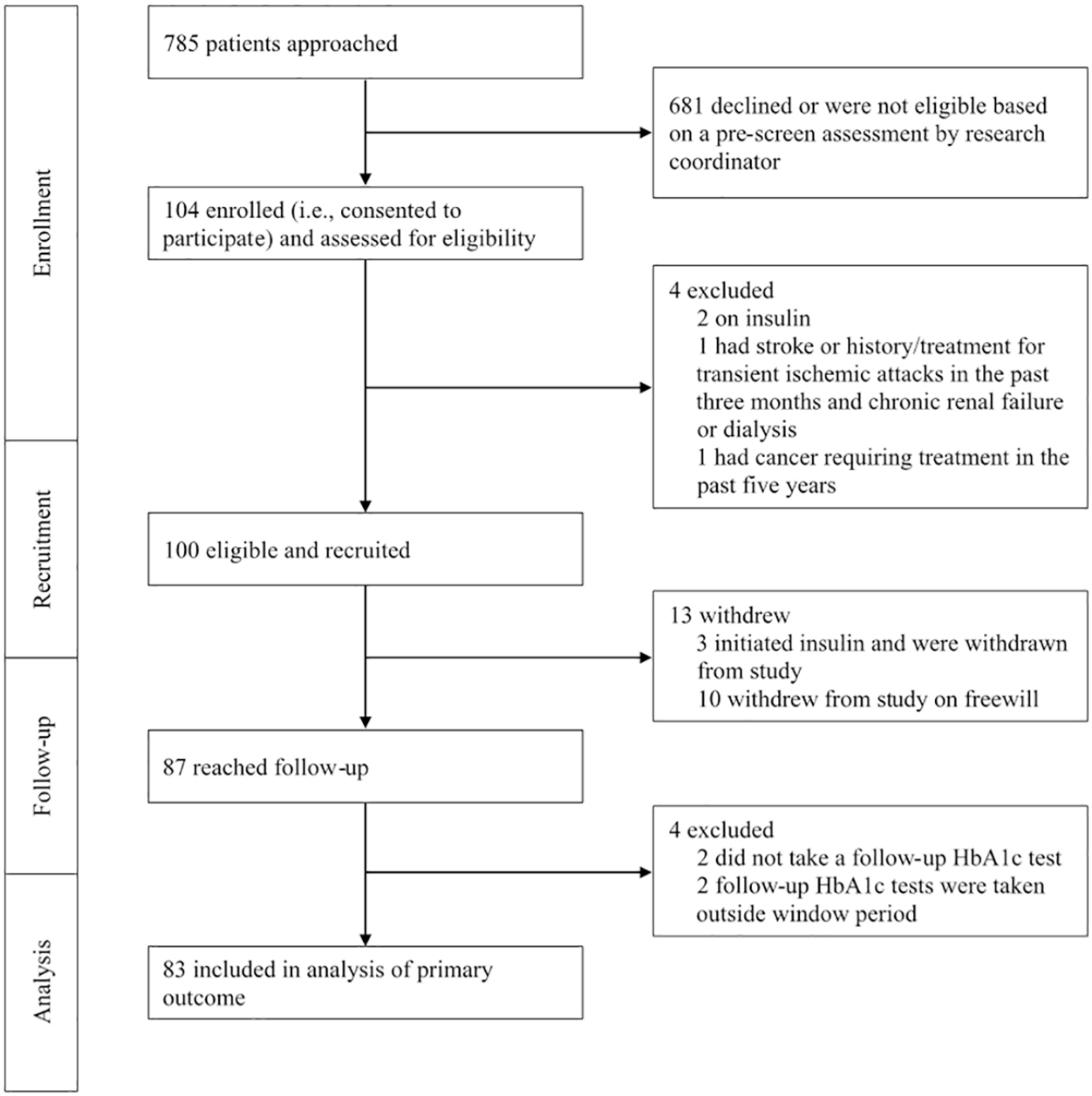


Table 2. Baseline characteristics of participants in the single-arm GlycoLeap feasibility study.

Characteristics	Total (n=100)	Completes (n=83)	*P* value
Age (years), mean (SD)	53.5 (9.6)	53.5 (9.7)	.98
Male, n (%)	50 (50)	44 (53)	.67
Weight (kg), mean (SD)	79.7 (16.8)	79.2 (15.6)	.87
Height (cm), mean (SD)	163.1 (8.9)	163.5 (9.3)	.75
Body mass index (kg/m^2), mean (SD)	29.8 (5.0)	29.5 (4.6)	.68
Systolic blood pressure (mm Hg), mean (SD)	132.1 (11.6)	132.5 (11.8)	.83
Diastolic blood pressure (mm Hg), mean (SD)	74.7 (10.6)	74.9 (10.5)	.88
HbA_{1c}[a] (%), mean (SD)	8.8 (1.6)	8.9 (1.7)	.96
Years diagnosed with diabetes, mean (SD)	9.3 (7.3)	8.8 (6.3)	.56
On oral medication for diabetes treatment, n (%)	98 (98)	81 (98)	.85
Ethnicity, n (%)			**.86**
Chinese	45 (45)	38 (46)	—
Malay	29 (29)	22 (27)	—
Indian	18 (18)	15 (18)	—
Other	8 (8)	8 (10)	—
Highest completed education level, n (%)			**.71**
High school or lower	61 (61)	49 (59)	—
Precollege (A-levels/polytechnic diploma)	21 (21)	17 (20)	—
College graduate/postgraduate	18 (18)	17 (20)	—
Current marital status, n (%)			**>.99**
Never married	12 (12)	10 (12)	—
Married	82 (82)	68 (82)	—
Other	6 (6)	5 (6)	—
Monthly household income, n (%)			**.58**
Less than SGD 5000	10 (10)	6 (7)	—
SGD 5000-SGD 9999	2 (2)	2 (2)	—
≥SGD 10,000	1 (1)	1 (1)	—
Prefer not to say	87 (87)	74 (89)	—
Employment status, n (%)			**.78**
Working (full-/part-time)	69 (69)	59 (71)	—
Homemaker[b]	22 (22)	17 (20)	—
Retired and not working[c]	9 (9)	7 (8)	—

[a]HbA_{1c}: glycated hemoglobin.

[b]Individuals who are full-time housekeepers, regardless of prior employment status.

[c]Individuals who were previously employed until retirement and are typically past retirement age.

Implementation (Program Engagement)

Generally high participant engagement was observed for all components in the first week which then decreased progressively over time. Trends in program engagement for the total sample (n=100) are shown in Figure 2. At the end of the 24-week program, a third of the participants (33/100, 33.0%) completed at least one online health lesson. On average, participants finished 9.2 lessons. Throughout the intervention period, fewer than 20 participants logged their blood glucose measurements at least 4 times a week yet more than 25 participants logged their weight measurements at least once a week, which were the recommendations. On average, participants entered 2.1 meal logs and sent 2.8 messages to their health coach each week. The mean, median, minimum, and maximum total number of messages sent by participants to their health coaches throughout

the 24-week program were 66.9, 8.5, 0, and 799, respectively. For total number of messages sent by health coaches to participants throughout the entire study period, the mean, median, minimum, and maximum were 234.9, 69, 16, and 2754, respectively.

In total, two of the 100 participants did not engage with any of the 5 evaluated components, and 14 engaged with at least one component every week throughout the intervention period. Thirteen participants engaged with the same component(s) every week throughout the intervention period: health lessons (2 participants), blood glucose monitoring (3 participants), weight monitoring (6 participants), meal logging (7 participants), and health coach messaging (5 participants).

Self-reported changes in diabetes self-care activities and lifestyle behaviors at follow-up from baseline are show in Table 3 for the 80 participants who completed the follow-up survey. Participants reported monitoring their blood glucose on more days in the week before the follow-up assessment compared to the week before starting the intervention (2.3 days [95% CI 1.9-2.7] vs 0.6 days [95% CI 0.2-1.0], *P*<.001). Similarly, 68 out of 80 participants (85%, *P*<.001) declared positive changes in diet due to app engagement. On average, participants reported consuming the recommended servings of fruit and vegetables on more days in the past week at follow-up compared to baseline (3.7 days [95% CI 3.1-4.2] vs 1.3 days [95% CI 0.8-1.8], *P*<.001) and high fat food on fewer days (1.6 days [95% CI 1.2-2.0] vs 2.3 days [95% CI 1.9-2.7], *P*=.003). Although 30 out of 80 participants (38%) claimed to have increased their weekly average level of moderate-to-vigorous physical activity due to app engagement, there was no statistically significant difference in the number of days where participants reported performing at least 30 minutes of continuous activity in the past week at follow-up compared to baseline (3.9 days [95% CI 3.4-4.4] vs 3.4 days [95% CI 2.7-4.0], *P*=.14).

Figure 2. Proportion of program engagement by week for the total sample (n=100). Percentage of participants who (A) completed at least one lesson, (B) logged 1, 2, 3, or ≥4 glucose measurements, (C) logged at least one weight log, (D) logged at least one meal log, or (E) sent at least one message to their health coach a week.

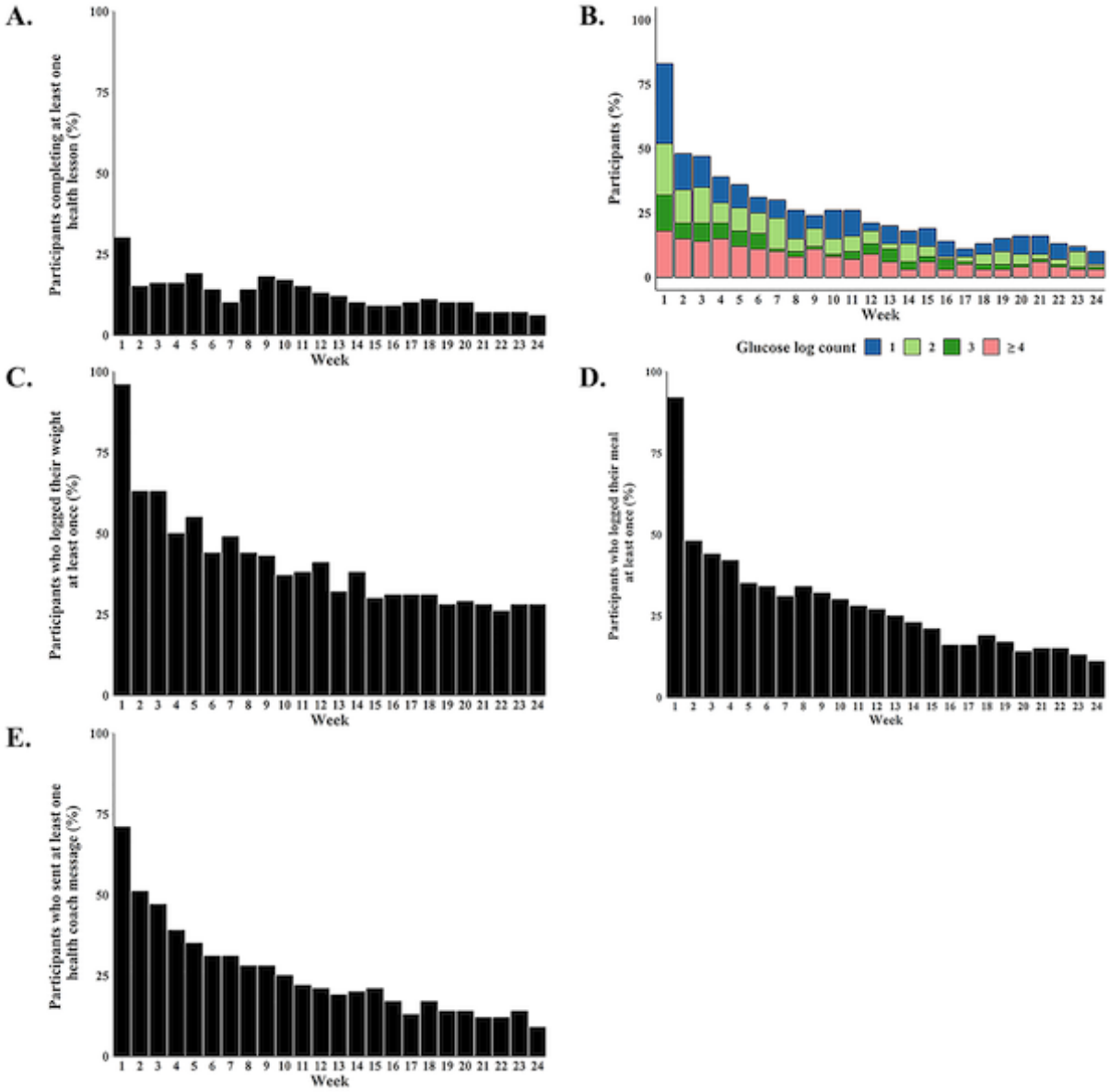


Table 3. Self-reported diabetes self-care activities and lifestyle behaviors at baseline and follow-up (n=80).

Category and behavior or activity	Baseline	Follow-up	Difference	*P* value
Blood glucose monitoring, mean (95% CI)				
Days that blood glucose was monitored in past week	0.6 (0.2 to 1.0)	2.3 (1.9 to 2.7)	1.7 (1.3 to 2.1)	<.001
Dietary habits				
Positive change in diet due to app engagement, n (%)	—[a]	68 (85)	N/A[b]	<.001
Days with fruit and vegetable consumption as per recommended servings in past week, mean (95% CI)	1.3 (0.8 to 1.8)	3.7 (3.1 to 4.2)	2.4 (1.6 to 3.1)	<.001
Days with high fat food consumption in past week, mean (95% CI)	2.3 (1.9 to 2.7)	1.6 (1.2 to 2.0)	–0.7 (–1.1 to –0.2)	.003
Physical activity, mean (95% CI)				
Days with at least 30 minutes of continuous activity including walking in past week	3.4 (2.7 to 4.0)	3.9 (3.4 to 4.4)	0.5 (–0.2 to 1.2)	.14
Change in average weekly level of moderate-to-vigorous physical activity due to app engagement, n (%)				**<.001**
Increased	—	30 (38)	N/A	
Decreased	—	0 (0)	N/A	
Stayed the same	—	50 (63)	N/A	

[a]Question was not present in the baseline survey as it asks for self-reported change due to app engagement.

[b]N/A: not applicable as question was not present in the baseline survey.

Effectiveness (Health Outcomes)

Changes in HbA_{1c} levels and weight are shown in Figure 3 and Table 4. Per-protocol analysis revealed that, on average, participants' HbA_{1c} levels were 1.3 percentage points lower at follow-up compared to baseline (7.6% [95% CI 7.2-7.9] vs 8.9% [95% CI 8.5-9.2], *P*<.001) and 49 of 83 participants (59%) achieved a ≥1 percentage point reduction in HbA_{1c} levels. The average duration between intervention start and follow-up HbA_{1c} tests was 24.2 weeks. Based on per-protocol analysis, participants achieved a weight loss of 2.3 kg at follow-up compared to baseline (77.3 kg [95% CI 74.0-80.7] vs 79.2 kg [95% CI 75.8-82.6], *P*<.001), and 17 out of 83 participants (20%) lost ≥5% of their initial body weight at baseline. Similar results and levels of statistical significance were obtained with the ITT analysis (Table 4).

Using the per-protocol approach, linear regression showed that HbA_{1c} decreased by an average of 1.0 percentage point more among those who logged their weight more (*P*=.007) (Multimedia Appendix 8). No statistically significant associations were found between change in weight and measures of program engagement (data not shown).

Figure 3. Distributions of health outcomes at baseline and follow-up: (A) HbA_{1c} for total sample (n=100), (B) HbA_{1c} for completes (n=83), (C) weight for total sample (n=100), and (D) weight for completes (n=83).

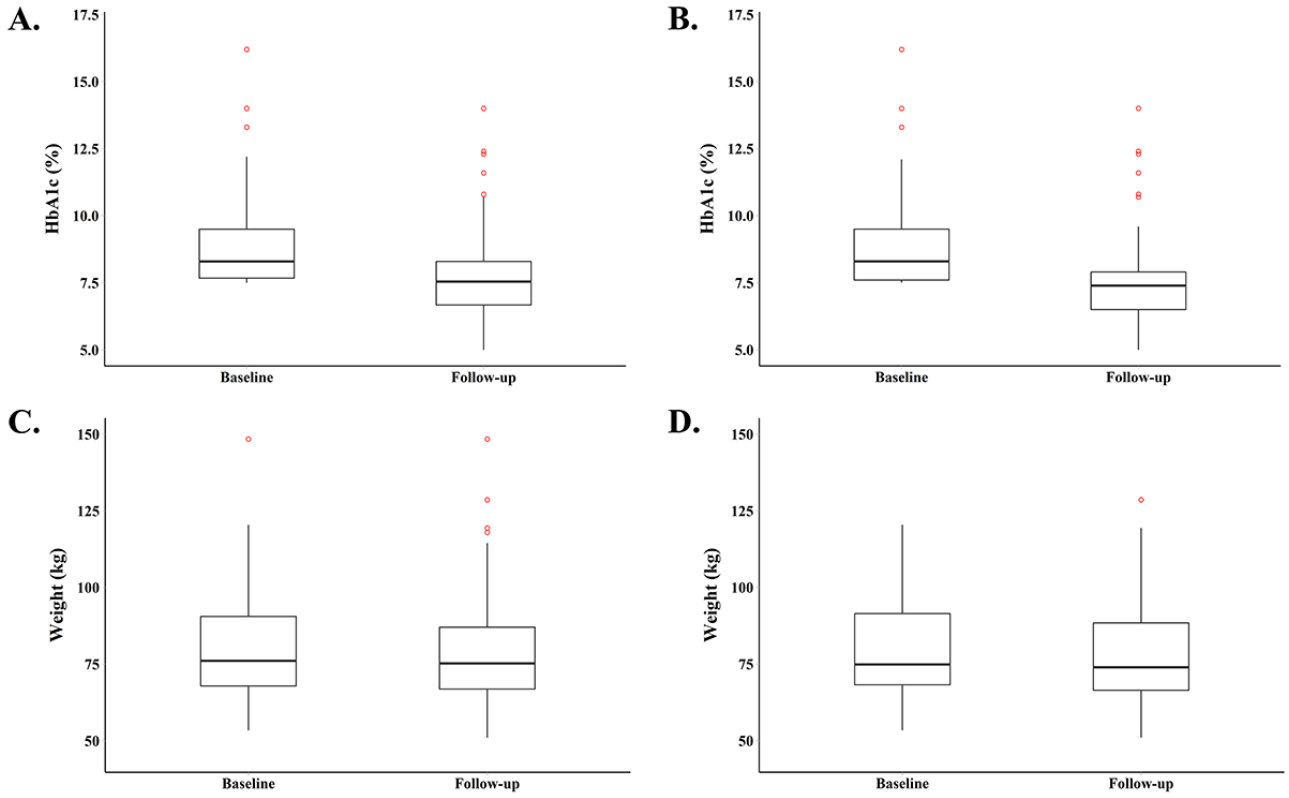


Table 4. Changes in health outcomes at follow-up compared to baseline for participants who reached follow-up.

Health outcomes and measures	Total (n=100)	*P* value	Completes (n=83)	*P* value
HbA_{1c}[a] change				
Percentage point change in HbA_{1c}, mean (95% CI)	–1.1 (–1.4 to –0.7)	<.001	–1.3 (–1.7 to –0.8)	<.001
Participants with ≥1 percentage point reduction, n (%)	49 (49)		49 (59)	
Weight change				
Weight change expressed as a percentage of baseline weight (%), mean (95% CI)	–2.0 (–2.8 to –1.2)	<.001	–2.3 (–3.3 to –1.4)	<.001
Participants with loss of ≥5% of initial baseline weight, n (%)	17 (17)		17 (21)	

[a]HbA_{1c}: glycated hemoglobin.

Maintenance (User Satisfaction)

Among the 80 participants who completed the follow-up survey, the average user-friendliness rating for various program components was 3.6 to 3.8 out of 5, where 5 = very easy (Multimedia Appendix 9). Program components received perceived usefulness scores of 3.3 to 3.5 out of 5, where 5 = very useful (Multimedia Appendix 10). Despite being given new glucometers to use for the study, participants gave average ratings of 3.8 and 3.5 out of 5 for glucometer user-friendliness and usefulness of performing blood glucose monitoring, respectively (Multimedia Appendices 9 and 10). Most participants gave the Glyco app an overall rating of good (41/80, 51%) or very good (18/80, 23%) (Multimedia Appendix 11). The majority of participants indicated that they either probably would (42/80, 53%) or definitely would (21/80, 26%) recommend the Glyco app to others for managing their diabetes. Only one participant probably would not recommend the app to others, and 20% (16/80) were neutral on this matter. Respondent data revealed that 21% (17/80) of participants stated that they were willing to purchase unlimited access to the Glyco app and health coaches for an annual fee of SGD100 (approximately US $73) but no participant said that they were willing to pay an annual fee of SGD200 (approximately US $147).

Discussion

Principal Findings

A reduction in HbA_{1c} levels of 1 percentage point has been shown to be associated with a 21%, 14%, 37%, and 21% decrease in risk of any end point related to diabetes, myocardial infarction, microvascular complications, and diabetes-related

death, respectively [28]. This study found that on average, participants achieved this level of improvement in HbA_{1c} level over the study period. Feasibility studies using mobile phone–based lifestyle management interventions for diabetes typically do not observe such large changes [9,11,29,30]. Additionally, almost a fifth of participants—all of whom were overweight (BMI >23 kg/m^2) at baseline—lost more than 5% of their initial body weight. Behavioral interventions have demonstrated that a weight loss of as little as 5% among overweight adults can result in health benefits [31]. Overall, these results are encouraging and suggest that the mHealth app has the potential to be clinically relevant in practice. However, we cannot rule out the possibility that factors unrelated to GlycoLeap could be positively influencing health outcomes.

Strengths and Limitations

In spite of relatively high levels of engagement across Glyco app components in the first week, the decreased usage over time advocates that more should be done to improve and sustain engagement. Nevertheless, despite low completion rates for the online lessons, self-reported increases in blood glucose monitoring frequency and improvements in dietary habits argue that a mobile intervention like GlycoLeap may be a viable strategy for patient education and behavior change. As the online lessons were administered on a different platform and required separate email access, this may have presented a barrier to access. A greater completion rate may have been achieved if the lessons were made available directly on the mobile app.

Unlike conventional T2DM lifestyle management programs administered in the primary care setting through in-person sessions, mHealth interventions comprising mobile phone apps like Glyco are highly scalable, requiring comparatively fewer manpower resources. Although we were unable to determine the true reach as defined by the RE-AIM framework [17,18], we attempted to obtain a proxy for reach at SingHealth Polyclinics–Tampines using the percentage of patients who consented to participate out of all patients approached. Despite an apparent low reach of 13.2%, this percentage does not accurately represent interest in GlycoLeap as patients who were deemed ineligible (eg, unable to speak English or did not have a mobile phone) during the prescreen assessment were not offered the opportunity to give informed consent and take the screener. Based on the anecdotal information from the research coordinator, most of the patients approached were keen to join the study. Hence, we anticipate that the actual willingness to participate in GlycoLeap if offered as an add-on to standard care will be higher than 13.2%, especially if GlycoLeap is made available in other languages such as Mandarin and mobile phone use becomes more prevalent.

This study also suggests that GlycoLeap might be a scalable intervention. The app received good ratings with 21% (17/80) of participants claiming they were willing to purchase unlimited access to use the app and health coaches at a modest fee. In addition, the GlycoLeap program received relatively high user-friendliness and user-satisfaction scores compared to other similar mobile phone apps for diabetes management [9,32].

Conclusions

These results suggest that GlycoLeap may be an effective strategy for helping some adults with T2DM attain better diabetic control and that it is feasible to integrate it within the primary care setting. However, given the study design, we cannot exclude the possibility that any health improvements were due to factors unrelated to the GlycoLeap program. Therefore, future efforts should assess the effectiveness and cost effectiveness of GlycoLeap using a randomized controlled trial of at least 12 months to evaluate longer term outcomes.

Acknowledgments

This work was supported by KKT Technology Pte Ltd (Holmusk). The sponsor funded the study, provided participants with the GlycoLeap program, collected program engagement data, and provided the program engagement data for analysis. The sponsor had involvement in the study design and the decision to submit the report for publication. The sponsor had no involvement in the analysis and interpretation of data, or writing of the report. We acknowledge Uma Rani, daughter of Singaram Samidurai, for recruiting patients and conducting follow-up assessments, Sankha Mukherjee for collating the program engagement data, and Mijie Li for contributing to manuscript writing.

Authors' Contributions

EAF was the principal investigator and is the corresponding author. DK, PSCG, TYY, NCT, and EAF were responsible for the study design; DK, RSML, TYY, and EAF for data gathering; RSML, YT, and EAF for data analysis; DK, RSML, YT, and EAF for interpretation of the results and drafting the manuscript. All authors read, edited, and approved the submitted version. As the corresponding author, EAF had full access to all study data and takes responsibility for data integrity and the accuracy of the data analysis.

Conflicts of Interest

EAF formerly served on the management committee for KKT Technology Pte Ltd (Holmusk), a company that developed and owns digital lifestyle change programs including GlycoLeap. TYY is the medical director at KKT Technology Pte Ltd (Holmusk) and receives a salary and stock options. TYY has no direct ties to the recruitment site or patients. The other authors have no conflicts.

Multimedia Appendix 1

Baseline survey.

[PDF File (Adobe PDF File), 92KB - mhealth_v7i5e12965_app1.pdf]

Multimedia Appendix 2

Follow-up survey.

[PDF File (Adobe PDF File), 121KB - mhealth_v7i5e12965_app2.pdf]

Multimedia Appendix 3

Glyco app home screen.

[PDF File (Adobe PDF File), 1MB - mhealth_v7i5e12965_app3.pdf]

Multimedia Appendix 4

Glyco app blood glucose monitoring.

[PDF File (Adobe PDF File), 645KB - mhealth_v7i5e12965_app4.pdf]

Multimedia Appendix 5

Glyco app weight monitoring.

[PDF File (Adobe PDF File), 619KB - mhealth_v7i5e12965_app5.pdf]

Multimedia Appendix 6

Glyco app meal log.

[PDF File (Adobe PDF File), 3MB - mhealth_v7i5e12965_app6.pdf]

Multimedia Appendix 7

Glyco app health coach correspondence.

[PDF File (Adobe PDF File), 1MB - mhealth_v7i5e12965_app7.pdf]

Multimedia Appendix 8

Output of linear regression with change in HbA_{1c} regressed on measures of program engagement.

[PDF File (Adobe PDF File), 175KB - mhealth_v7i5e12965_app8.pdf]

Multimedia Appendix 9

Average participant rating of the user-friendliness of different GlycoLeap program components, where 1=very difficult and 5=very easy.

[PDF File (Adobe PDF File), 66KB - mhealth_v7i5e12965_app9.pdf]

Multimedia Appendix 10

Average participant rating of the perceived usefulness of different GlycoLeap program components, where 1=not useful at all and 5=very useful.

[PDF File (Adobe PDF File), 50KB - mhealth_v7i5e12965_app10.pdf]

Multimedia Appendix 11

Distribution of overall rating of the Glyco app, where 1=very poor, 2=poor, 3=average, 4=good, and 5=very good.

[PDF File (Adobe PDF File), 39KB - mhealth_v7i5e12965_app11.pdf]

References

1. NCD Risk Factor Collaboration. Worldwide trends in diabetes since 1980: a pooled analysis of 751 population-based studies with 4.4 million participants. Lancet 2016 Apr 09;387(10027):1513-1530 [FREE Full text] [doi: 10.1016/S0140-6736(16)00618-8] [Medline: 27061677]
2. IDF Diabetes Atlas (8th edition): Western Pacific Fact Sheet.: International Diabetes Foundation; 2017. URL: https://www.idf.org/our-network/regions-members/western-pacific/members/113-singapore.html [accessed 2018-11-05] [WebCite Cache ID 73hL1Yguy]
3. Phan TP, Alkema L, Tai ES, Tan KHX, Yang Q, Lim W, et al. Forecasting the burden of type 2 diabetes in Singapore using a demographic epidemiological model of Singapore. BMJ Open Diabetes Res Care 2014;2(1):e000012 [FREE Full text] [doi: 10.1136/bmjdrc-2013-000012] [Medline: 25452860]
4. Pillay J, Armstrong MJ, Butalia S, Donovan LE, Sigal RJ, Vandermeer B, et al. Behavioral programs for type 2 diabetes mellitus: a systematic review and network meta-analysis. Ann Intern Med 2015 Dec 1;163(11):848-860. [doi: 10.7326/M15-1400] [Medline: 26414227]
5. Holmen H, Torbjørnsen A, Wahl AK, Jenum AK, Småstuen MC, Arsand E, et al. A mobile health intervention for self-management and lifestyle change for persons with type 2 diabetes, part 2: one-year results from the Norwegian randomized controlled Trial RENEWING HEALTH. JMIR Mhealth Uhealth 2014;2(4):e57 [FREE Full text] [doi: 10.2196/mhealth.3882] [Medline: 25499872]
6. Osborn CY, van Ginkel JR, Rodbard D, Heyman M, Marrero DG, Huddleston B, et al. One Drop | Mobile: an evaluation of hemoglobin A1c improvement linked to app engagement. JMIR Diabetes 2017 Aug 24;2(2):e21 [FREE Full text] [doi: 10.2196/diabetes.8039] [Medline: 30291059]
7. Quinn CC, Shardell MD, Terrin ML, Barr EA, Ballew SH, Gruber-Baldini AL. Cluster-randomized trial of a mobile phone personalized behavioral intervention for blood glucose control. Diabetes Care 2011 Sep;34(9):1934-1942 [FREE Full text] [doi: 10.2337/dc11-0366] [Medline: 21788632]
8. Bonoto BC, de Araujo VE, Godói IP, de Lemos LLP, Godman B, Bennie M, et al. Efficacy of mobile apps to support the care of patients with diabetes mellitus: a systematic review and meta-analysis of randomized controlled trials. JMIR Mhealth Uhealth 2017 Mar 01;5(3):e4 [FREE Full text] [doi: 10.2196/mhealth.6309] [Medline: 28249834]
9. Waki K, Fujita H, Uchimura Y, Omae K, Aramaki E, Kato S, et al. DialBetics: a novel smartphone-based self-management support system for type 2 diabetes patients. J Diabetes Sci Technol 2014 Mar 13;8(2):209-215 [FREE Full text] [doi: 10.1177/1932296814526495] [Medline: 24876569]
10. Hsu WC, Lau KHK, Huang R, Ghiloni S, Le H, Gilroy S, et al. Utilization of a cloud-based diabetes management program for insulin initiation and titration enables collaborative decision making between healthcare providers and patients. Diabetes Technol Ther 2016 Feb;18(2):59-67 [FREE Full text] [doi: 10.1089/dia.2015.0160] [Medline: 26645932]
11. Wayne N, Perez DF, Kaplan DM, Ritvo P. Health coaching reduces HbA1c in type 2 diabetic patients from a lower-socioeconomic status community: a randomized controlled trial. J Med Internet Res 2015;17(10):e224 [FREE Full text] [doi: 10.2196/jmir.4871] [Medline: 26441467]
12. Wayne N, Ritvo P. Smartphone-enabled health coach intervention for people with diabetes from a modest socioeconomic strata community: single-arm longitudinal feasibility study. J Med Internet Res 2014;16(6):e149 [FREE Full text] [doi: 10.2196/jmir.3180] [Medline: 24907918]
13. Castro SCM, Chiguluri V, Gumpina R, Abbott P, Madero EN, Payne M, et al. Outcomes of a digital health program with human coaching for diabetes risk reduction in a medicare population. J Aging Health 2017 Jan 01:898264316688791. [doi: 10.1177/0898264316688791] [Medline: 28553807]
14. Kempf K, Altpeter B, Berger J, Reuß O, Fuchs M, Schneider M, et al. Efficacy of the telemedical lifestyle intervention program Telipro in advanced stages of type 2 diabetes: a randomized controlled trial. Diabetes Care 2017 Dec;40(7):863-871. [doi: 10.2337/dc17-0303] [Medline: 28500214]
15. Grady M, Katz LB, Cameron H, Levy BL. Diabetes app-related text messages from health care professionals in conjunction with a new wireless glucose meter with a color range indicator improves glycemic control in patients with type 1 and type 2 diabetes: randomized controlled trial. JMIR Diabetes 2017 Aug 07;2(2):e19 [FREE Full text] [doi: 10.2196/diabetes.7454] [Medline: 30291092]
16. Zhou W, Chen M, Yuan J, Sun Y. Welltang: a smart phone-based diabetes management application improves blood glucose control in Chinese people with diabetes. Diabetes Res Clin Pract 2016 Jun;116:105-110. [doi: 10.1016/j.diabres.2016.03.018] [Medline: 27321324]
17. Glasgow RE, Vogt TM, Boles SM. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. Am J Public Health 1999 Sep;89(9):1322-1327. [Medline: 10474547]
18. Gaglio B, Shoup JA, Glasgow RE. The RE-AIM framework: a systematic review of use over time. Am J Public Health 2013 Jun;103(6):e38-e46. [doi: 10.2105/AJPH.2013.301299] [Medline: 23597377]
19. Toobert DJ, Hampson SE, Glasgow RE. The summary of diabetes self-care activities measure: results from 7 studies and a revised scale. Diabetes Care 2000 Jul;23(7):943-950 [FREE Full text] [Medline: 10895844]
20. Laranjo L. Social media and health behavior change. In: Inyed-Abdul S, Gabarron E, Lau AYS, editors. Participatory Health Through Social Media. 1st Edition. Cambridge: Academic Press; 2016:83-111.

21. Levine DM, Savarimuthu S, Squires A, Nicholson J, Jay M. Technology-assisted weight loss interventions in primary care: a systematic review. J Gen Intern Med 2015 Jan;30(1):107-117. [doi: 10.1007/s11606-014-2987-6] [Medline: 25134692]
22. Prochaska JO, DiClemente CC. Stages and processes of self-change of smoking: toward an integrative model of change. J Consult Clin Psychol 1983 Jun;51(3):390-395. [Medline: 6863699]
23. Prochaska JO, DiClemente CC, Norcross JC. In search of how people change: applications to addictive behaviors. Am Psychol 1992 Sep;47(9):1102-1114. [Medline: 1329589]
24. Miller W, Rollnick S. Motivational Interviewing: Preparing People to Change Addictive Behavior. 1st Edition. New York: The Guilford Press; 1991.
25. Rosenstock IM. The health belief model and preventive health behavior. Health Educ Behav 1974 Dec 01;2(4):354-386. [doi: 10.1177/109019817400200405]
26. Franz MJ, MacLeod J, Evert A, Brown C, Gradwell E, Handu D, et al. Academy of Nutrition and Dietetics nutrition practice guideline for type 1 and type 2 diabetes in adults: systematic review of evidence for medical nutrition therapy effectiveness and recommendations for integration into the nutrition care process. J Acad Nutr Diet 2017 Dec;117(10):1659-1679. [doi: 10.1016/j.jand.2017.03.022] [Medline: 28533169]
27. Holmen H, Wahl AK, Cvancarova SM, Ribu L. Tailored communication within mobile apps for diabetes self-management: a systematic review. J Med Internet Res 2017 Jun 23;19(6):e227 [FREE Full text] [doi: 10.2196/jmir.7045] [Medline: 28645890]
28. Stratton IM, Adler AI, Neil HA, Matthews DR, Manley SE, Cull CA, et al. Association of glycaemia with macrovascular and microvascular complications of type 2 diabetes (UKPDS 35): prospective observational study. BMJ 2000 Aug 12;321(7258):405-412 [FREE Full text] [Medline: 10938048]
29. Kim H, Choi W, Baek EK, Kim YA, Yang SJ, Choi IY, et al. Efficacy of the smartphone-based glucose management application stratified by user satisfaction. Diabetes Metab J 2014 Jun;38(3):204-210 [FREE Full text] [doi: 10.4093/dmj.2014.38.3.204] [Medline: 25003074]
30. Cui M, Wu X, Mao J, Wang X, Nie M. T2DM self-management via smartphone applications: a systematic review and meta-analysis. PLoS One 2016;11(11):e0166718 [FREE Full text] [doi: 10.1371/journal.pone.0166718] [Medline: 27861583]
31. Wing RR, Lang W, Wadden TA, Safford M, Knowler WC, Bertoni AG, et al. Benefits of modest weight loss in improving cardiovascular risk factors in overweight and obese individuals with type 2 diabetes. Diabetes Care 2011 Jul;34(7):1481-1486 [FREE Full text] [doi: 10.2337/dc10-2415] [Medline: 21593294]
32. Lithgow K, Edwards A, Rabi D. Smartphone app use for diabetes management: evaluating patient perspectives. JMIR Diabetes 2017 Jan 23;2(1):e2 [FREE Full text] [doi: 10.2196/diabetes.6643] [Medline: 30291051]

Abbreviations

BMI: body mass index
HbA_{1c}: hemoglobin A_{1c} (glycated hemoglobin)
ITT: intention-to-treat
mHealth: mobile health
T2DM: type 2 diabetes mellitus

Edited by G Eysenbach; submitted 28.11.18; peer-reviewed by M Grady, A Torbjørnsen; comments to author 24.01.19; revised version received 14.03.19; accepted 30.03.19; published 24.05.19.

Please cite as:
Koot D, Goh PSC, Lim RSM, Tian Y, Yau TY, Tan NC, Finkelstein EA
A Mobile Lifestyle Management Program (GlycoLeap) for People With Type 2 Diabetes: Single-Arm Feasibility Study
JMIR Mhealth Uhealth 2019;7(5):e12965
URL: http://mhealth.jmir.org/2019/5/e12965/
doi:10.2196/12965
PMID:31127720

Original Paper

Development and Testing of a Mobile App for Pain Management Among Cancer Patients Discharged From Hospital Treatment: Randomized Controlled Trial

Jing Yang[1,2*], MS; Lizhu Weng[2,3*], MS; Zhikui Chen[4], PhD; Hongfu Cai[1,2], MS; Xiaoyan Lin[5], PhD; Zhijian Hu[6], MS; Na Li[1,2], MS; Bijuan Lin[1,2], PhD; Bin Zheng[1,2], PhD; Qian Zhuang[1,2], MS; Bin Du[5], MS; Zhiyuan Zheng[2], MD; Maobai Liu[1,2], MS

[1]Department of Pharmacy, Fujian Medical University Union Hospital, Fuzhou, China

[2]School of Pharmacy, Fujian Medical University, Fuzhou, China

[3]Department of Pharmacy, Xiamen Maternity and Child Care Hospital, Xiamen, China

[4]Department of Ultrasound, Affiliated Union Hospital of Fujian Medical University, Fuzhou, China

[5]Department of Oncology, Affiliated Union Hospital of Fujian Medical University, Fuzhou, China

[6]Department of Information, Affiliated Union Hospital of Fujian Medical University, Fuzhou, China

[*]these authors contributed equally

Corresponding Author:
Maobai Liu, MS
Department of Pharmacy
Fujian Medical University Union Hospital
29 Xinquan Road, Gulou District
Fuzhou, 350001
China
Phone: 86 0591 86218372
Email: liumb0591@sina.com

Abstract

Background: The incidence of cancer pain increases in discharged patients because of discontinued standard treatments and reductions in medication adherence. Motivated by the need for better pain management in discharged patients, we developed a mobile phone app (Pain Guard) to provide continuous treatment information and feedback to discharged cancer patients suffering from pain.

Objective: The aim was to design, construct, and test the Pain Guard app in patients managing cancer pain, evaluate the total remission rate of pain and the improvement in quality of life (QoL) to improve pain management for cancer pain patients, and assess patient acceptance of the app.

Methods: This randomized controlled double-arm study involved 58 patients with cancer pain symptoms. Participants were randomly assigned to a group receiving care through the Pain Guard app (n=31) or to a control group (n=27) who received only traditional pharmaceutical care. In a pretest, participants were rated using a baseline cancer pain assessment and QoL evaluation. During treatment, the consumption levels of analgesic drugs were recorded every week. After a 4-week study period, another round of cancer pain assessment and QoL evaluation was conducted. The system's usability, feasibility, app compliance, and satisfaction were also assessed. Our primary outcome was remission rate of pain, and secondary outcomes were medication adherence, improvements in QoL, frequency of breakthrough cancer pain (BTcP), incidence of adverse reactions, and satisfaction of patients.

Results: All participants (N=58) successfully completed the study. There were no significant differences in baseline pain scores or baseline QoL scores between groups. At the end of the study, the rate of pain remission in the trial group was significantly higher than that in the control group (P<.001). The frequency of BTcP in the app group was considerably lower than that in the control group (P<.001). The rate of medication adherence in the trial group was considerably higher than that in the control group (P<.001). Improvements in global QoL scores in the trial group were also significantly higher than those in the control group (P<.001). The incidence of adverse reactions in the trial group (7/31) was lower than that in the control group (12/27), especially constipation, with significant differences (P=.01). The 31 participants in the trial group completed a satisfaction survey regarding

Pain Guard: 23 (74%) indicated that they were satisfied with receiving pharmaceutical care by Pain Guard, 5 (16%) indicated that they were somewhat satisfied, 2 (6%) indicated neutral feelings, and 1 (3%) indicated that they were somewhat dissatisfied; no participants indicated that they were very dissatisfied.

Conclusions: Pain Guard was effective for the management of pain in discharged patients with cancer pain, and its operability was effective and easily accepted by patients.

Trial Registration: Chinese Clinical Trials Registry ChiCTR1800016066; http://www.chictr.org.cn/showproj.aspx?proj=27153

(JMIR Mhealth Uhealth 2019;7(5):e12542) doi:10.2196/12542

KEYWORDS

cancer; pain management; quality of life; adherence

Introduction

Pain has been defined as the fifth vital sign following body temperature, pulse, respiration, and blood pressure [1]. Previous reports show that the incidence of pain is 30% to 50% in patients with early and midstage cancer and 75% to 90% in patients with advanced cancer [2]. Pain devastates the quality of life (QoL) of patients with cancer, impedes cancer recovery, interferes with activities of daily living, and results in long-term morbidity [3]. Poor pain management places a huge emotional burden on patients and their relatives and represents a significant cost burden to the health care system and families, with pain being the most common reason for patients with cancer to use emergency health services [4].

In 2011, the Chinese Ministry of Health initiated the Good Pain Management Ward Program to improve the management of cancer pain in hospitalized patients [5]. Despite existing guidelines to assess and manage pain [6], the management of cancer-related pain is suboptimal, and patients are regularly undertreated outside the hospital [7,8]. More than half of patients suffering from cancer pain are not treated adequately, especially discharged patients and remote patients with less access to health care [9].

According to data from the China Internet Network Information Center, as of June 2018, the number of mobile internet users in China had reached 788 million, and 98.3% of them used mobile phones to access the internet [10]. The use of mobile technology to develop low-cost and pragmatic patient-centered interventions is a key factor in reducing health care costs and advancing the science of symptom management [11].

Although there are 165,000 medical and health apps (categorized under mobile health [mHealth] or electronic health [eHealth]), most of these apps are not scientifically validated [12]. In a review of 279 pain-related apps, only one app was found to have undergone scientific evaluation [13]. There is a need to develop and test theory-based and evidence-based apps to better support patients with accessible pain care self-management.

Considering these points, we developed a mobile phone app, known as Pain Guard. With the existing medical system and insufficient medical resources in China, a pain management app can be useful and appreciated by patients and health care professionals, making it a good choice for the management of cancer pain in our country [14,15]. The purpose of our study was to develop and test Pain Guard for pain management among Chinese cancer patients discharged from hospital treatment.

Methods

Pain Guard Design

The Pain Guard app includes two opening screens: one opens to medical staff (the therapeutic interface) and the other opens to patients (the patient interface). The app in our study was designed to operate on the Android and iOS mobile operating systems to achieve an affordable, portable, and easy to use environment for patients. A total of 30 senior domestic pharmacy experts, clinical experts, nursing experts, and psychologists were included in the design process.

We used the Delphi method [16] to design the functions of Pain Guard to ensure that the modules and questionnaires included in Pain Guard were practical and scientifically sound. The construction of the questionnaires and the system module content indicators—at all levels of the entries, grades, and indicators of the app's function—underwent two rounds of expert consultation. The app's design was jointly accomplished by oncology clinical pharmacists and physicians, in collaboration with information technology (IT) engineers at the Union Hospital of Fujian Medical University and a music therapist. The design team adopted a modular approach consisting of several functional subsystems to facilitate speedy progress by multiple teams. After designing the system architecture, engineering work, such as programming and system integration, was outsourced to a professional IT company to produce an executable mobile phone app.

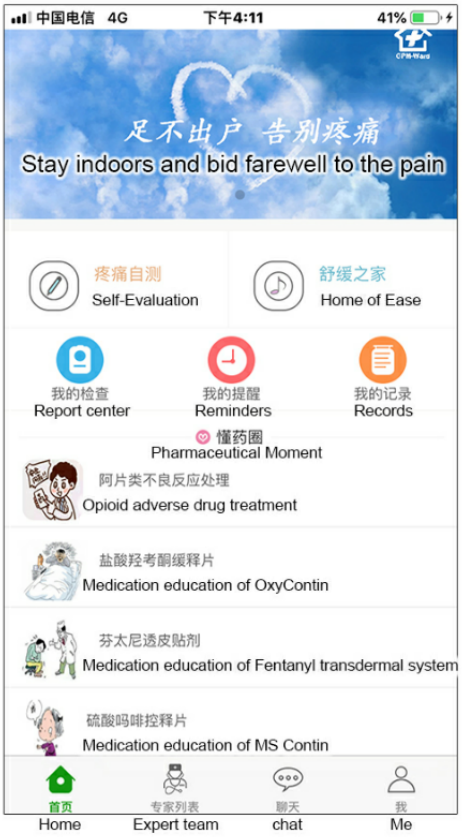
The therapeutic interface of the Pain Guard App was designed for checking a patient's medical history, which was recorded by the system, and for providing real-time patient consultation (Figure 1). Note that all screenshots of the app include translations that have been added in.

The Patient App

The patient interface of the app consisted of nine modules: self-evaluation, reminders, reports, records, real-time medication consultation, musical soothing treatment, pharmaceutical moments (a module for educating patients), team expert introduction, and my center (Figure 2).

Figure 1. Screenshot of the medical experts' search page.

Figure 2. Screenshot of the patient app home page.



Self-Evaluation Module

This module consisted of two parts: daily cancer pain assessment and breakthrough cancer pain (BTcP) assessment. The pain diary (Figure 3) was a questionnaire that consisted of 12 questions; this was used to track patients' self-reported pain data, including assessment of pain intensity, the location of the pain, the nature of the pain, the pain score, the frequency of BTcP, medication status, and any adverse reactions. A body map displayed on the mobile phone screen allowed the patient to choose the precise location of recently experienced cancer pain. The pain score was based on a numerical rating scale (NRS) from 1 to 10. The patients were asked to identify the highest severity of pain, using NRS scores, within the previous 24 hours, as well as to report the current pain score. The BTcP assessment included a questionnaire that contained the score, the location, and the duration of pain. An intervention alarm was designed to help patients manage their pain in real time. When moderate or severe BTcP (with a pain score greater than or equal to 4) occurred, the system would alert the patient, with a prompt to follow medication orders. An hour later, the patient would be prompted to reassess their BTcP. If the pain score still exceeded 4, the processing plan would repeat.

Reminders

This module was designed to remind patients to perform a BTcP reassessment and to view their examination and medication schedules so that they would remember to take their pain medicine on a regular basis (Figure 4).

Records

This module was designed so that patients could see their pain status and treatment history (Figure 4).

Report Center

This module was designed to enable patients to take photos of their inspection reports to be forwarded to the pain management team (Figure 4). These reports included blood tests (eg, blood routine, liver function, and kidney function), imaging data (eg, computed tomography and magnetic resonance imaging scans, and ultrasound), and other findings that may be associated with pain management that enable medical staff to monitor drug reactions.

Real-Time Medication Consultation

This module was designed to facilitate a real-time consultation session on pain management between the patient and the cancer pain management team (Figure 5).

Figure 3. Screenshot of the self-evaluation module, including the pain diary and breakthrough cancer pain assessment parts.

Figure 4. Screenshots of the reminds, records, and report center modules.

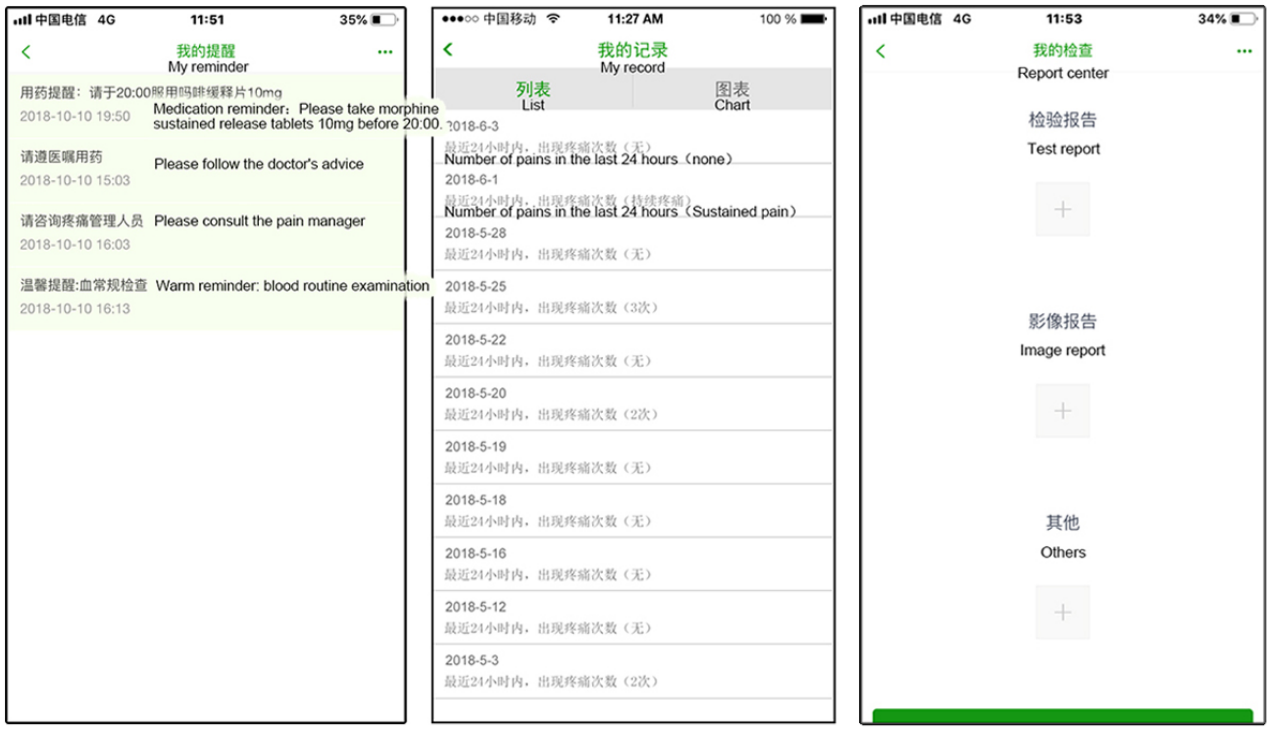


Figure 5. Screenshot of the real-time medication consultation module.

Soothing Music Treatment

This module consisted of a questionnaire and an intelligent filter push function. The content of the questionnaire and recommended music catalog were formulated by the Delphi expert correspondence method. A total of 56 sounds were selected: 20 nature, 10 piano music, 5 harp music, 5 easy-listening music, 5 Chinese music, 5 religious (Christian) music, and 6 religious (Buddhism) music. These sounds could be used free of charge. According to a brief questionnaire, the system would suggest music playlists and links to help patients relieve tension and anxiety, raise the threshold of pain, and increase comfort levels (Figure 6).

Pharmaceutical Moment

This module was designed to provide pharmacological knowledge for patients' self-learning, including information on antitumor drugs, analgesic drugs, and adverse reaction prevention (Figure 7).

My Center

This module was designed to survey satisfaction, collect feedback from patients for further improvement of the app, and to provide instructions for the usage and version information of the app (Figure 8).

Study Design

An experiment was designed to test the effectiveness of Pain Guard on cancer pain management. The experiment involved two groups: a Pain Guard trial group and a control group.

Pain Guard Study Group

After obtaining consent from all participants, clinical pharmacists conducted a standardized education session to teach the participants how to operate the mobile phone, use Pain Guard, properly assess pain, and use the rating system. The participants in the study group were asked to complete initial and final pain assessment questionnaires and QoL questionnaires on the mobile phones provided to them. Participants were encouraged to use Pain Guard as much as possible to record their pain status, at least once every day for 4 weeks.

Control Group

The control group received conventional pharmaceutical care. Initial and final pain and QoL assessment data were collected. Before the patient was discharged from the hospital, the clinical pharmacist conducted detailed medication education (including medication methods, prevention and treatment of adverse reactions, and precautions) and asked the patient to attempt to maintain a paper version of the pain diary.

Figure 6. Screenshot of the soothing music treatment module.

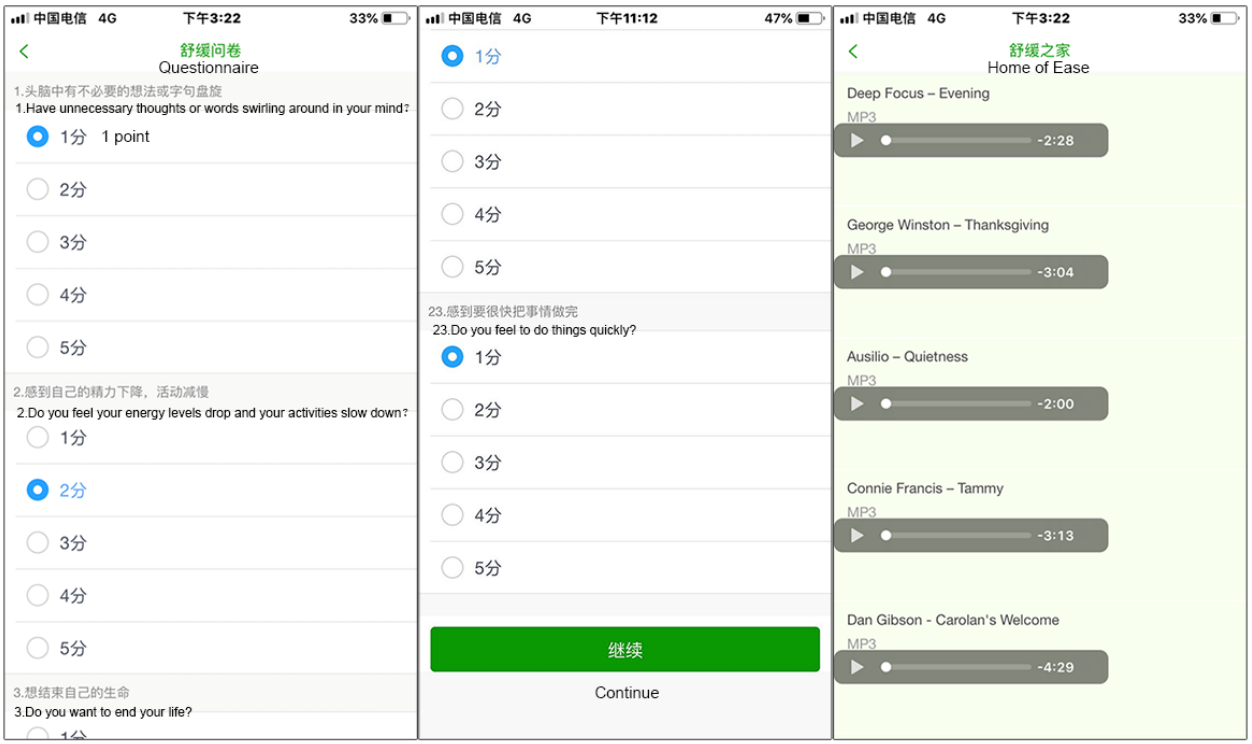


Figure 7. Screenshot of the pharmaceutical moment module.

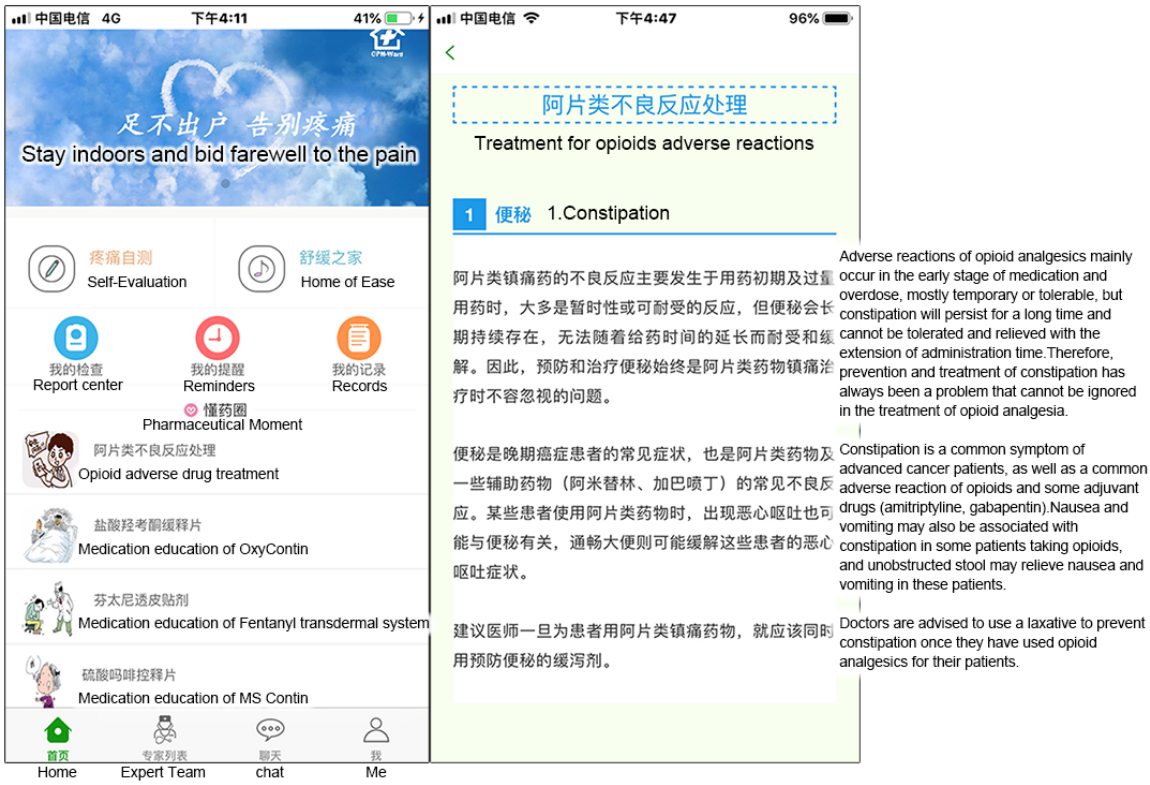
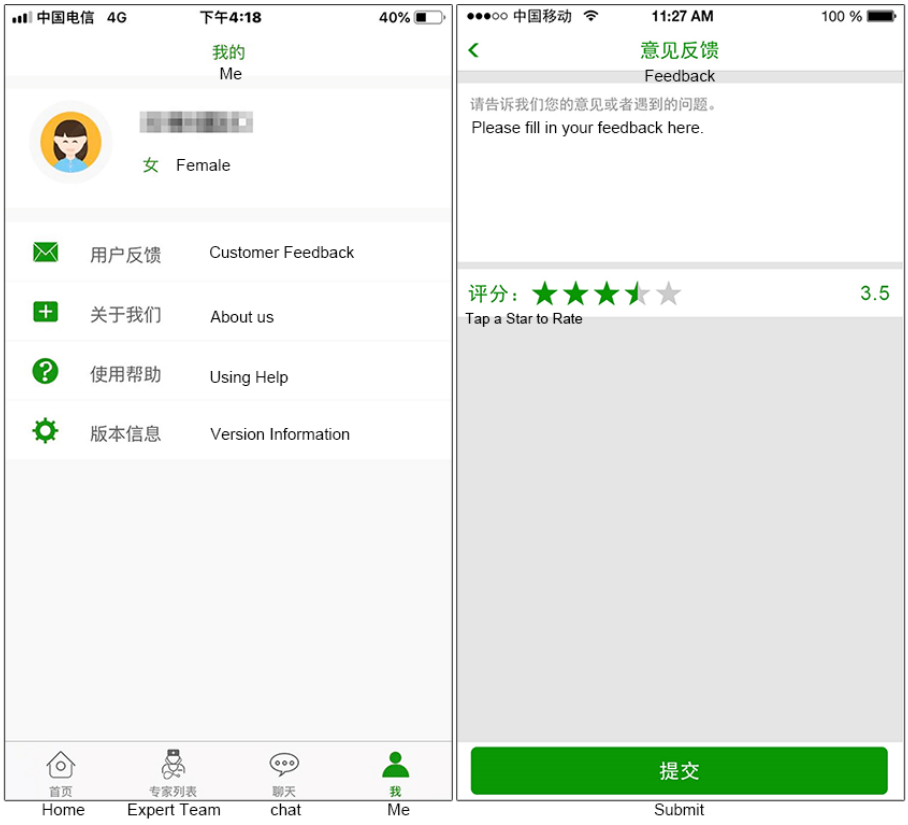


Figure 8. Screenshot of the My Center module.



Enrollment of the Participants

The participants had to meet the following screening criteria: (1) able to read Chinese and use a mobile phone, (2) aged between 18 and 75 years, (3) diagnosed with cancer and had self-reported cancer pain within a month before the study, and (4) could understand the study process and evaluation, agreed to participate in the trial, and signed the informed consent form. Exclusion criteria included (1) severe cognitive impairments; (2) hepatic insufficiency (alanine aminotransferase ≥2.5 × upper limits of normal [ULN], aspartate aminotransferase ≥2.5 × ULN, total bilirubin ≥ 1.5 × ULN), or renal insufficiency (serum creatinine ≥2.5 × ULN); (3) inability to complete the pain assessment; (4) participation in any other investigational therapies or other study protocols that may have an impact on pain intensity, which were the main outcomes of this study; (5) history of drug abuse, addiction, or severe alcoholism; and (6) opioid allergy.

Principle Objectives

The primary objective was to assess the effectiveness of pain management with Pain Guard. The secondary objectives were to evaluate the feasibility and changes in the quality of patients' lives, user satisfaction, incidence of adverse reactions, and medication adherence when using Pain Guard.

Measurement

Pain Assessment and Quality of Life Evaluation

All participants were asked to complete a general information questionnaire regarding pain management and assessment. A baseline pain assessment was conducted using an NRS in both groups. Furthermore, a baseline QoL evaluation was conducted using the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire-Core 30 in both groups. At the end of the trial, the pain assessment and QoL evaluation were repeated in both groups. Medication adherence was calculated by comparing the patient's daily dose with the physician's prescription and medication cycle [17]. The consumption of the analgesic drugs was recorded by a questionnaire on pain management in Fujian Medical University Union Hospital, Fuzhou City, China.

The feedback on Pain Guard was designed to evaluate satisfaction with the system. The questionnaire was completed by participants at the end of the study. Overall satisfaction ratings were displayed by a certain number of stars. The number of stars was on a scale of 1 to 5, with 1 star being the worst and 5 stars being the best. The data generated after the survey were used to evaluate patient satisfaction with Pain Guard. The questionnaire also contained an open-ended question in which participants were encouraged to give any other suggestions regarding improvements to Pain Guard.

Data Analysis

Due to the nature of this pilot work and the sample size, only limited statistical analyses were performed. Outcome evaluators were blinded to the data collection. The data were processed by using the R Statistical Software Package 3.5.1 (R Foundation for Statistical Computing, Vienna, Austria), and were tested for normality by using the Shapiro-Wilk test. Notably, the mean values of several groups were not normal distributed (P<.05), so the nonparametric test was used. The Wilcoxon-Mann-Whitney test was used to analyze differences in remission rate of pain, BTcP, medication adherence, and QoL between the trial and control groups.

User Statistics

The research was registered online at the Chinese Clinical Trials Registry on May 9, 2018. The ethical review of this study was approved by the Medical Ethics Committee of Fujian Medical University Union Hospital in Fuzhou City on May 4, 2018. This randomized controlled study of Pain Guard was conducted at the Oncology Center of Fujian Medical University Union Hospital in Fuzhou City.

Results

Participant Characteristics

A total of 58 participants (20 female, 38 male) were enrolled. All patients underwent treatment in which analgesia principles, titrations, maintenance, and safety for the conversion or rotation of drugs strictly followed the National Comprehensive Cancer Network Adult Cancer Pain Guidelines. A randomization scheme was generated by independent statistical personnel using a computer. All participants were then randomly assigned into two groups: the Pain Guard study group (n=31) and the control group (n=27). The study group had 14 (45%) females; the control group had 6 (22%) females. The participants' demographic information, as well as their disease characteristics, are summarized in Table 1.

We found no significant difference in the baseline pain scores between the two groups (Table 2). Over the 4-week study period, there was a significant difference in BTcP scores between the Pain Guard and control groups (median 3, interquartile range [IQR] 2-7 vs median 13, IQR 9.5-14, P<.001). The remission rate of pain was significantly different in the Pain Guard versus control group (median 50, IQR 45-63 vs median 0, IQR 0-25, P<001). The medication adherence of patients in the Pain Guard group was median 100 (IQR 98-100) compared with median 75 (IQR 62-89) in the control group, which was significantly different (P<.001).

The results shown in Multimedia Appendix 1 indicate that there was no significant difference in the baseline of all items of QoL scores between the two groups (P>.05). The study group, compared with the control group, scored significantly higher in nine QoL dimensions: cognitive functioning (W=768, P<001), emotional functioning (W=552.5, P=.03), social functioning (W=556, P=.03), sleeping disturbances (W=124, P<001), nausea and vomiting (W=272, P=.01), constipation (W=261, P=.008), fatigue (W=211.5, P=.001), pain (W=177, P<001), and global QoL (W=725.5, P<001). In the other dimensions, there was no significant difference. The total number of participants who completed the entire questionnaire was 58 (100%) at baseline and 58 (100%) at 4 weeks.

Table 1. Characteristics of participants at baseline (N=58).

Variable	Pain Guard (n=31)	Control (n=27)
Age (years), mean (SD)	51.10 (8.98)	53.96 (8.58)
Sex (female), n (%)	14 (45)	6 (22)
Primary diagnosis, n (%)		
Nasopharyngeal cancer	2 (6)	1 (4)
Cervical cancer	1 (3)	N/A
Esophagus cancer	N/A[a]	6 (22)
Stomach cancer	5 (16)	3 (11)
Column cancer	9 (29)	5 (19)
Lung cancer	7 (23)	9 (33)
Breast cancer	3 (10)	N/A
Ovarian cancer	1 (3)	N/A
Bladder cancer	2 (6)	N/A
Pancreatic cancer	N/A	1 (4)
Osteosarcoma	N/A	2 (7)
Soft tissue sarcoma	1 (3)	N/A
Therapeutic regimens, n (%)		
Oxycodone	7 (23)	4 (15)
Morphine	5 (16)	3 (11)
Methadone	13 (42)	12 (44)
Tramadol	6 (19)	8 (29.6)

[a]N/A: not available.

Table 2. Management outcome comparisons between the Pain Guard and control groups (N=58).

Management outcomes	Pain Guard (n=31), median (IQR[a])	Control (n=27), median (IQR)	W value[b]	*P* value
Baseline NRS[c]	4 (3-4)	4 (3-4)	495.5	.20
Remission rate of pain (%)	50 (45-63)	0 (0-25)	686.0	<.001
Frequency of breakthrough cancer pain	3 (2-7)	13 (10-14)	97.5	<.001
Medication adherence (%)	100 (98-100)	75 (62-89)	742.0	<.001

[a]IQR: interquartile range.

[b]W: Shapiro-Wilk test.

[c]NRS: numerical rating scale.

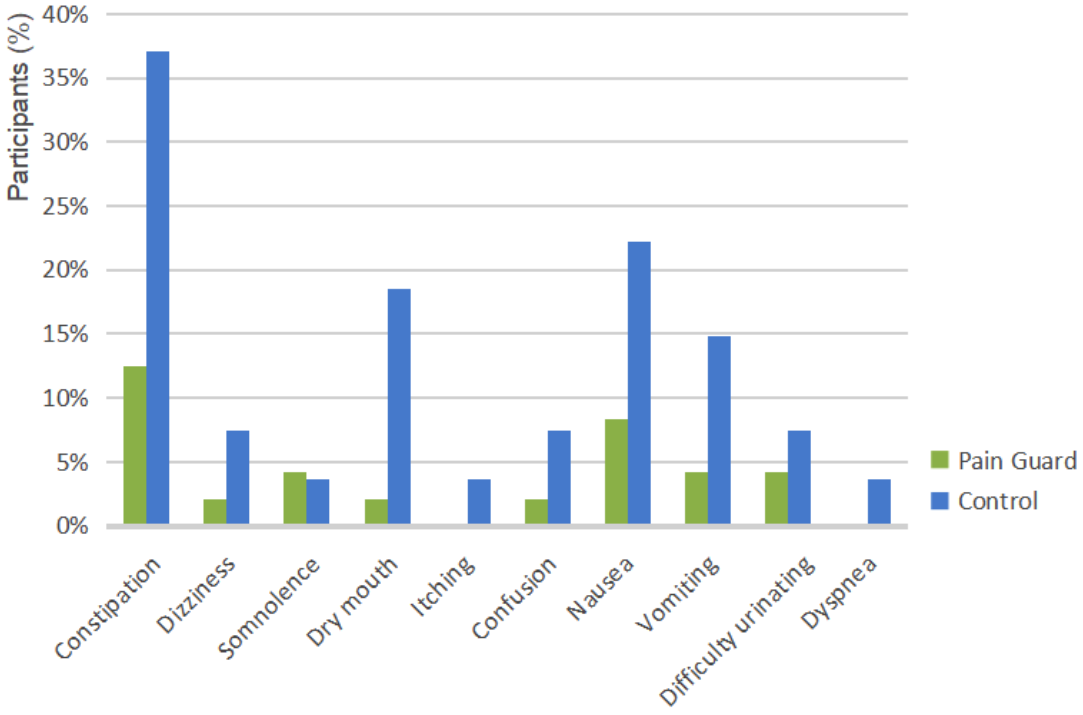
Adverse Reactions

The occurrence of adverse reactions (Figure 9), recorded by brief pain inventory, was lower in the trial group (7/31) than in the control group (12/27).

User Feedback

An evaluation was conducted to measure the participants' satisfaction with the Pain Guard app. Regarding their satisfaction, 23 (74%) participants in the trial group indicated they were very satisfied, 5 (16%) were somewhat satisfied, 2 (6%) were neutral, 1 (3%) was somewhat dissatisfied; none were very dissatisfied. Overall, the data suggest a high level of user satisfaction with Pain Guard.

Figure 9. Adverse reactions profile of participants during the study (N=58).



Discussion

Principal Results

This study aimed to provide continuous professional treatment for patients with pain after discharge from hospital to improve QoL. We built a seven-module app with the following five functions: (1) patients can report pain status, adverse drug reactions, and physical status at any time; (2) the management team can intervene and treat the patients, according to their reports, in a timely manner; (3) re-evaluation and medication reminders are available; (4) medication education is provided to patients; and (5) music therapy treatment can be administered in the patient's own home. Through a 4-week clinical trial investigating cancer pain management through Pain Guard, it was confirmed that Pain Guard significantly reduced the frequency of BTcP and improved pain relief and medication adherence in study participants with few adverse reactions. The satisfaction survey found that participants' acceptance and satisfaction with the app were high.

The cancer pain management team consisted of clinical pharmacists, physicians, and senior nurses. When the patient comes in for a consultation, the clinical pharmacist acted as the spokesperson for the pain management team. Simple problems were solved by the pharmacist, whereas complex problems were referred to the team by the pharmacist, and the corresponding expert would respond to the questions by consulting the patient's medical records. The pharmacist was in charge of communicating the solution to the patient.

Approximately 70% of patients with cancer pain are undertreated internationally [18] because physicians cannot access patients' current pain status, which is how physicians determine whether to change analgesic treatment, so the real-time messaging providing by the app plays the most important role in cancer pain management. If the drug adverse reaction and BTcP could not be treated timely and adequately, cancer pain may progress to a pain crisis [3]. An innovative real-time pain assessment mechanism and electronic reporting system were considered to be more effective in capturing pain data [19,20]. Our team collected detailed cancer pain reports through the app and provided personalized guidance for patients in real time.

Pain Guard was developed as a multidimensional tool, not only for real-time pain assessment but also for real-time drug adverse reaction treatment assessments and real-time messaging consultations. Such a mechanism enabled the team to convey clinical assistance and interventions to the patients. The Pain Guard app allowed patients to be able to instantaneously assess and report pain; thus, the team was able to provide prompt advice, which was not traditionally possible. From our study, the frequency of BTcP in the Pain Guard group was less than the control group. Through this app, BTcP could be reported in real time, standard treatment could be continued after the patient had been discharged, and the symptoms of cancer pain could be steadily controlled. Assessing adverse effects can lead to quicker responses, thus lowering levels of inconvenience experienced by patients [21]. The incidence of adverse reactions is reduced when symptoms are effectively controlled. For example, in one patient, constipation decreased from 37% to 10%. This was likely because the patient's bowel habits and diet structure were improved based on timely feedback and individualized education facilitated by the app.

According to our hospital's investigation, cancer patient's lack of pain knowledge can lead to poor control of cancer pain [22]. In the Pain Guard app, pharmacists can regularly educate patients to help them understand their pain and the drugs they take, which provides the necessary guidance for patients and ultimately improves patient adherence. The improved medication

adherence reinforces optimal pain management and improves QoL [23].

Music interventions may have a beneficial effect on anxiety, pain, fatigue, and QoL in people with cancer [24]. Furthermore, music represents an important intervention that is inexpensive, nontoxic, readily available, and can potentially minimize cancer pain, which helps reduce procedural pain [25].

According to a brief Music Therapy Self-Rating Scale [26], based on patient preferences and assessment outcomes, music can be suggested to patients intelligently to aid in the processing of thoughts and emotions and the improvement of symptom management.

Our app establishes a pain case system such that the patient's pain treatment information is available through a shareable information warehouse. The information warehouse collects basic information regarding patients, history of pain, history of treatment, and user feedback. After the patient is discharged from the hospital, authorized physicians and pharmacists can review the patient's history of pain management at any time and location. Thus, the pain management team can rapidly and effectively manage the patient's pain situation and achieve a seamless connection with the hospital.

Limitations

Because this was a pilot study, illustrative data analysis from the small sample collected over 2 months could only show apparent trends in the management of patients with cancer pain outcomes from baseline to poststudy.

Conclusions

Pain Guard succeeded in connecting hospital staff with discharged patients using a low-cost, conveniently implemented system to facilitate real-time pain recording and timely intervention among Chinese cancer patients. Pain Guard effectively enhanced the management of patients with cancer pain at home, reduced adverse reactions, improved patient medication adherence, and improved patient's QoL. In addition, the software operability was effective and easily accepted by patients.

Acknowledgments

This research was supported by the National Nature Science Foundation of China (71804025); Fujian Provincial Department of Science & Technology (grant #2017Y0035/2015J01471) of the People's Republic of China; Key Project for Youth Academic Talents (2015-ZQN-ZD-13) from Health and Family Planning Commission of Fujian Province.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Quality of life scores comparisons between Pain Guard and control groups.

[PDF File (Adobe PDF File), 292KB - mhealth_v7i5e12542_app1.pdf]

Multimedia Appendix 2

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 2MB - mhealth_v7i5e12542_app2.pdf]

References

1. Tompkins DA, Hobelmann JG, Compton P. Providing chronic pain management in the "Fifth Vital Sign" era: historical and treatment perspectives on a modern-day medical dilemma. Drug Alcohol Depend 2017 Apr 01;173 Suppl 1:S11-S21 [FREE Full text] [doi: 10.1016/j.drugalcdep.2016.12.002] [Medline: 28363315]
2. van den Beuken-van Everdingen MH, de Rijke JM, Kessels AG, Schouten HC, van Kleef M, Patijn J. High prevalence of pain in patients with cancer in a large population-based study in The Netherlands. Pain 2007 Dec 05;132(3):312-320. [doi: 10.1016/j.pain.2007.08.022] [Medline: 17916403]
3. Lara-Solares A, Ahumada Olea M, Basantes Pinos AD, Bistre Cohén S, Bonilla Sierra P, Duarte Juárez ER, et al. Latin-American guidelines for cancer pain management. Pain Manag 2017 Jul;7(4):287-298. [doi: 10.2217/pmt-2017-0006] [Medline: 28326952]
4. Jibb LA, Stevens BJ, Nathan PC, Seto E, Cafazzo JA, Johnston DL, et al. Perceptions of adolescents with cancer related to a pain management app and its evaluation: qualitative study nested within a multicenter pilot feasibility study. JMIR Mhealth Uhealth 2018 Apr 06;6(4):e80 [FREE Full text] [doi: 10.2196/mhealth.9319] [Medline: 29625951]
5. Yang Y, Ma Y, Huang Y, Zhao Y, Xu F, Tian Y, et al. The good pain management (GPM) ward program in China and its impact on Chinese cancer patients: the SYSUCC experience. Chin J Cancer 2014 Jul;33(7):323-329 [FREE Full text] [doi: 10.5732/cjc.014.10031] [Medline: 24874643]

6. Caraceni A, Hanks G, Kaasa S, Bennett MI, Brunelli C, Cherny N, European Palliative Care Research Collaborative (EPCRC), European Association for Palliative Care (EAPC). Use of opioid analgesics in the treatment of cancer pain: evidence-based recommendations from the EAPC. Lancet Oncol 2012 Feb;13(2):e58-e68. [doi: 10.1016/S1470-2045(12)70040-2] [Medline: 22300860]
7. Breivik H, Cherny N, Collett B, de Conno F, Filbet M, Foubert AJ, et al. Cancer-related pain: a pan-European survey of prevalence, treatment, and patient attitudes. Ann Oncol 2009 Aug;20(8):1420-1433. [doi: 10.1093/annonc/mdp001] [Medline: 19244085]
8. Deandrea S, Montanari M, Moja L, Apolone G. Prevalence of undertreatment in cancer pain. A review of published literature. Ann Oncol 2008 Dec;19(12):1985-1991 [FREE Full text] [doi: 10.1093/annonc/mdn419] [Medline: 18632721]
9. Li Y, Yu J, Tang L, Xu B, Fang F, Nie H, et al. Cancer pain management at home: voice from an underdeveloped region of China. Cancer Nurs 2013;36(4):326-334. [doi: 10.1097/NCC.0b013e318268f95b] [Medline: 23047798]
10. Cyberspace Administration of China. 2018. [The 42nd Statistical Report on the Development of China's Internet Network] URL: http://www.cac.gov.cn/2018-08/20/c_1123296882.htm [accessed 2019-04-28] [WebCite Cache ID 77z8LKcyk]
11. Fu MR, Axelrod D, Guth A, Scagliola J, Rampertaap K, El-Shammaa N, et al. A web- and mobile-based intervention for women treated for breast cancer to manage chronic pain and symptoms related to lymphedema: randomized clinical trial rationale and protocol. JMIR Res Protoc 2016 Jan 21;5(1):e7 [FREE Full text] [doi: 10.2196/resprot.5104] [Medline: 26795447]
12. Kessel KA, Vogel MM, Schmidt-Graf F, Combs SE. Mobile apps in oncology: a survey on health care professionals' attitude toward telemedicine, mHealth, and oncological apps. J Med Internet Res 2016 Dec 24;18(11):e312 [FREE Full text] [doi: 10.2196/jmir.6399] [Medline: 27884810]
13. Lalloo C, Jibb LA, Rivera J, Agarwal A, Stinson JN. "There's a pain app for that": review of patient-targeted smartphone applications for pain management. Clin J Pain 2015 Jun;31(6):557-563. [doi: 10.1097/AJP.0000000000000171] [Medline: 25370138]
14. Sun Y, Jiang F, Gu JJ, Wang YK, Hua H, Li J, et al. Development and Testing of an intelligent pain management system (IPMS) on mobile phones through a randomized trial among Chinese cancer patients: a new approach in cancer pain management. JMIR Mhealth Uhealth 2017 Jul 25;5(7):e108 [FREE Full text] [doi: 10.2196/mhealth.7178] [Medline: 28743681]
15. Thurnheer SE, Gravestock I, Pichierri G, Steurer J, Burgstaller JM. Benefits of mobile apps in pain management: systematic review. JMIR Mhealth Uhealth 2018 Oct 22;6(10):e11231 [FREE Full text] [doi: 10.2196/11231] [Medline: 30348633]
16. Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs 2000 Oct;32(4):1008-1015. [Medline: 11095242]
17. Kruse W. Patient compliance with drug treatment-new perspectives on an old problem. Clin Investig 1992 Feb;70(2):163-166. [Medline: 1600344]
18. Jan R, Wang J, Huang M, Tseng S, Su H, Liu L. An internet-based interactive telemonitoring system for improving childhood asthma outcomes in Taiwan. Telemed J E Health 2007 Jun;13(3):257-268. [doi: 10.1089/tmj.2006.0053] [Medline: 17603828]
19. Kessel KA, Vogel MM, Alles A, Dobiasch S, Fischer H, Combs SE. Mobile app delivery of the EORTC QLQ-C30 questionnaire to assess health-related quality of life in oncological patients: usability study. JMIR Mhealth Uhealth 2018 Feb 20;6(2):e45 [FREE Full text] [doi: 10.2196/mhealth.9486] [Medline: 29463489]
20. Jan R, Wang J, Huang M, Tseng S, Su H, Liu L. An internet-based interactive telemonitoring system for improving childhood asthma outcomes in Taiwan. Telemed J E Health 2007 Jun;13(3):257-268. [doi: 10.1089/tmj.2006.0053] [Medline: 17603828]
21. Kessel KA, Vogel MM, Schmidt-Graf F, Combs SE. Mobile apps in oncology: a survey on health care professionals' attitude toward telemedicine, mhealth, and oncological apps. J Med Internet Res 2016 Dec 24;18(11):e312 [FREE Full text] [doi: 10.2196/jmir.6399] [Medline: 27884810]
22. Lin XY, Yang J, Lai JH. [Analysis of factors influencing the compliance of pain treatment in cancer patients]. Chinese J Pain Med 2013;19(6):373-375 [FREE Full text] [doi: 10.3969/j.issn.1006-9852.2013.06.013]
23. Chou WC, Chen JS, Hung CY, Lu C, Shao Y, Chiou T, et al. A nationwide survey of adherence to analgesic drugs among cancer patients in Taiwan: prevalence, determinants, and impact on quality of life. Support Care Cancer 2018 Dec 14:1. [doi: 10.1007/s00520-018-4599-x] [Medline: 30552596]
24. Kwekkeboom KL. Music versus distraction for procedural pain and anxiety in patients with cancer. Oncol Nurs Forum 2003 Feb;30(3):433-440. [doi: 10.1188/03.ONF.433-440] [Medline: 12719743]
25. Keenan A, Keithley JK. Integrative review: effects of music on cancer pain in adults. Oncol Nurs Forum 2015 Nov;42(6):E368-E375. [doi: 10.1188/15.ONF.E368-E375] [Medline: 26488843]
26. Meadows A, Burns DS, Perkins SM. Measuring supportive music and imagery interventions: the development of the music therapy self-rating scale. J Music Ther 2015 Sep;52(3):353-375. [doi: 10.1093/jmt/thv010] [Medline: 26333953]

Abbreviations

BTcP: breakthrough cancer pain
IQR: interquartile range

NRS: numerical rating scale
QoL: quality of life
ULN: upper limits of normal

Edited by G Eysenbach; submitted 18.10.18; peer-reviewed by X Garcia-Eroles, N Khalili-Mahani, R Lee; comments to author 03.01.19; revised version received 24.02.19; accepted 25.03.19; published 29.05.19.

Please cite as:
Yang J, Weng L, Chen Z, Cai H, Lin X, Hu Z, Li N, Lin B, Zheng B, Zhuang Q, Du B, Zheng Z, Liu M
Development and Testing of a Mobile App for Pain Management Among Cancer Patients Discharged From Hospital Treatment: Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(5):e12542
URL: http://mhealth.jmir.org/2019/5/e12542/
doi:10.2196/12542
PMID:31144672

Original Paper

An App Detecting Dengue Fever in Children: Using Sequencing Symptom Patterns for a Web-Based Assessment

Tsair-Wei Chien[1*], MBA; Julie Chi Chow[2*], MD; Willy Chou[3,4*], MD

[1]Data Analyses & Statistics, Medical Research, Chi-Mei Medical Center, Tainan, Taiwan

[2]Pediatrics, Chi-Mei Medical Center, Tainan, Taiwan

[3]Physical Medicine and Rehabilitation, CHia-Li Campus, Chi-Mei Medical Center, Tainan, Taiwan

[4]Department of Recreation and Health-Care Management, Chia Nan University of Pharmacy, Tainan, Taiwan

[*]all authors contributed equally

Corresponding Author:
Willy Chou, MD
Physical Medicine and Rehabilitation
CHia-Li Campus
Chi-Mei Medical Center
901 Zhonghua Road, Yung-Kang District, Tainan
Tainan, 710
Taiwan
Phone: 886 62812811 ext 32000
Email: ufan0101@ms22.hinet.net

Abstract

Background: Dengue fever (DF) is one of the most common arthropod-borne viral diseases worldwide, particularly in South East Asia, Africa, the Western Pacific, and the Americas. However, DF symptoms are usually assessed using a dichotomous (ie, absent vs present) evaluation. There has been no published study that has reported using the specific sequence of symptoms to detect DF. An app is required to help patients or their family members or clinicians to identify DF at an earlier stage.

Objective: The aim of this study was to develop an app examining symptoms to effectively predict DF.

Methods: We extracted statistically significant features from 17 DF-related clinical symptoms in 177 pediatric patients (69 diagnosed with DF) using (1) the unweighted summation score and (2) the nonparametric *HT* person fit statistic, which can jointly combine (3) the weighted score (yielded by logistic regression) to predict DF risk.

Results: A total of 6 symptoms (family history, fever ≥39°C, skin rash, petechiae, abdominal pain, and weakness) significantly predicted DF. When a cutoff point of >–0.68 (*P*=.34) suggested combining the weighted score and the *HT* coefficient, the sensitivity was 0.87, and the specificity was 0.84. The area under the receiver operating characteristic curve was 0.91, which was a better predictor: specificity was 10.2% higher than it was for the traditional logistic regression.

Conclusions: A total of 6 simple symptoms analyzed using logistic regression were useful and valid for early detection of DF risk in children. A better predictive specificity increased after combining the nonparametric *HT* coefficient with the weighted regression score. A self-assessment using patient mobile phones is available to discriminate DF, and it may eliminate the need for a costly and time-consuming dengue laboratory test.

(JMIR Mhealth Uhealth 2019;7(5):e11461) doi:10.2196/11461

KEYWORDS

dengue fever; HT person mapping statistic; logistic regression; score summation; receiver operating characteristic curve

Introduction

Symptoms of Dengue Fever

Dengue fever (DF) is one of the most common arthropod-borne viral diseases worldwide [1], especially in South East Asia, Africa, the Western Pacific, and the Americas [2,3]. However, there is no accurate and speedy diagnostic screening test for DF at an early stage, as its signs and symptoms—for example, fever, headache, and myalgia—are similar to those of other illnesses [4-6]. Some studies [4,5] that used a univariate analysis report that the presumptive diagnosis of DF is imprecise. Multivariate logistic regressions also do not

significantly distinguish patients with dengue from those with other febrile illnesses [7]. The multivariate discrimination analyses reported sensitivity and a specificity 0.76 and an area under the receiver operating characteristic (ROC) curve (AUC) of 0.93, but costly laboratory tests (Dengue Duo Immunoglobulin M and Rapid Strips, Panbio, Queensland, Australia) [8-11] were needed before DF was serologically confirmed.

Assessment of Dengue Fever

DF symptoms are usually assessed using a dichotomous (ie, absent vs present) evaluation. The dependent variable (DF^+ vs DF^-) predicted using independent evaluations with a weighted summation score is more accurate than that predicted using simple evaluations with an unweighted summation score. So far, there has been no published study that has reported using the specific sequence of symptoms reported or observed in specific patients suspected of having DF. All published studies to date still report results using only a standard group of symptoms with an unweighted summation score, and they merely apply their results to a general group of patients who might have DF.

The HT Fit Statistic Applied to Detect Dengue Fever

The nonparametric HT fit statistic has been used in education and psychometrics to identify aberrant test respondents [12,13]. It is a transposed formulation of a scalability coefficient for items (eg, symptoms in this study), and it is the best among 36-person fit statistics for detecting abnormal behaviors [14].

Objectives

In this study, we used the *HT* coefficient combined with weighted and unweighted variables to examine whether these combinations provide a valid and reliable approach for the early detection of DF in children.

Methods

Sample and Clinical Symptoms

The sample of 177 pediatric patients (≤16 years old; DF^+: 69; DF^-: 108) was the same as in our previous paper [8] (see data in Multimedia Appendix 1). Guided by the literature [5-7], we collected 19 DF-related clinical symptoms from the patients' medical records to develop the initial set of items—designated as 0="absent" or 1="present"—to screen for DF infection: (1) personal history of DF, (2) family history of DF, (3) mosquito bites within the previous 2 weeks, (4) fever ≥39°C, (5) biphasic fever, (6) rash, (7) petechiae, (8) retroorbital pain, (9) bone pain (arthralgia), (10) headache, (11) myalgia, (12) abdominal pain, (13) anorexia, (14) occult hematuria, (15) stool occult blood, (16) cough, (17) sore throat, (18) soft (watery) stool, and (19) flushed skin. Data from these patients' charts were obtained and approved by the Research Ethics Review Board of the Chi-Mei Medical Center.

The HT Fit Statistic

HT is defined for the persons of a dichotomous dataset with *L* items (in columns) and *N* persons (in rows) [12-14], where X_{ni} is the scored (0,1) response of person *n* to item *i*, and $P_n=S_n/L$. Here, S_m is the raw score for person *m*, and S_n is the raw score for person *n*.

HT is the sum of the covariances between person *n* and the other persons divided by the maximum possible sum of those covariances so that the range of *HT* is from −1 to +1, see formula (1) in Figure 1. When the responses by person *n* are positively correlated with those of all the other persons, then *HT* (*n*) will be positive. In contrast, when the responses by person *n* are negatively correlated with those of all the other persons, then *HT* (*n*) will be negative. When person *n* 's responses are random, *HT* (*n*) will be close to zero [11]. We hypothesized that DF^+ patients have different *HT* coefficients than DF^- patients. All DF^+ group members were sequenced to the DF^- group members to obtain an *HT* coefficient using formula (1) in Figure 1.

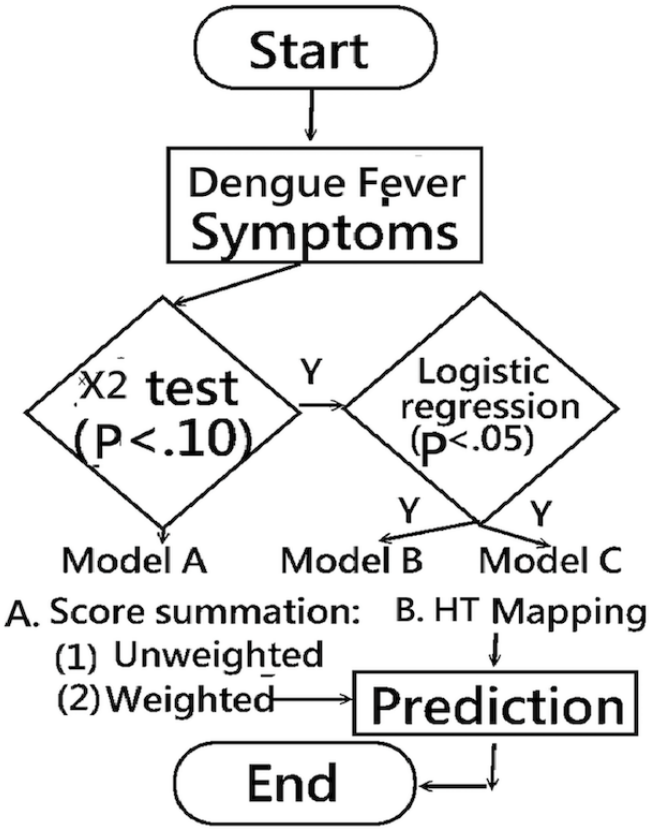
Figure 1. The equation of the HT fit statistic.

$$HT(n)=\frac{\sum_{m=1,m\neq n}^{N}\left(\left[\sum_{i=1}^{L}X_{ni}X_{mi}\right]/L-p_nP_m\right)}{\sum_{m=1,m\neq n}^{N}\left(\min\left[p_n(1-p_m),p_m(1-p_n)\right]\right)'} \quad (1)$$

Selecting Symptoms and Determining Predictor Variables

All symptoms were examined by the probability of Type 1 error using the following 3 steps in Figure 2 to determine predictor variables. First, each symptom was separately examined by the univariate approach using a Chi-square test and logistic regression, respectively, for identifying a significant association with DF. Second, 2 models (ie, the univariate and the multivariate approaches) were investigated for determining valid predictor variables associated with DF when the probability of Type 1 error was less than .05. Third, the predictor variables were used in a weighted combination for discriminating patients suspected with dengue virus infection.

Figure 2. Overall study concept and the flow chart.



Detecting Dengue Fever: A Comparison of Three Models

The efficacy of 3 models (A, B, and C) for detecting dengue fever was examined: (1) A comparison was made using univariate logistic regression in Model A to examine effects through the AUC, yielded by unweighted (ie, summed item) scores, weighted (ie, logistic regression) scores, and *HT* coefficients, respectively. (2) Multivariate logistic regression with the 3 aforementioned factors combined was used in Model B. (3) After selecting the significant variables in Model B, the combined predictive variables were analyzed using multivariate logistic regression in Model C to obtain effective weighted coefficients. (4) Finally, we wanted to use a single continuous variable yielded by the combined predictive variables in Model C to compare the AUC with the counterparts in Model A and C.

Moreover, we provide the F-measure for evaluating the predictive effect [15], which is calculated by following equations: precision=True Positives/(True Positives+False Positives); recall=True Positives/(True Positives+False Negatives); F-measure=(2×precision×recall)/(precision+recall).

Statistical Tools and Data Analyses

SPSS 15.0 for Windows (SPSS Inc) and MedCalc 9.5.0.0 for Windows (MedCalc Software) were used to calculate (1) the probability of false positives (Type 1 error) using a Chi-square test and logistic regression, (2) Youden J index (the higher, the better), AUC, sensitivity, specificity, and the cutoff point at maximal summations of specificity and sensitivity, (3) correlation coefficients among variables of unweighted, weighted, and *HT* scores.

Results

Demographic Characteristics of the Study Sample and the Likelihood of Dengue Fever

A total of 69 pediatric patients clinically diagnosed with DF and 108 pediatric patients with no evidence of DF infection were included in this study (Table 1). A Chi-square test and logistic regression analyses showed that only 6 symptoms (family history, fever ≥39°C, skin rash, petechiae, abdominal pain, and weakness) were significant for assessing the likelihood of DF (Table 2).

Table 1. Demographic characteristics of the study sample.

Demographical variables	Dengue fever (–)[a], n (%)	Dengue fever (+)[b], n (%)	Total, n (%)	*P* value[c]
Gender				
Female	47 (43.5)	29 (42)	76 (42.9)	.84
Male	61 (56.5)	40 (58)	101 (57.1)	—[d]
Age (years)				
0-4	48 (44.4)	11 (16.2)	59 (33.5)	.005
5-9	24 (22.2)	20 (29.4)	44 (25)	—
9-16	36 (33.3)	37 (54.4)	73 (41.5)	—

[a]Dengue fever (–): patients with a negative dengue fever strip test.

[b]Dengue fever (+): patients with a positive dengue fever strip test.

[c]*P* values were determined by the Chi-square test.

[d]Not applicable.

Table 2. Logistic analysis of symptoms for the patients suspected with dengue virus infection using the univariate approach.

Symptom variables and presence	Dengue fever (–)[a], n (%)	Dengue fever (+)[b], n (%)	Total, n (%)	Chi-square (df)	*P* value[c]	Logistic regression	
						Beta	*P* value
Family history							
No	79 (73.1)	40 (58.0)	119 (67.2)	3.7(2)	.053	1.35	.002
Yes	29 (26.9)	29 (42.0)	58 (32.8)	—[d]	—	—	—
High fever of 39°C							
No	87 (80.6)	37 (53.6)	124 (70.1)	13.3(2)	<.001	1.48	.048
Yes	21 (19.4)	32 (46.4)	53 (29.9)	—	—	—	—
Skin rash							
No	82 (75.9)	20 (29.0)	102 (57.6)	36.1(2)	<.001	2.63	.000
Yes	26 (24.1)	49 (71.0)	75 (42.4)	—	—	—	—
Petechiae							
No	106 (98.1)	60 (87.0)	166 (93.8)	7.3(2)	.007	2.34	.026
Yes	2 (1.9)	9 (13.0)	11 (6.2)	—	—	—	—
Abdominal pain							
No	104 (96.3)	53 (76.8)	157 (88.7)	14.1(2)	<.001	2.89	.000
Yes	4 (3.7)	16 (23.2)	20 (11.3)	—	—	—	—
Weak sense							
No	90 (83.3)	48 (69.6)	138 (78.0)	3.9(2)	.049	0.98	.048
Yes	18 (16.7)	21 (30.4)	39 (22.0)	—	—	—	—
Constant							
—	—	—	—	—	—	–3.28	—

[a]Dengue fever (–): patients with a negative dengue fever strip test.

[b]Dengue fever (+): patients with a positive dengue fever strip test.

[c]*P* values were determined by the Chi-square test and the Wald test of logistic regression.

[d]Not applicable.

Comparisons of the Area Under Receiver Operating Characteristic Curve for the Three Study Models

Comparisons of the AUCs for the 3 study models (A, B, and C) showed that the weighted variable (derived by the Logistic regression) and the *HT* coefficient could be jointly used for predicting DF risk using equation (2):

$$(\ Logit = -3.32 + 0.93 \times weighted\_score + 1.92 \times HT\_coefficient) \quad (2)$$

The risk probability can be computed using the transformed formula 3:

P=exp (log *it*)/ (1+exp(log *it*)) (3)

where *logit* denotes a unit of log odds.

A cutoff point of >–0.68 (*P*=.34) was determined using the combined predictive variables in Model C: sensitivity=0.91, specificity=0.76, AUC=0.88, and the highest F-measure=0.82 (see Figure 3 and Table 3). Predictive power was better: specificity was 10.2% (ie, 84.30–74.10, shown in Table 3) higher than when using traditional logistic regression, *that is, the independence variable=sum (weighted score for each symptom* x *the respective symptom response, 1 or 0, predicting the dependence variable, 1 or 0 for DF*). Even if AUC using the *HT* coefficient was slightly lower (0.72) than when using the unweighted (0.84) and the weighted (0.87) variables (*Table3*), and the *HT* coefficients related to the weighted and unweighted scores were 0.26 and 0.22, respectively, the weighted score had a higher correlation coefficient than the unweighted score to the *HT* coefficients, and the combined strategy of Model C or the single continuous variable yielded by the combined predictor variables (Table 3) are verified and available for use in practice. *More importantly, the sensitivity is more critical than the specificity in clinical settings, as we would not miss any 1 case with fatal diseases.*

Figure 3. Four models plotted by receiver operating characteristic curves. The Combined denotes Model C in this study (sensitivity=0.87, specificity=0.84, area under the receiver operating characteristic curve=0.91, F-measure=0.82).

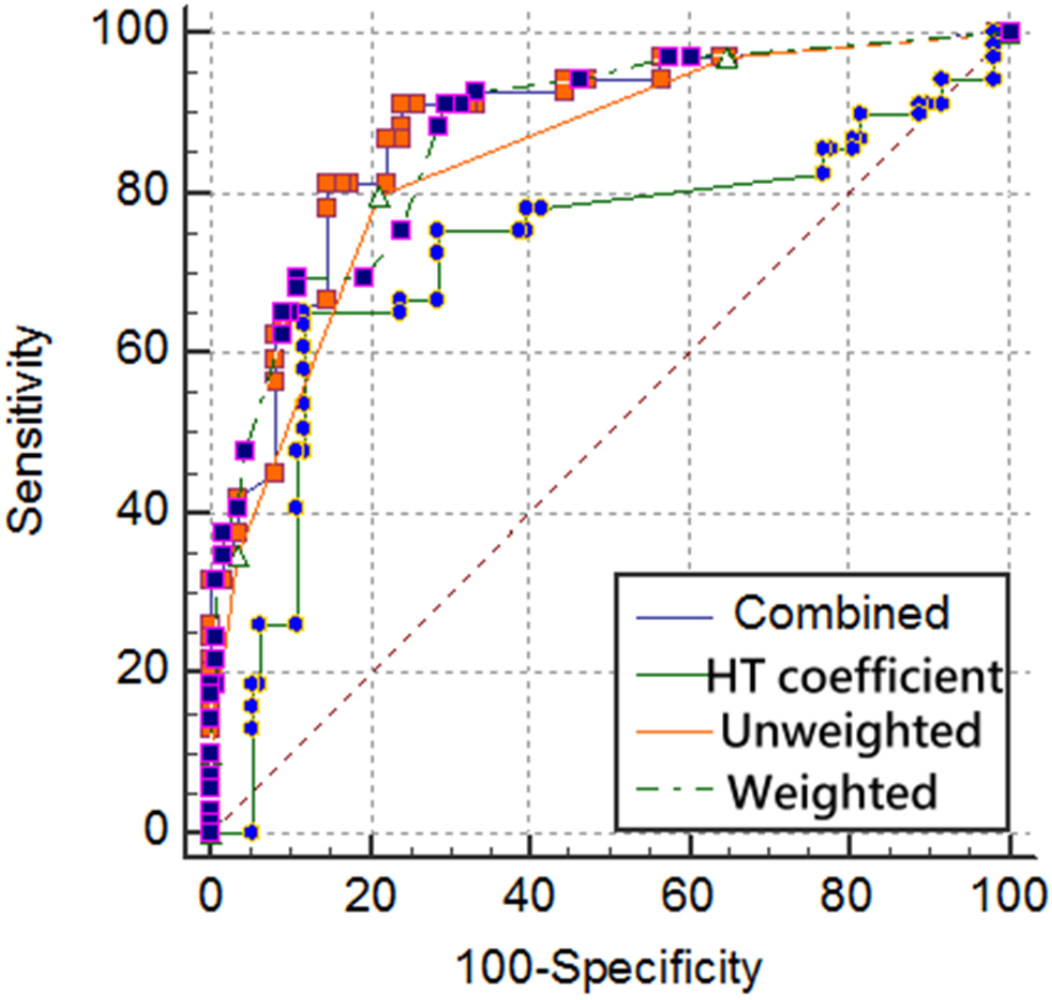


Table 3. Comparisons of area under receiver operating characteristic curve for the study models.

Approach and steps	Logistic regression		Receiver operating characteristic curve analysis					F-measure
	B[a]	*P* value	Area under receiver operating characteristic curve	Youden J[b]	Cut point	Sensitivity	Specificity	
Comparison of models								
Model A: *Univariate approach with a single variable compared with the dengue fever using Logistic regression and receiver operating characteristic analysis*								
Unweight[c]	1.60[d]	<.001	0.84	0.58	>1.00	79.7	78.7	—[e]
Weight[f]	0.97[d]	<.001	0.89	0.61	>–1.20	91.3	74.1	—
HT coefficient[g]	3.75[d]	<.001	0.72	0.53	>0.15	65.2	88	—
Model B: *Multivariate approach with combined these three variables in regressing the dengue fever using Logistic regression*								
Unweight	0.31	.595	—	—	—	—	—	—
Weight	0.77[d]	.014	—	—	—	—	—	—
HT coefficient	3.08[d]	.001	—	—	—	—	—	—
Constant	–1.03	.35	—	—	—	—	—	—
Model C *: Combined these 2 significant predictor variables using Logistic regression*								
Weight	0.919[d]	<.001	—	—	—	—	—	—
HT coefficient	2.962[d]	.001	—	—	—	—	—	—
Constant	–0.463	.751	—	—	—	—	—	—
A single continuous variable *yielded by the combined predictor variables in Model C*								
Combined[h]	1	<.001	0.91	0.71	>–0.68	87	84.3	—
The predictive effect: precision recall								
Unweight	—	.72	0.85	—	—	—	—	0.78
Weight	—	.93	0.65	—	—	—	—	0.77
HT coefficient	—	.78	0.82	—	—	—	—	0.8
The combined model	—	.87	0.78	—	—	—	—	0.82

[a]B: coefficient of logistic regression.

[b]Youden J index.

[c]Item-score summation method.

[d]*P*<.05.

[e]Not applicable.

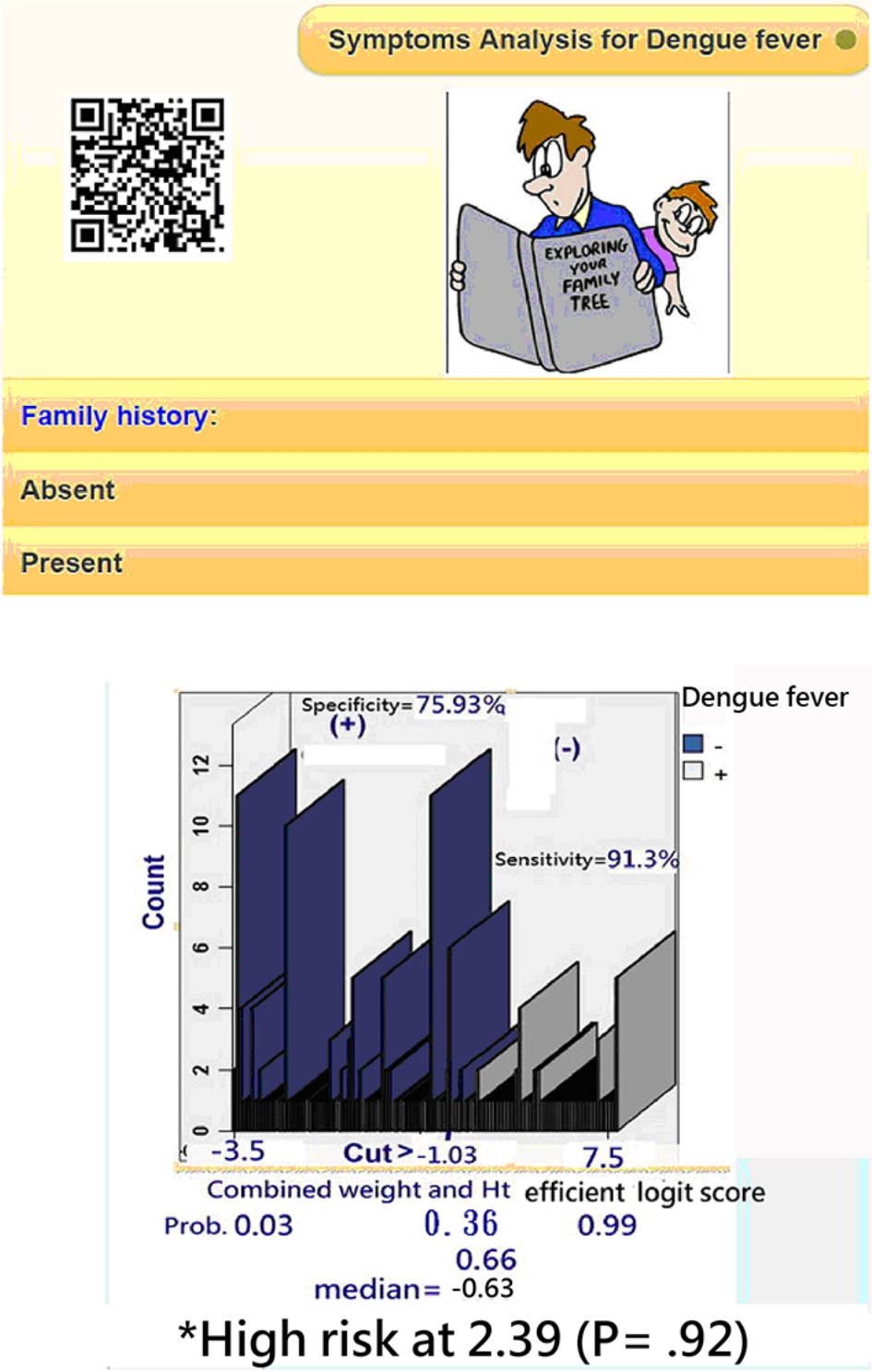
[f]Multiplying item score with the weighted regression coefficient.

[g]See Figure 1 for the HT equation

[h]Using the 2 combined variables to predict patient's dengue fever.

A snapshot on a mobile phone responding to questions (Figure 4, top) was generated, and the results for assessing whether the patient has DF (Figure 4, bottom) were determined, which indicated that patients suspected of having DF could directly scan the Quick Response Code to obtain their DF *logit* scores (or the risk probability) and examine whether these 6 symptoms are useful for predicting a high DF risk (>−1.03 *logits* or *P* ≥.26=exp(−1.03 logits)/(1+exp(-1.03 logits)). Interested readers are recommended to see the demonstration in Multimedia Appendix 2 using a MP4 video to display.

Figure 4. Snapshots on a mobile phone responding questions (top) and the result (bottom) for assessing the patient dengue fever.



Discussion

Principal Findings

We found that using the HT coefficient yielded predictions that were 10.2% more specific (ie, 84.30–74.10, shown in Table 3) than those of traditional logistic regression. The *HT* index is promising when the patient sequence symptom pattern is compared with the DF^+ group to detect dengue fever in children. It can be combined with the weighted summation score to jointly predict the DF risk and then report that risk on mobile phones.

The *HT* coefficient has been used in education and psychometrics to identify aberrant test respondents [12-14]. Although some have used item response theory fit statistics (eg, outfit mean square error >2.0) to select abnormal responses that indicate cheating, careless responding, lucky guessing, creative responding, or random responding [16], our literature review revealed no published papers that reported using the *HT* coefficient in medical settings, especially for detecting individual aberrant response patterns different from the study reference sample, or, like this study, identifying the DF risk by

comparing their sequence symptom pattern with that of the DF^+ group.

What This Knowledge Adds to What We Already Knew

A diagnosis of DF is usually confirmed by 3 steps: (1) observing DF-related symptoms, (2) testing laboratory data, such as white blood cells and platelets, and (3) serologically verifying DF using dengue Immunoglobulin M and Immunoglobulin G antibodies, polymerase chain reaction analysis, and virus isolation tests [8]. The latter 2 are relatively expensive. It is needed to develop a self-assessment approach (eg, scanning Quick Response Code, responding questions, and obtaining the DF risk on his/her smartphone), (1) helping patients for consultation at an earlier stage and (2) prompting doctors for sampling patient laboratory data when his/her DF risk reaches a cut point of P=.26 (=exp(−1.03 logits)/(1+exp(−1.03 logits)).

We found that the weighted score was a better predictor than the unweighted score (see Model A and Model B in Table 3). However, we still see so many scales in a medical setting using unweighted summation scores to determine the presence or absence of disease. Along with the mobile phones popularly used in the technical age, the way of obtaining the DF risk on mobile phones using the combined *HT* coefficient and weighted scores is available and worth recommending to health care providers to use for detecting the risk for DF.

Limitations and Future Study

This study has some limitations. First, the DF cut point based on the symptoms of this study sample might be biased toward that population. Moreover, we did not remove abnormal data when the *HT* coefficient was less than the critical value of 0.22, which best identifies aberrantly responding examinees [14]. Second, although the sample size was small, using the *HT* coefficient combined with the AUC yielded highly accurate discriminatory screening. However, this finding requires confirmation in prospective studies of other regions with a substantial incidence of DF. Third, the study sample size (=177) is too small to make the inference reliable and supportable. More DF patients collected in a study are required to be considered in the discernable future. Particularly, artificial intelligence (AI) has become increasingly prevalent in recent years.

Conclusions

Analyzing 6 simple symptoms using logistic regression is useful and valid for the early detection of DF risk in children. Combining the *HT* coefficient with the weighted score yields a prediction that is 10.2% more specific than that yielded by traditional logistic regression. A self-assessment app using patient mobile phones is available to help people suspected of having DF, and it might eliminate the need for costly and time-consuming laboratory tests.

Authors' Contributions

TWC conceived and designed the study, performed the statistical analyses, and was in charge of recruiting study participants. CC and TWC helped design the study, collected information, and interpreted data. WC monitored the research. All authors read and approved the final article. This research was supported by the grant Chi-Mei Foundation Hospital research CMFCR10593 from the Chi-Mei Medical Center.

Conflicts of Interest

None declared.

Multimedia Appendix 1
Data for the sample of 177 pediatric patients used in this study.
[XLSX File (Microsoft Excel File), 48 KB - mhealth_v7i5e11461_app1.xlsx]

Multimedia Appendix 2
How to run the check on DF online .
[MP4 File (MP4 Video), 4848 KB - mhealth_v7i5e11461_app2.mp4]

Multimedia Appendix 3
Response to the editors.
[DOCX File , 14 KB - mhealth_v7i5e11461_app3.docx]

References

1. World Health Organization. Geneva; 2002. Dengue and Dengue Haemorrhagic Fever. Fact sheet N°117 URL: http://apps.searo.who.int/PDS_DOCS/B5318.pdf [accessed 2019-05-06] [WebCite Cache ID 78A5N8kGU]
2. Henchal EA, Putnak JR. The dengue viruses. Clin Microbiol Rev 1990 Oct;3(4):376-396 [FREE Full text] [Medline: 2224837]
3. Gubler DJ. Dengue and dengue hemorrhagic fever. Clin Microbiol Rev 1998 Jul;11(3):480-496 [FREE Full text] [Medline: 9665979]

4. Phuong HL, de Vries PJ, Nga TT, Giao PT, Hung LQ, Binh TQ, et al. Dengue as a cause of acute undifferentiated fever in Vietnam. BMC Infect Dis 2006 Jul 25;6:123 [FREE Full text] [doi: 10.1186/1471-2334-6-123] [Medline: 16869969]
5. Nunes-Araújo FR, Ferreira M, Nishioka S. Dengue fever in Brazilian adults and children: assessment of clinical findings and their validity for diagnosis. Ann Trop Med Parasitol 2003 Jun;97(4):415-419. [doi: 10.1179/000349803235002263] [Medline: 12831527]
6. Hammond S, Balmaseda A, Pérez L. Differences in dengue severity in infants, children, and adults in a 3-year hospital-based study in Nicaragua. Am J Trop Med Hyg 2005;73:1063-1070. [Medline: 16354813]
7. Potts J, Rothman A. Clinical and laboratory features that distinguish dengue from other febrile illnesses in endemic populations. Trop Med Int Health 2008 Nov;13(11):1328-1340 [FREE Full text] [doi: 10.1111/j.1365-3156.2008.02151.x] [Medline: 18803612]
8. Lai W, Chien T, Lin H, Su S, Chang C. A screening tool for dengue fever in children. Pediatr Infect Dis J 2013 Apr;32(4):320-324. [doi: 10.1097/INF.0b013e31827e111e] [Medline: 23249912]
9. Lai W, Chien T, Lin H, Kan W, Su S, Chou M. An approach for early and appropriate prediction of dengue fever using white blood cells and platelets. HealthMED 2012;6(3):806-812 [FREE Full text]
10. Kittigul L, Suankeow K. Use of a rapid immunochromatographic test for early diagnosis of dengue virus infection. Eur J Clin Microbiol Infect Dis 2002 Mar;21(3):224-226. [doi: 10.1007/s10096-001-0691-z] [Medline: 11957027]
11. Vaughn D, Nisalak A, Kalayanarooj S, Solomon T, Dung NM, Cuzzubbo A, et al. Evaluation of a rapid immunochromatographic test for diagnosis of dengue virus infection. J Clin Microbiol 1998 Jan;36(1):234-238 [FREE Full text] [Medline: 9431954]
12. Sijtsma K. A coefficient of deviant response patterns. Kwantitatieve Methoden 1986;7:131-145 [FREE Full text]
13. Linacre J. A comment on the HT person fit statistic. Rasch Meas Trans 2012;26(1):1358 [FREE Full text]
14. Karabatsos G. Comparing the aberrant response detection performance of thirty-six person-fit statistics. Appl Meas Educ 2003;16(4):277-298 [FREE Full text] [doi: 10.1207/S15324818AME1604_2]
15. Lin C, Hsu C, Lou Y, Yeh S, Lee C, Su S, et al. Artificial intelligence learning semantics via external resources for classifying diagnosis codes in discharge notes. J Med Internet Res 2017 Dec 6;19(11):e380 [FREE Full text] [doi: 10.2196/jmir.8344] [Medline: 29109070]
16. Linacre J. Optimizing rating scale category effectiveness. J Appl Meas 2002;3(1):85-106 [FREE Full text] [Medline: 11997586]

Abbreviations

AI: artificial intelligence
AUC: area under receiver operating characteristic curve
DF: dengue fever
ROC: receiver operating characteristic

Edited by G Eysenbach; submitted 01.07.18; peer-reviewed by S Su, I Yoo; comments to author 08.10.18; revised version received 02.12.18; accepted 07.04.19; published 31.05.19.

Please cite as:
Chien TW, Chow JC, Chou W
An App Detecting Dengue Fever in Children: Using Sequencing Symptom Patterns for a Web-Based Assessment
JMIR Mhealth Uhealth 2019;7(5):e11461
URL: https://mhealth.jmir.org/2019/5/e11461
doi:10.2196/11461
PMID:31152525

Original Paper

A Mobile Prenatal Care App to Reduce In-Person Visits: Prospective Controlled Trial

Kathryn I Marko[1], MD; Nihar Ganju[1], MD; Jill M Krapf[2], MD; Nancy D Gaba[1], MD; James A Brown[3], MS; Joshua J Benham[1], MD; Julia Oh[4], PhD; Lorna M Richards[5], MA; Andrew C Meltzer[5], MD, MS

[1]The George Washington University School of Medicine and Health Sciences, Washington, DC, United States

[2]OB Hospitalist Group, Baylor All-Saints Medical Center, Fort Worth, TX, United States

[3]University of Pittsburgh School of Medicine, Pittsburgh, PA, United States

[4]The Jackson Laboratory, Farmington, CT, United States

[5]Department of Emergency Medicine, George Washington University School of Medicine and Health Sciences, Washington, DC, United States

Corresponding Author:
Andrew C Meltzer, MD, MS
Department of Emergency Medicine
George Washington University School of Medicine and Health Sciences
2120 L Street NW
Suite 450
Washington, DC, 20037
United States
Phone: 1 202 741 2952
Email: ameltzer@mfa.gwu.edu

Abstract

Background: Risk-appropriate prenatal care has been asserted as a way for the cost-effective delivery of prenatal care. A virtual care model for prenatal care has the potential to provide patient-tailored, risk-appropriate prenatal educational content and may facilitate vital sign and weight monitoring between visits. Previous studies have demonstrated a safe reduction in the frequency of in-person prenatal care visits among low-risk patients but have noted a reduction in patient satisfaction.

Objective: The primary objective of this study was to test the effectiveness of a mobile prenatal care app to facilitate a reduced in-person visit schedule for low-risk pregnancies while maintaining patient and provider satisfaction.

Methods: This controlled trial compared a control group receiving usual care with an experimental group receiving usual prenatal care and using a mobile prenatal care app. The experimental group had a planned reduction in the frequency of in-person office visits, whereas the control group had the usual number of visits. The trial was conducted at 2 diverse outpatient obstetric (OB) practices that are part of a single academic center in Washington, DC, United States. Women were eligible for enrollment if they presented to care in the first trimester, were aged between 18 and 40 years, had a confirmed desired pregnancy, were not considered *high-risk*, and had an iOS or Android smartphone that they used regularly. We measured the effectiveness of a virtual care platform for prenatal care via the following measured outcomes: the number of in-person OB visits during pregnancy and patient satisfaction with prenatal care.

Results: A total of 88 patients were enrolled in the study, 47 in the experimental group and 41 in the control group. For patients in the experimental group, the average number of in-person OB visits during pregnancy was 7.8 and the average number in the control group was 10.2 (P=.01). There was no statistical difference in patient satisfaction (P>.05) or provider satisfaction (P>.05) in either group.

Conclusions: The use of a mobile prenatal care app was associated with reduced in-person visits, and there was no reduction in patient or provider satisfaction.

Trial Registration: ClinicalTrials.gov NCT02914301; https://clinicaltrials.gov/ct2/show/NCT02914301 (Archived by WebCite at http://www.webcitation.org/76S55M517)

(JMIR Mhealth Uhealth 2019;7(5):e10520) doi:10.2196/10520

KEYWORDS

prenatal care; mobile applications; patient monitoring; patient safety; patient satisfaction; technological innovations; controlled clinical trial; mobile health

Introduction

Background

In the United States, there are nearly 4 million live births each year. This makes prenatal care one of the most widely utilized preventative care health services [1,2]. Despite its widespread use, the effectiveness and organization of standard prenatal care has been debated [3-5]. Rigorous scientific evidence of the effectiveness of standard prenatal care, including effects on maternal and infant outcomes, health-related behaviors, and health care costs is limited [2]. In the early mid-1980s, an expert panel recommended stratifying women into high- and low-risk categories, with high-risk women receiving more intensive prenatal care and low-risk women following a reduced visit frequency schedule [6]. The rationale for this recommendation was that unnecessary visits for low-risk patients consume health care resources that could be applied more judiciously to women with higher-risk pregnancies. However, despite these recommendations, the standard model of prenatal care with high-frequency visits has persisted. Almost a third of low-risk women receive more visits than recommended by the American College of Obstetricians and Gynecologists (ACOG) [7]. The barriers cited by providers for not reducing visits for low-risk pregnancies include the following: concern about decreased patient satisfaction, need for frequent weight and blood pressure monitoring, and concern about reduced transmission of educational information regarding health and lifestyle choices during pregnancy.

Goal of This Study

The primary objective of this study was to determine if a mobile prenatal care app facilitates a reduced in-person prenatal care visit schedule while maintaining patient and provider satisfaction. Mobile health apps have the potential to address many of the perceived barriers to reducing in-person visits [8]. The application of mobile phone technology has been shown to improve disease management for diabetes self-care activities, HIV infection medication adherence, and sickle cell anemia medication adherence [9-11]. We hypothesize that a similar approach using a mobile app and connected monitoring devices may also be beneficial to manage prenatal care. A pilot trial with 8 patients was performed. This trial ensured the platform and devices functioned appropriately, any blood pressure and weight triggers were identified and managed, a reduced visit schedule was achieved on a small scale, and the platform was satisfactory to patients. The feedback from the pilot guided this investigation [12].

Methods

Study Design

This pragmatic controlled trial compared an experimental group that received a mobile app for prenatal care and an integrated Wi-Fi-connected blood pressure and weight scale with a control group that received usual care. Institutional Review Board (IRB) approval has been obtained from the George Washington University (GWU) IRB (IRB#: 015422).

Study Setting

The educational components and clinical triggers were developed and refined at GWU in Washington, DC, United States, working in conjunction with a local mobile health technology firm 1EQ and their product Babyscripts [13]. The clinical trial occurred in 2 obstetric (OB) and gynecology (GYN) offices in the United States: one in downtown Washington, DC, and one in suburban Maryland. Prenatal care is provided at both locations by OB and GYN physicians and nurse midwives. Low-risk women are cared for by obstetrician-gynecologists at both locations, and all deliveries take place at the GWU hospital in Washington, DC. Enrollment occurred between July 2015 and March 2016.

Inclusion and Exclusion Criteria

Eligible participants were women aged between 18 and 40 years, presenting for a first trimester (up to 13 weeks gestational age) verification of pregnancy or new OB visit, and who were considered low-risk. Low-risk was defined as a singleton pregnancy with no previous diagnosis of essential hypertension, diabetes, renal disease, collagen vascular disease, maternal substance abuse, or other previously documented condition posing a high risk of poor pregnancy outcome. Dropout criteria for the study included the antepartum diagnosis of fetal abnormalities, placenta previa, intrauterine growth restriction, pregnancy-induced hypertension, gestational diabetes, or premature rupture of membranes. (Multimedia Appendix 1) Participants were also required to regularly use a mobile phone and be fluent in English.

Allocation

Allocation into the experimental group versus the control group was based on the operating system of the patient's mobile phone. Patients with iOS-based mobile phones (ie, iPhones) were allocated to the experimental group; patients with Android or Windows operating systems were allocated to the control group. This allocation system was chosen as a practical solution to the challenges of randomization and blinding. In addition, the app had yet to be developed on the Android platform. Allocation was concealed until after consent was obtained.

Table 1. Sample alternative prenatal care schedules.

Schedule	New OB[a] visit	Week number												Total visits
		16	20	24	28	30	32	34	36	37	38	39	40	
Traditional prenatal care schedule (1=visit)	1	1	1	1	1	1	1	1	1	1	1	1	1	13
Babyscripts prenatal care schedule (0=no visit, 1=visit)	1	1	0	1	1	0	1	0	1	0	1	1	1	9

[a]OB: obstetric.

Study Arms

Participants who were assigned to the experimental group were instructed to download the mobile app and set up the connected monitoring devices at the time of enrollment. Participants incurred no additional costs for the app or connected devices. Participants allocated to the control arm were not given access to the mobile app. The experimental group was also placed on an alternative prenatal care schedule of 8 expected visits as compared with the typical 12-14 expected visits in the control group (Table 1). The experimental visit schedule was based on the Department of Defense-Veterans Affairs Uncomplicated Pregnancy Guidelines [14]. For all participants, prenatal care met established guidelines for the management of uncomplicated pregnancies.

Data Collection

Patient demographics were obtained by self-report at the time of enrollment. Detailed patient characteristics were also collected at the time of enrollment, including age, race, insurance status, and ethnicity. The primary outcome of the study was the number of in-person prenatal care visits as assessed by patient chart review. As a secondary outcome, all patients were evaluated for satisfaction with their prenatal care experience and pregnancy outcomes. (Multimedia Appendix 2) Patient satisfaction surveys were administered to participants at gestational weeks 16, 20, 25, 30, and 35 and 2 weeks postpartum. Satisfaction survey consisted of 16 questions using a 4-point Likert scale modified from the hospital consumer assessment of health care providers and systems survey instrument [15] In addition, 6 questions specific to the Babyscripts platform were submitted to participants in the experimental group. Participants who completed all 6 surveys and responded to all related questionnaires were compensated with a US$20 Amazon gift card at completion of the study. An additional secondary outcome was obtained via a chart review for significant clinical outcomes—preeclampsia, eclampsia, neonatal intensive care unit admissions, stillbirth, neonatal mortality, other serious outcome per investigator judgment, and route of delivery including cesarean delivery rate. Chart reviews were completed using trained abstractors with defined data collection sheets.

Statistical Analysis

We compared outcomes between study arms using *t* tests for continuous outcomes and chi-square tests for dichotomous outcomes. For differences in the baseline characteristics between the 2 study groups, we used a hierarchical generalized logistic or linear regression model that includes an indicator for the study arm. We conducted descriptive heterogeneity of treatment effect analyses by age, gender, race/ethnicity, and highest attained maternal education level.

Description of Mobile App

The Babyscripts app was designed with 2 major goals: (1) to deliver educational content via a mobile app and (2) to remotely monitor blood pressure and weight. The educational content was based upon ACOG standards and refined by a committee of 4 board-certified obstetricians at the GWU School of Medicine. Input from a variety of other stakeholders including patients, midwives, and administrators was also obtained. The mobile app sends educational content to the expectant mother at gestation-appropriate times throughout the pregnancy. This information encompasses material covering pregnancy progression; preexisting risk hazards such as alcohol intake, smoking, or drug abuse; advice to address these risk hazards; dietary and nutritional content; breastfeeding information; guidelines for appropriate weight gain; and warning signs for pregnancy complications. In addition, the app integrates with a Wi-Fi-connected scale and blood pressure cuff to provide both feedback and alerts depending on the readings. The alerts were created to provide early warnings to patients and providers about aberrant data points with the hope of providing early detection of hypertensive disorders of pregnancy and abnormal weight gain, indicating an increased risk of gestational diabetes, nutritional deficiency, or edema associated with preeclampsia.

Results

User Statistics

A total of 181 women met the inclusion criteria and were screened for enrollment. Of those, 118 met screening criteria and agreed to participate. Furthermore, 60 women were allocated to the experimental group and 58 women were allocated to the control group (Figure 1). Of those, 13 participants in the experimental group and 17 in the control group discontinued involvement in the study because they transferred care (n=11), developed high-risk characteristics (n=7), experienced a miscarriage (n=5), requested to no longer participate (n=2), or other reasons (n=5). Ultimately, 47 participants in the experimental group and 41 in the control group were retained in the study until completion and were analyzed. There was no significant difference in baseline characteristics (Table 2).

Figure 1. CONSORT flow diagram.

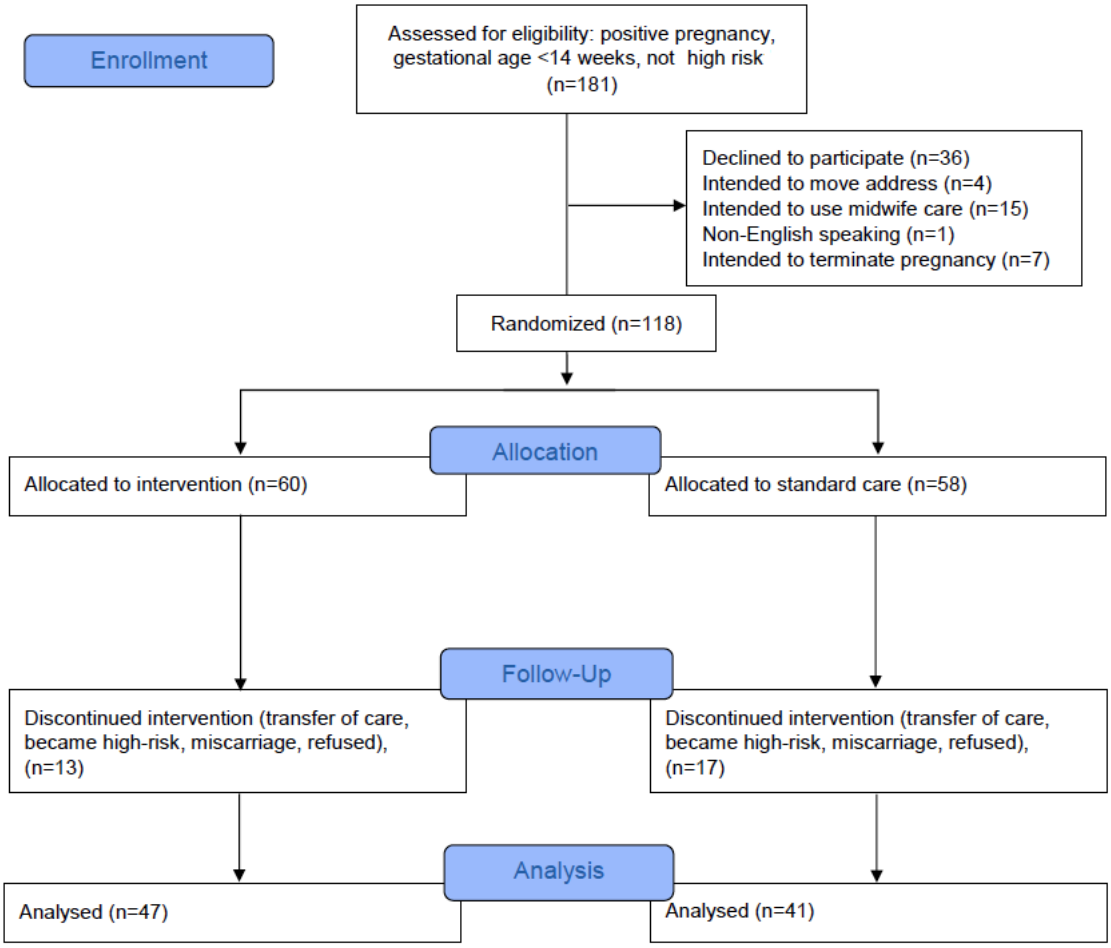


Table 2. Patient characteristics and number of visits per group.

Characteristics	Babyscripts (n=47)	Standard care (n=41)
Age at screening (years), mean (SD)	33.0 (3.3)	32.2 (3.2)
Body mass index at screening[a], mean, kg/m2, (SD)	22.9 (3.2)	24.9 (4.0)
African-American, n (%)	14 (30)	13 (32)
Hispanic, n (%)	3 (6)	3 (7)
College graduate, n (%)	45 (96)	36 (88)
Gravidity per patient, n (%)		
One	25 (53)	22 (53)
Greater than one	22 (47)	19 (46)
Parity per patient, n (%)		
Zero	30 (64)	25 (61)
Greater than one	17 (36)	16 (39)
Number of in-person prenatal care visits[a], mean (SD)	7.9 (1.8)	10.2 (1.8)

[a]*P*<.001.

Evaluation Outcomes

The experimental group had significantly fewer prenatal care visits (7.9 visits) than patients in the control group (10.2 visits; *P*=.01 Table 2; Figure 2). Patient satisfaction measured over several intervals demonstrated no difference in satisfaction between the experimental and control group. Satisfaction scores were aggregated for all 16 questions that were asked to participants in both the control and experimental groups at each time point during gestation. For data visualization, the scores

were normalized from 0 to 1 and then compared for statistical differences at each time point. Provider surveys demonstrated aggregate scores demonstrating highly perceived quality and satisfaction with the virtual care platform. There was no statistical difference in patient satisfaction (P>.05) or provider satisfaction (P>.05) in either group (Figure 3).

Although maternal and fetal outcomes were tracked, the study was not powered to demonstrate the effect of the virtual care platform on maternal or fetal outcomes. We identified 1 adverse fetal event of stillbirth at 38 weeks gestation. Investigation of the case revealed that the patient had an uncomplicated pregnancy and was compliant with the prenatal care schedule of the experimental group. Her obstetrician saw her in person within 3 days of the fetal demise at which time normal fetal movement and fetal heart tones were identified. Workup revealed a likely fetal-maternal hemorrhage unrelated to the study protocol.

Figure 2. Total number of prenatal clinical visits per group. Babyscripts (Brx) versus controls. *P* value is based on Wilcoxon ranked-sum test.

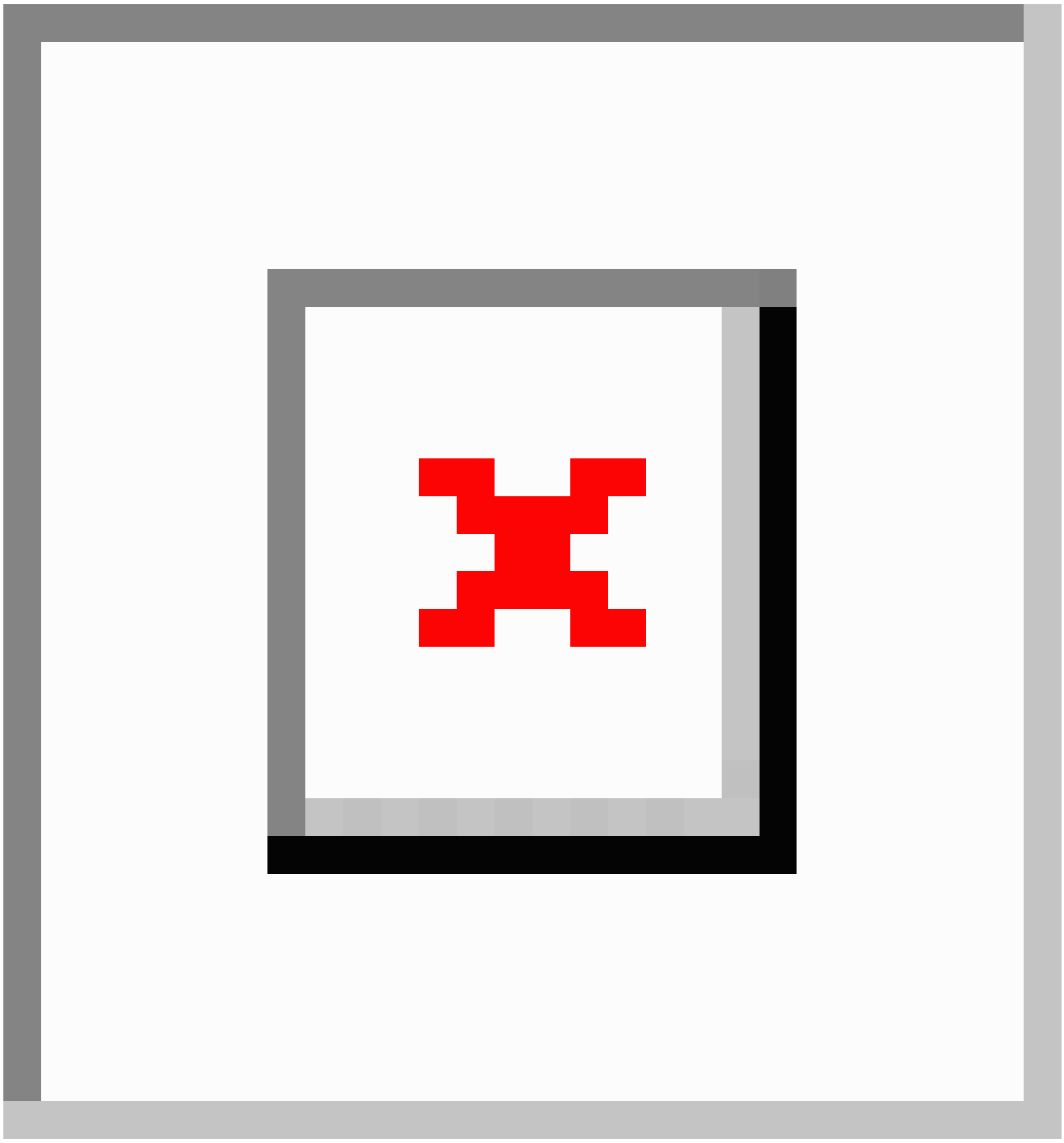
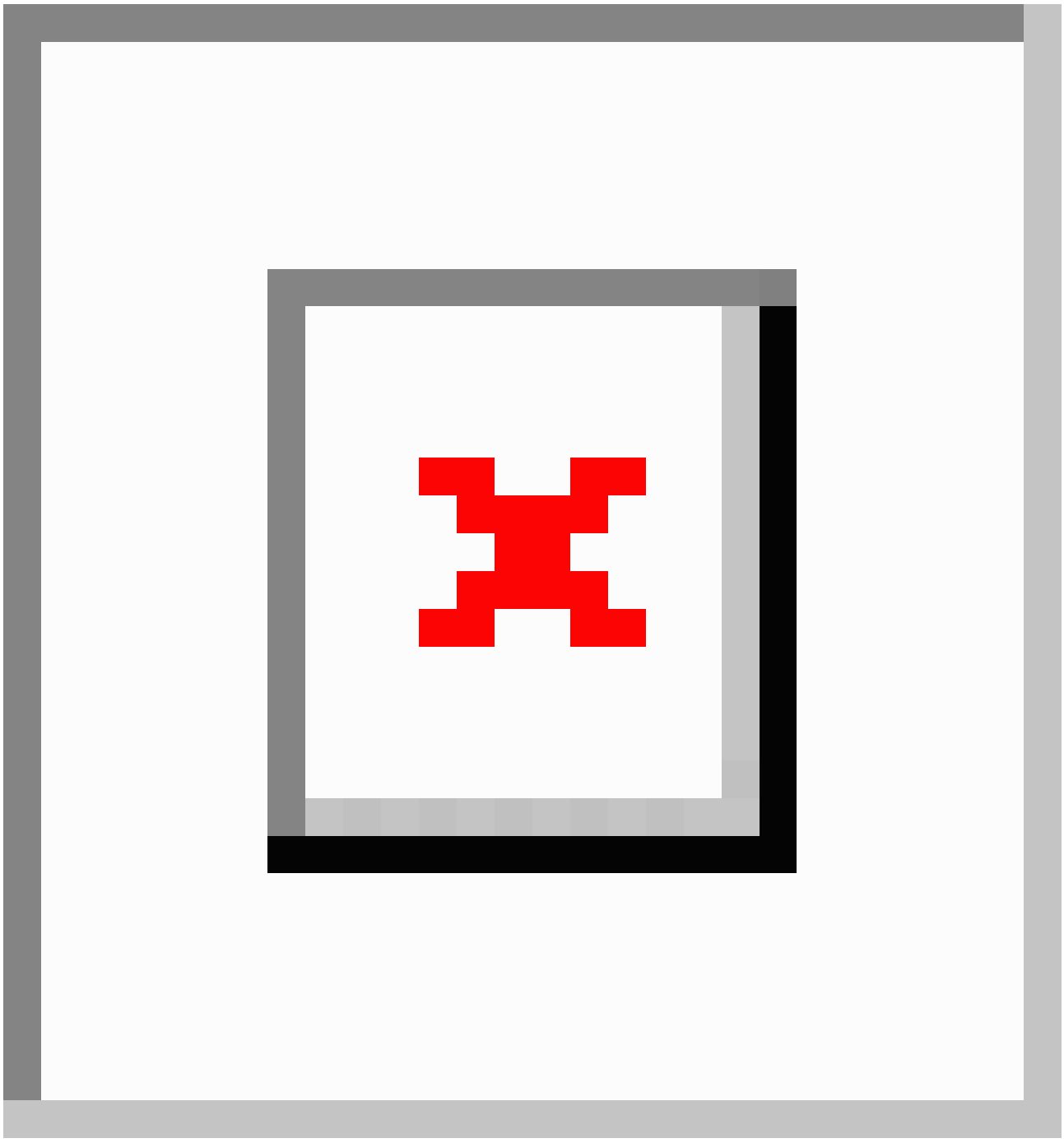


Figure 3. Survey data for patient satisfaction with overall prenatal care between Babyscripts (Brx) and control groups. For each time point, Brx and control patients are given a survey. The total numeric scores for all questions at that time point is calculated. All numeric scores are then normalized to 1 to standardize between different numbers of survey questions. SOC: standard of care.



Discussion

Principal Findings

Pregnant women represent a promising target for digital health apps. Unlike digital health apps that target chronically ill or elderly populations, pregnant women are a young and healthy population. A mobile health app that targets pregnant women may facilitate the integration of prenatal care into other aspects of their family and professional life. In addition, pregnancy is a unique period of life when healthy behaviors including exercise, diet, and sleep take on greater importance. As such, women are highly engaged with their health care decisions during pregnancy and may be more receptive to educational programs that can be delivered through a mobile health app. Finally, the majority of prenatal care visits are scheduled to exchange educational information with the patient or weight and blood pressure data with the provider. Both of these exchanges are especially amenable to communication via mobile technology or remote monitors. Ultimately, the app does not replace in-person visits but may replace some of the current activities that occur at each visit. If the app can communicate basic educational components of prenatal care, the in-person visits may allow for more individualized discussions.

As part of this study, several important elements emerged as critical to the success of a mobile prenatal care app. First, the initial assessment is critical to identifying high-risk versus normal-risk on initial assessment; second, accurate communication of patient data to the provider is necessary to assess early signs of pregnancy complications; third, educational information must be provided at the appropriate time during pregnancy; fourth, educational information should be targeted to individual patient (eg, not all women need regular reminders about the importance of smoking cessation); and fifth, a clear explanation that the role of the mobile app is to augment and not replace the obstetrician or midwife.

Limitations

There are several limitations to our study. First, the use of a quasi-randomization scheme where participants were allocated by type of phone creates potential confounders, as it is possible that iOS users differ from non-iOS users. However, there was no significant difference in age, race, education level, gravity, or parity between the cohorts. We did not collect data on differences in household income, a possible confounder. The second risk concerns the possibility of contamination across groups. It is possible that by reducing visits in the experimental group, physicians were more likely to reduce the visits in all patients. However, given there was a statistically significant difference in visit schedule, this was unlikely to confound this cohort. Third, the mobile prenatal care app was prescribed as part of the reduced care schedule, and it is unknown if a reduced care schedule might have been effective without the mobile care app or with a different solution. Finally, we have limited information regarding emergency visits or phone calls that may have occurred outside of the chart review or the direct patient calls. It is possible that there were some therapeutic interactions missed in both the control and experimental group.

Conclusions

In conclusion, satisfaction was unchanged and visits were reduced through the use of the prenatal mobile care app, Babyscripts. Future studies will look for predictors of adverse clinical outcomes in a variety of populations in hopes of mitigating risk of adverse events.

Acknowledgments

There was no external funding for this study. Funding was provided by the 1EQ Inc and George Washington-Medical Faculty Associates.

Authors' Contributions

KIM, JMK, NG, NDG, JAB, JO, LMR, and ACM made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data. KIM, JMK, NG, NDG, JAB, JO, LMR, and ACM were involved in drafting the manuscript or revising it critically for important intellectual content. KIM, JMK, NG, NDG, JAB, JO, LMR, and ACM provided final approval of the version to be published. KIM, JMK, NG, NDG, JAB, JO, LMR, and ACM agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Conflicts of Interest

JO and ACM declare a competing interest and hold a position in Babyscripts. KIM and NDG are unpaid consultants. JAB holds stock in the company 1EQ, and was employed by 1EQ at the time of the study. The remaining authors declare no conflicts of interest.

Multimedia Appendix 1

Excluded high-risk obstetrical conditions (n=84).

[PDF File (Adobe PDF File), 30KB - mhealth_v7i5e10520_app1.pdf]

Multimedia Appendix 2

Patient Satisfaction Survey, weeks 14, 20, 25, 30, 35, and postpartum.

[PDF File (Adobe PDF File), 150KB - mhealth_v7i5e10520_app2.pdf]

References

1. Martin JA, Hamilton BE, Osterman MJ, Driscoll AK, Mathews TJ. Births: final data for 2015. Natl Vital Stat Rep 2017 Jan;66(1):1 [FREE Full text] [Medline: 28135188]
2. Alexander GR, Kotelchuck M. Assessing the role and effectiveness of prenatal care: history, challenges, and directions for future research. Public Health Rep 2001;116(4):306-316 [FREE Full text] [doi: 10.1093/phr/116.4.306] [Medline: 12037259]
3. Villar J, Carroli G, Khan-Neelofur D, Piaggio G, Gülmezoglu M. Patterns of routine antenatal care for low-risk pregnancy. Cochrane Database Syst Rev 2001(4):CD000934. [doi: 10.1002/14651858.CD000934] [Medline: 11687086]

4. Dowswell T, Carroli G, Duley L, Gates S, Gülmezoglu AM, Khan-Neelofur D, et al. Alternative versus standard packages of antenatal care for low-risk pregnancy. Cochrane Database Syst Rev 2010 Oct 6(10):CD000934 [FREE Full text] [doi: 10.1002/14651858.CD000934.pub2] [Medline: 20927721]
5. Walker DS, Day S, Diroff C, Lirette H, McCully L, Mooney-Hescott C, et al. Reduced frequency prenatal visits in midwifery practice: attitudes and use. J Midwifery Womens Health 2002;47(4):269-277. [doi: 10.1016/S1526-9523(02)00259-3] [Medline: 12138935]
6. Krans EE, Davis MM. Preventing low birthweight: 25 years, prenatal risk, and the failure to reinvent prenatal care. Am J Obstet Gynecol 2012 May;206(5):398-403. [doi: 10.1016/j.ajog.2011.06.082] [Medline: 21889122]
7. VanderWeele TJ, Lauderdale DS, Lantos JD. Medically induced preterm birth and the associations between prenatal care and infant mortality. Ann Epidemiol 2013 Jul;23(7):435-440 [FREE Full text] [doi: 10.1016/j.annepidem.2013.04.010] [Medline: 23726822]
8. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
9. Estepp JH, Winter B, Johnson M, Smeltzer MP, Howard SC, Hankins J. Improved hydroxyurea effect with the use of text messaging in children with sickle cell anemia. Pediatr Blood Cancer 2014 Nov;61(11):2031-2036. [doi: 10.1002/pbc.25177] [Medline: 25132074]
10. Perera AI, Thomas MG, Moore JO, Faasse K, Petrie KJ. Effect of a smartphone application incorporating personalized health-related imagery on adherence to antiretroviral therapy: a randomized clinical trial. AIDS Patient Care STDS 2014 Nov;28(11):579-586 [FREE Full text] [doi: 10.1089/apc.2014.0156] [Medline: 25290556]
11. Kim YJ, Rhee SY, Byun JK, Park SY, Hong SM, Chin SO, et al. A smartphone application significantly improved diabetes self-care activities with high user satisfaction. Diabetes Metab J 2015 Jun;39(3):207-217 [FREE Full text] [doi: 10.4093/dmj.2015.39.3.207] [Medline: 26124991]
12. Marko KI, Krapf JM, Meltzer AC, Oh J, Ganju N, Martinez AG, et al. Testing the feasibility of remote patient monitoring in prenatal care using a mobile app and connected devices: a prospective observational trial. JMIR Res Protoc 2016 Nov 18;5(4):e200 [FREE Full text] [doi: 10.2196/resprot.6167] [Medline: 27864167]
13. Babyscripts. URL: http://www.getbabyscripts.com [accessed 2016-07-27] [WebCite Cache ID 6jJbesFzc]
14. US Department of Veterans Affairs. 2009. VA/DoD Clinical Practice Guideline for Pregnancy Management URL: https://www.healthquality.va.gov/guidelines/WH/up/mpg_v2_1_full.pdf [accessed 2019-02-27] [WebCite Cache ID 76Vu4Ti55]
15. Hospital Consumer Assessment of Healthcare Providers and Systems. Baltimore, MD: Centers for Medicare & Medicaid Services URL: https://www.hcahpsonline.org/ [accessed 2019-03-06] [WebCite Cache ID 76gYyLf5J]

Abbreviations

ACOG: American College of Obstetricians and Gynecologists
GWU: George Washington University
GYN: gynecology
IRB: Institutional Review Board
OB: obstetric

Edited by G Eysenbach; submitted 28.03.18; peer-reviewed by P van Montfort, R Pryss; comments to author 18.08.18; revised version received 07.01.19; accepted 26.01.19; published 01.05.19.

Please cite as:
Marko KI, Ganju N, Krapf JM, Gaba ND, Brown JA, Benham JJ, Oh J, Richards LM, Meltzer AC
A Mobile Prenatal Care App to Reduce In-Person Visits: Prospective Controlled Trial
JMIR Mhealth Uhealth 2019;7(5):e10520
URL: https://mhealth.jmir.org/2019/5/e10520/
doi:10.2196/10520
PMID:31042154

Original Paper

Change in Waist Circumference With Continuous Use of a Smart Belt: An Observational Study

Myeonggyun Lee[1], MS; Jaeyong Shin[2,3,4], MD, MPH, PhD

[1]Division of Biostatistics, Department of Population Health and Environmental Medicine, New York University School of Medicine, New York City, NY, United States

[2]Department of Preventive Medicine, Ajou University School of Medicine, Suwon, Republic of Korea

[3]Institute of Health Services, Graduate School, Yonsei University, Seoul, Republic of Korea

[4]Department of Policy Analysis and Management, College of Human Ecology, Cornell University, Ithaca, NY, United States

Corresponding Author:
Jaeyong Shin, MD, MPH, PhD
Department of Preventive Medicine
Ajou University School of Medicine
164 Worldcup-ro, Yeongtong-Gu
Suwon, 16499
Republic of Korea
Phone: 82 312197455
Fax: 82 312195084
Email: drshin@ajou.ac.kr

Abstract

Background: Health insurers and policymakers are trying to prevent and reduce cardiovascular diseases due to obesity. A smart belt that monitors activity and waist circumference is a new concept for conquering obesity and may be a promising new strategy for health insurers and policymakers.

Objective: This preliminary study evaluated whether the use of a smart belt was associated with a decrease in waist circumference.

Methods: In the manufacturer's database, there were data on a total of 427 men at baseline. A total of 223, 81, and 27 users kept using the smart belt for 4, 8, and 12 weeks, respectively. Paired *t* tests and repeated measures analysis of variance (ANOVA) were used to identify the change in waist circumference at specified time intervals (at 4, 8, and 12 weeks). In addition, a linear mixed model was used to incorporate all users' waist circumference data at each time point. Preexisting data on waist circumference and self-reported demographics were obtained from the manufacturer of the smart belt (WELT Corporation, South Korea).

Results: Compared with baseline, the waist circumference (cm) decreased significantly at all time points: –0.270 for week 4, –0.761 for week 8, and –1.972 for week 12 (all $P<.01$). Although each paired *t* test had a different sample size because of loss to follow-up, the differences between baseline and each subsequent week increased. Equal continuous reduction in waist circumference was observed with the ANOVA and mixed model analysis (beta=–0.158 every week).

Conclusions: The smart belt is a newly developed, wearable device that measures real-time steps, sedentary time, and waist circumference. In this study, we showed that wearing the smart belt was associated with reducing waist circumference over 12 weeks. This direct-to-consumer smart health device may contribute toward reducing the risk of obesity and related conditions and controlling increasing health costs for health insurers.

(JMIR Mhealth Uhealth 2019;7(5):e10737) doi:10.2196/10737

KEYWORDS

smart health care; wearable device; obesity; internet of things; mHealth; digital health care; lifestyle modification; metabolic syndrome

Introduction

Noncommunicable diseases (NCDs) are one of the leading causes of death worldwide [1-3]. According to the World Health Organization, of the 56.4 million global deaths in 2015, 39.5 million (70%) were due to NCDs [4]. Currently, the burden of NCDs is growing faster than our ability to combat them because of the obesity epidemic [5-7]. In the United States, the

percentage of the national medical expenditures devoted to treating obesity-related illness in adults increased from 6.13% in 2001 to 7.91% in 2015 [8].

Due to a change to a Western lifestyle, including a high-calorie diet, and the aging society in Korea, the epidemiology of NCDs in Korea has dramatically risen. According to national statistics, the age-adjusted prevalence rate of metabolic syndrome in Korea has increased from 21.6% in 2007 to 26.5% in 2015 among men, whereas there has been slight decrease among women from 20.3% to 17.6% during the same period [9]. Based on forecasting models including these variables, the 2020 and 2030 estimates for obesity prevalence in Korea are 47% and 62% for men and 32% and 37% for women, respectively [10]. Therefore, it is necessary to control unhealthy lifestyles in Korea.

Several insurers and health policymakers have attempted to prevent NCDs and their related lifestyles [11-13]. However, changing lifestyles to prevent NCDs is not easy. Thus, new concepts to conquer NCDs have been introduced and are being tested. For example, the Centers for Disease Control and Prevention (CDC) approved some digital health programs as part of the National Diabetes Prevention Program. Diabetes can be controlled and prevented by lifestyle modification. Therefore, a key part of the National Diabetes Prevention Program is the lifestyle change program, which aims to prevent or delay type 2 diabetes [14,15]. The CDC recognizes lifestyle change programs that meet certain standards and show that they can achieve results. These standards include following an approved curriculum, being facilitated by a trained lifestyle coach, and submitting data each year to show that the program has had an impact. However, only a few wearable devices included in the programs satisfy these standards. In addition, most of them measure daily activities and self-reported diet and body mass index (BMI) [16].

The WELT smart belt was launched in 2016 by the WELT Corporation, which is a health care technology company. The smart belt looks like a normal belt but can monitor a wide range of health data using a mobile phone app. It can measure an individual's waist circumference, overeating habits, number of steps, and sedentary time with the tracking technology stored in the belt buckle (see the image in Figure 1) to help improve the individual's health and the effectiveness and the efficiency of the health care system [17]. From the collected information, the device provides users with their daily activity score categorized into three groups: best, good, and poor. All real-time information is presented on a mobile phone app, and comparisons between daily activities and the previous days' activities are possible. However, the device includes no feedback or notification that encourages exercise or warns of overeating.

This preliminary study aims to evaluate whether the waist circumference of smart belt users decreased with use. Moreover, this device has a distinct measurement approach, which other smart devices do not have. The waist circumferences can be measured every 30 minutes, 24 hours per day, and these measurements are saved into a database as a longitudinal concept. Thus, to the best of our knowledge about the smart belt, and because none of the previous studies have shown this type of user waist data from the smart wearable device, evaluating its effectiveness is necessary.

Figure 1. The WELT smart belt and mobile phone app.

Methods

Study Population and Variables

This study included smart belt users who downloaded the app on their mobile phones and provided information on their weight, height, and age. When they downloaded the app, the WELT cooptation asked them to monitor their health variables without identification. The app measures the waist circumference of the user wearing the smart belt at 30-minute intervals and provides information on waist circumference, steps, and sedentary time. Preexisting data were obtained from the WELT Corporation from July 2016. A total of 451 male users were registered in the database initially, and 24 who had missing values or were outliers for height, weight, and initial waist circumference were excluded (Figure 2). This study was approved by the institutional review board at Yonsei University in Korea (approval number: Y2018-0264-001).

In the database, 504 subjects were registered initially, and we excluded 77 who were female (n=53), missing data (n=10), or outliers in height, weight, and initial waist circumference (n=14). Finally, we obtained a total number of 427 users' data.

The daily and weekly waist circumferences (cm) of the users were calculated using an average value of the waist circumference data measured every 30 minutes (Figure 3). If a person did not wear the smart belt, the circumference was recorded as a missing value in our dataset. Therefore, we knew when the users started to wear the smart belt and when they took it off. The waist circumference measurements were obtained when it was worn. We calculated the average waist circumferences without missing values. The average waist circumference after the first week of wearing the smart belt and using the app was considered the baseline measurement. This study reports the waist circumference of users during the first 12 weeks of using the smart belt.

Figure 2. Flowchart for study participants. Preexisting data were obtained from the WELT Corporation. BMI: body mass index.

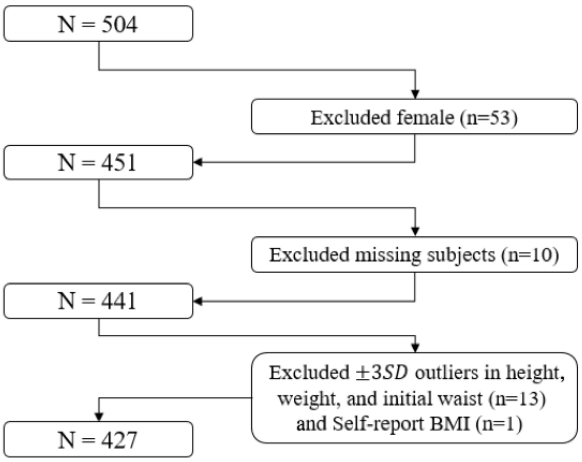
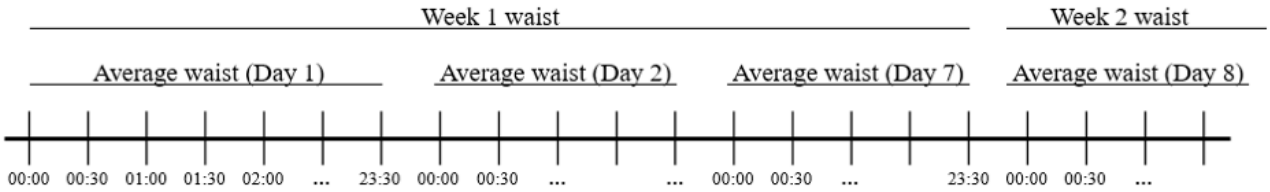


Figure 3. Calculations of daily and weekly waist circumferences. The first week was the baseline measure.



Statistical Analysis

From the 427 user datasets, descriptive statistics were calculated for the users' baseline characteristics, and follow-up waist circumference measurements were reported as means with standard deviations. There were losses to follow-up on wearing time; thus, the sample size decreased at 12 weeks compared with baseline.

A paired *t* test is generally performed to compare different conditions among the same subjects; this approach has been used elsewhere in public health research [18]. In this study, a paired *t* test was performed to identify the change in waist circumference between baseline and each follow-up week for the same subjects.

As a multivariate approach, a repeated measures analysis of variance (ANOVA), used to distinguish the influence of time on waist circumference change [19], was used to identify the change in waist circumference by considering time intervals. We used the following time intervals: every week, biweekly, and at 4-week intervals. Although this method extracts the contribution of subjects from the error term, some subjects were excluded from the entire analysis if the values for even one time point were missing.

A linear mixed model analysis, which incorporated all subjects' waist circumference data at each time point, was conducted. A linear mixed model was used to identify the trend of the findings of previous public health studies by Xu [20], De Onis et al [21], and Brodersen et al [22]. The following models were considered:

using only week time (model 1), using week time and BMI (model 2), and using adjusted time and baseline variables (model 3). Model 3 consisted of covariates from model 2 as well as age and wearing time (days) as an adjustment and attachment, respectively.

Based on the duration of continuous smart belt use, the users were divided into three groups: initial only (used the belt in the first week), passive users (used the belt between 2 and 5 weeks), and active users (used the belt for more than 5 weeks). ANOVA was used to compare the demographic characteristics and initial waist circumference of the three groups. Significance tests were two-sided, with the significance level set at .05. Multiple comparison problems were not considered because this was an observational study and our purpose was to identify the trend in waist circumference change in subjects who wore the smart belt. All statistical analyses were performed using SAS 9.4 software (SAS Institute Inc, Cary, NC, USA).

Results

Table 1 shows the descriptive characteristics of the smart belt users. To calculate statistics, we divided the followed-up users as (1) all cases, (2) completed users up to the first 4 weeks, (3) completed users up to the first 8 weeks, and (4) completed users up to 12 weeks. For example, if one user dropped out at week 6, data from that user was used for "up to 4 weeks" but not "up to 8 weeks" or later. From the demographics, the population was representative of Korean, middle-aged, male, obese office workers based on age, weight, height, and BMI.

Average waist circumference decreased a mean 0.158 (SD 0.043) cm every week from baseline to the last followed-up week (Figure 4). To compare each week with the baseline, a paired *t* test was conducted. Since users who abnormally increased or decreased their waist size could be outliers, we excluded those outliers with a cutoff of 3 standard deviations from the mean for each comparison. Mean waist circumference decreased at all time points (all weeks) compared to the baseline measurement, with *P* values for the difference between baseline and each time point of $P<.05$ except for weeks 2 and 5 (Figure 5). Although each paired *t* test had a different sample size because of the loss to follow-up, the magnitude of difference between the baseline value and the value at each subsequent week until the final follow-up week increased (Figure 5). With the decreasing trend of waist size in Figure 5, the increment of waist size at week 7 in Figure 4 could be caused by outliers with cutoffs 3 standard deviations from the mean.

Table 1. Descriptive statistics of smart belt users. Except for waist circumference, other characteristics, including age, weight, height, and body mass index (BMI), were self-reported at baseline when the user first started using the app.

Characteristic	All (N=427), mean (SD)	Used up to 4 weeks, (n=223), mean (SD)	Used up to 8 weeks, (n=81), mean (SD)	Used up to 12 weeks, (n=27), mean (SD)
Age (years)	42.29 (11.83)	44.45 (12.35)	45.17 (12.62)	43.11 (12.97)
Weight (kg)	79.80 (12.86)	79.65 (12.72)	78.65 (13.06)	75.07 (12.19)
Height (m)	1.75 (0.07)	1.74 (0.07)	1.75 (0.08)	1.72 (0.07)
BMI (kg/m^2)	25.89 (3.47)	25.96 (3.38)	25.54 (3.48)	25.25 (3.40)
Waist (cm)				
Week 1 (n=427)	89.64 (9.04)	89.38 (8.79)	89.41 (8.71)	89.79 (7.29)
Week 2 (n=303)	89.46 (9.19)	89.48 (8.94)	89.46 (8.66)	89.66 (7.95)
Week 3 (n=276)	89.48 (9.04)	89.28 (8.99)	89.31 (8.64)	89.43 (7.87)
Week 4 (n=239)	89.43 (9.04)	89.23 (9.04)	89.18 (8.76)	89.48 (7.77)
Week 5 (n=195)	89.20 (9.07)		88.85 (9.09)	89.23 (7.82)
Week 6 (n=165)	89.51 (8.71)		89.20 (8.71)	89.03 (7.95)
Week 7 (n=146)	89.94 (8.33)		88.65 (8.92)	87.86 (8.76)
Week 8 (n=134)	88.98 (9.22)		88.37 (9.19)	87.55 (9.14)
Week 9 (n=103)	87.88 (8.97)			87.60 (8.97)
Week 10 (n=94)	88.72 (8.81)			87.48 (9.09)
Week 11 (n=74)	87.53 (9.83)			87.10 (9.63)
Week 12 (n=55)	88.32 (9.27)			86.89 (9.68)

Figure 4. Trend of waist circumference by week. The blue solid line is the mean value of the waist size for each week; the orange line is a linear trend of the waist size by week.

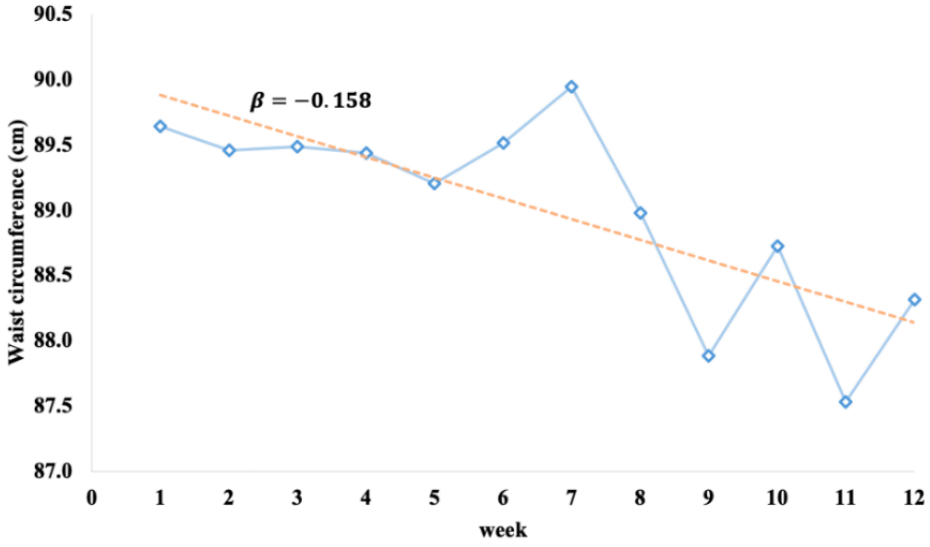
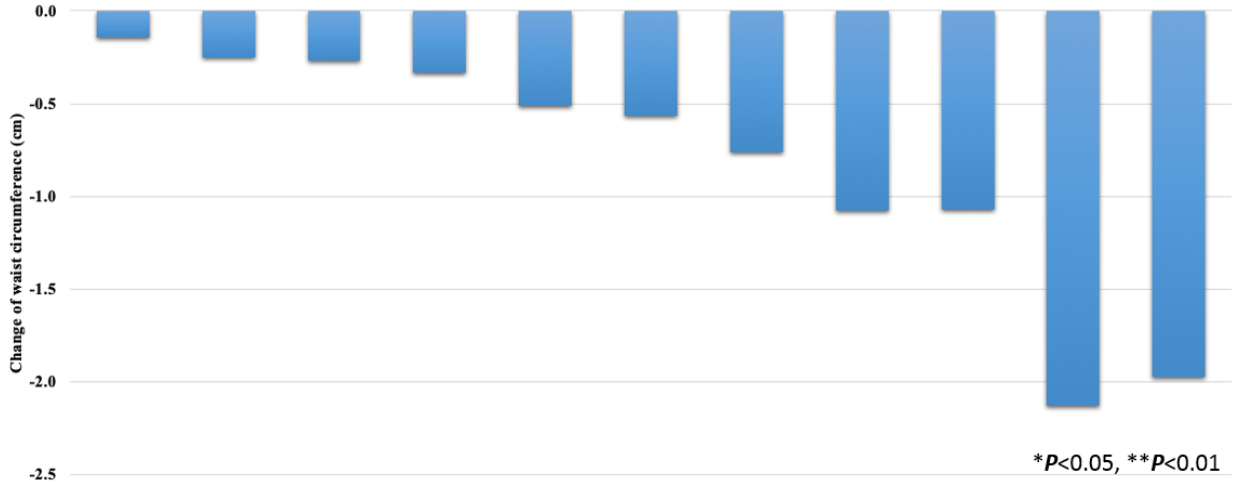


Figure 5. Reduction in waist circumferences of users each week over 12 weeks compared to baseline using paired *t* test.



Week	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12
N	299	270	231	189	159	139	129	99	91	72	54
Difference	-0.145	-0.247**	-0.270*	-0.334	-0.510**	-0.565**	-0.761*	-1.075*	-1.068*	-2.124**	-1.972*
95% CI	(-0.29,0.00)	(-0.42,-0.07)	(-0.53,-0.01)	(-0.67,0.00)	(-0.86,-0.16)	(-0.98,-0.15)	(-1.34,-0.18)	(-2.01,-0.14)	(-2.02,-0.11)	(-3.65,-0.60)	(-3.64,-0.30)

Change in waist circumference was analyzed using repeated measures ANOVA. We considered three models at different time intervals (Table 2). Model 1 shows that the differences in waist circumference at weeks 2, 3, and 4 were mean –0.157 (SD 1.260), mean –0.254 (SD 1.364), and mean –0.325 (SD 1.816) cm, respectively. The differences decreased by week. In the case of model 3 as a four-week interval, decreasing waist circumference was more distinct; there was an average decrease of 1.435 (SD 4.442) cm between baseline and week 12. Since the objective was to compare each week as a group, we included all users in which they wore the smart belt each week and the data were recorded, even though they may not have followed up at all time points. Although Table 1 shows the completed users (eg, the 27 users for “used up to 12 weeks” means that they completely followed up), we used users who were followed up at weeks 1, 4, 8, and 12 in the repeated measures ANOVA. Thus, the number of users could be different. Although each model had a different sample size because of loss to follow-up, all three models showed a *P* value less than .05 under the null hypothesis that the differences of the waist circumference would not differ for each week compared to baseline. Thus, all three models had a statistically significant change in waist circumference, with the mean waist circumference decreasing by week.

Repeated measures ANOVA could be used to determine the change in waist circumference; however, in this study, only the data of users with complete information at the measuring points (ie, 4, 8, and 12 weeks) were included. Thus, an additional analysis that incorporated all the observed waist circumferences was necessary to identify overall waist circumference reduction. Table 3 shows the results of the linear mixed model using all datasets. The results show a decrease in waist circumference regardless of adjustments.

Table 2. Differences in waist circumference using repeated ANOVA for three models.[a]

Model	Difference from baseline, mean(SD)	*P* value	Cohen *f*
Model 1: weeks 1, 2, 3, 4 (n=215)		.002	0.1532
Week 2	–0.157 (1.260)		
Week 3	–0.254 (1.364)		
Week 4	–0.325 (1.816)		
Model 2: weeks 1, 2, 4, 8 (n=111)		.003	0.2065
Week 2	–0.157 (1.260)		
Week 4	–0.254 (1.364)		
Week 8	–0.325 (1.816)		
Model 3: weeks 1, 4, 8, 12 (n=41)		.03	0.2744
Week 4	–0.173 (1.402)		
Week 8	–0.762 (3.584)		
Week 12	–1.435 (4.442)		

[a] Note that each model has a different sample size due to loss at follow-up.

Table 3. Adjusted models for waist circumference using the linear mixed model.[a]

Variable	Model 1			Model 2			Model 3		
	Estimate	Standard error	*P* value	Estimate	Standard error	*P* value	Estimate	Standard error	*P* value
Intercept	89.939	0.442	<.001	64.302	3.018	<.001	63.455	3.381	<.001
Week	–0.142	0.026	<.001	–0.141	0.026	<.001	–0.143	0.026	<.001
BMI[b]				0.990	0.116	<.001	0.999	0.116	<.001
Age							0.006	0.034	.87
Wearing period (days)							0.008	0.010	.39

[a]Change in waist circumference was analyzed through repeated ANOVA using three models with different time intervals. Note that each model had a different sample size because they had a loss to follow-up.

[b]BMI: body mass index.

We also investigated users who used the smart belt for as long as possible. Here, ANOVA was used to compare the following three groups to investigate the difference in characteristics: initial only, passive users, and active users (Multimedia Appendix 1). There were no statistically significant differences in BMI and waist circumference at baseline, but the group of active users was older than the other groups (age variable was the only significant variable).

Discussion

In our study, waist circumference tended to decrease among smart belt users. In the paired *t* test analysis, the users who wore it longer had greater reductions in waist circumference. The same trend was also found using the linear mixed models. A reduction in waist circumference was observed after each week of device use.

The recent Patient Protection and Affordable Care Act has incentivized physical activity counseling by primary care physicians, which may increasingly affect not only physicians but also health insurers and accountable care organizations. The insurers want to monitor the health behaviors of the insured and to predict the risky health outcomes. Therefore, it is necessary to evaluate the efficiency of newly developed wearable smart health devices and determine how these devices can improve health outcomes.

Various insurers and health-related companies are tapping into wearable technology for their customers to help them stay in shape and to offer their customers wellness discounts and other benefits if they meet a required number of fitness goals [17,23-25]. For example, health insurance start-up provider Oscar provides their customers with a Misfit fitness band to track their physical movements and day-to-day activities—it even monitors the users' sleep. Based on these data, customers are rewarded once they meet their fitness goals [24]. UnitedHealthcare and Fitbit reward their users with up to US $1500 for activities completed using their device [23]. The companies can track their enrollees' physical activity and monitor their diet; therefore, they can provide appropriate feedback through personalized wellness programs. Optimal

Health's program differs slightly regarding the incentives because the total reward is paid out incrementally as the participants complete each of the four steps in the challenge, which includes downloading the app and completing a personal health assessment questionnaire, tracking nutrition, wearing the devices, and doing physical activity [23]. Another major health insurance provider, Humana, also provides a 10% annual premium discount to their members if they use fitness tracking devices [23].

Cadmus-Bertram et al [26] performed randomized trials of a Fitbit-based physical activity intervention with 51 inactive postmenopausal women with a BMI greater than 25.0 kg/m^2 for 16 weeks. Fitbit devices and app-based weight control services were provided to 21 participants in the treatment group (mobile phone app plus Fitbit), whereas the control group used simple pedometers (Fitbit). The participants were encouraged to perform physical activity. The treatment group showed an improvement in moderate-to-vigorous physical activity at 16 weeks compared to those at baseline. However, no statistically significant difference in physical activity between the treatment and control groups was observed.

Jakicic et al [27] compared a standard behavioral weight loss intervention (n=233) and a technology-enhanced weight loss intervention (n=237). The technology-enhanced weight loss intervention group was equipped with commercially available wearable devices, which included Web interface technology. No significant difference between the two groups at 24 months was observed. In these well-designed randomized controlled trials, no reduction in body weight between the group using smart devices and the control group was observed.

According to a systematic review of 22 studies consisting of 16,476 and 14,475 participants in the intervention and control groups, respectively [28], Web-based physical activity interventions had a significantly positive effect on increasing physical activity and daily walking steps among the general population at the initial stage. However, the effect appeared to depend on the design of the study, age of the participants, and duration of the study.

Although controversies regarding the effectiveness and efficiency of smart health devices exist, we believe that our results provide scientific evidence for the effectiveness of the smart belt. Most of the users in our study were obese (BMI >25 kg/m^2), middle-aged Korean males. Korea has the second-longest working hours according to the Organization for Economic Co-operation and Development (OECD); on average, Koreans work 393 more hours per year than their counterparts in the OECD [29]. Their length of working hours is approximately 1.6 times that of Dutch workers. Thus, exercising regularly is not easy for middle-aged Korean workers. As they may need to check their daily activities and increase them, they would most likely purchase a smart belt for themselves or receive it as a gift from someone, which in turn provides motivation for the smart belt users to lose weight. We could assume that this motivation may be greater in more obese users. Moreover, on the basis of our baseline results, active users who continued to use the smart belt for more than 5 weeks were older than other users. Older people might have more medical conditions and therefore be more interested in their health status, which may explain the high retention in this age group.

However, we still had a fundamental question about the reduction in waist circumference without specific intervention in this study. We speculated that wearing a smart belt device would have a positive effect on one's health behaviors. Choi et al [30] investigated a 12-week mobile health (mHealth) physical activity intervention for feasibility and potential efficacy. Thirty pregnant women were randomized to either the intervention (mobile phone app plus digital intervention program) or control (digital intervention program alone) group. The difference in weekly mean steps per day was not statistically significant between the groups, which could be because the active control condition (included the use of a wearable activity monitor and final goal steps that were similar to those in the intervention group) may itself have had substantial effects. Even though there is no study on the smart belt, we assume that wearing smart devices including the smart belt possibly motivates users to reduce their waist circumference.

The number of steps taken is a good indicator of an individuals' sedentary level; this information will provide more details on the characteristics of people who use a smart belt. Although we did not include the results for step count in this study, the mean steps per hour while wearing the device increased over time (see Multimedia Appendix 2). For example, the differences in the mean steps per hour from baseline to weeks 2, 4, and 8 were 220.6, 541.7, and 862.9, respectively. Therefore, it seems that participants in this study had increased activity levels during the 12-week period of observation.

This study has several limitations. First, there was no control group in this study. Therefore, the study participants were motivated individuals who were interested in trying a novel product to help with a lifestyle change such as weight loss. Hence, future studies need to include participants who do not wear the smart belt as a comparison control group. Second, some important variables may have been omitted (for example, regular exercise hours or health conditions that may affect daily activity). In this study, we used only the data collected from the wearable device and self-reported data; that is, this cohort could be more self-driven and determined to change their behavior or lifestyle to lose weight compared to the general population. Thus, this cohort might also seek other weight loss interventions, such as an exercise program, which could contribute as confounding factors during the period of using the smart belt. However, those confounding factors also have some limitations in this study; for example, it is uncommon for a waist belt to be worn with sports attire or active wear when users exercise. Since it might be unclear whether participants wore the smart belt all day long, including when they exercised, further research collecting more accurate data on physical activity levels from users is necessary. Third, the retention rate was low. In our study, the follow-up status of the users depended on their usage of the smart belt. Therefore, our results may be applied mainly to active smart belt users. Moreover, it is limited that there are different characteristics of how long the smart belt was used continuously among participants. Fourth, we have to be careful about reverse causation. The relative long-term smart belt users

may have higher unmeasured motivation to reduce waist circumference. However, we were unable to measure the amount of motivation in this study. Thus, a comprehensive measurement of the motivation to decrease waist circumference would be a potential strength in future studies.

Our study also has strengths compared with previous research. Although this study is insufficient to fully validate the effect of the smart belt due to flaws in study design and unmeasured confounding factors, to our knowledge no other studies have reported on the smart belt and its effectiveness. Through this observational study, we examined the association between a new smart wearable belt device and a reduction in waist circumference. Lastly, the smart belt could be used daily by male workers in Korea. Since a dress code is important in the strict and conservative organizational culture in Korea [31], most male employees in Korea usually wear the belt during the weekdays. Therefore, we expect that the smart belt would not involve having to remember to wear an additional device because almost everyone in this demographic has to wear a belt anyway.

A smart belt is a newly developed wearable device that measures real-time steps, sedentary time, and waist circumference. In our study, we showed the smart belt may have an association with a reduction in waist circumference among users through the various quantitative approaches in an observational study. Based on this result, further studies including other confounding factors (eg, change in lifestyle habits and health programs) and a control group to make comparisons should be considered. Therefore, the benefit of this research is that it may be used as the foundation for future related studies. This direct-to-consumer smart health device may contribute toward reducing the risk for NCDs and controlling the increasing health costs for health insurers, but a randomized controlled trial is necessary to further measure its effectiveness.

Acknowledgments

WELT Corporation provided the deidentified life logs of the users of the device and app. However, the corporation was not involved in analyzing the data nor in the writing of the manuscript. The authors thank the Editor and the referees for their helpful comments and suggestions. This research was supported by National IT Industry Promotion Agency (NIPA) grant funded by the Korean government (MSIT: Ministry of Science & ICT) in 2017 (NIPA-C0601-17-1008, Healthcare Wearable Smart Belt).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Comparing the baseline characteristics among the three groups.

[PDF File (Adobe PDF File), 53KB - mhealth_v7i5e10737_app1.pdf]

Multimedia Appendix 2

Difference between average steps per hour compared to baseline week.

[PDF File (Adobe PDF File), 50KB - mhealth_v7i5e10737_app2.pdf]

References

1. Alwan A. Global Status Report on Noncommunicable Diseases 2010. Geneva: The World Health Organization; 2011.
2. Alwan A, Maclean DR, Riley LM, d'Espaignet ET, Mathers CD, Stevens GA, et al. Monitoring and surveillance of chronic non-communicable diseases: progress and capacity in high-burden countries. Lancet 2010 Nov 27;376(9755):1861-1868. [doi: 10.1016/S0140-6736(10)61853-3] [Medline: 21074258]
3. Lopez AD, Murray CC. The global burden of disease, 1990-2020. Nat Med 1998 Nov;4(11):1241-1243. [doi: 10.1038/3218] [Medline: 9809543]
4. World Health Statistics 2016: Monitoring Health For The Sustainable Development Goals (sdgs) (world Health Statistics Annual). Geneva: World Health Organization; 2016. URL: https://www.who.int/sorry/ [accessed 2019-04-16] [WebCite Cache ID 77g4fG2aI]
5. Ogden CL, Carroll MD, Kit BK, Flegal KM. Prevalence of childhood and adult obesity in the United States, 2011-2012. JAMA 2014 Feb 26;311(8):806-814 [FREE Full text] [doi: 10.1001/jama.2014.732] [Medline: 24570244]
6. Kopelman PG. Obesity as a medical problem. Nature 2000 Apr 06;404(6778):635-643. [doi: 10.1038/35007508] [Medline: 10766250]
7. Wang YC, McPherson K, Marsh T, Gortmaker SL, Brown M. Health and economic burden of the projected obesity trends in the USA and the UK. Lancet 2011 Aug 27;378(9793):815-825. [doi: 10.1016/S0140-6736(11)60814-3] [Medline: 21872750]
8. Biener A, Cawley J, Meyerhoefer C. The impact of obesity on medical care costs and labor market outcomes in the US. Clin Chem 2018 Jan;64(1):108-117 [FREE Full text] [doi: 10.1373/clinchem.2017.272450] [Medline: 29097513]

9. Huh JH, Kang DR, Jang J, Shin J, Kim JY, Choi S, et al. Metabolic syndrome epidemic among Korean adults: Korean survey of Cardiometabolic Syndrome (2018). Atherosclerosis 2018 Oct;277:47-52. [doi: 10.1016/j.atherosclerosis.2018.08.003] [Medline: 30172084]
10. Baik I. Forecasting obesity prevalence in Korean adults for the years 2020 and 2030 by the analysis of contributing factors. Nutr Res Pract 2018 Jun;12(3):251-257 [FREE Full text] [doi: 10.4162/nrp.2018.12.3.251] [Medline: 29854331]
11. Kohro T, Furui Y, Mitsutake N, Fujii R, Morita H, Oku S, et al. The Japanese national health screening and intervention program aimed at preventing worsening of the metabolic syndrome. Int Heart J 2008 Mar;49(2):193-203 [FREE Full text] [Medline: 18475019]
12. Diabetes Prevention Program Research Group, Knowler WC, Fowler SE, Hamman RF, Christophi CA, Hoffman HJ, et al. 10-year follow-up of diabetes incidence and weight loss in the Diabetes Prevention Program Outcomes Study. Lancet 2009 Nov 14;374(9702):1677-1686 [FREE Full text] [doi: 10.1016/S0140-6736(09)61457-4] [Medline: 19878986]
13. Centers for Disease Control and Prevention. Health and economic costs of chronic diseases URL: https://www.cdc.gov/chronicdisease/about/costs/index.htm [accessed 2019-04-10] [WebCite Cache ID 77WPDJaNk]
14. Diabetes Prevention Program (DPP) Research Group. The Diabetes Prevention Program (DPP): description of lifestyle intervention. Diabetes Care 2002 Dec;25(12):2165-2171 [FREE Full text] [Medline: 12453955]
15. Lee W, Lee S. National Health Screening Program of Korea. J Korean Med Assoc 2010;53(5):363-370. [doi: 10.5124/jkma.2010.53.5.363]
16. Burkhauser R, Cawley J. Beyond BMI: the value of more accurate measures of fatness and obesity in social science research. J Health Econ 2008 Mar;27(2):519-529. [doi: 10.1016/j.jhealeco.2007.05.005] [Medline: 18166236]
17. Roman D, Conlee K. The Digital Revolution Comes to US Healthcare: Technology, Incentives Align to Shake up the Status Quo. New York: Goldman Sachs; 2015 Jun 29. URL: https://tinyurl.com/yyuu7go5 [accessed 2019-04-10] [WebCite Cache ID 77WPtSuo6]
18. Flechtner-Mors M, Ditschuneit H, Johnson T, Suchard M, Adler G. Metabolic and weight loss effects of long-term dietary intervention in obese patients: four-year results. Obes Res 2000 Aug;8(5):399-402 [FREE Full text] [doi: 10.1038/oby.2000.48] [Medline: 10968732]
19. Winer B. Multifactor experiments having repeated measures on the same elements. In: Statistical Principles In Experimental Design. London: Mcgraw-hill Publishing Co; 1971:514-603.
20. Xu R. Measuring explained variation in linear mixed effects models. Stat Med 2003 Nov 30;22(22):3527-3541. [doi: 10.1002/sim.1572] [Medline: 14601017]
21. de Onis M, Blössner M, Borghi E, Frongillo E, Morris R. Estimates of global prevalence of childhood underweight in 1990 and 2015. JAMA 2004 Jun 02;291(21):2600-2606. [doi: 10.1001/jama.291.21.2600] [Medline: 15173151]
22. Brodersen N, Steptoe A, Boniface D, Wardle J. Trends in physical activity and sedentary behaviour in adolescence: ethnic and socioeconomic differences. Br J Sports Med 2007 Mar;41(3):140-144 [FREE Full text] [doi: 10.1136/bjsm.2006.031138] [Medline: 17178773]
23. Sisolak P. Policygenius. 2017 Jan 11. Can a fitness tracker help you save on health insurance? URL: https://www.policygenius.com/blog/can-a-fitness-tracker-help-you-save-on-health-insurance/ [accessed 2019-04-10] [WebCite Cache ID 77WQJYbZD]
24. Bertoni S. Forbes. 2014 Dec 08. Oscar Health using Misfit Wearables to reward fit customers URL: https://www.forbes.com/sites/stevenbertoni/2014/12/08/oscar-health-using-misfit-wearables-to-reward-fit-customers/ [accessed 2019-04-10] [WebCite Cache ID 77WQTbOrB]
25. Anderson K, Burford O, Emmerton L. Mobile health apps to facilitate self-care: a qualitative study of user experiences. PLoS One 2016;11(5):e0156164 [FREE Full text] [doi: 10.1371/journal.pone.0156164] [Medline: 27214203]
26. Cadmus-Bertram L, Marcus B, Patterson R, Parker B, Morey B. Randomized trial of a Fitbit-based physical activity intervention for women. Am J Prev Med 2015 Sep;49(3):414-418 [FREE Full text] [doi: 10.1016/j.amepre.2015.01.020] [Medline: 26071863]
27. Jakicic J, Davis K, Rogers R, King W, Marcus M, Helsel D, et al. Effect of wearable technology combined with a lifestyle intervention on long-term weight loss: the IDEA randomized clinical trial. JAMA 2016 Sep 20;316(11):1161-1171 [FREE Full text] [doi: 10.1001/jama.2016.12858] [Medline: 27654602]
28. 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 Feb 17;12(1):e4 [FREE Full text] [doi: 10.2196/jmir.1376] [Medline: 20164043]
29. OECD. Health at a Glance 2017. Paris: OECD Publishing; 2017.
30. Choi J, Lee JH, Vittinghoff E, Fukuoka Y. mHealth physical activity intervention: a randomized pilot study in physically inactive pregnant women. Matern Child Health J 2016 May;20(5):1091-1101 [FREE Full text] [doi: 10.1007/s10995-015-1895-7] [Medline: 26649879]
31. Williamson L. BBC News. 2012 Jun 21. South Korea dress code: too hot for suits, but legs are taboo URL: https://www.bbc.com/news/magazine-18507075 [accessed 2019-04-10] [WebCite Cache ID 77WR5PR6u]

Abbreviations

NCD: noncommunicable disease
ANOVA: analysis of variance
CDC: Centers for Disease Control and Prevention
BMI: body mass index
OECD: Organization for Economic Co-operation and Development

Edited by G Eysenbach; submitted 12.04.18; peer-reviewed by LK Chai, B Price, K Ng; comments to author 17.08.18; revised version received 27.11.18; accepted 12.12.18; published 02.05.19.

Please cite as:
Lee M, Shin J
Change in Waist Circumference With Continuous Use of a Smart Belt: An Observational Study
JMIR Mhealth Uhealth 2019;7(5):e10737
URL: http://mhealth.jmir.org/2019/5/e10737/
doi:10.2196/10737
PMID:31045500

Original Paper

Development of the Top Tips Habit-Based Weight Loss App and Preliminary Indications of Its Usage, Effectiveness, and Acceptability: Mixed-Methods Pilot Study

Nathalie Kliemann[1,2], PhD; Helen Croker[2], PhD; Fiona Johnson[2], PhD; Rebecca J Beeken[2,3], PhD

[1]Nutritional Epidemiology Group, International Agency for Research on Cancer, World Health Organization, Lyon, France

[2]Department of Behavioural Science and Health, University College London, London, United Kingdom

[3]Leeds Institute of Health Sciences, University of Leeds, Leeds, United Kingdom

Corresponding Author:
Rebecca J Beeken, PhD
Leeds Institute of Health Sciences
University of Leeds
Level 10, Worsley Building
Clarendon Way
Leeds, LS2 9NL
United Kingdom
Phone: 44 113 343 3228
Email: r.beeken@leeds.ac.uk

Abstract

Background: The Ten Top Tips (10TT) is an intervention based on the habit formation theory that promotes a set of weight management behaviors alongside advice about repetition in a consistent context. Overall, 3 studies have demonstrated that the 10TT can support individuals to lose weight when delivered in a leaflet format. Delivery of 10TT via new technology such as a mobile app could potentially improve its effectiveness and make it more convenient, appealing, and wide reaching.

Objective: This study aimed to provide preliminary indications of the usage, effectiveness, and acceptability of an Android app of the 10TT intervention (Top Tips only app) and a second version including self-regulatory strategies for dealing with tempting foods (Top Tips plus app).

Methods: The 3-month pilot randomized adults with overweight or obesity to (1) Top Tips only app, (2) Top Tips plus app, or (3) waiting list condition. Automated data from app users were collected. Validated questionnaires assessed self-regulatory skills, weight loss (kg), and behaviors at baseline and 3 months. Users' feedback on their experience using the app was assessed using open questions.

Results: A total of 81 participants took part in the pilot; 28 participants were randomized to the Top Tips only app, 27 to the Top Tips plus app, and 26 to the waiting list condition. On average, participants viewed a mean of 43.4 (SD 66.9) screens during a mean of 24.5 (SD 44.07) log-ins and used the app for 124.2 (SD 240.2) min over the 3-month period. Participants randomized to the Top Tips only app reported the greatest improvement in self-regulatory skills (mean 0.59, SD 1.0), weight loss (mean 4.5 kg, SD 5.2), and adherence to the target behaviors (mean 0.59, SD 0.49) compared with the Top Tips plus ($mean_{self\text{-}regulation}$ 0.15, SD 0.42; $mean_{weight}$ −1.9, SD 3.9; and $mean_{behaviors}$ 0.29, SD 0.29) and waiting list condition ($mean_{self\text{-}regulation}$ −0.02, SD 0.29; $mean_{weight}$ −0.01, SD 0.51; and $mean_{behaviors}$ 0.08, SD 0.38). Participants who reported the largest improvements, on average, viewed pages 2 to 3 times more, had 2 to 3 times more log-ins, logged their weight 2 to 3 times more, and achieved the tips more than those who reported smaller changes in these outcomes. According to users' feedback, engagement with the app could be increased by making the app more interactive and allowing more tailoring.

Conclusions: This study suggests that the Top Tips app could potentially be a useful intervention for promoting eating self-regulatory skills, weight loss, and weight management behaviors among adults with overweight or obesity. Future research should develop the app further based on user feedback and test it in larger sample sizes.

Trial Registration: ISRCTN Registry ISRCTN10470937; http://www.isrctn.com/ISRCTN10470937 (Archived by Webcite at http://www.webcitation.org/76j6rQibI)

(JMIR Mhealth Uhealth 2019;7(5):e12326) doi:10.2196/12326

KEYWORDS

weight loss; habits; mobile apps; mHealth

Introduction

Background

Interest is growing in lifestyle interventions that utilize a habit-based approach as they have the potential to promote lasting weight loss and healthy dietary behaviors and are easily scalable. Habit-based interventions promote the repetition of target behaviors in a consistent context to make the behaviors more automatic and habitual [1-3]. These interventions may also promote self-regulatory skills (eg, goal setting, planning, self-monitoring, and feedback on performance) to translate the intended behavior into action and override unwanted automated responses [1,4].

Although weight loss interventions based on habit theory are still scarce [5,6], recent studies using this approach have shown promising results. A total of 3 studies have explored the delivery of a paper-based weight loss intervention that encourages habit formation for a set of target health behaviors, called Ten Top Tips (10TT). In all 3 studies, adults affected by overweight or obesity who received the 10TT leaflet lost significantly more weight (between 1.7 and 3.3 kg) than those allocated to a control group [5,7,8]. A recent study also found that the weight loss promoted by the 10TT was mediated by improvements in both self-regulatory skills and automaticity of the target behaviors [9]. However, the paper format of the 10TT is becoming outdated, and the use of new technology such as mobile apps could potentially encourage engagement among a wider range of users and improve the effectiveness of this intervention.

According to 2 meta-analyses, mobile app interventions can lead to significantly greater weight loss compared with other interventions such as paper-based interventions and counseling and education lesson–based interventions [10-12]. The retention rates of weight loss interventions may also be greater when they are technology-based compared with paper-based format [2]. There are also some indications that the use of new technologies may help people to form healthy habits and break unhealthy ones. Brief technology-based interventions promoting self-regulatory practice have been effective at improving this capacity [13-17]. This, in turn, may help people to form habits. However, most technology-based weight loss interventions currently do not support habit formation [18], and those available are not typically based on theory or evidence [19,20].

Delivering the 10TT weight loss intervention via a mobile app has the potential to be novel, effective, convenient, appealing, cost-effective, and wide reaching. Developing a mobile app version of the 10TT also offers an opportunity for testing out additional components, which could enhance the effectiveness of the intervention. Evidence suggests that strategies such as engaging in pleasant imagery tasks [21], developing intention implementations [22,23], and attention bias [24] could potentially help people to deal with tempting food and, therefore, break existing unhealthy habits. There is also some evidence to support the use of mobile phone apps to break habits through developing implementation intentions and to reduce cravings for unhealthy food through the use of imagery tasks [25]. The addition of a self-regulatory training element to help people deal with food cravings could reduce unhealthy food intake, in addition to the established effects of the 10TT on increasing healthy food intake.

Objectives

Therefore, this study developed an Android app of the 10TT intervention (Top Tips only app) and a second version that included self-regulatory strategies for dealing with tempting foods (Top Tips plus app). The aim was to provide preliminary indications of the usage, effectiveness, and acceptability of the 2 apps.

Methods

Initial Development of the Top Tips App

The development of the Top Tips apps was completed through an iterative process over a period of 1 year, involving 3 main phases: (1) initial development, (2) user testing, and (3) pilot testing.

Both the content and format of the Top Tips apps were developed based on (1) the 10TT leaflet [7], (2) the principles of habit theory, (3) empirical evidence from the field of weight loss and behavioral nutrition, and (4) the experience of the developers in designing health apps for behavior change. They were designed for Android devices as users of these devices tend to have greater socioeconomic variability compared with iPhone operation system users [26]. The team of researchers and app developers met regularly during the development process and agreed to keep the Top Tips apps simple, including only the essential features (see Multimedia Appendix 1), to allow a flexible development process and also because of budget constraints. Although the branding was kept in line with the 10TT leaflet, some necessary changes were made to develop a coherent, well-structured, and attractive app that maximized engagement with the target population.

To encourage habit formation, the apps advised users to make context-specific plans to turn each tip into a habit and adjust these whenever needed. Example plans for each tip were provided. The app also asked users to track their weight in kg and adherence to the tips each day. The apps provided automatic updates of how many times each tip was achieved per week as well as daily reminders to promote engagement with the app. A total of 9 notifications were designed related to different functions of the app, for example, “Don’t worry if you forgot to log anything last week, it’s easy to add to past days—why don’t you start now?”. A random notification was sent each day in the evenings as it was anticipated that this would be the most likely time people would log their adherence to the target

behaviors and review their plans. However, participants could turn this function off if they wanted.

The Top Tips plus app included an additional tip targeting self-regulatory strategies to resist tempting food. This new tip was developed based on the current evidence for reducing unhealthy food cravings and avoiding lapses [22-24]. The tip promoted visual imagery and distraction strategies to avoid cues that elicit urges to eat unhealthy foods, which may increase the likelihood of resisting tempting food. The additional tip also provided examples of forming coping plans using these strategies. In line with the other tips, users were required to make their own coping plans to resist unhealthy food and monitor their progress every day, assessing whether they experienced food cravings and whether they could resist them (Multimedia Appendix 2).

User Testing

The Top Tips only app and the Top Tips plus app were tested with a small convenience sample of adults who owned an Android phone. The user testing aimed to assess preliminary functionality and usability of the Top Tips apps. A total of 8 (63%[5/8] female) people took part in this study, of which 4 tested the Top Tips only app and 4 tested the Top Tips plus app. Participants were invited to download the latest version of the app and were given an individual passcode. They were instructed to enter at least one plan, log completed tips and their weight, check the content of the app for spelling errors, and provide feedback on their experience of the apps and any technical flaws.

Overall, participants reported that they liked the app and found it neat, user-friendly, and attractive. Although the app worked well for most participants, the following issues were raised: (1) inability to enter decimals to their weight in kg, (2) technical difficulties for 2 specific types of Android phones, and (3) difficulty in understanding how to use the app because there was no tutorial. These technical and weight recording issues were fixed in the final versions of the apps. To assist participants with downloading and navigating the app, a PDF document with instructions and a tutorial lasting less than 3 min explaining how to use the app were developed for each version (Multimedia Appendices 3 and 4).

The final versions of the 2 Top Tips apps were released on the Google store for pilot testing. Screenshots of the tips, planning, daily tracking, and automatic feedback features for the Top Tips only app are shown in Figure 1, whereas Figure 2 shows these features for the Top Tips plus app.

Figure 1. Screenshots of the Top Tips only app.

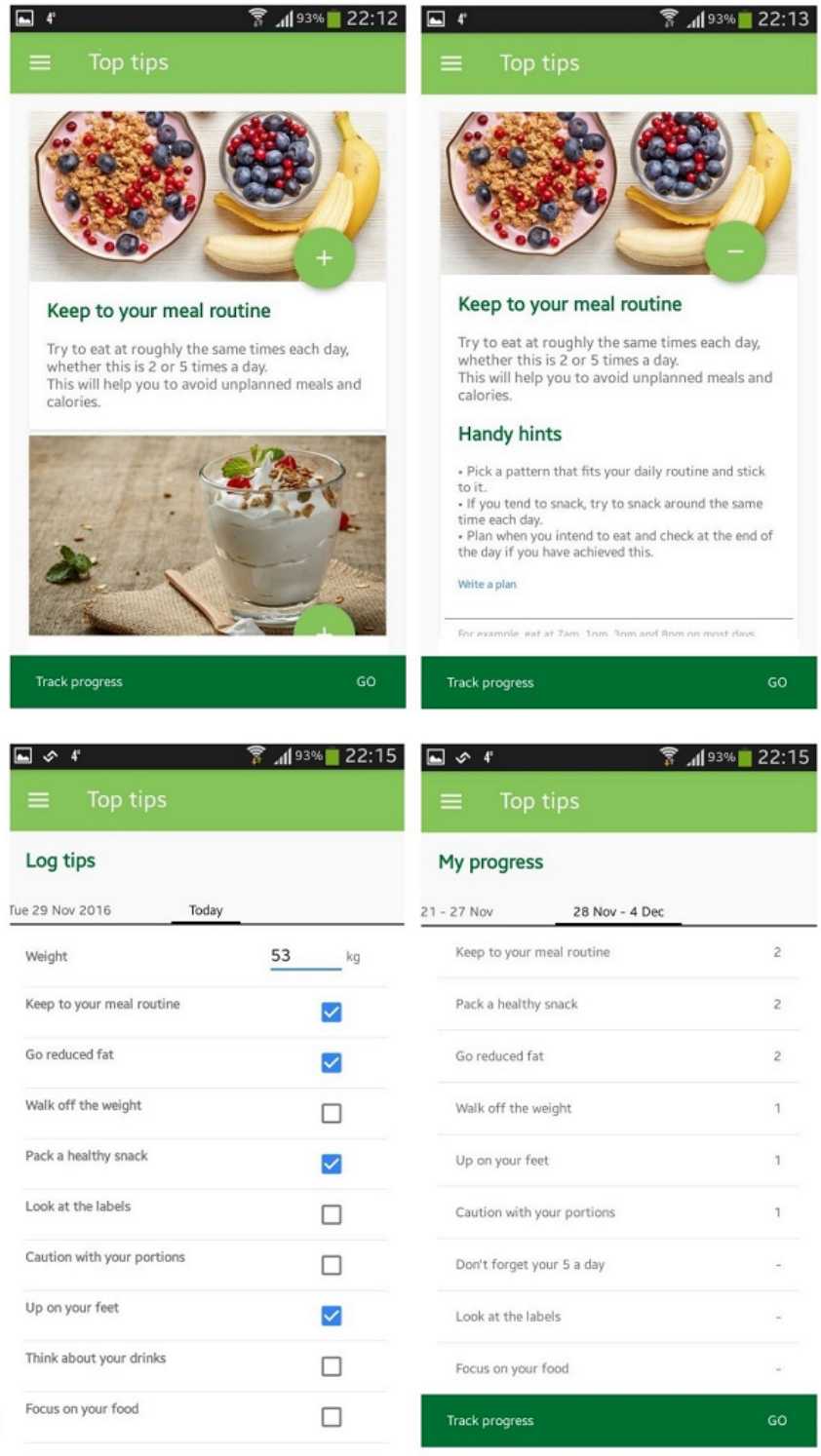
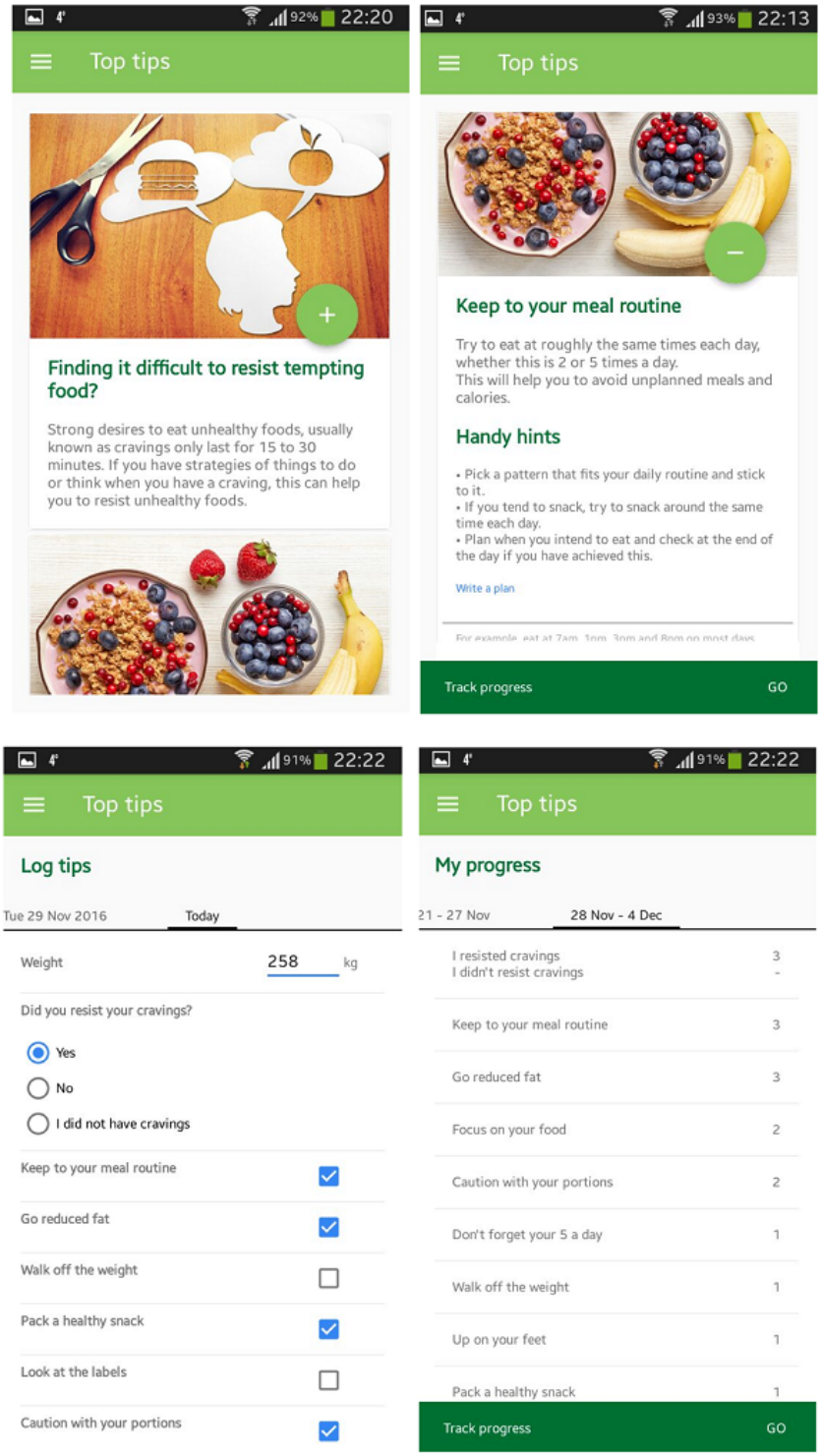


Figure 2. Screenshots of the Top Tips plus app.



Pilot Testing

The pilot was a 3-arm, individually randomized, controlled study in an online sample of adults with overweight and obesity comparing (1) Top Tips only app, (2) Top Tips plus app (including an additional tip on dealing with tempting foods), and (3) no intervention control group (waiting list). The active intervention period was 3 months, and follow-up data were collected at the end of this period. The study was approved by the University College London Ethics Committee (study ID: 5766/003). It also followed the Consort extension for pilot studies (Multimedia Appendix 5)

Participants and Recruitment

Participants were eligible to take part in the study if they (1) were adults (18 years or older) from the United Kingdom, (2) owned an Android mobile phone (3) could read English fluently, and (4) were overweight or obese (body mass index [BMI] ≥25

kg/m^2). Participants were excluded if they (1) were unable to provide informed consent, (2) were pregnant or breastfeeding, (3) were expecting to undergo bariatric surgery in the following 3 months or were recovering from a bariatric surgery, or (4) were on a strict weight loss treatment, such as meal replacements. No upper age limit was established in line with the 10TT leaflet trial [7].

Potential participants were invited via recruitment posters, social media, recruitment websites, and snowball sampling via personal contacts. A research website was also set up to provide interested participants with additional information about the study. Interested participants were invited to fill out an online survey where they were screened for eligibility. Recruitment took place over 2 months, from the beginning of January to the beginning of March 2017.

Eligible participants who gave informed consent and completed the baseline questionnaire were individually randomized to 1 of the 3 group conditions: (1) Top Tips only app, (2) Top Tips plus app, and (3) waiting list. Randomization was performed using the Minimpy software (Sourceforge) [27] and was stratified by gender, age, and BMI classification.

Sample Size

A rule of thumb for the sample size of pilot studies is to have around 25 participants per randomization group for a small standardized effect size of 0.2 [28]. As this study involved 3 experimental groups, we aimed to recruit at least 75 participants.

Procedure

After randomization, participants randomized to 1 of the 2 app groups received an email with instructions about the intervention and a passcode to access the app. Participants were instructed to use the app every day for 3 months, which is the period usually required to form habits [1,2]. Participants randomized to the control condition received an email explaining that they had been allocated to the waiting list group and that they would receive access to the weight loss app in 3 months' time. Before randomization and at 3-month follow-up, all participants were requested to complete an online questionnaire. To promote completion of the postintervention assessment, all participants had the chance to enter a draw to win 1 of 3 £20 High Street vouchers. However, they were only informed about the prize draw at the end of the intervention, to ensure that only participants who were motivated to lose weight and to improve their diet were recruited to the study. Participants from the Top Tips apps conditions were also invited to answer qualitative questions at the end of the online follow-up questionnaire to further explore their experience of using the apps.

Measures

Automated data from Top Tips only and Top Tips plus users were collected over the 3-month intervention period to assess usage patterns. This included data on the number of log-ins, pages viewed, plans made, and the total time spent on the apps in minutes. Information was also collected on the number of times that weight was logged and each tip was achieved. Information was collected from the Top Tips plus users on the number of times they resisted food cravings, number of times they did not resist, and number of times they did not have food cravings.

The online questionnaires at baseline and follow-up collected information on sociodemographics, self-regulation, anthropometrics, and dietary behaviors. Participants were asked to report their gender, age, ethnicity, marital status, education, and employment status. Due to the small sample size, all variables were categorized into 2 groups. Ethnicity was categorized as *white* or *other* (black, Asian, mixed, or other). Marital status was categorized as *married* (or living as married) and *not married or other* (single, separated, divorced, or widowed). Education was categorized as *nondegree* or *degree*. Employment status was categorized as *paid work* and *unpaid work or other* (unemployed, homemaker, voluntary work, disabled or too ill to work, student, or retired).

Eating self-regulatory skills were assessed using the 5-item Self-Regulation of Eating Behavior Questionnaire [29]. Total mean score and changes over 3 months were calculated. Weight and height were self-reported. For those who did not complete the follow-up questionnaire, their last weight logged on the app was used. Changes in weight in kg over 3 months were calculated. BMI was also calculated and then categorized into overweight (25-29.9 kg/m^2) or obese (30 kg/m^2 or over).

Frequency questions were based on those used within a previous 10TT randomized controlled trial (RCT) [7] and assessed the 10 eating and activity behaviors plus self-weighing targeted in the Top Tips apps. For some of these behaviors, more than 1 frequency question was used to better assess adherence to the behavior. For example, for the *look at labels* behavior, 2 questions were generated to ask about how often people looked at labels when (1) preparing food and (2) buying food. A total of 16 questions were used to assess the frequency of carrying out each of the target behaviors over the previous 2 weeks on a 5-point Likert scale. The overall mean score for the 16 behaviors was calculated as well as the mean change from baseline to 3-month follow-up. Dietary intake was assessed in more detail using validated food frequency questionnaires. For example, fat intake was assessed using the dietary fat scale from the validated Dietary Instrument for Nutrition Education [30], which was adapted to broaden the range of ethnically diverse foods and the main components of the UK diet. Fruit and vegetable intake was assessed using 2 validated food frequency questions [31], measuring intake on a 7-point response scale. Similarly, 2 food frequency questions assessed sweet snack (SS) intake, including foods such as chocolates, sweets biscuits, cakes, buns, pastries, and ice-cream. In addition, 4 frequency questions assessed the consumption of sugary sweetened drinks (SSD) intake, including nondiet fizzy drinks, sugar-containing squashes, milkshakes, and hot chocolate. The response options ranged from 1 (never or rarely) to 7 (3 or more times a day). Following the study by McGowan et al (2012), answers were recoded to represent daily intake, for example, 2 to 3 times a week was coded as 0.36. The mean scores for fruit and vegetable, SS, and SSD frequencies were calculated as well as the mean change from baseline to 3-month follow-up.

To assess acceptability of the Top Tips apps, 8 open questions relating to users' experience of using the apps were included in

the online follow-up questionnaire but stated as optional (Multimedia Appendix 6). This included their overall views toward the app, if there was anything that they disliked or found hard to use, if there was anything they liked or found easy to use, if there was anything that they were expecting to see but did not, how the app could be improved, and if they had any other comments they would like to make.

Statistical Analyses

For this pilot study, usage pattern and users' feedback were the primary variables of interest, but the impact of the intervention on self-regulatory skills, weight loss, and behaviors was also explored. All analyses were conducted on an intention-to-treat basis, with participants analyzed based on assigned randomization group [32]. Descriptive analyses were used to characterize the sample by study arm. Baseline differences between those who downloaded and did not download the app were tested using Chi-square tests for categorical variables and *t* tests or Mann-Whitney tests for continuous variables. Descriptive analyses were used to show the usage pattern of the Top Tips apps. Mean, SD, median, and total range were reported for each usage metric.

Exploratory descriptive analyses were performed to obtain an early indication of the effect of the Top Tips apps on eating self-regulatory skills, weight, and behaviors, including dietary intake. Initially, a completer analysis was performed using the complete data at baseline and follow-up for each outcome. Participants with more than 20% of missing data at baseline for the self-regulation and target behaviors questionnaires and with any missing data for dietary intake questions were excluded from the analyses. When there were up to 20% missing data for the self-regulation and target behaviors questionnaires, the individual median score was imputed. Within-group changes from baseline to 3 months were described for each outcome, including 95% CI, and the Cohen effect size was calculated. Sensitivity analysis using the last observation carried forward approach was performed to investigate the potential effect of missing responses on effect sizes. Analysis of variance was also conducted using the imputed data. Descriptive analyses were used to assess the relationships between overall app usage and changes in eating self-regulatory skills, weight, and target behaviors over 3 months. For this analysis, the levels of change in self-regulatory skills, weight, and target behaviors were categorized into 2 groups using ranked percentiles: (1) percentile <75 represented medium-to-small changes and (2) percentile ≥75 represented large changes. Rank percentiles were used to categorize data into low and high as the data were skewed.

Users' feedback on their experience using the app was analyzed using thematic analysis [33]. This method identifies and reports patterns (themes) within data. All quantitative analyses were undertaken using SPSS version 24.0 (SPSS Inc).

Results

Participants and Recruitment

A total of 201 adults were interested in the intervention and were assessed for eligibility. Of these, 120 were excluded because they had a BMI <25.0 kg/m^2 (n=10), did not own an Android mobile phone (n=81), or did not complete the baseline questionnaire (n=29). A total of 81 participants were eligible to take part in the study; 28 participants were randomized to the Top Tips only app, 27 to the Top Tips plus app, and 26 to the waiting list group. Multimedia Appendix 7 displays the flow diagram of study participation over the 3-month study period.

Table 1 shows the baseline characteristics of the participants, which appeared similar across the 3 study arms. The majority of the participants were female (approximately 90%) and white (approximately 84%). Approximately two-thirds had a degree (approximately 74%) and half were married (approximately 54%) and were in paid work (approximately 59%). Overall mean age was 42.4 (SD 13.4) and BMI was 34.3 kg/m^2 (SD 7.0). The Top Tips app was downloaded by 60% (17/28) of the participants randomized to the Top Tips only condition and by 70% (19/27) of those randomized to the Top Tips plus condition. Those who did not download the app were not significantly different at baseline with regard to any of the sociodemographic variables from those who downloaded the app.

Table 1. Baseline characteristics of the condition groups.

Characteristics	Top Tips only (n=28)	Top Tips plus (n=27)	Waiting list (n=26)
Gender, n (%)			
Female	24 (86)	25 (93)	24 (92)
Age (years), mean (SD)	43.6 (13)	44.0 (14)	40.6 (13)
Ethnic group, n (%)			
White	23 (82)	23 (85)	22 (85)
Other[a]	5 (18)	4 (15)	4 (15)
Marital status, n (%)			
Married[b]	15 (54)	15 (55)	14 (54)
Not married or other[c]	13 (46)	12 (44)	12 (46)
Education, n (%)			
Nondegree[d]	6 (21)	8 (30)	7 (27)
Degree[e]	22 (79)	19 (70)	18 (69)
Missing[f]	—	—	3.8 (1)
Employment situation, n (%)			
Paid work[g]	20 (71)	15 (56)	13 (50)
Unpaid work or other[h]	8 (29)	12 (44)	12 (46)
Missing[f]	—	—	1 (4)
Weight status, n (%)			
Overweight[i]	8 (29)	7 (26)	7 (27)
Obese[j]	20 (71)	20 (74)	19 (73)
Body mass index, mean (SD)	33.7 (7)	35.0 (8)	34.0 (7)
Eating self-regulatory skills[k], mean (SD)	2.81 (0.57)	2.87 (0.69)	2.85 (0.51)

[a]Black, Asian, mixed, or other.

[b]Married or living as married.

[c]Single, separated, divorced, or widowed.

[d]Primary/secondary school or O level/GCSEs/A levels or technical/trade certificate/diploma.

[e]Degree or postgraduate degree.

[f]No response.

[g]Employed full-time/employed part-time/self-employed.

[h]Unemployed/full-time homemaker/unpaid or voluntary work/disabled or too ill to work/student/retired.

[i]BMI from 25.0 kg/m^2 to 29.9 kg/m^2.

[j]BMI 30.0 kg/m^2 or over.

[k]Eating self-regulatory skills assessed using the Self-Regulation of Eating Behavior Questionnaires.

Usage Pattern

Usage pattern for each Top Tip app and overall is presented in Table 2. Although there was significant variability between participants, on average, participants viewed a mean of 43.4 (SD 66.9) screens during a mean of 24.5 (SD 44.07) log-ins and used the app for 124.2 (SD 240.2) min over the 3-month intervention. Plans were made on average 4.6 (SD 3.9) times, weight was logged around 8.3 (SD 15.9) times, and tips were achieved on average 10.1 (SD 21.2) times over the course of the intervention. Participants randomized to the Top Tips only condition seemed to have used the app twice as much as those randomized to the Top Tips plus condition.

Table 2. Usage pattern per app and overall.

Usage pattern	Top Tips only (n=17)	Top Tips plus (n=19)	Overall (n=36)	
	Mean (SD)	Mean (SD)	Mean (SD)	Minimum-Maximum observations
Number of screens viewed	56.6 (94.9)	32.2 (24.9)	43.4 (66.9)	2-283
Number of log-ins	33.8 (63.1)	16.5 (13.5)	24.5 (44.07)	1-253
Cumulative minutes using the app	162.1 (296.5)	92.1 (181.4)	124.2 (240.2)	0.01-1200.8
Number of plans made	4.8 (3.9)	4.4 (3.9)	4.65 (3.9)	0-11
Number of times weight was logged	9.9 (20.8)	6.8 (10.3)	8.3 (15.9)	0-74
Number of times tips were achieved	14.0 (29.0)	6.7 (9.9)	10.1 (21.2)	0-102
Number of times each tip was achieved				
Keep to your meal routine	0.06 (0.24)	0.35 (0.81)	0.22 (0.63)	0-3
Go reduced fat	0.18 (0.52)	0.25 (0.55)	0.22 (0.53)	0-2
Walk off the weight	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0-0
Pack a healthy snack	0.12 (0.48)	0.15 (0.36)	0.14 (0.42)	0-2
Look at the labels	0.18 (0.53)	0.25 (0.71)	0.22 (0.63)	0-3
Caution with your portions	12.4 (28.4)	4.4 (6.9)	8.1 (20.0)	0-100
Up on your feet	0.06 (0.24)	0.10 (0.31)	0.08 (0.27)	0-1
Think about your drinks	0.06 (0.24)	0.10 (0.31)	0.08 (0.27)	0-1
Focus on your food	0.29 (0.47)	0.35 (0.59)	0.32 (0.53)	0-2
Don’t forget your 5-a-day	0.65 (1.0)	0.85 (1.75)	0.76 (1.46)	0-7
Extra: Cravings were resisted[a]	—[b]	4.6 (6.8)	4.6 (6.8)	0-25
Extra: Cravings were not resisted[c]	—[b]	3.2 (4.6)	3.2 (4.6)	0-16

[a]Number of times people resisted their food cravings.

[b]No data for the Top Tips only app as it did not have these extra tips.

[c]Number of times participants did not resist their food cravings.

The tip most frequently achieved was *Caution with portions* (mean 8.1, SD 20.0), followed by *don’t forget your 5 a day* (mean 0.76, SD 1.46) and *Focus on your food* (mean 0.32, SD 0.53). The tip least achieved was *Walk off the weight*, which was not achieved by any participant during the entire intervention. This pattern was found in both apps. Regarding the tip on how to resist tempting food within the Top Tips plus app, participants logged success (mean 4.6, SD 6.8) more times than failure (mean 3.2, SD 4.6) for their attempts to resist tempting food.

Postintervention Effect on Eating Self-Regulatory Skills, Weight, and Behaviors

Baseline and follow-up data for each outcome per group are illustrated in Figure 3, whereas changes over time are shown in Table 3. Eating self-regulatory skills increased the most in the Top Tips only group (mean 0.59, SD 1.0), followed by the Top Tips plus group (mean 0.15, SD 0.42). No changes were found for the waiting list group (mean −0.02, SD 0.29). These changes represented a medium-sized effect for the Top Tips only and small-sized effect for the Top Tips plus condition, which were in line with the effect sizes found in the sensitivity analysis (Multimedia Appendix 8).

Figure 3. Baseline and follow-up weight and behaviors per condition group. F&V: Fruit and Vegetables.

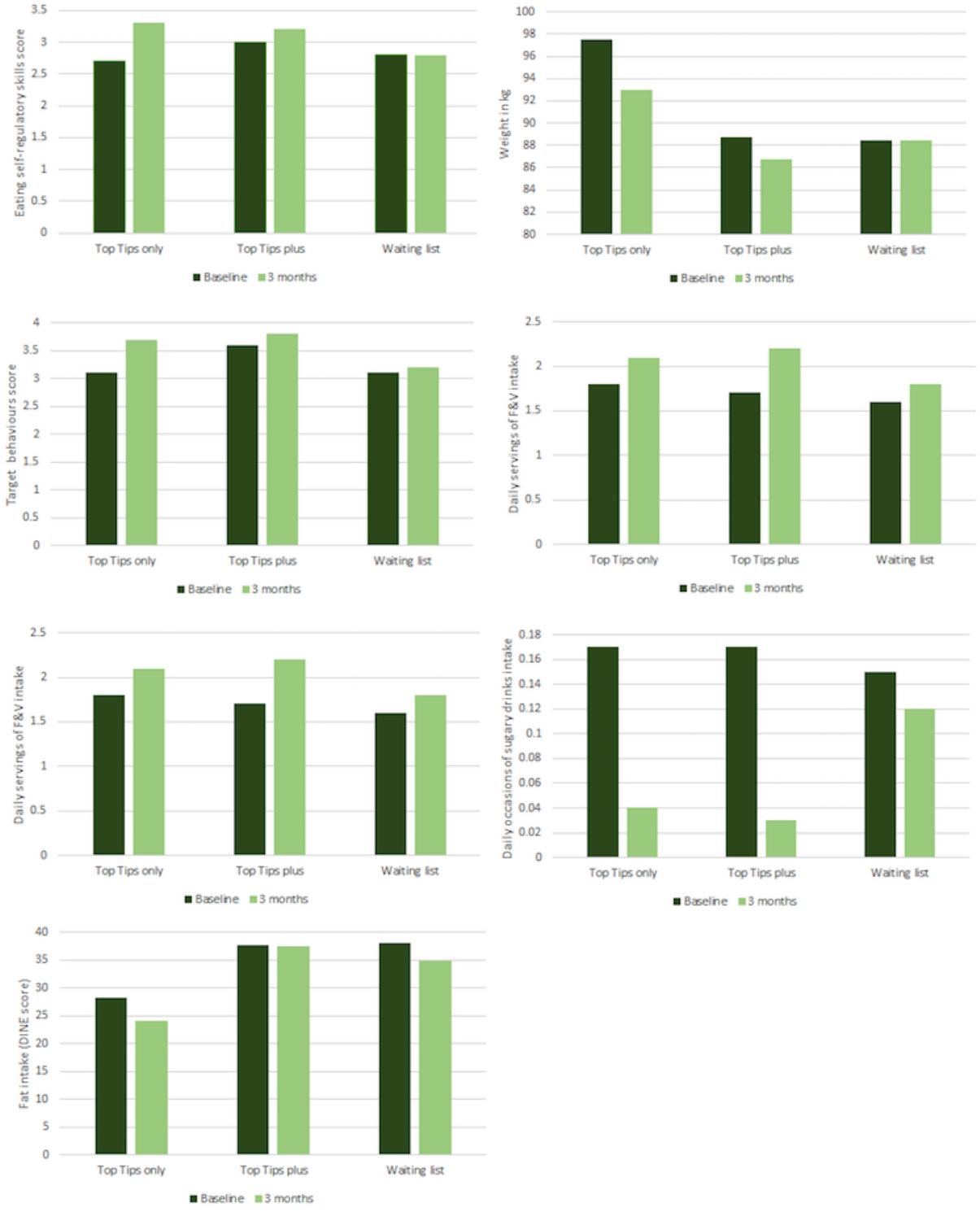


Table 3. Preliminary indication of the effect of the Top Tips apps on weight loss and behaviors.

Changes over 3 months	Top Tips only (n=17)			Top Tips plus (n=19)			Waiting list (n=26)		
	n (%)	Mean (95% CI)	*d* [a]	n (%)	Mean (95% CI)	*d* [a]	n (%)	Mean (95% CI)	*d* [a]
Self-regulation[b]	5	0.59 (−0.76 to 1.94)	0.53	8	0.15 (−0.20 to 0.50)	0.35	18	−0.02 (−0.16 to 0.12)	0.06
Weight in kg	7	−4.50 (−0.93 to 0.27)	0.80	11	−1.90 (−4.4 to 0.43)	0.05	8	−0.15 (−5.24 to 4.95)	0.002
Target behaviors[c]	6	0.59 (0.06 to 1.11)	1.10	8	0.29 (0.05 to 0.53)	1.00	20	0.08 (−0.09 to 0.26)	0.22
Fruit and vegetable intake[d]	5	0.26 (−0.65 to 1.18)	0.35	8	0.42 (−0.39 to 1.23)	0.43	18	0.22 (−0.29 to 0.73)	0.21
Sweet snacks intake[e]	5	−0.04 (−0.17 to 0.08)	0.45	8	−0.29 (−0.84 to 0.26)	0.43	18	−0.18 (−0.44 to 0.07)	0.35
Sugary sweetened drinks intake[f]	5	−0.07 (−0.17 to 0.04)	0.78	8	−0.13 (−0.41 to 0.14)	0.40	18	−0.03 (−0.08 to 0.02)	0.25
Fat intake[g]	5	−4.20 (−14.1 to 5.74)	0.52	8	−0.12 (−10.84 to 10.59)	0.01	18	−3.10 (−7.20 to 0.87)	0.04

[a]Cohen *d* effect size.

[b]Eating self-regulatory skills assessed using the Self-Regulation of Eating Behavior Questionnaire, scores ranged from 1 (never) to 5 (always).

[c]Overall mean score for the frequency of the 16 target behaviors, scores ranged from 1 (none of the time) to 5 (all of the time).

[d]Fruit and vegetable intake in servings per day.

[e]Daily occasions of sweet snacks intake.

[f]Daily occasions of sugary sweetened drinks intake.

[g]Score for the Dietary Instrument for Nutrition Education questionnaire. Cutoffs: <30 low fat, 30-40 medium fat, and >40 high fat.

The results also suggest that over the 3-month period, weight loss was greater among those who received the Top Tips only (mean −4.5 kg, SD 5.2), followed by those who received the Top Tips plus (mean −1.9 kg, SD 3.9), and no weight loss was observed among those allocated to the waiting list (mean −0.01 kg, SD 0.51). This represented a large-sized effect for the Top Tips only and a medium-sized effect for the Top Tips plus, but according to the sensitivity analysis, the effect on weight loss was small for both app conditions (Multimedia Appendix 8).

Similarly, the Top Tips only app appeared to promote a greater increase in adherence to the target behaviors (mean 0.59, SD 0.49) than the Top Tips plus app (mean 0.29, SD 0.29), whereas no changes were observed in the waiting list (mean 0.08, SD 0.38) condition. These changes represented a large-sized effect for both app conditions. However, the sensitivity analysis suggested that the effect of both apps on adherence to the target behaviors represented a medium-sized effect (Multimedia Appendix 8).

Regarding the effect on dietary intake, the Top Tips plus group experienced the greatest increase in fruit and vegetable intake (mean 0.42, SD 0.97) and decrease in SS (mean −0.29, SD 0.66) and SSD (mean −0.13, SD 0.33) intake (all representing a medium-sized effect) compared with Top Tips only and waiting list groups. With respect to fat intake, the Top Tips only group reported the greatest changes (mean −0.42, SD 8.01, representing a medium-sized effect) compared with the Top tips plus and waiting list groups. Sensitivity analyses suggested a small effect size for all dietary changes (Multimedia Appendix 8).

None of the changes in the outcomes between the condition groups were found to be significant in the sensitivity analysis (Multimedia Appendix 8). However, this should be interpreted with caution as this study was not powered to detect significant differences.

Relationships Between App Usage and Changes in Eating Self-Regulatory Skills, Weight, and Behaviors

Table 4 shows the relationships between the Top Tips apps usage and changes in self-regulatory skills, weight, and adherence to target behaviors. The results suggest that participants with the greatest changes for these outcomes, on average, viewed pages 2 to 3 times more, had 2 to 3 times more log-ins, logged their weight 2 to 3 times more, and achieved the tips more than those who showed smaller changes in these outcomes. Moreover, participants who reported the greatest changes in eating self-regulatory skills, weight, and adherence to target behaviors made on average 1, 2, and 3 plans less than those with smaller changes, respectively. App usage in minutes was also higher among those with greater improvements for eating self-regulatory skills (approximately 500% higher) and target behaviors (approximately 140% higher) than those who made smaller changes. In contrast, those who lost more weight used the apps about 15% less than those who lost less weight over the course of the intervention.

Table 4. App usage per level of changes in self-regulatory skills, weight, and target behaviors over 3 months (data from both Top Tips apps).

Outcome/app feature	Percentile <75[a], changes over 3 months			Percentile ≥75[a], changes over 3 months		
	n	Mean (SD)	Median	n	Mean (SD)	Median
Self-regulation[b]						
All participants[c]	8	−0.06 (0.58)	−0.10	5	0.92 (0.58)	0.80
Number of screens viewed	8	41 (26)	37	5	84 (113)	41
Number of log-ins	8	22 (14)	18	5	38 (45)	22
Cumulative minutes using the app	8	48.6 (57.8)	15.9	5	241.6 (339.1)	64.6
Number of plans made	8	6 (4)	7	5	5 (5)	8
Times weight was logged	8	7 (4)	6	5	24 (32)	6
Times tips were achieved	8	7.7 (4.3)	5	5	23.2 (30.6)	9
Weight[d]						
All participants[c]	6	1.13 (2.09)	0.97	12	−4.97 (4.05)	−3.65
Number of screens viewed	6	62 (63)	6	12	85 (101)	12
Number of log-ins	6	26 (21)	21	12	53 (73)	32
Cumulative minutes using the app	6	212.9 (315.7)	50.3	12	184.8 (350.6)	60.6
Number of plans made	6	7 (4)	8	12	5 (4)	6
Times weight was logged	6	9 (9)	6	12	21 (24)	10
Times tips were achieved	6	10.3 (11.2)	7	12	25.5 (33.2)	10
Target behaviors[e]						
All participants[c]	7	0.12 (0.26)	0.25	7	.73 (.26)	.62
Number of screens viewed	7	38 (20)	39	7	109 (122)	72
Number of log-ins	7	18 (11)	19	7	69 (89)	44
Cumulative minutes using the app	7	175.8 (310.6)	35.9	7	242.0 (438.3)	64.6
Number of plans made	7	7 (4)	7	7	4 (5)	2
Times weight was logged	7	7 (4)	6	7	25 (28)	10
Times tips were achieved	7	7.5 (5.13)	7	7	32.4 (39.9)	10

[a]Changes to the outcome over 3 months categorized according to the percentile, that is, <75=medium to low changes and ≥75=greater changes.

[b]Changes in eating self-regulatory skills assessed using the Self-Regulation of Eating Behavior Questionnaire

[c]Data from Top Tips only and Top Tips Plus participants.

[d]Changes in weight in kg.

[e]Changes in the overall mean score for the frequency of the 16 target behaviors.

Acceptability Feedback

A total of 8 participants gave feedback on their experience using the Top Tips apps. Of these, 75% were female (n=7). Two participants complained about technical issues: one participant reported an issue in downloading the app and they were, therefore, unable to follow the intervention, and the other participant was unable to access the daily tips.

Participants' overall views toward the app were both positive and negative. Some participants mentioned that they did not find the app useful and found it unoriginal and boring. Others said the app was well designed and helped them to track their diet plan:

> *It is very well designed. It helps you to keep track of your weight loss goals.* [Male, 30 years old]
>
> *I didn't find it particularly helpful.* [Female, 43 years old]
>
> *Helped me focus on my diet plan.* [Female, 57 years old]
>
> *Boring, unoriginal and old hat.* [Female, 58 years old]

Participants also commented on what they liked and found easy to use. The way the tick boxes were designed to track their adherence to the target behaviors was considered effective and easy to use. Some participants also mentioned that they liked the daily reminders and the possibility of setting their own plans:

> *[I liked the] daily weight reminder.* [Female, 57 years old]
>
> *The way you have to tick boxes. Easily and effective. It helps you to build new eating habits.* [Male, 30 years old]

In contrast, some participants said they disliked the reminder, as they found it annoying. The lack of interactivity was also mentioned as a negative aspect of the Top Tips app and the fact that they could not tailor the app to their personal needs:

> *Lack of any interactivity.* [Female, 50 years old]
>
> *Absence of feedback support.* [Female, 58 years old]
>
> *Couldn't delete the goals I didn't like.* [Female, 43 years old]

With respect to users' expectations, some participants said that the app was just not what they expected. Some were expecting the app to include a food diary and allow them to tailor the goals. Some were also expecting that the app would involve more complex information related to weight loss:

> *Just wasn't what I expected.* [Female, 56 years old]
>
> *New ideas motivating information on metabolism and food and exercise. Your app had the standard I would expect from a gcse student.* [Female, 58 years old]
>
> *Ability to tailor goals more.* [Female, 43 years old]

Finally, participants made some suggestions for improving the Top Tips app. Some participants suggested the inclusion of recipes and the use of different strategies to remind people about the tips apart from the daily notifications, such as emails. They also suggested the inclusion of food diaries to track their dietary intake:

> *Reminders should come up at different times—and also try different strategies (notifications, e-mail, etc.). It could [also] include healthy recipes to help people cook healthy food at home.* [Male, 30 years old]
>
> *Something more like my fitness pal.* [Female, 50 years old]

Discussion

Principal Findings

The findings of this study suggested that the Top Tips habit-based app could potentially be a useful intervention for promoting eating self-regulatory skills, weight loss, and healthy behaviors among adults with overweight or obesity. The usage patterns indicated those who engaged more with the app also reported greater changes in self-regulatory skills, weight, and adherence to target behaviors. Although there are hundreds of commercially available mobile phone apps designed to help people lose weight or form habits, most of these are neither theory nor evidence based [19]. Recently, a study assessed the feasibility of the Habit app for weight loss problem solving [34]. However, this study did not include all the self-regulatory components required to form habits, such as self-monitoring, and neither did it assess the effect of the app intervention on automaticity or dietary behaviors. Therefore, this is one of the first studies to provide some indications of the usage and effect of a weight loss app based on the habit formation theory.

The app was expected to be accessed every day over 3 months to log achieved tips and current weight, but it was accessed on average 25 times. This is in agreement with a systematic review that suggested that most mobile app interventions for weight loss interventions tend to have high attrition and participants tend to disengage from the app after the first month [10]. Furthermore, on average, people made plans for half of the tips, and most of the tips were achieved less than 3 times over the course of the intervention. The exception was the tip on *Caution with portions*, which was achieved on average 8 times, and weight, which was also logged 8 times on average. The tip *walk off the weight* was not achieved by anyone. However, this does not mean people did not increase the number of steps because of this intervention, as they might have increased but not reached the 10,000 steps recommended per day. The integration of an electronic activity monitor to the app could help to better understand the effect of the intervention on activity behaviors. Overall, this usage pattern suggests that there is room for improvement regarding engagement with the app.

According to users' feedback, engagement with the app could be increased by making the app more interactive, enabling more tailoring to personal needs, and including more resources for weight loss (eg, recipes). The app's simplicity and design should be maintained, as these were aspects considered positive by users. A recent study highlighted these features as important for keeping users engaged, alongside other features such as feedback function, ability to change design to suit own preference, and tailored information [10]. In-depth focus groups with the target population could also help to better understand the aspects necessary to improve the app. In addition, a better understanding of engagement should also be considered by using the recently published conceptual framework of engagement with digital behavior change interventions (DBCIs), which states that engagement may be influenced by the DBCI itself, the context of use, mechanisms of action of the DBCI, and the target behavior [35]. Finally, considering that only 8 participants opted to answer the feedback questions, future studies should consider including closed feedback questions to increase the response rate and improve our understanding of users' experiences of these apps.

There was significant interest in the study: over 2 months, 201 people signed up for the study. Of these, 81 adults with overweight or obesity were excluded because they did not have an Android phone. Among those randomized to the app conditions, about one-third did not download the app, suggesting a reduction in interest before even beginning the intervention. The responses for the follow-up online questionnaire were also very low for the app conditions (approximately 25%). In contrast, the follow-up response for the waiting list was high (77%), possibly because of the fact that completing the follow-up questionnaire was a condition for subsequent access to the Top Tips app. Future studies should improve the instruction materials and test other strategies to reduce the dropout for the intervention conditions. The inclusion of face-to-face or telephone support before and after technology-based interventions may increase retention and

engagement as well as weight loss [11]. In addition, making the app available for iPhone operation system could increase the reach of the intervention and also improve retention [36]. Dual phone-computer access could also help increase retention, as it is valued by users [37]. These were not possible to implement in this study because of budget constraints but could be addressed in future studies.

The majority of participants in this pilot were white, female, and highly educated. This corroborates with findings from weight management interventions, which tend to underrepresent men [38]. Mobile phone ownership also tends to be higher among those more affluent and educated [39,40]. However, the use of mobile phones has been increasing among lower SES populations, reducing social inequalities for access to evidence-based health apps [39]. Future studies should consider recruitment through clinical settings and targeted strategies to try and achieve a more socioeconomically balanced sample.

Although this study was not powered to detect changes or to explore mechanisms of action, the direction of the results observed in this pilot is a preliminary indication that the app worked as expected. For example, participants in both app group conditions improved their eating self-regulatory skills, whereas no changes were observed in the waiting list group. Users more engaged with the self-regulatory components of the app improved their eating self-regulatory skills to a greater extent, had greater adherence to the target behaviors, and lost more weight than those who engaged less. This is in line with recent evidence suggesting that nutrition and weight loss interventions using self-regulation components tend to be more effective [41,42]. Frequency of app use has also been related to higher success in changing diet and activity behaviors [43]. In contrast with these results, the time of using the apps was greater among those who lost less weight. This may reflect the fact that some people left their app open but were not necessarily using it, suggesting that this is a less accurate tool for measuring engagement.

Both the Top Tips only (mean −4.5 kg) and the Top Tips plus (mean −1.9 kg) apps appeared to promote a greater weight loss than the waiting list (mean −0.01 kg). This is in line with findings from a pilot RCT using the 10TT leaflet in a UK community-based sample, which showed that the 10TT group lost 2 kg, whereas the waiting list group lost 0.4 kg over 8 weeks [5]. Similarly, a recent RCT using the 10TT with adults who had overweight or obesity in Australia found a weight loss of 3.3 kg in the 10TT group compared with 0.4 kg in the control group over 12 weeks [8]. In addition, a previous RCT in patients with obesity from UK primary care showed that those who received the 10TT leaflet lost 1.68 kg over 3 months, and this was maintained over 2 years [44]. Regarding the effect on target behaviors and dietary intake, both app groups reported changes in the expected direction, which were in general greater than those reported by the waiting list group. Although the effects of the Top Tips app on weight loss and behaviors are promising, they should be interpreted with caution because of the small sample size. Future studies should also test the longer-term effects of the Top Tips app.

It was not possible to draw conclusions regarding any differences in impact between the app conditions. The Top Tips plus app appeared to promote greater absolute decreases in SS and SSD intake compared with the Top Tips only app, as expected. However, in contrast, the absolute changes in self-regulatory skills, target behaviors, and weight loss appeared greater in the Top Tips only condition. This may reflect differences in usage between the apps, as participants in the Top Tips only group used the app almost twice as much as the Top Tips plus participants. The greater usage patterns among those using the Top Tips only app may simply reflect the small sample size of this study, which increases the chance of false positives and can inflate effect sizes. Therefore, the potential additive effect of this new tip on how to deal with tempting food needs further examination using larger samples sizes. Future studies testing the Top Tips app would also benefit from a variety of experimental designs that tease out the main active components within the intervention, for example, a sequential multiple assignment trial or a multiphase optimization strategy design [45].

Limitations

This study had a number of limitations. The sample size was small and the study was not powered to detect differences in the outcomes, thus limiting the conclusions that can be drawn from the results. Allocation bias might also have affected the results, as people were not blinded to their condition. Ethnic minorities, men, and people from lower SES backgrounds were underrepresented. There are also limitations related specifically to the measure used to assess dietary intake. The frequency questions lacked portion size information and did not allow the calculation of overall energy intake. This may have limited the accuracy of the data collected and the understanding of changes in dietary intake. However, the unannounced and self-administered features of these questions combined with the fact that they captured habitual behaviors are important strengths of these measures [46]. Furthermore, given that the measures used in this study were all self-reported, changes in self-regulatory skills, weight, adherence, and dietary intake may represent the individuals' perception of change rather than actual change. Future studies should aim to use real-time mobile-based assessment of nutrition, physical activity, and behaviors, as this may reduce participant burden and bias [47].

Conclusions

The findings of this paper suggest that an app version of the 10TT habit-based program may potentially enhance self-regulatory skills and promote healthy dietary behaviors and weight loss. Although engagement was moderate, the results indicated that absolute changes in the outcomes increased with app usage, suggesting that it worked better among those who did engage. According to users, the Top Tips app could be improved and engagement encouraged through more interactivity and weight loss resources and by enabling tailoring. Future research should seek to develop the app further and test it in larger, more diverse samples using designs that enable the main active components within the intervention to be examined.

Acknowledgments

Researchers involved in this study were supported by Cancer Research UK (CRUK), Yorkshire Cancer Research (YCR), and Coordination for the Improvement of Higher Education Personnel (CAPES/Brazil). The views expressed in this paper are those of the authors and not necessarily those of CRUK, YCR, and CAPES/Brazil. The app development was partially funded by the charity Weight Concern and partially funded by White October, a Web design and development company.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Top Tips app features.

[PDF File (Adobe PDF File), 14KB - mhealth_v7i5e12326_app1.pdf]

Multimedia Appendix 2

Additional tips on resisting tempting foods.

[PDF File (Adobe PDF File), 40KB - mhealth_v7i5e12326_app2.pdf]

Multimedia Appendix 3

Tutorial for the Top Tips only app.

[PDF File (Adobe PDF File), 9KB - mhealth_v7i5e12326_app3.pdf]

Multimedia Appendix 4

Tutorial for the Top Tips plus app.

[PDF File (Adobe PDF File), 9KB - mhealth_v7i5e12326_app4.pdf]

Multimedia Appendix 5

Consort checklist for pilot studies.

[PDF File (Adobe PDF File), 103KB - mhealth_v7i5e12326_app5.pdf]

Multimedia Appendix 6

Qualitative questions on users' experience.

[PDF File (Adobe PDF File), 23KB - mhealth_v7i5e12326_app6.pdf]

Multimedia Appendix 7

Flow diagram of participation during the 3-month study period.

[PDF File (Adobe PDF File), 56KB - mhealth_v7i5e12326_app7.pdf]

Multimedia Appendix 8

Sensitivity analysis.

[PDF File (Adobe PDF File), 26KB - mhealth_v7i5e12326_app8.pdf]

References

1. Lally P, Gardner B. Promoting habit formation. Health Psychol Rev 2013 May;7(sup1):S137-S158. [doi: 10.1080/17437199.2011.603640]
2. Lally P, van Jaarsveld CH, Potts HW, Wardle J. How are habits formed: modelling habit formation in the real world. Eur J Soc Psychol 2009 Jul 16;40(6):998-1009. [doi: 10.1002/ejsp.674]
3. Gardner B, Sheals K, Wardle J, McGowan L. Putting habit into practice, and practice into habit: a process evaluation and exploration of the acceptability of a habit-based dietary behaviour change intervention. Int J Behav Nutr Phys Act 2014 Oct 30;11:135 [FREE Full text] [doi: 10.1186/s12966-014-0135-7] [Medline: 25359007]

4. Nederkoorn C, Houben K, Hofmann W, Roefs A, Jansen A. Control yourself or just eat what you like? Weight gain over a year is predicted by an interactive effect of response inhibition and implicit preference for snack foods. Health Psychol 2010 Jul;29(4):389-393. [doi: 10.1037/a0019921] [Medline: 20658826]
5. Lally P, Chipperfield A, Wardle J. Healthy habits: efficacy of simple advice on weight control based on a habit-formation model. Int J Obes (Lond) 2008 Apr;32(4):700-707. [doi: 10.1038/sj.ijo.0803771] [Medline: 18071344]
6. Gardner B. A review and analysis of the use of 'habit' in understanding, predicting and influencing health-related behaviour. Health Psychol Rev 2015;9(3):277-295 [FREE Full text] [doi: 10.1080/17437199.2013.876238] [Medline: 25207647]
7. Beeken RJ, Croker H, Morris S, Leurent B, Omar R, Nazareth I, et al. Study protocol for the 10 Top Tips (10TT) trial: randomised controlled trial of habit-based advice for weight control in general practice. BMC Public Health 2012 Aug 16;12:667 [FREE Full text] [doi: 10.1186/1471-2458-12-667] [Medline: 22898059]
8. Cleo G, Glasziou P, Beller E, Isenring E, Thomas R. Habit-based interventions for weight loss maintenance in adults with overweight and obesity: a randomized controlled trial. Int J Obes (Lond) 2019 Feb;43(2):374-383. [doi: 10.1038/s41366-018-0067-4] [Medline: 29686382]
9. Kliemann N, Vickerstaff V, Croker H, Johnson F, Nazareth I, Beeken RJ. The role of self-regulatory skills and automaticity on the effectiveness of a brief weight loss habit-based intervention: secondary analysis of the 10 top tips randomised trial. Int J Behav Nutr Phys Act 2017 Dec 05;14(1):119 [FREE Full text] [doi: 10.1186/s12966-017-0578-8] [Medline: 28870208]
10. Flores MG, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
11. Schippers M, Adam PC, Smolenski DJ, Wong HT, de Wit JB. A meta-analysis of overall effects of weight loss interventions delivered via mobile phones and effect size differences according to delivery mode, personal contact, and intervention intensity and duration. Obes Rev 2017 Dec;18(4):450-459. [doi: 10.1111/obr.12492] [Medline: 28187246]
12. Carter MC, Burley VJ, Nykjaer C, Cade JE. Adherence to a smartphone application for weight loss compared to website and paper diary: pilot randomized controlled trial. J Med Internet Res 2013;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
13. Norman GJ, Kolodziejczyk JK, Adams MA, Patrick K, Marshall SJ. Fruit and vegetable intake and eating behaviors mediate the effect of a randomized text-message based weight loss program. Prev Med 2013 Jan;56(1):3-7 [FREE Full text] [doi: 10.1016/j.ypmed.2012.10.012] [Medline: 23085329]
14. Lange D, Richert J, Koring M, Knoll N, Schwarzer R, Lippke S. Self-regulation prompts can increase fruit consumption: a one-hour randomised controlled online trial. Psychol Health 2013;28(5):533-545. [doi: 10.1080/08870446.2012.751107] [Medline: 23282217]
15. Poddar KH, Hosig KW, Anderson-Bill ES, Nickols-Richardson SM, Duncan SE. Dairy intake and related self-regulation improved in college students using online nutrition education. J Acad Nutr Diet 2012 Dec;112(12):1976-1986. [doi: 10.1016/j.jand.2012.07.026] [Medline: 23044035]
16. Kattelmann KK, Bredbenner CB, White AA, Greene GW, Hoerr SL, Kidd T, et al. The effects of young adults eating and active for health (YEAH): a theory-based web-delivered intervention. J Nutr Educ Behav 2014;46(6):S27-S41. [doi: 10.1016/j.jneb.2014.08.007] [Medline: 25457733]
17. Crane MM, Ward DS, Lutes LD, Bowling JM, Tate DF. Theoretical and behavioral mediators of a weight loss intervention for men. Ann Behav Med 2016 Dec;50(3):460-470 [FREE Full text] [doi: 10.1007/s12160-016-9774-z] [Medline: 26842133]
18. Klasnja P, Consolvo S, Pratt W. How to Evaluate Technologies for Health Behavior Change in HCI Research. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. 2011 May 07 Presented at: CHI '11; May 07-12, 2011; Vancouver, Canada p. 3063-3072. [doi: 10.1145/1978942.1979396]
19. Thomas JG, Bond DS. Review of innovations in digital health technology to promote weight control. Curr Diab Rep 2014;14(5):485. [doi: 10.1007/s11892-014-0485-1] [Medline: 24664797]
20. DiFilippo KN, Huang W, Andrade JE, Chapman-Novakofski KM. The use of mobile apps to improve nutrition outcomes: a systematic literature review. J Telemed Telecare 2015 Jul;21(5):243-253. [doi: 10.1177/1357633X15572203] [Medline: 25680388]
21. Knäuper B, Pillay R, Lacaille J, McCollam A, Kelso E. Replacing craving imagery with alternative pleasant imagery reduces craving intensity. Appetite 2011 Aug;57(1):173-178. [doi: 10.1016/j.appet.2011.04.021] [Medline: 21569806]
22. Gollwitzer P, Sheeran P. Implementation intentions and goal achievement: a meta-analysis of effects and processes. Adv Exp Soc Psychol 2006 May 07;38:69-119. [doi: 10.1016/S0065-2601(06)38002-1]
23. Verplanken B, Wood W. Interventions to break and create consumer habits. J Public Policy Mark 2006 Mar;25(1):90-103. [doi: 10.1509/jppm.25.1.90]
24. Kakoschke N, Kemps E, Tiggemann M. Attentional bias modification encourages healthy eating. Eat Behav 2014 Jan;15(1):120-124. [doi: 10.1016/j.eatbeh.2013.11.001] [Medline: 24411764]
25. Stawarz K, Cox A, Blandford A. Beyond self-tracking and reminders: designing smartphone apps that support habit formation. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015 Apr 18 Presented at: CHI '15; April 18-23, 2015; Seul, Republic of Korea p. 2653-2662. [doi: 10.1145/2702123.2702230]

26. Smith A. Pew Research Center. 2019. Smartphone ownership - 2013 update URL: http://www.pewinternet.org/2013/06/05/smartphone-ownership-2013 [WebCite Cache ID 75mh8f7ox]
27. Saghaei M. An overview of randomization and minimization programs for randomized clinical trials. J Med Signals Sens 2011 Jan;1(1):55-61 [FREE Full text] [Medline: 22606659]
28. Whitehead AL, Julious SA, Cooper CL, Campbell MJ. Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. Stat Methods Med Res 2016 Jun;25(3):1057-1073 [FREE Full text] [doi: 10.1177/0962280215588241] [Medline: 26092476]
29. Kliemann N, Beeken RJ, Wardle J, Johnson F. Development and validation of the Self-Regulation of Eating Behaviour Questionnaire for adults. Int J Behav Nutr Phys Act 2016 Aug 02;13:87 [FREE Full text] [doi: 10.1186/s12966-016-0414-6] [Medline: 27484457]
30. Roe L, Strong C, Whiteside C, Neil A, Mant D. Dietary intervention in primary care: validity of the DINE method for diet assessment. Fam Pract 1994 Dec;11(4):375-381. [Medline: 7895964]
31. Cappuccio FP, Rink E, Perkins-Porras L, McKay C, Hilton S, Steptoe A. Estimation of fruit and vegetable intake using a two-item dietary questionnaire: a potential tool for primary health care workers. Nutr Metab Cardiovasc Dis 2003 Feb;13(1):12-19. [Medline: 12772433]
32. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. Br Med J 1999 Sep 11;319(7211):670-674 [FREE Full text] [Medline: 10480822]
33. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
34. Pagoto S, Tulu B, Agu E, Waring ME, Oleski JL, Jake-Schoffman DE. Using the habit app for weight loss problem solving: development and feasibility study. JMIR Mhealth Uhealth 2018 Jun 20;6(6):e145 [FREE Full text] [doi: 10.2196/mhealth.9801] [Medline: 29925496]
35. Perski O, Blandford A, West R, Michie S. Conceptualising engagement with digital behaviour change interventions: a systematic review using principles from critical interpretive synthesis. Transl Behav Med 2017 Dec;7(2):254-267 [FREE Full text] [doi: 10.1007/s13142-016-0453-1] [Medline: 27966189]
36. Localytics. 2012. App retention increasing: iPhone ahead of Android URL: http://info.localytics.com/blog/app-user-loyalty-increasing-ios-beats-android [accessed 2019-01-29] [WebCite Cache ID 75mhZyjbP]
37. Tang J, Abraham C, Stamp E, Greaves C. How can weight-loss app designers' best engage and support users? A qualitative investigation. Br J Health Psychol 2015 Feb;20(1):151-171. [doi: 10.1111/bjhp.12114] [Medline: 25130682]
38. Pagoto SL, Schneider KL, Oleski JL, Luciani JM, Bodenlos JS, Whited MC. Male inclusion in randomized controlled trials of lifestyle weight loss interventions. Obesity (Silver Spring) 2012 Jun;20(6):1234-1239. [doi: 10.1038/oby.2011.140] [Medline: 21633403]
39. Ofcom. 2016. The Communications Market Report URL: https://www.ofcom.org.uk/__data/assets/pdf_file/0024/26826/cmr_uk_2016.pdf [WebCite Cache ID 75mhy8UjZ]
40. Ernsting C, Dombrowski SU, Oedekoven M, O Sullivan JL, Kanzler M, Kuhlmey A, et al. Using smartphones and health apps to change and manage health behaviors: a population-based survey. J Med Internet Res 2017 Dec 05;19(4):e101 [FREE Full text] [doi: 10.2196/jmir.6838] [Medline: 28381394]
41. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: a meta-regression. Health Psychol 2009 Nov;28(6):690-701. [doi: 10.1037/a0016136] [Medline: 19916637]
42. Dombrowski SU, Knittle K, Avenell A, Araújo-Soares V, Sniehotta FF. Long term maintenance of weight loss with non-surgical interventions in obese adults: systematic review and meta-analyses of randomised controlled trials. Br Med J 2014;348:g2646 [FREE Full text] [Medline: 25134100]
43. Safran NJ, Madar Z, Shahar DR. The impact of a web-based app (eBalance) in promoting healthy lifestyles: randomized controlled trial. J Med Internet Res 2015;17(3):e56 [FREE Full text] [doi: 10.2196/jmir.3682] [Medline: 25732936]
44. Beeken RJ, Leurent B, Vickerstaff V, Wilson R, Croker H, Morris S, et al. A brief intervention for weight control based on habit-formation theory delivered through primary care: results from a randomised controlled trial. Int J Obes (Lond) 2017 Dec;41(2):246-254 [FREE Full text] [doi: 10.1038/ijo.2016.206] [Medline: 27867204]
45. Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. Am J Prev Med 2007 May;32(5 Suppl):S112-S118 [FREE Full text] [doi: 10.1016/j.amepre.2007.01.022] [Medline: 17466815]
46. Walton J. Dietary assessment methodology for nutritional assessment: a practical approach. Top Clin Nutr 2015;30(1):33-46. [doi: 10.1097/TIN.0000000000000018]
47. Bruening M, van Woerden I, Todd M, Brennhofer S, Laska MN, Dunton G. A mobile ecological momentary assessment tool (devilSPARC) for nutrition and physical activity behaviors in college students: a validation study. J Med Internet Res 2016 Jul 27;18(7):e209 [FREE Full text] [doi: 10.2196/jmir.5969] [Medline: 27465701]

Abbreviations

10TT: 10 Top Tips

BMI: body mass index
DBCIs: digital behavior change interventions
RCT: randomized controlled trial
SS: sweet and salty snacks
SSD: sugary sweetened drinks

Edited by G Eysenbach; submitted 26.09.18; peer-reviewed by G Cleo, D Ben Neriah; comments to author 27.10.18; revised version received 29.01.19; accepted 10.02.19; published 10.05.19.

Please cite as:
Kliemann N, Croker H, Johnson F, Beeken RJ
Development of the Top Tips Habit-Based Weight Loss App and Preliminary Indications of Its Usage, Effectiveness, and Acceptability: Mixed-Methods Pilot Study
JMIR Mhealth Uhealth 2019;7(5):e12326
URL: https://mhealth.jmir.org/2019/5/e12326/
doi:10.2196/12326
PMID:31094352

Original Paper

Using Smartphone-Based Psychoeducation to Reduce Postnatal Depression Among First-Time Mothers: Randomized Controlled Trial

Ko Ling Chan[1], PhD; Wing Cheong Leung[2], MD; Agnes Tiwari[3], PhD; Ka Lun Or[4], PhD; Patrick Ip[5], MPH

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

[2]Department of Obstetrics & Gynaecology, Kwong Wah Hospital, Hong Kong, China (Hong Kong)

[3]School of Nursing, Hong Kong Sanatorium & Hospital Limited, Hong Kong, China (Hong Kong)

[4]Department of Industrial and Manufacturing Systems Engineering, The University of Hong Kong, Hong Kong, China (Hong Kong)

[5]Department of Paediatrics and Adolescent Medicine, The University of Hong Kong, Hong Kong, China (Hong Kong)

Corresponding Author:
Ko Ling Chan, PhD
Department of Applied Social Sciences
The Hong Kong Polytechnic University
Hung Hom
Kowloon
Hong Kong,
China (Hong Kong)
Phone: 852 2766 5709
Fax: 852 2773 6558
Email: koling.chan@polyu.edu.hk

Abstract

Background: Smartphone-based psychoeducation interventions may be a low-cost, user-friendly alternative to resource-consuming, face-to-face antenatal classes to educate expectant mothers.

Objective: This study aimed to empirically examine whether such an intervention would lead to reduced postnatal depression, anxiety, or stress and result in a better health-related quality of life.

Methods: A single-blind randomized controlled trial was conducted in Hong Kong. All first-time expectant mothers with less than 24 weeks of gestation remaining and attending the antenatal clinic at a public hospital were included. Participants were assigned to the intervention group or the control group by drawing lots. The lots, presented in sealed opaque envelopes, were randomly designated as “intervention” or “control” by stratified randomization. The intervention, a psychoeducational mobile app, was provided in addition to the treatment as usual (TAU) services from the hospital. Follow up with participants took place at 4 weeks postpartum. The primary outcome was the difference in the levels of antenatal and postnatal depression, assessed by the Edinburgh Postnatal Depression Scale (EPDS). The intention-to-treat approach was employed in the analyses.

Results: The final sample was 660 expectant mothers ($n_{intervention}$=330 and $n_{control}$=330). The mean difference in EPDS scores between the two groups was −0.65 (95% CI −1.29 to 0.00; *P*=.049) after adjusting for confounding factors. Associations were found between participation in the intervention and reduced depression, and attendance in TAU classes and increased stress levels.

Conclusions: The smartphone-based intervention plus TAU services was effective in reducing postnatal depression at 4 weeks postpartum compared with a control condition of TAU only, making this a cost-effective alternative to TAU education for expectant mothers. Limitations of the study included the short postpartum period after which the follow-up assessment was conducted and the inclusion of first-time mothers rather than all mothers.

Trial Registration: HKU Clinical Trials Registry HKUCTR-2024; http://www.hkuctr.com/Study/Show/34f62a2f6d594273a290491827206384

(JMIR Mhealth Uhealth 2019;7(5):e12794) doi:10.2196/12794

KEYWORDS

smartphone technology; postnatal depression; psychoeducation; randomized controlled trial

Introduction

Background

The potential positive effects of engaging parents in parental and infant health have increased research interest globally. Antenatal education classes, which provide an opportunity to teach expectant mothers about known complications and the support needed during pregnancy and childbirth, are effective for preparing expectant couples for parenthood. Research has revealed that antenatal classes are useful for reducing postnatal distress among first-time mothers (eg, at 6 weeks postpartum in a study by Matthey et al [1]), reducing the risks of postnatal depression (PND) [2], increasing mothers' awareness of pain relief techniques [3], and reducing parental stress and anxiety symptoms [4]. Among these negative health consequences, PND affects approximately 13% of the population as found in a meta-analysis by O'hara and Swain [5]. Review studies have observed that an antenatal intervention can be effective for preventing PND and, in turn, prevent future negative consequences for the mother, child, and family [6].

Despite the well-documented effects of such traditional antenatal interventions, the amount of face-to-face support available for expectant parents is often limited, owing to a lack of resources, social support, and personnel [7]. To face the increasing burdens of the health care system in the past decade, advanced computer-based technologies have been considered as effective means of providing support to health care consumers, given their popularity and easy accessibility [8]. More specifically, smartphone technology has been increasingly demonstrated as a useful tool to disseminate health information, by providing psychoeducation and/or timely interactions via short message service text messaging or two-way communication between patients and health professionals [9]. There has been evidence showing that therapist audio recordings and remote communication between patients and service providers offered in mobile apps are flexible, portable, and relatively low cost, making them an ideal choice for patients who are restricted by time and location constraints and cannot partake in traditionally used services [10]. Recent empirical research has also shown its effectiveness in psychoeducation, skills training, and two-way communication with health professionals via smartphone apps to promote well-being among mental health patients [1], and meta-analytic studies have also demonstrated that smartphone-based interventions may reduce depression and anxiety [11,12].

Objectives of the Study

Due to the previous successful application of smartphone technology in health care, the aim of the present study was to test the effectiveness of a smartphone-based intervention to educate expectant parents during pregnancy. A randomized controlled trial (RCT) was conducted to determine if a smartphone-based intervention can be a low-cost alternative to traditional face-to-face antenatal classes for educating expectant parents and promoting health outcomes. Compared with treatment as usual (TAU) where approximately 50% of the expectant parents could be enrolled in a nurse-led antenatal class while the remainder would receive information booklets in print form, a smartphone-based intervention can deliver all essential materials that would be taught in class. Materials comprised small video clips and short passages for parents to read or watch at their convenience. Parents could also interact with health professionals using the specifically designed platform whenever they had an inquiry. The intervention was expected to provide parenthood education without the time or space constraints. With better preparation, expectant parents were hypothesized to show lower levels of PND as well as postnatal anxiety and stress postintervention versus those who received only TAU or standard antenatal services in hospitals. In addition, this study also explored the overall effects of this smartphone-based intervention versus traditional TAU classes on promoting postnatal health and reducing maternal PND, anxiety, and stress.

Methods

Study Design

A single-blind RCT was conducted to compare an intervention group, where participants used a smartphone-based app for antenatal education, and a control TAU group that received antenatal services from Kwong Wah Hospital (KWH), a major public hospital in Hong Kong, during 2017 and 2018. This study obtained ethical approval from the Institutional Review Board of the Hospital Authority (Kowloon West Cluster) Research Ethics Committee and was registered under the Clinical Trials Registry of Clinical Trials Centre of the University of Hong Kong. No important changes to methods and outcomes were made after the commencement of the study.

Participants

First-time pregnant women with less than 24 weeks of gestation remaining and who attended the antenatal clinic at KWH were recruited to participate in the study. To ensure the representativeness of the sample, no limit was set on the ethnicity or nationality of the participants. Eligible participants were first-time expectant mothers receiving regular antenatal care services at KWH, were able to read and understand Chinese or English, and were willing to consent to the terms of the study. The nursing staff at KWH provided assistance in the identification of potential participants during the recruitment procedure. Expectant mothers were excluded if they were unable to give informed written consent or communicate with the interviewers. The participation of their partners was encouraged but was not a requirement for participation.

Randomization and Masking

Participants were assigned to the intervention or control group using lots in a ratio of 1:1. The lots, presented in sealed opaque envelopes, were randomly designated as "intervention" or "control" by stratified randomization, with random numbers generated from statistical software. Each participant received

a sealed envelope with information indicating the group they were assigned to. The recruitment and randomization procedures were conducted by different groups of researchers, and the outcome assessors were blinded to the allocation in each group.

Intervention and Control

The intervention involved the provision of access to a smartphone app, the iParent app (hereinafter referred to as the app), to the selected expectant mothers (n=330) between their first visit to the antenatal clinic and childbirth in addition to the TAU services they received at KWH. All materials presented in the app were equivalent to those offered in the face-to-face nurse-led antenatal TAU classes at KWH. For example, there were articles about nutrition, infant caring, and vaccine injections for infants, as well as videos demonstrating what expectant mothers might face when delivering their baby. The major difference was the mode of information delivery: information was delivered off-site, organized by topic to increase usability, and open to access at any time and anywhere for assigned users. Furthermore, a platform within the app allowed users to ask questions related to pregnancy, childbirth, and infant health and care. All questions were answered by obstetricians via private, direct messages within the app and then shared in the Frequently Asked Questions module of the app, if permitted by the user and after personal identifiable information was removed. This function dealt only with regular and nonemergent consultations.

The app was designed by the authors and was tested and updated through various trials. For the intervention group, they were given a specific login name and password (which participants were encouraged to change when using the app for the first time) for registration and were permitted free access to the app. Use parameters such as the frequencies of logins and the time spent using the app were recorded. Prompts would be made via emails if the participants in the intervention group did not log in to the app during the study period.

The control group (n=330) received standard services provided by KWH only. They could enroll in the 4-session, nurse-led antenatal classes. The control group participants were not allowed to access the app.

Procedure

At their first visit to the antenatal clinic at KWH, eligible participants were informed about the study details and participant rights in a private room in the hospital. In particular, they were ensured that refusal to participate would not affect any service they received at KWH. Eligible expectant mothers who agreed to participate provided written consent and completed the baseline T1 survey that assessed their antenatal depression, anxiety and stress levels, health-related quality of life (QoL), and demographic characteristics. Participants also indicated their preferred methods of contact at later stages of the study.

Around 4 weeks after childbirth, participants were contacted for the follow-up T2 survey. The follow-up T2 survey included scales assessing their PND, anxiety and stress levels, health-related QoL, and usage of the app and antenatal classes provided by the hospital.

Although there was no potential risk associated with the intervention, medical and health professionals of the antenatal clinic at KWH were available for assistance during the RCT.

Outcomes

The primary outcome was the difference between the level of depression before and after the RCT (ie, depression during pregnancy and PND) among the expectant mothers. The validated Chinese version of the 10-item Edinburgh Postnatal Depression Scale (EPDS) was used to detect depressive symptoms [13]. Total scores ranged from 0 to 30, with a higher score indicating more severe depression levels. The Cronbach alphas of the Chinese EPDS in this study were .85 at baseline survey and .84 at follow up, indicating a good reliability.

Secondary outcomes included differences in anxiety levels, stress levels, and health-related QoL before and after the RCT. Anxiety and stress levels were assessed with the Anxiety and Stress subscales of the Depression Anxiety Stress Scale (DASS) [14]. Each subscale has 7 items. Total scores ranged from 0 to 21, with higher scores indicating higher levels of anxiety or stress. The Cronbach alphas of the Anxiety subscale were .78 and .74 at baseline and follow up, respectively, whereas those of the Stress subscale were .85 and .83 at the 2 time points, respectively. Health-related QoL was measured by the 12-item Short Form Health Survey (SF-12) [15]. Scores on the SF-12 were computed as 2 separate composite scores: one for Physical Component Summary (PCS) and the other for Mental Component Summary (MCS). Both PCS and MCS ranged from 0 to 100, with higher scores indicating better health-related QoL. The reliability of SF-12 was good in this study, with Cronbach alphas of .87 and .85 at baseline and follow up, respectively.

We also recorded selected demographic characteristics and antenatal service utilization among the participants. These demographic characteristics included the following: age, education level, employment status, marital status, household income, and presence of diagnosed chronic health conditions or mental illness. The use of the app and other relevant services (eg, antenatal classes and other resources about pregnancy such as books and websites) was documented by self-reported responses from study participants (ie, through the use of yes/no questions).

Statistical Analysis

All statistical analyses were conducted and underscored by the intention-to-treat principle. All missing data were treated by the last observation carried forward imputation method. All statistical analyses were conducted with IBM SPSS 24.0 using a significance level of .05.

With reference to a previous psychological intervention for 6 weeks PND [16], a sample size of at least 276 participants was required for both the intervention and control groups in this study. The final sample in this study was 330 participants in each group.

The analysis of covariance (ANCOVA) was used to test the hypotheses of this study. In particular, the effectiveness of the intervention was first examined by comparing the primary

outcomes between groups, adjusted for demographic characteristics and the use of the app and relevant antenatal services. Similar analyses were also conducted on the secondary outcomes.

The effects of the intervention and the attendance of antenatal classes were then tested using the ANCOVA, which compared primary and secondary outcomes, adjusted for demographics, service usage, and baseline T1 scores.

This RCT was registered with the HKU Clinical Trials Registry (HKUCTR-2024).

Role of the Funding Source

The funder of this study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in this study and had final responsibility for the decision to submit for publication.

Results

In the recruitment stage, 803 first-time expectant mothers were assessed for eligibility and 660 were included for the baseline T1 survey and randomization (response rate=82.2%). Of the 660 participants, 330 were randomly assigned to the intervention group and 330 were allocated to the control group. At the follow-up T2 survey after the intervention, the retention rates were 66.1% for the intervention group and 68.2% for the control group. More details on the sampling procedures are presented in the Consolidated Standards of Reporting Trials flow diagram (see Figure 1). No visible harm or unintended effect was noted in the groups. There was no statistically significant difference in the demographic characteristics between the participants who completed the study and those who had withdrawn (all $P>.05$).

Demographic characteristics of the participants at baseline are shown in Table 1. All participants were Chinese. The mean ages of the intervention group and control group were 31.3 years (SD 4.6) and 31.2 years (SD 4.5), respectively. There was no significant between-group difference in any demographic characteristic measured in this study (all $P>.05$).

Table 2 presents the mean scores, standard deviation (SD), internal reliabilities, and results of the between-group comparisons of the primary and secondary outcomes (N=660). The mean EPDS score of the intervention group dropped from 7.3 (SD 4.6) to 5.3 (SD 4.4) after the RCT and that of the control group dropped from 7.2 (SD 4.6) to 5.9 (SD 4.7) after receiving TAU care. The mean difference between groups was −0.65 (95% CI −1.29 to 0.00; $P=.049$), which was a significant difference after adjusting for the baseline T1 mean scores and demographic characteristics.

The results in this study did not show significant between-group differences in the secondary outcomes tested (ie, the mean scores of anxiety, stress, and health-related QoL before and after the RCT). Both of the mean scores on the Anxiety subscale and Stress subscale of the DASS dropped in both groups from baseline to follow up, although the differences between groups were not significant after the adjustment of the baseline scores and demographic factors ($P=.94$ and $P=.74$, respectively). Nonsignificant mean differences between groups were also found in the SF-12 scores, although the 2 composite scores increased from baseline to post-RCT follow up in both groups ($P=.54$ and $P=.81$, respectively).

Figure 1. Consolidated Standards of Reporting Trials flow diagram of the randomized controlled trial.

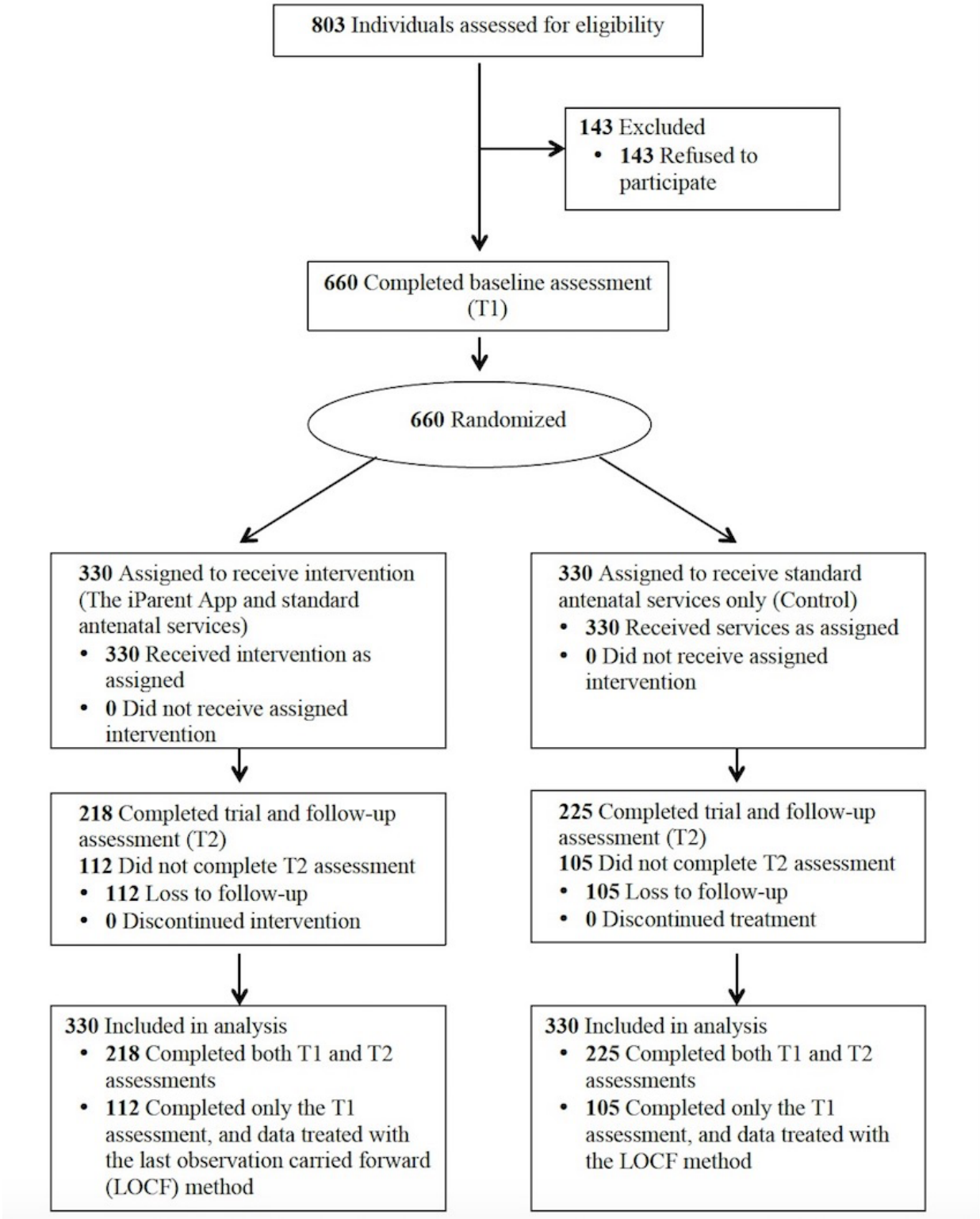


Table 1. Selected demographic characteristics of the mothers.

Characteristics	Intervention (n=330)	Control (n=330)	*P* value
Age at baseline (years), mean (SD)	31.3 (4.6)	31.2 (4.5)	.84
Educational attainment, n (%)			.20
Junior high school or below	27 (8.2)	41 (12.4)	.20
High school	110 (33.3)	109 (33.0)	.20
University, higher institution, or above	191 (57.9)	179 (54.2)	.20
Missing	2 (0.6)	1 (0.3)	.20
Employed, n (%)	268 (81.2)	256 (77.6)	.28
Marital status, n (%)			.21
Married	280 (84.8)	282 (85.5)	.21
Living or cohabiting as couple	12 (3.6)	20 (6.1)	.21
Single and never married	28 (8.5)	23 (7.0)	.21
Separated, divorced, or widowed	1 (0.3)	2 (0.6)	.21
Missing	9 (2.7)	3 (0.9)	.21
Diagnosed with chronic health conditions[a], n (%)	17 (5.2)	24 (7.3)	.40
Diagnosed with mental illness, n (%)	2 (0.6)	0 (0)	.24
Household income (monthly), n (%)			.98
HKD[b] 15,000 (approximately US $1911) or below	46 (13.9)	48 (14.5)	.98
HKD 15,001 to HKD 39,999 (approximately US $1911-$5097)	123 (37.3)	121 (36.7)	.98
HKD 40,000 (approximately US $5097) or above	148 (44.8)	146 (44.2)	.98
Missing	13 (3.9)	15 (4.5)	.98
Receiving social security assistance	16 (4.8)	11 (3.3)	.22

[a]A condition that is long-lasting or recurrent, for example, diabetes mellitus, chronic obstructive pulmonary disease, asthma, heart failure, cancer, arthritis, and chronic renal failure.

[b]HKD: Hong Kong Dollar.

Table 2. Mean (SD) values of the outcome measures pre- (baseline, T1) and postintervention (follow up, T2) and the between-group mean differences.

Outcome	Cronbach alpha	Intervention (n=330), mean (SE)		Control (n=330), mean (SE)		Adjusted mean difference between groups[a] (95% CI)	*P* value
		Baseline (T1)	Follow up (T2)	Baseline (T1)	Follow up (T2)		
Edinburgh Postnatal Depression Scale[a,b]	.84-.85	7.3 (4.6)	5.3 (4.4)	7.2 (4.6)	5.9 (4.7)	−0.65 (−1.29 to 0.00)	.049
DASS[c] (Anxiety)[d]	.74-.78	3.1 (2.7)	1.9 (2.1)	2.8 (2.6)	1.8 (2.3)	0.01 (−0.30 to 0.32)	.94
DASS (Stress)[e]	.83-.85	3.9 (3.5)	3.0 (3.1)	3.8 (3.4)	2.9 (3.1)	0.07 (−0.35 to 0.50)	.74
SF-12[f,g] (Physical Component Summary)	.85-.87	45.5 (7.0)	48.8 (6.9)	46.4 (7.0)	48.8 (7.2)	0.32 (−0.71 to 1.34)	.54
SF-12 (Mental Component Summary)	.85-.87	48.6 (8.9)	51.2 (8.4)	48.2 (8.9)	51.2 (9.0)	−0.15 (−1.39 to 1.09)	.81

[a]Estimated between-group difference (intervention – control) after giving birth (follow up, T2), adjusted for baseline (T1) values.

[b]Possible ranges from 0 to 30, with higher scores indicating more severe depressive symptoms.

[c]DASS: Depression Anxiety Stress Scale.

[d]Range: 0 to 21 (with higher scores indicating more severe anxiety symptoms).

[e]Range: 0 to 21 (with higher scores indicating more severe stress symptoms).

[f]SF-12: 12-Item Short Form Health Survey.

[g]Range: 0 to 100 (with higher scores indicating better health-related quality of life).

Table 3 summarizes the findings on the effects of participation in the intervention and attendance in the nurse-led antenatal classes on the primary and secondary outcomes (N=443). Among these 443 participants who completed the whole study, 218 were in the intervention group and 225 were in the control group. Concerning the attendance in the antenatal classes, 272 of the 443 participants had enrolled in the classes whereas the remaining 171 had not. The mean scores of the EPDS were significantly different between the intervention group and control group (estimated marginal means for the intervention group was 4.2 [SE 0.3]; estimated marginal means for the control group was 5.1 [SE 0.3]; *P*=.01), with the follow-up scores of EPDS lower in the intervention group. Relevant results are noted in Figure 2. On the contrary, those of the Stress subscale of the DASS were significantly different between the participants who attended the nurse-led antenatal classes versus those who did not (estimated marginal means was 2.8 [SE 0.2] for those who attended and 1.9 [SE 0.2] for those who did not; *P*=.04). These findings demonstrated that participants who attended TAU classes reported higher levels of stress than those who did not attend the classes at follow up.

These findings indicated that participation in the intervention was significantly associated with decreased depression levels, whereas attending TAU nurse-led antenatal classes was significantly associated with increased stress levels after the study period.

Table 3. Mean (SD) values of the outcome measures pre- (baseline, T1) and postintervention (follow up, T2), and between-group differences between intervention and antenatal class participation.

Outcome and time of assessment	Usage of intervention, mean (SE)		Usage of antenatal classes, mean (SE)		*P* value[a] (intervention)	*P* value[a] (antenatal classes)
	Intervention (n=218)	Control (n=225)	Attended (n=272)	Not attended (n=171)		
Edinburgh Postnatal Depression Scale[b]						
Baseline (T1)	7.3 (4.7)	7.0 (4.4)	7.2 (4.6)	7.1 (4.4)	.01	.47
Follow up (T2)	4.3 (3.9)	5.2 (4.3)	5.0 (4.2)	4.3 (4.0)	—[c]	—
Estimated marginal means (SE)	4.2 (0.3)	5.1 (0.3)	5.0 (0.3)	4.3 (0.3)	—	—
DASS[d] (Anxiety)[e]						
Baseline (T1)	3.2 (2.9)	2.7 (2.5)	3.1 (2.8)	2.7 (2.6)	.30	.06
Follow up (T2)	1.4 (1.7)	1.1 (1.6)	1.4 (1.8)	1.0 (1.3)	—	—
Estimated marginal means (SE)	1.3 (0.1)	1.1 (0.1)	1.4 (0.1)	1.0 (0.1)	—	—
DASS (Stress)[f]						
Baseline (T1)	3.9 (3.4)	3.7 (3.4)	3.9 (3.4)	3.8 (3.4)	.99	.04
Follow up (T2)	2.5 (2.7)	2.4 (2.8)	2.8 (2.9)	1.9 (2.5)	—	—
Estimated marginal means (SE)	2.4 (0.2)	2.4 (0.2)	2.8 (0.2)	1.9 (0.2)	—	—
SF-12[g] (Physical Component Summary)[h]						
Baseline (T1)	45.9 (6.9)	46.6 (7.2)	46.3 (7.2)	46.2 (6.9)	.18	.81
Follow up (T2)	50.9 (5.9)	50.0 (7.1)	50.6 (6.5)	50.1 (6.6)	—	—
Estimated marginal means (SE)	50.8 (0.5)	50.0 (0.4)	50.6 (0.4)	50.2 (0.5)	—	—
SF-12 (Mental Component Summary)[h]						
Baseline (T1)	48.3 (9.0)	48.8 (8.5)	48.9 (9.0)	48.1 (8.3)	.25	.63
Follow up (T2)	52.1 (8.2)	53.2 (8.0)	52.0 (8.1)	53.7 (7.9)	—	—
Estimated marginal means (SE)	52.3 (0.6)	53.3 (0.5)	52.0 (0.5)	53.6 (0.6)	—	—

[a]*P* values obtained with analysis of covariance, with baseline (T1) scores being adjusted.

[b]Range: 0 to 30 (with higher scores indicating more severe depressive symptoms).

[c]Not applicable.

[d]DASS: Depression Anxiety Stress Scales.

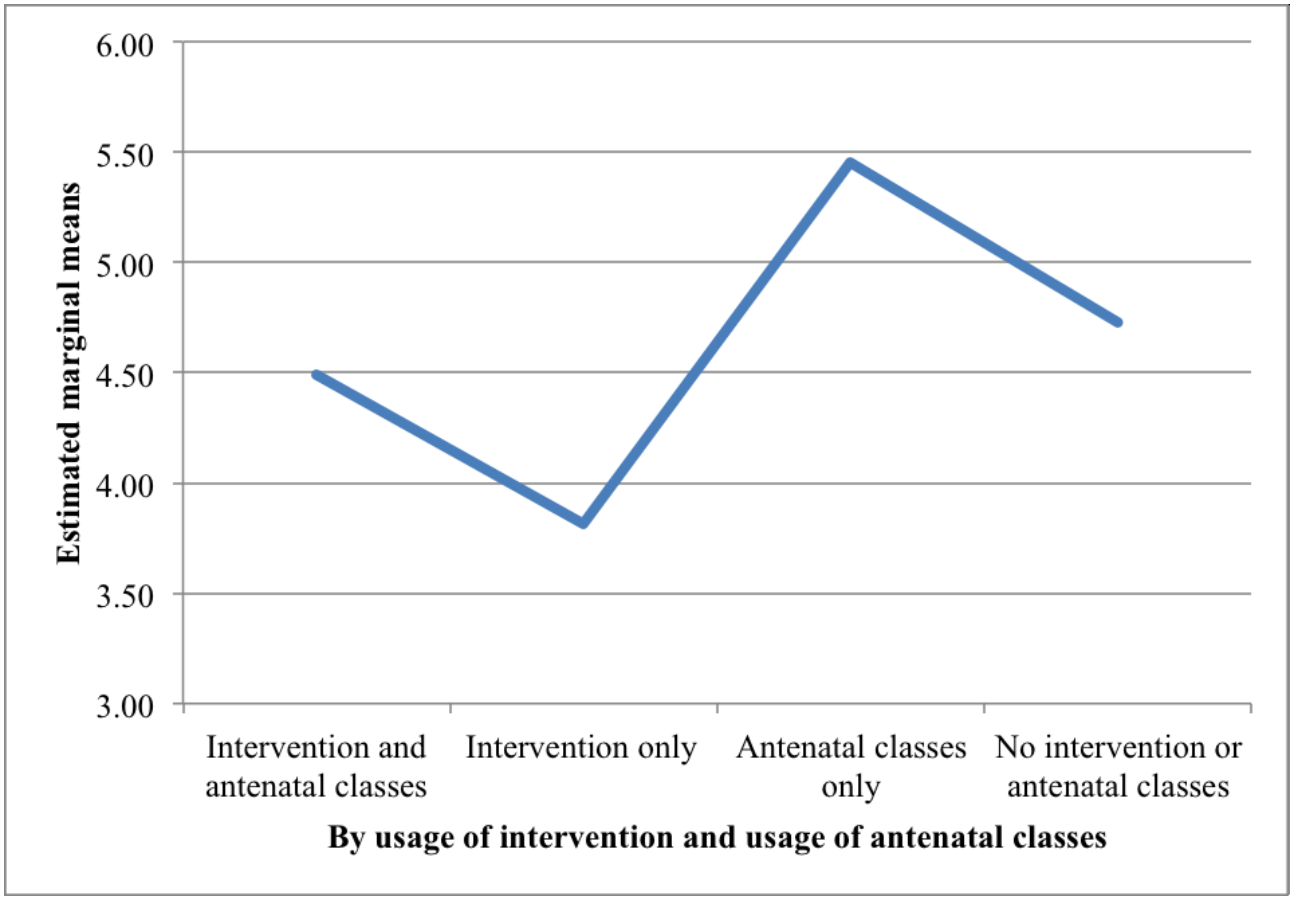
[e]Range: 0 to 21 (with higher scores indicating more severe anxiety symptoms).

[f]Range: 0 to 21 (with higher scores indicating more severe stress symptoms).

[g]SF-12: 12-Item Short Form Health Survey.

[h]Range: 0 to 100 (with higher scores indicating better health-related quality of life).

Figure 2. Estimated marginal means of the Edinburgh Postnatal Depression Scale scores postintervention (follow up, T2), adjusted for baseline (T1) scores by usage of intervention and antenatal classes.



Discussion

Our findings provide empirical support for the relative effectiveness of a smartphone-based intervention in reducing PND among first-time mothers at 4 weeks postpartum. Compared with those who received only standard TAU services, mothers who had received the intervention and used the iParent app in addition to standard care reported significantly lower levels of PND when baseline depression and other confounding factors were adjusted for. This finding compares favorably with previous research showing that computer- or smartphone-based technologies are effective in promoting user health [9-12,16,17]. More specifically, this study sheds light on the potential generalizability of the relative effectiveness of these technologies in reducing depression among patients' diagnoses of mental illness to other health care service users such as pregnant women. Such smartphone-based technologies do not only help manage and/or monitor depressive symptoms among diagnosed patients but might also be useful for preventing or reducing depression among the general population.

Our findings also show that the intervention, more so than the traditional face-to-face TAU classes, was significantly associated with reduced levels of PND when adjusted for baseline depression during pregnancy. Through the iParent platform, expectant parents had essential information on antenatal care, postnatal care, infant care, and child protection at their fingertips. The portability and mobility of the app and the popularity of its use may indeed facilitate the engagement of parents in parenthood preparation with less time and space constraints [18], which may in turn promote parental health and a successful pregnancy, since asking questions or discussions with health professionals about relevant issues can be done easily through the app [7]. Another possible reason for the effectiveness of the smartphone-based intervention may be its provision of unlimited access to essential information; compared with the traditional face-to-face TAU classes that are normally one-off events, the app allowed for continuous access to information. A third possible mechanism for the effectiveness of the intervention might be the increase of a sense of connection among the mothers when they were using the app to interact with the health professionals. Normally, new mothers, who need time for physical recovery, might feel isolated for the first few weeks after giving birth, but iParent could provide them an opportunity to stay connected with others. This sense of connection might increase the level of perceived social support from professionals and even from peers (ie, other mother users of the app who could ask questions and share experiences in the app), which might in turn reduce PND [19].

The encouraging evidence found herein supports the relative effectiveness of iParent and warrants the potential use of other advanced technologies as viable low-cost alternatives to face-to-face TAU services in educating and monitoring the nonemergent cases of health care service users to reduce the burden to the health care system and relevant professionals.

This study observed reductions in postnatal anxiety and stress levels after the intervention, although the differences were not

statistically significant. This diverges from the results shown in a meta-analysis by Firth et al [12]. Our preliminary results indicated that the smartphone-based intervention might not be sufficiently effective to add extra value on top of the standard TAU services in reducing anxiety and stress among first-time mothers. However, further empirical evidence is needed before establishing a conclusion. Indeed, researchers have suggested that the design of the RCT itself might affect the results obtained. For example, effect sizes of the smartphone-based interventions are found to be greater when compared with waitlist controls than active controls [12], of which the latter condition was employed in this RCT.

A surprising finding in this study was the association noted between attendance of face-to-face antenatal classes and levels of postnatal stress. After adjustment for baseline stress, first-time mothers who had attended the classes reported higher levels of postnatal stress than did those who did not attend any. This discovery adds some evidence to the claim that the effects of TAU antenatal classes might not be universal across all populations. For example, a Spanish study found increased levels of stress during childbirth among immigrants but decreased levels among a local sample after attending antenatal classes [20]. It was suggested that the group receiving antenatal classes may have increased awareness of problems and difficulties that may arise during the postnatal period, making them more likely to show higher stress levels [21]. Other researchers have suggested that differences across individuals might be due to gender, age, education level, and the different expectations about participating in the classes [22]. Future studies should assess expectant parents' expectations on the education they receive either in such classes or through the intervention as a covariate to further explore the underlying effect of the program on parental stress.

To our knowledge, this is the first RCT study of a smartphone-based psychoeducation intervention for first-time expectant mothers. However, the study has several limitations, and the present findings should not be generalized unequivocally. First, these outcomes were measured up to 4 weeks postpartum, and outcomes during childbirth were not assessed. It is possible that these effects might change across time, so future studies should extend the study period to include more precise longitudinal tracking of such outcomes. Second, this study only analyzed health outcomes among first-time mothers. Health outcomes of fathers and infants are also useful indicators of the effectiveness of these kinds of interventions for pregnancy and childbirth. Thus, future research should also include an assessment of paternal and child outcomes. Third, similar to studies involving self-reported data, our study is subjective to reporting bias and recall error. Future research could employ more and/or complementary objective assessments such as journals, secondary clinical data, and obstetricians' reports. Finally, this study only employed first-time mothers residing in Hong Kong, and previous local research has suggested the possibility of ethnic differences in the effectiveness of antenatal education on health outcomes [20]. Future research may extend such RCT initiatives to other populations in the world to offset cultural bias.

Using an RCT, the conclusions of this study suggest that smartphone-based psychoeducation plus standard antenatal services at hospitals in Hong Kong can be effective in reducing PND at 4 weeks postpartum among first-time mothers when compared with a control group receiving standard services only.

Acknowledgments

The authors appreciate the contribution of the team of midwives at KWH for assisting with data collection. This study was funded by the Health and Medical Research Fund, the Government of the Hong Kong Special Administrative Region (project #15161101).

Authors' Contributions

All authors contributed to the design and conceptualization of the study. KLC oversaw the whole study process, interpreted the results, and prepared the manuscript with input from all other authors. WCL provided assistance on the recruitment of participants in the hospital. AT contributed to the design of the RCT. KLO contributed to the analyses and interpretation of the results. PI contributed to the design of the RCT and the interpretation of the results.

Conflicts of Interest

None declared.

References

1. Matthey S, Kavanagh DJ, Howie P, Barnett B, Charles M. Prevention of postnatal distress or depression: an evaluation of an intervention at preparation for parenthood classes. J Affect Disord 2004 Apr;79(1-3):113-126. [doi: 10.1016/S0165-0327(02)00362-2] [Medline: 15023486]
2. Bowen A, Baetz M, Schwartz L, Balbuena L, Muhajarine N. Antenatal group therapy improves worry and depression symptoms. Isr J Psychiatry Relat Sci 2014;51(3):226-231 [FREE Full text] [Medline: 25618288]
3. Fabian HM, Rådestad IJ, Waldenström U. Childbirth and parenthood education classes in Sweden. Women's opinion and possible outcomes. Acta Obstet Gynecol Scand 2005 May;84(5):436-443. [doi: 10.1111/j.0001-6349.2005.00732.x] [Medline: 15842207]

4. Milgrom J, Schembri C, Ericksen J, Ross J, Gemmill AW. Towards parenthood: an antenatal intervention to reduce depression, anxiety and parenting difficulties. J Affect Disord 2011 May;130(3):385-394. [doi: 10.1016/j.jad.2010.10.045] [Medline: 21112641]
5. O'hara MW, Swain A. Rates and risk of postpartum depression—a meta-analysis. Int Rev Psychiatry 2009 Jul 11;8(1):37-54. [doi: 10.3109/09540269609037816]
6. Dennis CL. Psychosocial and psychological interventions for prevention of postnatal depression: systematic review. Br Med J 2005 Jul 2;331(7507):15. [doi: 10.1136/bmj.331.7507.15] [Medline: 15994688]
7. Renkert S, Nutbeam D. Opportunities to improve maternal health literacy through antenatal education: an exploratory study. Health Promot Int 2001 Dec;16(4):381-388. [doi: 10.1093/heapro/16.4.381] [Medline: 11733456]
8. Free C, Phillips G, Galli L, Watson L, Felix L, Edwards P, et al. The effectiveness of mobile-health technology-based health behaviour change or disease management interventions for health care consumers: a systematic review. PLoS Med 2013;10(1):e1001362 [FREE Full text] [doi: 10.1371/journal.pmed.1001362] [Medline: 23349621]
9. Kim H, Jeong H. A nurse short message service by cellular phone in type-2 diabetic patients for six months. J Clin Nurs 2007 Jun;16(6):1082-1087. [doi: 10.1111/j.1365-2702.2007.01698.x] [Medline: 17518883]
10. Luxton DD, McCann RA, Bush NE, Mishkind MC, Reger GM. mHealth for mental health: integrating smartphone technology in behavioral healthcare. Prof Psychol Res Pr 2011;42(6):505-512. [doi: 10.1037/a0024485]
11. Firth J, Torous J, Nicholas J, Carney R, Pratap A, Rosenbaum S, et al. The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. World Psychiatry 2017 Oct;16(3):287-298 [FREE Full text] [doi: 10.1002/wps.20472] [Medline: 28941113]
12. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. J Affect Disord 2017 Aug 15;218:15-22 [FREE Full text] [doi: 10.1016/j.jad.2017.04.046] [Medline: 28456072]
13. Lee DT, Yip SK, Chiu HF, Leung TY, Chan KP, Chau IO, et al. Detecting postnatal depression in Chinese women. Validation of the Chinese version of the Edinburgh Postnatal Depression Scale. Br J Psychiatry 1998 May;172:433-437. [doi: 10.1192/bjp.172.5.433] [Medline: 9747407]
14. Taouk M, Lovibond P, Laube R. Psychometric Properties of a Chinese Version of the 21-Item Depression Anxiety Stress Scales (DASS 21). Sydney: Cumberland Hospital; 2001.
15. Lam CL, Tse EY, Gandek B. Is the standard SF-12 health survey valid and equivalent for a Chinese population? Qual Life Res 2005 Mar;14(2):539-547. [doi: 10.1007/s11136-004-0704-3] [Medline: 15892443]
16. Hind J, Sibbald SL. Smartphone applications for mental health—a rapid review. WURJHNS 2015:1-9. [doi: 10.5206/wurjhns.2014-15.16]
17. Or C, Tao D. Usability study of a computer-based self-management system for older adults with chronic diseases. JMIR Res Protoc 2012;1(2):e13 [FREE Full text] [doi: 10.2196/resprot.2184] [Medline: 23612015]
18. Office of the Communications Authority. Key communications statistics URL: https://www.ofca.gov.hk/en/media_focus/data_statistics/key_stat/ [WebCite Cache ID 73mDQI68O]
19. Pfeiffer P, Heisler M, Piette J, Rogers M, Valenstein M. Efficacy of peer support interventions for depression: a meta-analysis. Gen Hosp Psychiatry 2011;33(1):29-36 [FREE Full text] [doi: 10.1016/j.genhosppsych.2010.10.002] [Medline: 21353125]
20. Artieta-Pinedo I, Paz-Pascual C, Grandes G, Remiro-Fernandezdegamboa G, Odriozola-Hermosilla I, Bacigalupe A, et al. The benefits of antenatal education for the childbirth process in Spain. Nurs Res 2010;59(3):194-202. [doi: 10.1097/NNR.0b013e3181dbbb4e] [Medline: 20421842]
21. Urizar GG, Muñoz RF. Impact of a prenatal cognitive-behavioral stress management intervention on salivary cortisol levels in low-income mothers and their infants. Psychoneuroendocrinology 2011 Nov;36(10):1480-1494 [FREE Full text] [doi: 10.1016/j.psyneuen.2011.04.002] [Medline: 21641117]
22. Ahldén I, Ahlehagen S, Dahlgren LO, Josefsson A. Parents' expectations about participating in antenatal parenthood education classes. J Perinat Educ 2012;21(1):11-17 [FREE Full text] [doi: 10.1891/1058-1243.21.1.11] [Medline: 23277726]

Abbreviations

ANCOVA: analysis of covariance
DASS: Depression Anxiety Stress Scale
EPDS: Edinburgh Postnatal Depression Scale
KWH: Kwong Wah Hospital
MCS: Mental Component Summary
PCS: Physical Component Summary
PND: postnatal depression
QoL: quality of life
RCT: randomized controlled trial
SF-12: 12-item Short Form Health Survey
TAU: treatment as usual

Edited by G Eysenbach; submitted 12.11.18; peer-reviewed by S Kita, A Torbjørnsen, A Pratap, M Madsen Thompson, S Kita, N Reeves; comments to author 14.02.19; revised version received 17.03.19; accepted 29.03.19; published 14.05.19.

Please cite as:
Chan KL, Leung WC, Tiwari A, Or KL, Ip P
Using Smartphone-Based Psychoeducation to Reduce Postnatal Depression Among First-Time Mothers: Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(5):e12794
URL: http://mhealth.jmir.org/2019/5/e12794/
doi:10.2196/12794
PMID:31094354

Original Paper

Validation of the Mobile App–Recorded Circadian Rhythm by a Digital Footprint

Yu-Hsuan Lin[1,2,3,4], PhD, MD; Bo-Yu Wong[1], BS; Yuan-Chien Pan[1,5], MS; Yu-Chuan Chiu[6], MD; Yang-Han Lee[7], PhD

[1]Institute of Population Health Sciences, National Health Research Institutes, Miaoli, Taiwan

[2]Department of Psychiatry, National Taiwan University Hospital, Taipei, Taiwan

[3]Department of Psychiatry, College of Medicine, National Taiwan University, Taipei, Taiwan

[4]Institute of Health Behaviors and Community Sciences, College of Public Health, National Taiwan University, Taipei, Taiwan

[5]Department of Psychology, National Taiwan University, Taipei, Taiwan

[6]Department of Psychiatry, MacKay Memorial Hospital, Taipei, Taiwan

[7]Department and Graduate School of Electrical Engineering, Tamkang University, New Taipei City, Taiwan

Corresponding Author:
Yu-Hsuan Lin, PhD, MD
Institute of Population Health Sciences
National Health Research Institutes
35 Keyan Road, Zhunan
Miaoli, 35053
Taiwan
Phone: 886 37 246166 ext 36383
Fax: 886 37 586 261
Email: yuhsuanlin@nhri.org.tw

Abstract

Background: Modern smartphone use is pervasive and could be an accessible method of evaluating the circadian rhythm and social jet lag via a mobile app.

Objective: This study aimed to validate the app-recorded sleep time with daily self-reports by examining the consistency of total sleep time (TST), as well as the timing of sleep onset and wake time, and to validate the app-recorded circadian rhythm with the corresponding 30-day self-reported midpoint of sleep and the consistency of social jetlag.

Methods: The mobile app, Rhythm, recorded parameters and these parameters were hypothesized to be used to infer a relative long-term pattern of the circadian rhythm. In total, 28 volunteers downloaded the app, and 30 days of automatically recorded data along with self-reported sleep measures were collected.

Results: No significant difference was noted between app-recorded and self-reported midpoint of sleep time and between app-recorded and self-reported social jetlag. The overall correlation coefficient of app-recorded and self-reported midpoint of sleep time was .87.

Conclusions: The circadian rhythm for 1 month, daily TST, and timing of sleep onset could be automatically calculated by the app and algorithm.

(JMIR Mhealth Uhealth 2019;7(5):e13421) doi:10.2196/13421

KEYWORDS

circadian rhythm; sleep; smartphone; mobile applications

Introduction

Background

Human beings, like other animals and plants, have a biological clock that helps to prepare their physiology for the fluctuations of the day. This regular adaptation is referred to as the circadian rhythm. Since the 1970s, scientists have investigated the molecular mechanisms controlling the circadian rhythm [1-7]. Chronic circadian dysregulation has recently been implicated in the increased risk of cancer, neurodegenerative disorders, metabolic disorders, and inflammation [8]. Circadian disruption

has also been associated with several psychiatric disorders, such as bipolar disorder, major depression, and schizophrenia. However, although research has led to a better understanding of the biological basis of the circadian rhythm, most clinical studies are conducted either in the artificial settings of a laboratory (eg, polysomnography), which are not scalable for administration to a large population [9], or through subjective self-report questionnaires with the value reduced by biases [10-12].

Nowadays, human circadian rhythms could be observed from their digital footprint. Digital footprint refers to data rising from day-to-day interactions with newer technologies such as smartphones [13]. Real-time and passively collected data can provide a long-term recording of the circadian rhythm in a naturalistic setting and contribute toward self-awareness or clinical applications, such as sleep diary and social jetlag estimation. In addition, smartphone ownership has shown to not be affected by socioeconomic status [14]. Given the convenience of smartphones, health-related mobile apps might serve as a *digital lifeline*, particularly in rural and low-income regions, helping mental health care professionals with medical intervention and behavioral modification [15]. The widespread use and deep reach of smartphones in modern life motivate the use of smartphones to measure behaviors in an affordable, reliable, and unobtrusive way.

Objective

This pilot study had proposed that the longest nonusage episodes during night-time could represent actual sleep time [16]. A previous study also preliminarily validated that the consistency between app-recorded and self-reported sleep time was 83.0%. However, this validation was based on 14-day app-recorded data with 1 self-reported weeknight and weekend night sleep time. The current version of the app, *Rhythm*, with 2 major algorithm revisions is hypothesized to improve the consistency of app-recorded and self-reported sleep time. In addition, the app-recorded sleep indicators, especially as midpoint of sleep time, can be used to infer a relative long-term pattern of the circadian rhythm. The specific aims of this study were (1) to validate the app-recorded sleep time with daily self-reports by examining the consistency of total sleep time (TST), as well as the timing of sleep onset and wake time, and (2) to validate the app-recorded circadian rhythm with the corresponding 30-day self-reported midpoint of sleep and the consistency of social jetlag.

Methods

Participants and Procedure

A total of 28 college students (13 men, mean age 20.8 (SD 1.6) years, range 17 to 23) were recruited. The sleep time data were collected by the mobile app, *Rhythm*, from February to June 2018. After informed consent, the participants were asked to install Rhythm for at least 30 days. Data collected on the first day and the last day were excluded owing to the incomplete nature of data on those dates. The app-recorded sleep data of the first 30 days of each participant were selected to be analyzed. The 30-day data consisted of about 16.9 days weekday data and 13.1 days weekend data.

This app automatically estimated sleep onset and wake time via an algorithm daily, and participants received a notification from this app to show their sleep onset and wake time last night at 21:00 every night. Then participants were asked to adjust the sleep onset and wake time as their self-reported sleep time. If the difference of sleep time between self-reported and app-recorded was greater than 2 hours, the researchers would confirm the self-reported sleep time with participants via a phone call. The participants were blind to the purpose of the confirmation process. The study was approved by the Institutional Review Board of National Health Research Institutes. All clinical investigations were conducted according to the principles expressed in the Declaration of Helsinki.

Measures

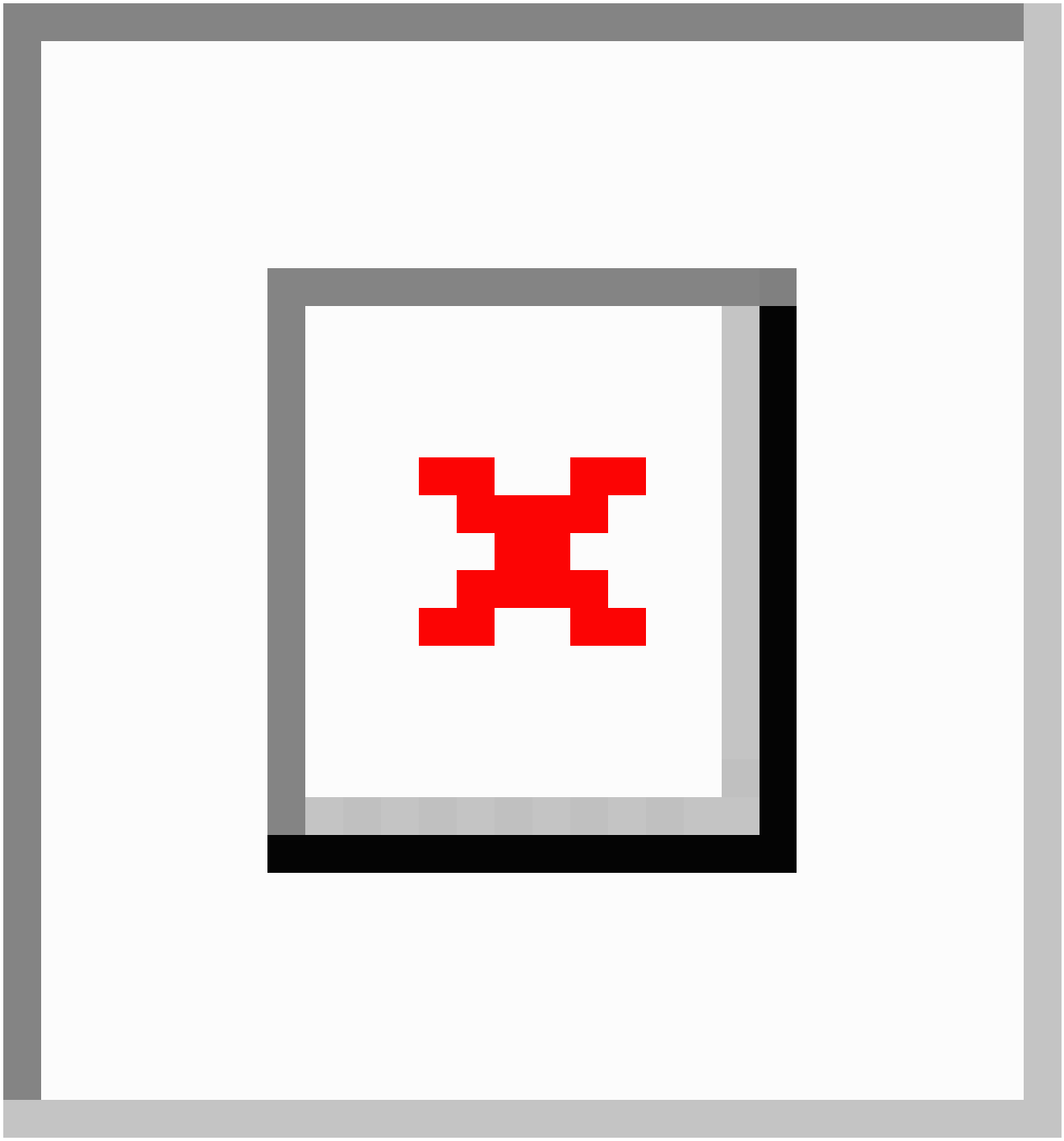
Designing the App, Rhythm

The app, *Rhythm*, automatically recorded smartphone behaviors, especially the notifications and screen-on and screen-off timing. This app collects data in the background without interrupting smartphone operation or impacting battery life (less than 1%) [16-18]. The app saves all recorded behavior data in a log file and routinely uploads to the database every midnight (01:00) and the following noon (12:00). The sleep indicators were calculated, and the server would send a notification with their sleep onset, wake time, and TST at 21:00. We choose 21:00 as the time to send the notification because these sleep indicators could not be calculated in real time on their smartphones. In addition, participants might easily ignore these notifications during their working hours.

Smartphone use from screen-on to the successive screen-off was defined as one episode of use. This app calculated the daily total duration of smartphone usage episodes. The usage episode with no notification within 1 min before screen-on was classified into proactive use. The upper right box of Figure 1 shows a sample of reactive use and proactive use. Only proactive usage episodes were included in the calculation of sleep time. To avoid a high number of frequent notifications confounding proactive use as reactive use, which has a notification within 1 min before screen-on, all notifications from the app which presented more than 500 times per day were excluded. In contrast to these proactive usage episodes, the event from screen-off to screen-on was defined as the nonusage episode.

Figure 1 shows 3 examples to demonstrate the algorithm of identifying sleep time via smartphone use data. For the first example, the maximal nonusage episode between 22:00 and the following 10:00 is defined as *sleep time*. The other 2 examples demonstrated an additional algorithm to identify the sleep onset or wake time not located between 22:00 and 10:00. First, the dummy screen-off (sleep onset) at 22:00 and screen-on (wake time) at 10:00 were labeled to the nonusage episodes with the screen-off before 22:00 and screen-on after 10:00. Second, if the maximal nonusage episode between 22:00 and the following 10:00 was labeled a dummy screen-off at 22:00 or screen-on at 10:00, the original screen-off before 22:00 or screen-on after 10:00 was resumed as the timing of sleep onset and wake time, respectively. Finally, the sleep onset, wake time, midpoint of sleep, and TST could be identified.

Figure 1. Definitions of reactive use and proactive use and the rules of identifying sleep time via smartphone use data.



Validation of the App-Recorded Sleep Indicators and Circadian Rhythm

Validation of the App-Recorded Sleep Indicators

A paired *t* test was used to examine the differences between app-recorded and self-reported indicators, namely sleep onset, wake time, midpoint of sleep time, and TST. Figure 2 shows the consistency of the app-recorded and self-reported sleep time by an overlap ratio [19], and the overlap ratio can mathematically be expressed as follows:

$$\text{Overlap Ratio} = \text{Overlap}(TST_x) / [(TST_{app} + TST_{self}) / 2] \times 100\%$$

The app-recorded and self-reported TST can be calculated as TST_{app} and TST_{self}. The overlapping TST between app-recorded and self-reported TST is defined as TST_x. In Figure 2, the overlap ratio for the example is:

$$(23{:}36 \sim 06{:}00) / \{[(23{:}00 \sim 06{:}00) + (23{:}36 \sim 06{:}49)] / 2\} = 90.0\%$$

In addition, a paired *t* test was used to compare these overlap ratios on weeknights and weekend nights.

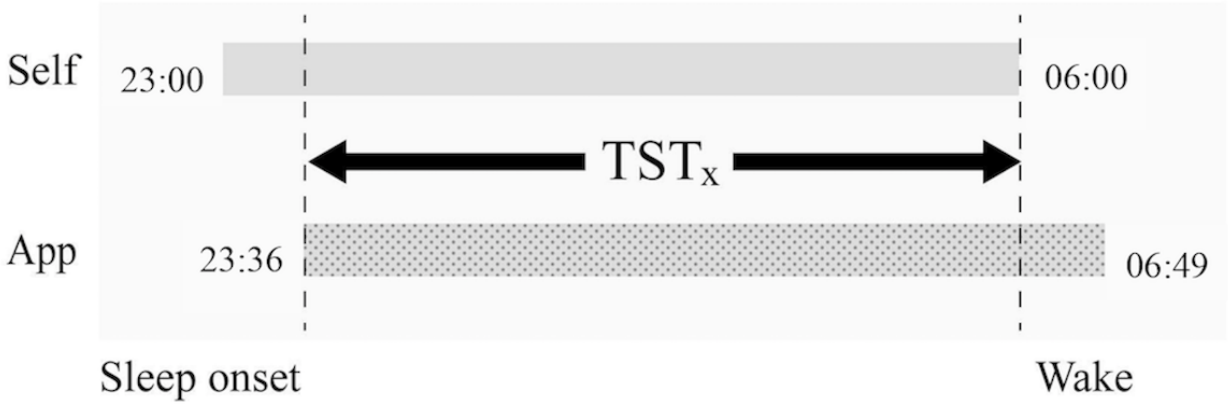
Validation of the Circadian Rhythm

The Pearson correlation coefficient of app-recorded and self-reported midpoint of sleep time within a participant's 30-day data was calculated along with the average midpoint of

sleep differences between weeknights and weekend nights, and these differences are social jetlag [20,21]. Then, a paired *t* test was used to examine the difference between app-recorded and self-reported social jetlag.

All statistical assessments were 2-tailed, and $P<.05$ was considered to be statistically significant. Statistical analyses were performed using SPSS version 18.0 software (SPSS Inc).

Figure 2. Definition of sleep onset time, midpoint of sleep, and wake time. TST: total sleep time.



Results

The average total daily smartphone use duration of the participants was 5.73 (SD 3.42) hours. Table 1 shows that there is no significant difference between app-recorded and self-reported midpoint of sleep time. App-recorded sleep onset had a 242.9-second delay from self-reported onset with a borderline *P* value (.053). App-recorded wake time had a significant 623.7-second advance to self-reported wake time ($P=.001$). App-recorded TST had a significantly shorter 866.6 seconds than self-reported time ($P<.001$). However, the overlap ratio of app-recorded and self-reported was 90.4%. There was no significant difference of overlap ratio between weeknights and weekend nights ($P=.213$).

Figure 3 shows 28 participants' average app-recorded and self-reported midpoint of sleep time and the correlation coefficients from Day 1 to Day 30. These daily correlation coefficients ranged from .70 to .95, and the overall coefficient between app-recorded and self-reported midpoint of sleep time was .87 ($N=840$). There is no significant difference ($P=.140$) between app-recorded (34.4 [SD 52.5] min) and self-reported social jetlag (27.0 [SD 49.8] min).

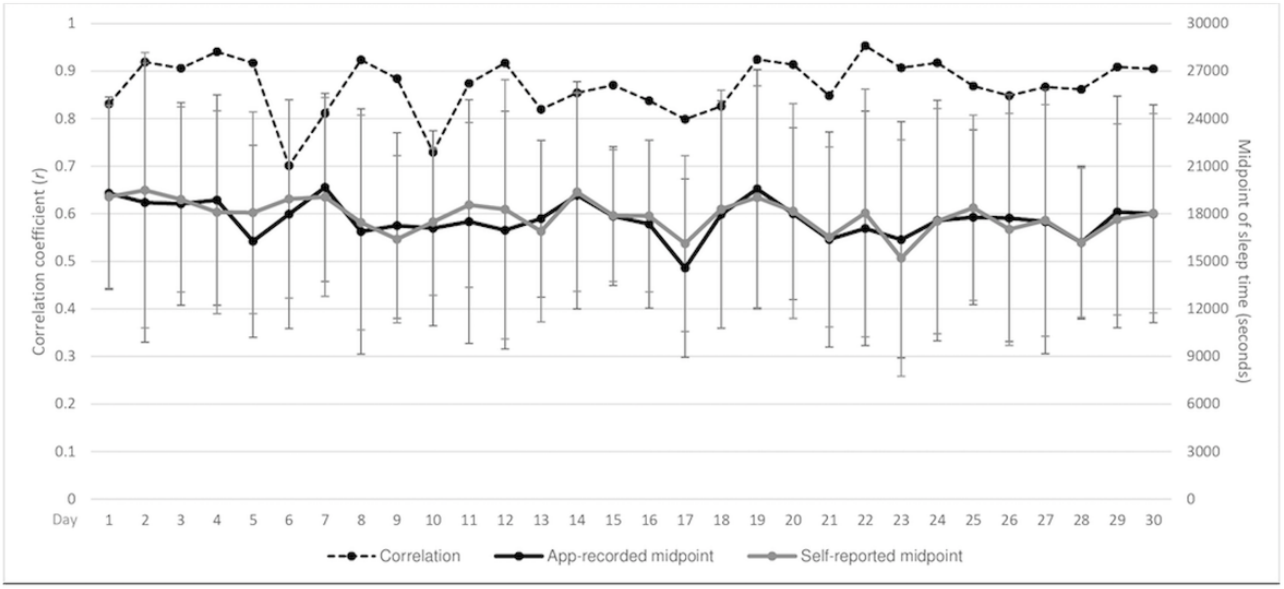
Table 1. The consistency between app-recorded and self-reported sleep indicator.

Sleep indicators	Elapsed time since midnight (seconds)					
	App-recorded		Self-reported		t_{27}	*P* value
	Average time	Mean (SD)	Average time	Mean (SD)		
Sleep onset	01:26:43	5202.53 (7093.71)	01:22:40	4959.64 (6662.52)	1.94	.053
Wake	08:21:44	30104.20 (9037.06)	08:32:08	30727.93 (8417.06)	−3.35	.001[a]
Midpoint	04:54:13	17653.27 (6707.71)	04:57:24	17843.74 (6362.29)	−1.64	.103
Total sleep time	06:55:02	24901.67 (9165.46)	07:09:28	25768.30 (8279.86)	−4.02	<.001[b]

[a]$P<.05$.

[b]$P<.001$.

Figure 3. Circadian rhythm fluctuations for app-recorded and self-reported midpoint of sleep time with 1 SD error bar.



Discussion

Principal Findings

After a review of the literature, this is the first study to validate an innovative approach to automatically record 1-month circadian rhythm and sleep time by a mobile app. The high correlation (r=.87) between app-recorded and self-reported circadian rhythm was validated by 28 participants and their 30-day sleep-wake cycles with a total of 840 pairs of app-recorded data and self-reports. Collecting data passively from a person's smartphone (app) may be more informative than self-reports. This data collection method can provide continuous monitoring over longitudinal periods. In addition, the value of self-reports might be reduced by biases [13], whereas data collection by this automatically operated app is less prone to biases, as the process of collecting data from everyday interactions with technology is unlikely to induce reactivity. In addition, of note is that the temporal resolution of self-reported sleep time is usually 1 hour [22]. Compared with the temporal resolution of 1 second in this study, these app-generated parameters increased the temporal resolution 3600-fold compared with the resolution in conventional epidemiology surveys.

Literature Review/Related Work

There have been several mobile apps on the market to measure sleep automatically via smartphone sensors. Best Effort Sleep [23] uses a sensor-based inference algorithm that combines smartphone usage patterns along with environmental cues such as light and ambient sound to infer a user's sleep duration. Similarly, Toss 'N' Turn [24] also collects sound, light, movement, screen state, app usage, and battery status to classify sleep state and quality. The systems, iSleep [25] and wakeNsmile [26], use a built-in phone microphone to detect body movement and sounds such as cough and snoring to predict sleep phases. However, such apps typically assess sleep time or sleep phases but these mobile apps do not take the circadian rhythm into consideration.

These mobile sensing-based algorithms with less power consumption would advantage from delineating the circadian rhythm from a long consecutive sleep recording. Only a couple of mobile apps compute the sleep time and circadian rhythm solely based on smartphone usage patterns. The pilot study we performed identified proactive smartphone screen-on and screen-off patterns to estimate sleep time and achieved 83% accuracy [19]. UbiComp [27] similarly showed that smartphone usage patterns were able to detect sleep duration as well as symptoms of sleep deprivation. Although these mobile sensing-based apps were validated to assess sleep time, this is the first study to validate both sleep time and circadian rhythm for 30 days with corresponding day-by-day self-reports.

Strengths

Technically, 2 additional criteria were added to improve the algorithm of sleep time estimation with the consistency of 90.4% in this study. First, the extended tracing of sleep onset before 22:00 and wake time after 10:00 could identify the relative irregular circadian rhythm compared with the other methods that limited sleep time between 22:00 and 10:00. Second, the exclusion of frequent notifications elaborated the differentiation

from proactive and reactive use and delineated the sleep time more precisely. These 2 improvements in the present algorithm promoted the accuracy of TST estimation from 83.0% [19] to 90.4%. Owing to the extended tracing of the circadian rhythm and much precise sleep time identification, the correlation coefficient of app-recorded and self-reported 1-month circadian rhythm was .87. In addition, the app-recorded and self-reported social jetlag was similar, without any significant difference. These findings showed that using a smartphone to record passive data, the timing of screen-on and screen-off, and notifications could automatically calculate the sleep time and circadian rhythm for 1 month.

Despite no significant difference between the app-recorded and self-reported sleep time midpoints, the app-recorded wake time was significantly earlier than self-reports by 10.4 min (623.7 seconds) and app-recorded sleep onset was longer than self-reports by 4.0 min (242.9 seconds) with a borderline *P* value (.053). The later sleep and earlier wake timing of app records counteracted the differences in midpoints of sleep time, the indicators of circadian rhythm. Therefore, the 1-month circadian rhythm based on the midpoint of sleep correlation of app records and self-reports reached .87 in this study. The study results revealed that it is feasible to monitor the circadian rhythm automatically by this app that we developed. However, the app-recorded TST might be 14.4 min (866.6 seconds) shorter than the self-reported time despite the consistency of sleep time being 90.4%.

The differences of sleep and wake timing between app records and self-reports might have resulted from some habitual bedtime behaviors, such as bedtime smartphone use before sleep and lying in bed after waking. In addition, distorted time perception [18] also play an important role in these timing differences, especially in participants with average daily smartphone use of 5.84 (SD 2.92) hour/day in this study. The participants might use the smartphone at waking up and might have stayed in bed for an average 10.4 min to be aware that they actually woke up. In this condition, the app could estimate time in bed by self-reported TST and sleep time by app-recorded TST. Therefore, the sleep efficiency, defined as the ratio of sleep to time in bed, could be estimated by TST_{app}/TST_{self}. Although the gold standard method to define the sleep and wake timing is polysomnography, this app-recorded sleep-wake cycle in a naturalistic environment has provided a more cost-efficient and convenient way to continuously delineate the circadian rhythm.

Limitations

There are several methodological limitations that should be noted when interpreting the study's findings. First, the study utilized a selected sample with excessive smartphone use (average daily smartphone use duration: 5.84 (SD 2.92) hour/day) and late chronotype (average sleep onset at 01:26:43 and TST: 6.92 hours). A previous study had demonstrated the association between excessive internet use and late chronotype [10]. In addition, both excessive smartphone use and late chronotype might limit the ability to generalize these findings because this algorithm to estimate sleep patterns depended on participants' smartphone events. Using smartphone use data combined with an activity wristband that gathers the whole day's activities and physiological indicators could improve the reliability of computing the circadian rhythm for participants owning such devices. Second, the app was based on the Android operating system. Various versions applicable to other operating systems such as iOS and Windows should be developed in the future. In addition, the app failed to record notifications in 1 smartphone brand. However, it is essential to calculate the sleep time by screen-on, screen-off, and notifications. Third, this algorithm to determine sleep time could not account for sleep interruptions with proactive smartphone use or shift workers' daytime sleep. This algorithm should be adjusted and validated for patients with sleep disturbance and shift workers. Finally, although these app-recorded sleep indicators were validated by participants' daily self-reports, it would be better to use the current gold standard to assess circadian rhythm, actigraphy, to validate the app-recorded sleep time in a future study.

Conclusions

In conclusion, this study validated the algorithm of sleep estimation and circadian rhythm by using the app, *Rhythm*, that can collect passive data from naturalistic settings. The circadian rhythm for 1 month, daily TST, and timing of sleep onset could be automatically calculated by the app.

Acknowledgments

This study was supported by a grant (06A1-PHPP30-014) from the National Health Research Institutes, Taiwan. The supporters had no role in the design, analysis, interpretation, or publication of this study. The authors thank Ms Hsiao-Han Lin and Ms Wei-Wen Sun for their excellent technique assistance. Ms Lin and Ms Sun declared no conflict of interest in this study.

Conflicts of Interest

None declared.

References

1. Bargiello TA, Jackson FR, Young MW. Restoration of circadian behavioural rhythms by gene transfer in Drosophila. Nature 1984 Dec;312(5996):752-754. [doi: 10.1038/312752a0]
2. Hardin PE, Hall JC, Rosbash M. Feedback of the Drosophila period gene product on circadian cycling of its messenger RNA levels. Nature 1990 Feb 8;343(6258):536-540. [doi: 10.1038/343536a0] [Medline: 2105471]
3. Liu X, Zwiebel LJ, Hinton D, Benzer S, Hall JC, Rosbash M. The period gene encodes a predominantly nuclear protein in adult Drosophila. J Neurosci 1992 Jul;12(7):2735-2744 [FREE Full text] [Medline: 1613555]

4. Price JL, Blau J, Rothenfluh A, Abodeely M, Kloss B, Young MW. double-time is a novel Drosophila clock gene that regulates PERIOD protein accumulation. Cell 1998 Jul 10;94(1):83-95 [FREE Full text] [Medline: 9674430]
5. Siwicki KK, Eastman C, Petersen G, Rosbash M, Hall JC. Antibodies to the period gene product of Drosophila reveal diverse tissue distribution and rhythmic changes in the visual system. Neuron 1988 Apr;1(2):141-150. [Medline: 3152288]
6. Vosshall LB, Price JL, Sehgal A, Saez L, Young MW. Block in nuclear localization of period protein by a second clock mutation, timeless. Science 1994 Mar 18;263(5153):1606-1609. [Medline: 8128247]
7. Zehring WA, Wheeler DA, Reddy P, Konopka RJ, Kyriacou CP, Rosbash M, et al. P-element transformation with period locus DNA restores rhythmicity to mutant, arrhythmic Drosophila melanogaster. Cell 1984 Dec;39(2 Pt 1):369-376. [Medline: 6094014]
8. Hirota T, Kay S. Identification of small-molecule modulators of the circadian clock. Methods Enzymol 2015;551:267-282. [doi: 10.1016/bs.mie.2014.10.015] [Medline: 25662461]
9. Chang A, Aeschbach D, Duffy JF, Czeisler CA. Evening use of light-emitting eReaders negatively affects sleep, circadian timing, and next-morning alertness. Proc Natl Acad Sci U S A 2015 Jan 27;112(4):1232-1237 [FREE Full text] [doi: 10.1073/pnas.1418490112] [Medline: 25535358]
10. Lin Y, Gau SS. Association between morningness-eveningness and the severity of compulsive Internet use: the moderating role of gender and parenting style. Sleep Med 2013 Dec;14(12):1398-1404. [doi: 10.1016/j.sleep.2013.06.015] [Medline: 24157101]
11. Roenneberg T, Wirz-Justice A, Merrow M. Life between clocks: daily temporal patterns of human chronotypes. J Biol Rhythms 2003 Feb;18(1):80-90. [doi: 10.1177/0748730402239679] [Medline: 12568247]
12. Roenneberg T, Kuehnle T, Juda M, Kantermann T, Allebrandt K, Gordijn M, et al. Epidemiology of the human circadian clock. Sleep Med Rev 2007 Dec;11(6):429-438. [doi: 10.1016/j.smrv.2007.07.005] [Medline: 17936039]
13. Bidargaddi N, Musiat P, Makinen V, Ermes M, Schrader G, Licinio J. Digital footprints: facilitating large-scale environmental psychiatric research in naturalistic settings through data from everyday technologies. Mol Psychiatry 2017 Dec;22(2):164-169 [FREE Full text] [doi: 10.1038/mp.2016.224] [Medline: 27922603]
14. Blumenstock J, Cadamuro G, On R. Predicting poverty and wealth from mobile phone metadata. Science 2015 Nov 27;350(6264):1073-1076 [FREE Full text] [doi: 10.1126/science.aac4420] [Medline: 26612950]
15. Anthes E. Mental health: there's an app for that. Nature 2016 Apr 7;532(7597):20-23. [doi: 10.1038/532020a] [Medline: 27078548]
16. Lin Y, Lin Y, Lin S, Lee Y, Lin P, Chiang C, et al. To use or not to use? Compulsive behavior and its role in smartphone addiction. Transl Psychiatry 2017 Dec 14;7(2):e1030 [FREE Full text] [doi: 10.1038/tp.2017.1] [Medline: 28195570]
17. Lin Y, Lin P, Chiang C, Lee Y, Yang CH, Kuo TB, et al. Incorporation of mobile application (app) measures into the diagnosis of smartphone addiction. J Clin Psychiatry 2017 Jul;78(7):866-872. [doi: 10.4088/JCP.15m10310] [Medline: 28146615]
18. Lin Y, Lin Y, Lee Y, Lin P, Lin S, Chang L, et al. Time distortion associated with smartphone addiction: identifying smartphone addiction via a mobile application (app). J Psychiatr Res 2015 Jun;65:139-145. [doi: 10.1016/j.jpsychires.2015.04.003] [Medline: 25935253]
19. Lin Y, Wong B, Lin S, Chiu Y, Pan Y, Lee Y. Development of a mobile application (App) to delineate. J Psychiatr Res 2018 Dec 8;110:9-15. [doi: 10.1016/j.jpsychires.2018.12.012] [Medline: 30611008]
20. Roenneberg T, Allebrandt KV, Merrow M, Vetter C. Social jetlag and obesity. Curr Biol 2012 May 22;22(10):939-943 [FREE Full text] [doi: 10.1016/j.cub.2012.03.038] [Medline: 22578422]
21. Wittmann M, Dinich J, Merrow M, Roenneberg T. Social jetlag: misalignment of biological and social time. Chronobiol Int 2006 Jul;23(1-2):497-509. [doi: 10.1080/07420520500545979] [Medline: 16687322]
22. Hale L, Guan S. Screen time and sleep among school-aged children and adolescents: a systematic literature review. Sleep Med Rev 2015 Jun;21:50-58 [FREE Full text] [doi: 10.1016/j.smrv.2014.07.007] [Medline: 25193149]
23. Chen Z, Lin M, Chen F, Lane N, Cardone G, Wang R, et al. Unobtrusive Sleep Monitoring using Smartphones. In: Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare. 2013 May 23 Presented at: PervasiveHealth'13; May 5-8, 2013; Venice, Italy p. 145-152 URL: https://eudl.eu/pdf/10.4108/icst.pervasivehealth.2013.252148 [doi: 10.4108/icst.pervasivehealth.2013.252148]
24. Min JK, Doryab A, Wiese J, Amini S, Zimmerman J, Hong JI. Toss 'n' turn: smartphone as sleep and sleep quality detector. In: Proceedings of the SIGCHI conference on human factors in computing systems.: ACM; 2014 Apr Presented at: CHI'14; AprIL 26-May 1, 2014; Toronto, Ontario, Canada p. 477-486 URL: https://www.cs.utah.edu/~wiese/publications/tossnturn.pdf [doi: 10.1145/2556288.2557220]
25. Hao T, Xing G, Zhou G. iSleep: unobtrusive sleep quality monitoring using smartphones. In: Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems.: ACM; 2013 Presented at: SenSys'13; November 11-15, 2013; Roma, Italy URL: https://cse.buffalo.edu/~lusu/cse721/papers/iSleep%20Unobtrusive%20Sleep%20Quality%20Monitoring%20using%20Smartphones.pdf [doi: 10.1145/2517351.2517359]
26. Krejcar O, Jirka J, Janckulik D. Use of mobile phones as intelligent sensors for sound input analysis and sleep state detection. Sensors (Basel) 2011;11(6):6037-6055 [FREE Full text] [doi: 10.3390/s110606037] [Medline: 22163941]

27. Abdullah S, Matthews M, Murnane EL, Gay G, Choudhury T. Towards circadian computing: "early to bed and early to rise" makes some of us unhealthy and sleep deprived. In: Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing.: ACM; 2014 Presented at: UbiComp'14; September 13-17, 2014; Seattle, Washington p. 673-684 URL: http://pac.cs.cornell.edu/pubs/clockwise-ubicomp-2014.pdf [doi: 10.1145/2632048.2632100]

Abbreviations

TST: total sleep time

Edited by G Eysenbach; submitted 17.01.19; peer-reviewed by M Ponum, C Fernández; comments to author 06.03.19; revised version received 20.03.19; accepted 24.03.19; published 16.05.19.

Please cite as:
Lin YH, Wong BY, Pan YC, Chiu YC, Lee YH
Validation of the Mobile App–Recorded Circadian Rhythm by a Digital Footprint
JMIR Mhealth Uhealth 2019;7(5):e13421
URL: https://mhealth.jmir.org/2019/5/e13421/
doi:10.2196/13421
PMID:31099340

Original Paper

A Focused Review of Smartphone Diet-Tracking Apps: Usability, Functionality, Coherence With Behavior Change Theory, and Comparative Validity of Nutrient Intake and Energy Estimates

Giannina Ferrara[1*], BA; Jenna Kim[2*]; Shuhao Lin[3*], BS; Jenna Hua[4], BSc (Hons), RD, MPH, PhD; Edmund Seto[5], AB, MS, PhD

[1]Global Burden of Disease, Institute for Health Metrics and Evaluation, University of Washington, Seattle, WA, United States

[2]Paul G Allen School of Computer Science & Engineering, University of Washington, Seattle, WA, United States

[3]Department of Kinesiology and Nutrition, University of Illinois, Chicago, Chicago, IL, United States

[4]Department of Medicine, Stanford Prevention Research Center, Stanford University, Stanford, CA, United States

[5]Department of Environmental & Occupational Health Sciences, University of Washington, Seattle, WA, United States

[*]these authors contributed equally

Corresponding Author:
Edmund Seto, AB, MS, PhD
Department of Environmental & Occupational Health Sciences
University of Washington
1959 NE Pacific Street, Campus Box 357234
Seattle, WA, 98195
United States
Phone: 1 206 543 1475
Email: eseto@uw.edu

Abstract

Background: Smartphone diet-tracking apps may help individuals lose weight, manage chronic conditions, and understand dietary patterns; however, the usabilities and functionalities of these apps have not been well studied.

Objective: The aim of this study was to review the usability of current iPhone operating system (iOS) and Android diet-tracking apps, the degree to which app features align with behavior change constructs, and to assess variations between apps in nutrient coding.

Methods: The top 7 diet-tracking apps were identified from the iOS iTunes and Android Play online stores, downloaded and used over a 2-week period. Each app was independently scored by researchers using the System Usability Scale (SUS), and features were compared with the domains in an integrated behavior change theory framework: the Theoretical Domains Framework. An estimated 3-day food diary was completed using each app, and food items were entered into the United States Department of Agriculture (USDA) Food Composition Databases to evaluate their differences in nutrient data against the USDA reference.

Results: Of the apps that were reviewed, LifeSum had the highest average SUS score of 89.2, whereas MyDietCoach had the lowest SUS score of 46.7. Some variations in features were noted between Android and iOS versions of the same apps, mainly for MyDietCoach, which affected the SUS score. App features varied considerably, yet all of the apps had features consistent with Beliefs about Capabilities and thus have the potential to promote self-efficacy by helping individuals track their diet and progress toward goals. None of the apps allowed for tracking of emotional factors that may be associated with diet patterns. The presence of behavior change domain features tended to be weakly correlated with greater usability, with R^2 ranging from 0 to .396. The exception to this was features related to the Reinforcement domain, which were correlated with less usability. Comparing the apps with the USDA reference for a 3-day diet, the average differences were 1.4% for calories, 1.0% for carbohydrates, 10.4% for protein, and −6.5% for fat.

Conclusions: Almost all reviewed diet-tracking apps scored well with respect to usability, used a variety of behavior change constructs, and accurately coded calories and carbohydrates, allowing them to play a potential role in dietary intervention studies.

(JMIR Mhealth Uhealth 2019;7(5):e9232) doi:10.2196/mhealth.9232

KEYWORDS

diet; nutrition assessment; behavior and behavior mechanisms

Introduction

Background

A large number of apps focused on health and fitness have emerged on the smartphone market. In 2017, a total of 325,000 mobile health (mHealth) apps were available in major app stores, and the number of users of mHealth apps will continue to rise in the upcoming years [1]. These apps have the potential to facilitate tracking of health-related behaviors and weight management [2]. Within this group of apps, diet-tracking apps are very popular, with some downloaded as much as 50 million times (based on MyFitnessPal for Android market, April 2017). Tracking the consumption of certain foods and drinks may potentially help individuals achieve an improved understanding of their dietary patterns [3]. Using a diet-tracking app may improve self-monitoring, goal setting, and knowledge and develop self-efficacy—all of which are key behavior change constructs [2-6]. However, it remains unclear how many of the current diet-tracking apps employ such features that are consistent with behavior change theory. In past reviews of diet-related health apps, it was found that adoption of behavior change theory tended to be quite poor [7-11].

Given the current obesity epidemic within the United States [12] and in many other countries around the world [13], there is a great need for effective theory-driven tools to help individuals manage weight. Aside from tracking energy intake, other aspects of diet may be important to monitor. For instance, shifts in diet toward greater intake of processed foods, meals outside of the home, and greater amounts of oil and sugar-added foods have been implicated as potential causes of the global obesity epidemic [14]. Within the United States, shifts in portion sizes may be responsible for increased energy intake and obesity [15]. Moreover, diet tracking may be particularly useful for those who are at risk of or who already need to manage specific diet-related health issues, such as a need to monitor carbohydrate intake for metabolic disease syndrome [16], sodium intake for hypertension [17], or food elimination for irritable bowel syndrome [18] or allergies [19]. Therefore, the ability for apps to accurately track food intakes may be useful for understanding the changing diet trends within certain populations, as well as diet patterns relevant to an individual's health. Yet, few studies have been conducted evaluating the accuracy of diet-tracking apps. For example, a single user recording a 3-day diet in a study found that the accuracy in tracking energy intake among apps was only fair in terms of total calories and amounts of macro- and micronutrients compared with a gold standard nutrient coding and that there was large variability among apps [2].

In addition to the importance of accurate diet measures, usability is an important aspect of quality of diet-tracking apps. Usability encompasses multiple dimensions of user interaction with an app, which includes ease of use, complexity, need for training and support, and willingness to continue use. Various weight loss apps have been scored for usability in single-user studies [2], with apps generally receiving good to very good ratings of usability. Moreover, acceptability of mobile apps (ie, the willingness of people to use the apps) for health interventions has generally been found to be high among study participants [10]. However, usability and acceptance may change, particularly as the number of apps focused on diet-tracking increases, and apps become more complex with additional new features.

Objectives

The goal of our study was to expand upon the existing literature that evaluates health apps in general, with a more focused review of the top Android and iPhone operating system (iOS) diet-tracking apps to evaluate multiple aspects of quality and applicability for use in behavior change studies. This review includes ratings of app usability and coherence of app features with behavior change theory using a 50-item list adapted from the Theoretical Domains Framework (TDF). We further assessed each app's ability to code calorie, carbohydrate, fat, and protein intake compared with the United States Department of Agriculture (USDA) reference coding.

Methods

Selection of Diet-Tracking Apps

Apps were identified using terms *diet tracking* and *diet app* in both Android and iTunes stores in April 2017. iTunes returned the top 100 iPhone apps in both cases but did not display less relevant items. Android Play store returned 245 results. The following categories were used to screen the apps: (1) ability to record dietary intake of users, (2) free of charge, (3) availability for both iOS and Android devices, and (4) high popularity ranking in app store (based on an algorithm that looks at average app store ratings, rating and review volume, download counts, and app usage statistics generated by the App Store and Google Play [20,21]). We selected apps that were free of charge because of their wider user base and number of downloads. We selected the top 7 apps (Table 1). Beyond these 7, the remaining apps were either much less popular with few reviews or appeared in only 1 of the app stores. We refer to some of the apps by their producer names for simplicity in the results below.

Table 1. Characteristics of the diet-tracking apps used in the United States.

App name (year established)	Producer (country of origin)	Rating (Android/iOS/ratings)	Downloads (Android)[a]
Calories Counter (2018)	FatSecret (Australia)	4.4/4.5/1300	10,000,000
LifeSum (2018)	LifeSum (Sweden)	4.3/4.5/22,600	5,000,000
MyPlate Calorie Tracker (2017)	Livestrong (United States)	4.6/4.5/7900	500,000
Argus (2015)	Azumio (United States)	4.3/4.5/22,600	100,000
Lose It! (2008)	FitNow (United States)	4.4/4.0/156,600	5,000,000
Calorie Counter & Diet Tracker (2009)	MyFitnessPal (United States)	4.6/4.5/490,000	50,000,000
MyDietCoach (2014)	InspiredApps (United States)	4.4/4.5/5700	10,000,000

[a]No download statistics available for iOS.

Evaluation Criteria

Usability was scored according to the System Usability Scale (SUS)—a 10-item questionnaire developed by Brooke in 1986 [22], which has been used to evaluate usability for a variety of electronic devices and systems, including health-related smartphone apps [23]. The SUS is a valid and widely used de facto standard for assessing usability [24-26]. As described by Brooke, the SUS aims to assess the degree to which a system is *fit for purpose*. Its questions assess multiple aspects of usability, including ease of use and complexity, and learning and expertise required to use a system. Each question is graded on a 5-point Likert scale ranging from *Strongly agree* to *Strongly disagree*. A composite usability measure is calculated by summing the scores for the odd numbered questions and *5 minus* the score for each of the even numbered questions and multiplying the result by 2.5. The resulting SUS measure ranges from 0 to 100, with a measure above 68 being considered *above average* [27].

Functionality was evaluated by the TDF developed and validated by Cane et al [28]. Developed for behavioral change research, the TDF groups 112 theoretical constructs into 14 domains: Knowledge, Skills, Social/Professional Role and Identity, Beliefs about Capabilities, Optimism, Beliefs about Consequences, Reinforcement, Intentions, Goals, Memory, Attention and Decision Processes, Environmental Context and Resources, Social Influences, Emotions, and Behavioral Regulation, which could be integrated with health behavior change theories, such as the Transtheoretical Model/Stages of Change [29]. The TDF was validated using word sort and clustering exercises by behavior change experts. On the basis of the domains in the TDF, we created a checklist of 50 questions that quantify the presence of diet-tracking app features that relate to specific TDF domains. The questions were developed a priori of using the apps. Each question was designed iteratively by the researchers after reviewing TDF domains and subdomains, discussing the intended meaning of the domain, and how it might manifest as an app feature. Furthermore, the wording of each question was discussed to ensure clarity and ease of scoring. Despite careful wording, because the presence of a feature may not be clear, if reviewers were discordant in their response for a feature, the discordance was discussed, the app was re-reviewed, and consensus was determined by the reviewers.

As all of the apps were targeted for general consumer use and not tied to particular professional services (eg, patient care programs and nutritional services), we focused the Social/Professional Role and Identity domain on characterizing user identity. For example, the question *Does the app make use of Avatars?* relates to a feature that helps the user establish an *Identity* within the app. Similarly, for other domains, we identified specific app features. The question *Does the app provide any encouraging messages?* relates to *Optimism*, and the question *Does the app reward the user in some way (eg, stars, accolades, and achievements) for using of the app?* relates to *reinforcement*. The full list of app functionality questions is in Multimedia Appendix 1.

App Evaluation

Among the authors of this study, 3 authors who were undergraduate students trained in Nutritional Sciences at the time of the assessment downloaded and evaluated each of the apps for a 2-week period. Using their own smartphones, 2 of the authors used the iOS version of the apps, whereas the third author used the Android version of the apps. After using each app, each researcher independently rated the usability and functionality of the app according to the evaluation criteria. The mean SUS measure was computed for each app, as well as mean scores for individual SUS items. The presence or absence of specific app features (functionality) were noted and compared for each app between the 3 users, and discordances were reviewed and discussed to come to a consensus. Responses on questions related to app features were grouped, and positive responses were summed to create a score. Descriptive statistics for these scores were computed for each domain in the TDF. Correlations of SUS score were calculated among the 3 authors, and mean app SUS was compared with app functionality.

Finally, each user evaluated the nutrient coding of the apps using their 3-day record of all foods consumed using USDA. They recorded all foods consumed in real time for 3 consecutive days that included 2 weekdays and a weekend day (Thursday, Friday, and Saturday) during the 2-week period for each app. At the end of each day, the nutrient intakes (total calories, protein, fat, and carbohydrate) from all food consumed as estimated by each app were noted. The resulting nutrient measures were averaged across the 3 days for each user and each app. The 3-day diet records were also coded by each researcher separately using the USDA Food Composition

Database [30]. Portion sizes were estimated based on the researchers' prior nutritional training. Mean differences between each app's 3-day average and the USDA reference was computed across the 3 users.

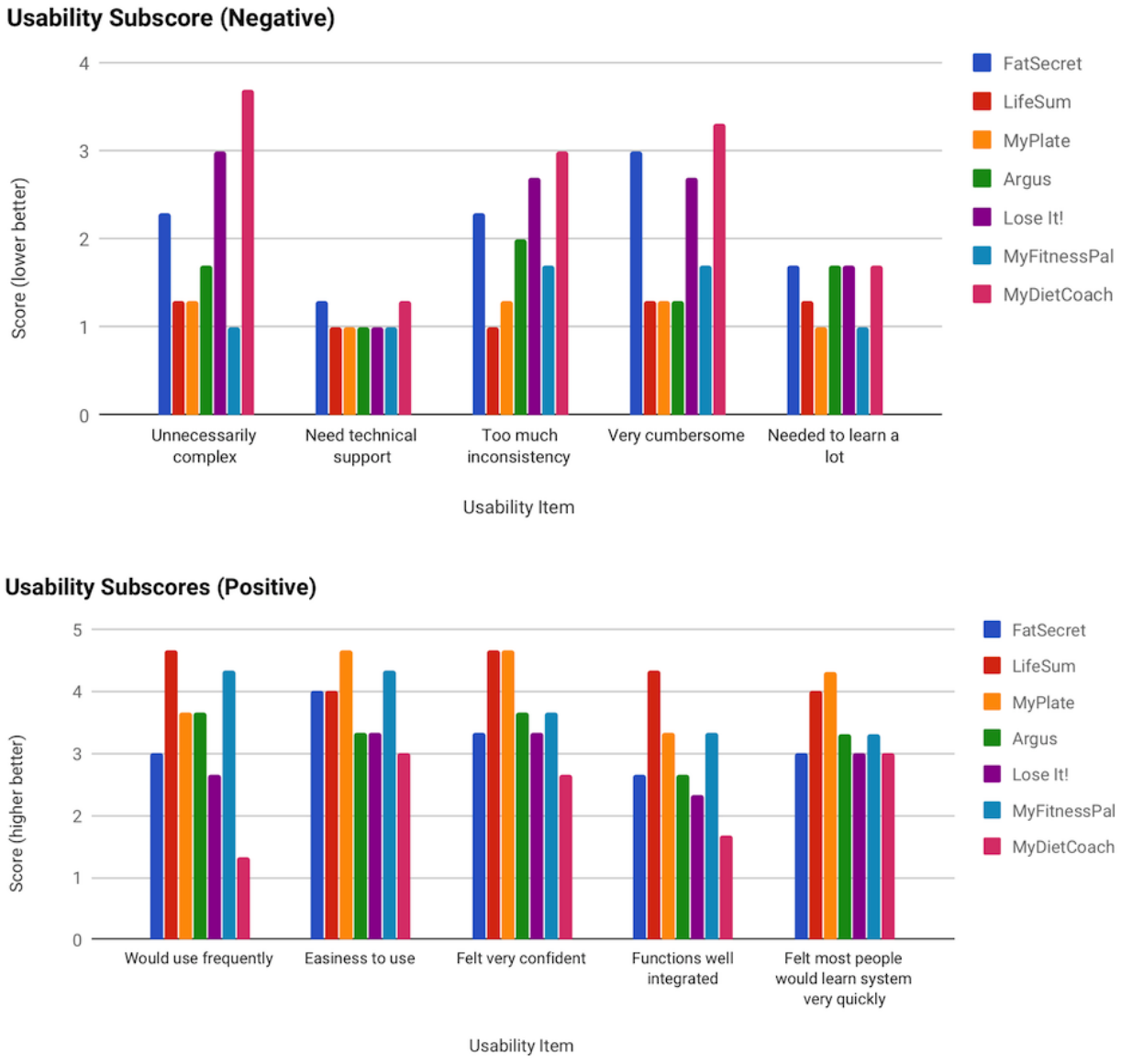
As an additional assessment of coding accuracy, 3 common example food items were input into each app and the USDA database to examine consistency in estimates of calories and macronutrients. A medium banana, a plain Nature Valley granola bar, and a Big Mac from McDonald's were chosen as examples of a common and popular fruit, a packaged food, and a fast food item, respectively. For the 2 processed foods, food label nutrient data were also recorded.

Results

Usability

Figure 1 illustrates both the positive and negative aspects of usability from the SUS. Detailed usability subscores (averages of the reviewers) and the aggregate SUS score for each app are reported in Multimedia Appendix 2. The users' usability scores were consistent and positively correlated with each other and ranged from moderate to high correlation (Pearson correlations of 0.66, 0.84, and 0.89). High contrast scores were not observed. However, there were some notable differences between the iOS and Android versions of the same app; for example, the iOS version of FatSecret had a slightly different user interface and had additional features that made it unnecessarily complex compared with the Android version, which resulted in larger variation in this app's final SUS score compared with the other apps.

Figure 1. Positive items (top) and negative items (bottom) of usability.



Out of the 7 apps in the study, LifeSum was found to be the most user-friendly app based on the SUS scale with an average SUS score of 89.2, whereas MyDietCoach was found to be the least user-friendly app with an average SUS score of 46.7. Comparing the usability subscores among all 7 apps, LifeSum consistently had high scores among the positive usability subcategories, such as for the following items: *Would use frequently*, *Easiness to use*, and *Felt very confident*. Conversely, MyDietCoach had consistently low usability subscores among these usability subcategories.

For the items that represented aspects of negative usability, such as *Unnecessarily complex*, *Needs technical support*, and *Too much inconsistency*, lower scores in this category indicated that the app was more usable. LifeSum had the lowest scores among almost all of those items, which indicates that the app is more user friendly than the other apps that were tested. MyDietCoach had higher scores among these items, which suggests that the app is less usable than the others. These findings are consistent with the other usability subscores that asked positive usability questions.

Generally, all of the apps scored well in terms of ease of use and did not require considerable amount of learning or technical support. Of particular note, many of the apps have features that greatly improve the ease of entering food items. For instance, 6 out of the 7 apps utilize bar code scanning for input of packaged food, and all of the 7 apps are able to remember recent or frequent food items for quick input. LoseIt! also has a feature that attempts to recognize a food item from a photo.

Functionality

Each app was evaluated with a 50-question checklist to identify features that could potentially change the users' behavior based on the TDF. Some apps showed feature discrepancies between the Android and iOS versions. The biggest functionality discrepancy was found in MyDietCoach, whose iOS version allowed diet recording and analysis for free, whereas the Android version required payment to unlock the feature. A feature was counted as present if it was available in at least one of the iOS or Android versions. The complete feature scoring of the apps can be found in Multimedia Appendix 1.

Figure 2 illustrates the number of apps that had features within a particular TDF domain (for domains with multiple feature checklist questions, the average number of apps across the questions in the domain is reported). Notably, all apps had features for the *Belief about Capabilities* domain, which emphasizes building self-efficacy through tracking of progress and working toward goals. Most of the apps also had features within the *Social/Professional Role and Identity* domain, which tries to tailor to the users' requirements by establishing a user identity through account registration, use of avatars, and by tracking user-specific profile information.

With respect to the domains that were less featured in the apps, none of the apps had an *Emotion* feature. For example, the apps did not track the effect of diet on mood or stress nor did they try to explicitly track users' guilt associated with eating certain foods. None of the apps allowed users to note flavor, which is an important aspect of taste, nor did they track users' hunger or satiety. The next least prominent behavior change domain was the *Beliefs about Consequences* domain, which was only present in 1 app, LifeSum, which had a Health Test feature that assessed and challenged users about their diet beliefs and knowledge.

Figure 2. Number of apps that contain specific TDF domain features (for domains with multiple feature checklist questions, the average number of apps across the questions in the domain is reported). TDF: Theoretical Domains Framework.

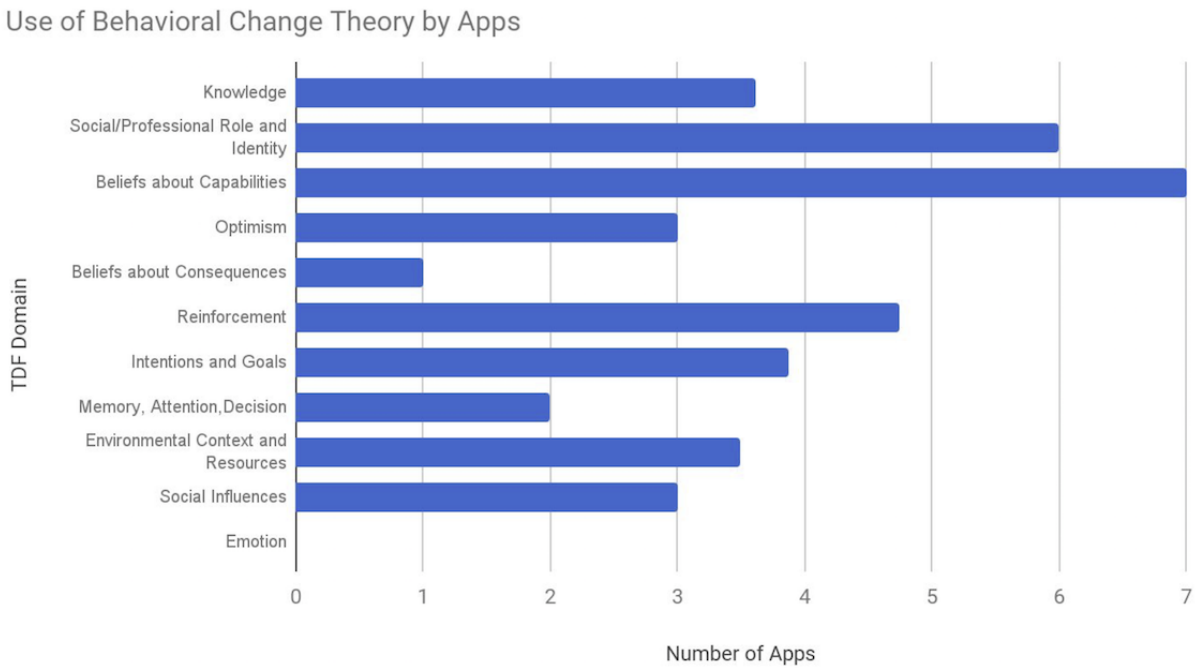


Figure 3 illustrates which behavior change domain features are present in each app (linear models are presented in Multimedia Appendix 3). The apps are ordered on the x-axis in increasing SUS score to illustrate the relationship between each feature and overall usability. Generally, there was little correlation between app behavioral domain features and usability (R^2: 0-.396), with *Beliefs about Consequences* and *Memory, Attention, and Decision* being the features most correlated with greater usability (R^2 of .396 and .262, respectively).

Reinforcement was the only domain that was negatively correlated with usability, although the correlation was not strong (R^2=.094). Notably 1 of the apps, MyDietCoach, had a number of features that related to reinforcement, including rewarding the user (eg, stars, accolades, and achievements), having game-like functions, and providing occasional reminders. However, the app did not score as well as others in terms of usability, illustrating that feature richness does not necessarily relate to greater usability.

Figure 4 provides example screen captures that illustrate how some of the apps implement features related to behavioral constructs. The 2 apps that ranked high on usability, MyFitnessPal and LifeSum, were the only apps that featured the *Memory, Attention, and Decision* domain. A decision-making feature in an app provides the user with judgments about the quality of their diet choices. Figure 4 shows an example of this judgment from LifeSum, which provides different facial *emoji* icons that are related to the calorie content of food items recorded. Similarly, a screen from MyFitnessPal lets the user know that "this food has lots of vitamin C." The app with the highest SUS score, LifeSum, was the only app that featured the *Beliefs about Consequences* domain. Figure 4 illustrates how it uses a *Health Test*, which assesses and challenges the user's belief and knowledge about a healthy diet.

Figure 3. Individual feature correlation with usability. MDC: MyDietCoach; LI: Lose It!; FS: FatSecret; A: Argus; MFP: MyFitnessPal; MP: MyPlate; LS: LifeSum.

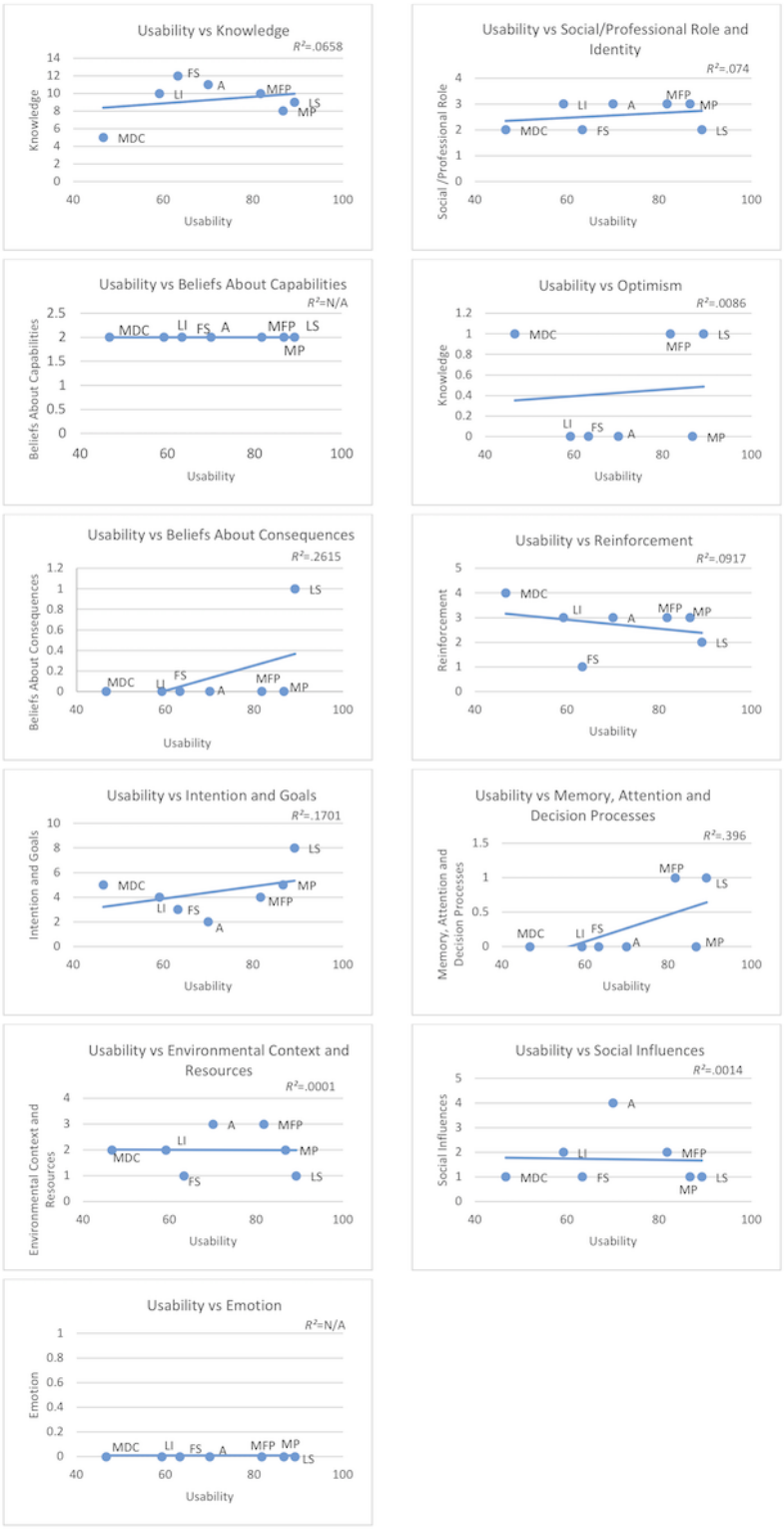
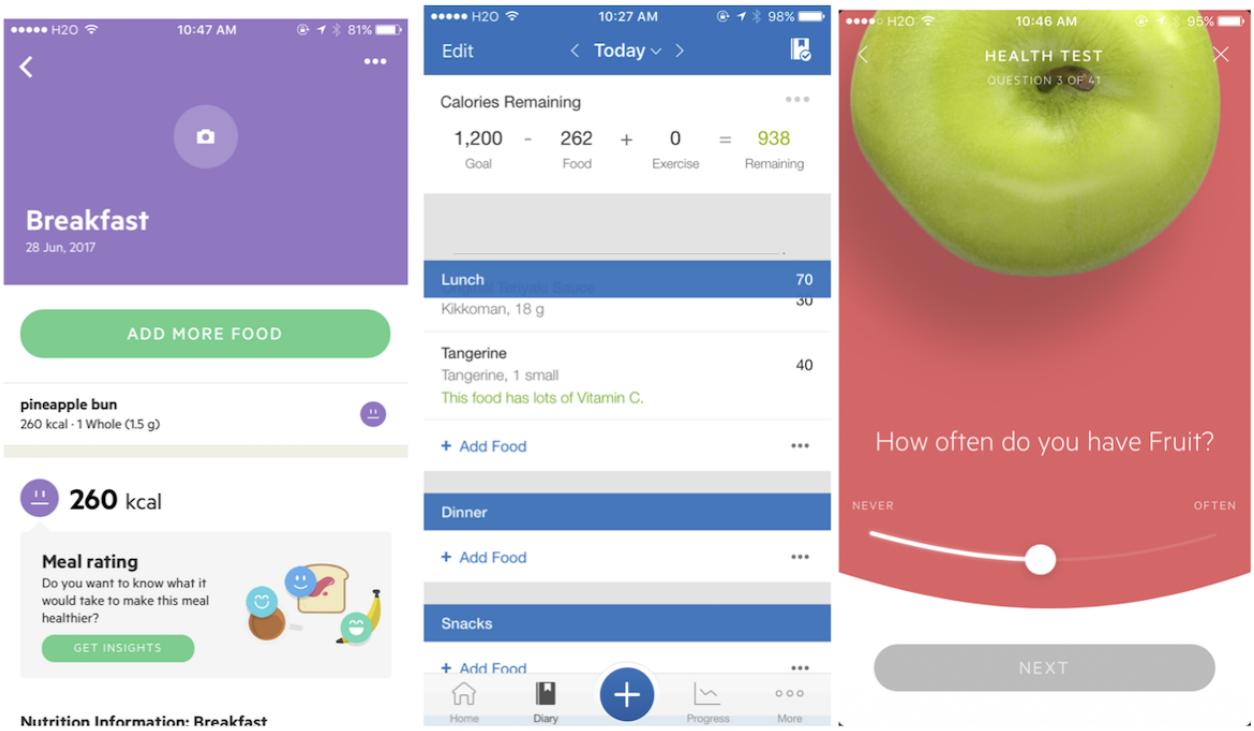


Figure 4. Meal rating feature with emojis in LifeSum (left), feedbacks on nutrient content for food items in MyFitnessPal (center), and Health test in LifeSum (right).



Consistency Between Apps' Nutrient Intake Estimates

On the basis of the subjects' 3-day intakes, we observed differences in each app's estimated averages of 3-day total calorie (kcal) and macronutrient intakes (g) compared with calorie and macronutrient estimates generated by using the USDA Food Composition Database (Figure 5). As there was no free diet-tracking feature in the Android version of MyDietCoach, only the 2 iOS users' data were used for that app. The average total calorie intake was relatively similar among apps, with an average difference of 1.4% compared with the USDA (mean difference 9.6 [SD 50.3] kcal). LifeSum had the highest deviation, an overestimate of 7.29% (SD 70.3) kcal compared with the USDA. On average, the apps only slightly overestimated carbohydrate intake by a difference of 1.0% compared with the USDA (mean difference .8 [SD 4.7] g). With the exception of Lose It!, which slightly underestimated protein intake, most of the apps tended to greatly overestimate intake compared with the USDA (difference of 10.4% and mean difference 3.2 [SD 5.2] g). On average, the apps underestimated fat intake by a difference of 6.5% compared with the USDA. The app that tended to be most accurate in coding calorie and macronutrients relative to the USDA reference was MyFitnessPal, whereas the least accurate was LifeSum.

As an additional assessment of coding accuracy, a ripe medium banana, a plain Nature Valley granola bar, and a Big Mac from McDonald's were coded using each app and the USDA database (Table 2). We found high consistency for caloric content of a medium banana across the apps and the USDA database with average difference of 3.7 kcal (3.5%) compared with USDA. All apps showed consistent estimates except for LifeSum. LifeSum showed higher calories and macronutrient. This may be due to portion size variations, as the only banana item listed in that app was for a weight of 130 g, which may be larger than a *medium banana* portion size provided by the other apps. A large difference was found for the Nature Valley granola bar, with average caloric difference of −8.4 kcal (−4.2%). Interestingly, most apps showed better consistency with the bar's food label (190 kcal) rather than the USDA database (203 kcal). The food label indicated not only slightly lower calories but also lower protein (3 g vs 4 g) and fat content (7 g vs 9 g) compared with the USDA database. For the granola bar, again LifeSum showed the greatest discrepancy in calories compared with USDA. McDonald's Big Mac had the largest difference among the 3 food items, with average difference of −7.9% in calories. All the apps tended to underestimate calories and macronutrients compared with the USDA. Nutrient data found on McDonald's website also had lower values for this food item than USDA's data. In total, 2 of the apps were consistent with the nutrition data provided by McDonald's, whereas the others were not.

Figure 5. Difference in macronutrient measurements from 3-day diets in 1 week. USDA: United States Department of Agriculture.

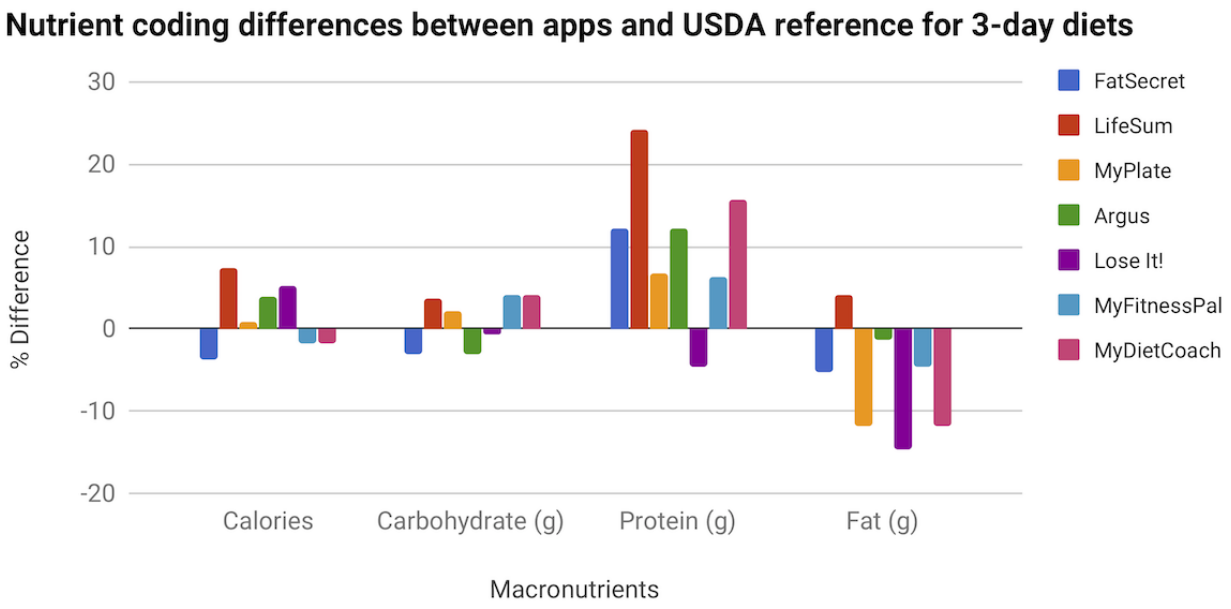


Table 2. Example of nutrient coding differences between apps for 3 food items.

Food item	FatSecret	LifeSum	MyPlate	Argus	Lose It!	MyFitnessPal	MyDietCoach	United States Department of Agriculture	Food label
Banana (1 medium, ripe)[a]									
Calories (kcal)	105	131	105	105	105	105	105	105	—[b]
Carbohydrate (g)	26.95	29.7	26.95	27	27	27	26.95	27	—
Protein (g)	1.29	1.4	1.29	1.3	1.3	1.3	1.29	1	—
Fat (g)	0.39	0.7	0.39	0.4	0.4	0.4	0.39	0	—
Nature Valley (plain, crunchy, 1 package/2 bar)[c]									
Calories (kcal)	190	220	190	192	190	190	190	203	190
Carbohydrate (g)	29	29	29	27.1	29	29	29	28	29
Protein (g)	4	4	4	3.4	3	4	4	4	3
Fat (g)	6	11	6	7.2	7	6	6	9	7
McDonald's Big Mac (1 sandwich)[d]									
Calories (kcal)	530	518	540	563	540	550	530	585	540
Carbohydrate (g)	47	45	47	44	46	46	47	48	47
Protein (g)	24	24	25	25.9	25	25	24	27	25
Fat (g)	27	29	28	32.8	28	29	27	32	28

[a]LifeSum only had 1 banana (130 g) food item; Argus and MyFitnessPal had multiple verified banana items.

[b]Not applicable.

[c]LifeSum had multiple food items for *Nature valley crunchy*; the first was used.

[d]MyPlate and Argus contained multiple food items for Big Mac; the first was used. MyFitnessPal had multiple food items for Big Mac; the fourth one with the verified restaurant label that was not the United Kingdom or Canadian version was used.

Discussion

Principal Findings

Our review found that current diet-tracking apps generally scored well in terms of usability with 4 apps having SUS scores of 70 or above. Usability may affect compliance and the willingness of users to use these apps in behavior change studies [2]. In this study, for the 3-day diet record, we entered food items into the apps as well as the USDA website and found that the apps were easier to use compared with traditional diet coding

approaches. Notably, the apps make use of features such as barcode scanning, photo entry, lists of frequently entered food items, and *auto-completion* —a short list of suggestions of food items by entering only a few letters of a food item word, which greatly increase the ease in which items can be entered. The databases in some of the apps may provide more convenient nutrient coding. For example, 1 of our coders tried entering a *tofu rice bowl* —an integrated food item made up of multiple components. For the USDA database, coders would need to estimate the portion of, and input each individual component, including condiments into the database, whereas the single *tofu rice bowl* item was already in the app's database. If users take the time to enter each component of a composite food item, and do so completely, energy and nutrient composition may be more accurately assessed than simply selecting a generic composite food item because there may be large variations in how different restaurants make the dish.

In evaluating apps, we noticed some inconsistencies between iOS and Android versions of the same app, which can affect both usability and features. These differences may exist because the iOS and Android platforms and their underlying user interfaces are inherently different. The operating systems also have different feature sets and application programming interfaces. Furthermore, the apps are typically coded in different programming languages (Java for Android and Objective C for iOS). Although there has been progress in the development of application frameworks, such as those that leverage HTML5, that allow for cross-platform app development, not all the diet-tracking apps may rely on these frameworks or the frameworks may still allow for platform-specific design choices that affect use. Due to the potential difference in iOS versus Android versions, researchers should carefully evaluate both versions of a diet-tracking app to ensure that they have similar features before using the app for a behavior change study.

Our findings differ slightly from the previous diet-related app review study [11], which found that the apps generally lacked adherence to behavior change theory. In our study, we found that some elements of behavior change theory are beginning to be implemented in some of the most popular diet-tracking apps. For instance, all of the apps we reviewed promote self-efficacy by allowing users to track their diet progress and work toward simple personal goals. Described by Bandura in 1977 [31], self-efficacy is a key behavioral theory that describes a person's belief and expectations that they can accomplish certain tasks, which could include meeting dietary goals (eg, weight loss, balanced diets, and following certain diet patterns). However, despite the presence of features in apps that could potentially be used to improve self-efficacy, the evidence regarding the effectiveness for diet-tracking apps is still limited. Future studies, particularly those conducted in real-world contexts will need to evaluate whether users' self-efficacy in meeting their dietary goals is enhanced by using certain apps.

The presence of behavior change features generally tended to be correlated with higher usability. This is encouraging, as apps may be selected by researchers for their studies because they adhere to known and specific behavior change constructs. At the same time, having specific behavior change functions does not necessarily make the apps less enjoyable to use for study participants. However, not all the features were positively correlated with usability. Notably, the *Reinforcement* domain, which in our case, related to user rewards such as giving stars, accolades, achievements, etc, was slightly negatively associated with usability. Although gamification is generally thought of as a powerful behavior change strategy used within technology [32], it may be that game-like features within diet-tracking apps may detract from the core functionality of recording diet and observing nutrition estimates. Further research may be needed to explore the role of reinforcement from diet-tracking apps in intervention studies.

The goal of this study was not to evaluate the effectiveness of specific behavior change features in altering diet; however, we did observe considerable variations between the diet-tracking apps, which make selecting particular apps for intervention studies an important consideration. No app is inherently better than another in terms of features, but instead, certain apps might be more appropriate for use within particular studies because they are stronger in specific behavior change domains than others and better fit an overall intervention strategy. Behaviorists may choose certain apps because they reinforce the domains of focus for their health education and other interactions with subjects. In doing so, having their subjects use an appropriate app could be a complement to in-person work. The app may bridge the gaps between in-person visits, allowing the subjects to explore their own diet activity [8,33,34]. Despite these opportunities, a survey conducted of the Australian, New Zealand, and British dietetic association members (ie, registered dietitians and practitioners) found that although diet app usage is high, the apps have yet to be fully integrated into nutritional care practice or within behavior change programs [35]. Figure 3 and the detailed feature checklist results in Multimedia Appendix 1 may help behaviorists identify appropriate apps. The American Academy of Nutrition and Dietetics regularly reviews individual apps, which may be an additional resource [36].

Although it is important to acknowledge variations between the apps, the presence and absence of features within certain apps are noteworthy. For instance, LifeSum scored the highest in usability and had a unique feature—a Health Test. We did not evaluate the accuracy of their test, but the concept of integrating routine knowledge/belief assessments, while tracking its effect on changes in the users' diet tracking, has the potential to be a useful intervention tool. We also noticed that features related to emotions were missing from all the apps despite considerable research, which has identified the associations between both positive and negative emotions and diet [37-41]. Moreover, tracking of potential upstream determinants and downstream effects of diet (eg, hunger, satiety, guilt, stress, happiness, and taste) tend be to lacking from these diet-tracking apps but would potentially be useful to track to assess associations with diet patterns. These missing behavioral domains could be an important future area for diet-tracking app development.

As part of our testing, we evaluated the accuracy of a 3-day diet record compared with the USDA reference. Generally, we observed that tracking of calories and carbohydrates closely matched estimates from the USDA database; however, there were large inconsistencies between apps with respect to protein

and fat estimates. This is somewhat surprising as proteins and fats are both components of caloric intake. We suspect the underestimation of fat may be due to the difficulties in estimating oils used in cooking, particularly for restaurant meals. These differences could have implications for those with specific health issues. For example, for individuals with cardiovascular disease who rely on such apps to track their fat intake, underestimation of their fat intake could contribute to high blood cholesterol levels and exacerbate their health conditions. This led us to take a closer look at select food items to better understand how the apps may vary from the USDA reference. Notably 2 of our items, the granola bar and fast food sandwich, have food labels, and we found that the USDA nutritional values differed from the labeling. In some cases, we observed that although the apps were not consistent with USDA, they were consistent with the food label. We found the unlabeled item, the banana, to be more consistent with USDA than the other 2 items. Perhaps food reformulations may make some databases, including USDA's, more easily outdated compared with unprocessed food items. Furthermore, because some of the apps rely upon user suggestions for nutrient content, it may be that more users submit requests for updated caloric content for certain macronutrients for food items.

Interestingly, the app with highest usability, LifeSum, had the greatest nutrient coding inconsistency compared with the USDA reference. This presents a potential challenge for interventions, which rely upon accurate estimates from apps, instead of having trained staff recode the food items using a standard reference database. We conducted additional research to learn about the database used by LifeSum and found that it uses a series of databases, which includes the USDA database, as well as MyNetDiary, UK Food Standards Agency (United Kingdom), Bundeslebensmittelschlüssel (Germany), Livsmedelverket (Sweden), and food contributed by users (new food items created from users by entering nutrient details) [42]. LifeSum's headquarters is located in Sweden, and perhaps because of this, it may be more accommodating of European users. The app's nutrient coding may be more accurate when compared with the European databases, although we did not perform this evaluation. As another example, we researched the development of MyFitnessPal (the app we found to be most consistent with USDA's database). The app initially started with its creator entering food items, and later, relied upon user crowdsourcing (ie, users entering the nutritional content of food items into their database) [43]. Consistent with other studies, despite underestimation of nutrients, MyFitnessPal had the highest correlation with the USDA database [44-46]. Of note, the study by Chen et al is 1 of the only studies in which participants recruited from the general population of smartphone users in Australia, who had not used an app (MyFitnessPal), were asked to use the app to record their dietary intake in a close to real-world context (ie, they were asked to install and use the app, were recruited specifically for the study, which also included phone-based 24-hour recalls) and found that daily energy intake was significantly underestimated by 445 kcal [47], whereas studies similar to ours [45], which rely upon users with nutritional training to use the app are likely to observe more complete and accurate results. Teixeira et al [46] recruited non-nutrition major university students in a study comparing paper versus app (MyFitnessPal) in Brazil and found that while many nutrients were underestimated compared with paper records, there was moderate correlation between both methods. Griffiths et al [45] who relied upon research staff to dietary recall data into different tracking apps also observed significant differences in app-reported nutrient levels compared with the standard they used for their research (the Nutrition Data System for Research). Due to these inconsistencies, Chen et al [47] recommended the use of apps with guidance from dietitians if more accurate dietary data are required.

With the inclusion of elements of behavior change theory into popular diet-tracking apps, there may be reasons to incorporate apps into interventions that leverage these features. However, there could also be disadvantages of using smartphone diet/nutrition apps. These include increased screen time, usability and acceptance issues, as well as concerns about users' privacy [48]. In addition, there is some emerging evidence that suggests that some diet-tracking apps may be frequently used by those with eating disorders, and their use may be perceived as contributing to their disorder [49]. For individuals with eating disorders, caution need to be exercised for the use of diet/nutrition apps, which may not be appropriate replacements for clinical treatment or medical monitoring [48,50].

Limitations

There are potential limitations of our research that are worth noting. First, the apps were reviewed by 3 college-aged authors who are technology-savvy and trained in nutritional science. This was necessary to conduct an accurate review of the apps' features and accurate diet coding. However, their usability scores may not generalize to populations who do not have training in nutrition. Due to digital entitlement [51], certain populations do not have access to smart devices, and it cannot be assumed that all users will be similarly comfortable with using these apps. We did not conduct any inter- or intrarater reliability assessment for the 3 reviewers and the measurement tools used. Future studies should include both inter- and intrarater reliability assessments and more users.

Of the 7 apps we evaluated, 5 originated from the United States. These apps may be using US food databases only, which may not be applicable in other countries because of the differences in food processing, regulations, and policies.

There may also be issues generalizing some of the 3-day diet findings to what might occur in a more general population. We conducted diet assessment of 3 consecutive days because of its relative ease and less reviewer burden. Consecutive days may limit the variation in the food intake compared with nonconsecutive days; however, we did include 2 weekdays and 1 weekend day to capture more variety in the authors' food intake. The authors' diet coding may be more accurate than what users in general may be able to achieve, as the authors have a better idea of the caloric content of certain foods, allowing them to quickly catch coding errors. In addition, they may be able to estimate serving sizes better compared with normal users. Another potential issue that may affect broader generalization is that the average caloric intake of the 3 authors during the 3-day period was lower than the average caloric intake for US adults, which is 2091 kcal per day for 2007-2010

[30]. If the caloric intake was higher on the 3 days, the discrepancies between the apps and the USDA reference may potentially have been higher. Furthermore, energy underreporting is common with diet assessment, which may compound the problem of underestimation of food intake, particularly with apps as has been found by others [47]. However, other researchers have noted that although self-reported energy intake may not accurately reflect true energy intake, self-report methods can still provide valuable information about foods and beverages consumed by populations, which can be used to inform nutritional policy and associations between diet and disease [52]. The origin and types of food items (Mexican banana vs banana from other origins) may contribute to their nutrient and caloric variabilities. However, we did not include this information in our nutrient coding.

This is the first use of a behavior change checklist to assess diet-tracking apps, and although questions were developed through an iterative process to align them with a behavior change framework, the checklist has not been thoroughly validated. Despite these limitations, our review provides a broad assessment of the potential use of the current generation of diet-tracking apps for diet intervention studies. Given the popularity of these apps, further research to evaluate the effectiveness of interventions that use these apps is warranted.

Conclusions

This study showed that the 7 most popular diet-tracking apps on Apple App Store and Google Play are feature-rich and easy to use. The apps incorporate features consistent with many behavior change domains, notably promoting self-efficacy through tracking diet and progress toward goals. Although the relatively large deviations in coding of protein and fat compared with the USDA reference deserve further examination, the apps performed similarly to the USDA in coding calories and carbohydrates. Together, these aspects allow these diet-tracking apps to be useful for a wide range of dietary intervention studies.

Acknowledgments

Funding for this study was provided by NIH NIMHD grant U54 MD011240 and NIEHS grant 5P30 ES007033-23. JH was supported by NIH NHLBI T32 HL007034.

Authors' Contributions

GF, JK, and SL contributed to study design, data collection, data analyses, and writing. JH contributed to study design and writing. ES contributed to study design, data analyses, and writing.

Conflicts of Interest

None declared.

Multimedia Appendix 1
App feature checklist and corresponding Theoretical Domains Framework (TDF) domain and construct (discordant indicates differences between iOS and Android versions).
[DOCX File , 28 KB - mhealth_v7i5e9232_app1.docx]

Multimedia Appendix 2
System Usability Scale scores.
[DOCX File , 15 KB - mhealth_v7i5e9232_app2.docx]

Multimedia Appendix 3
Correlation coefficients, standard errors, and *P* values, and R2 for linear models of usability versus TDF domain features.
[DOCX File , 14 KB - mhealth_v7i5e9232_app3.docx]

References

1. Pohl M. Research 2 Guidance.: research2guidance; 2017. 325,000 mobile health apps available in 2017 - Android now the leading mHealth platform URL: https://research2guidance.com/325000-mobile-health-apps-available-in-2017/ [accessed 2018-08-08] [WebCite Cache ID 71Ze5E0an]
2. Chen J, Cade J, Allman-Farinelli M. The most popular smartphone apps for weight loss: a quality assessment. JMIR Mhealth Uhealth 2015 Dec 16;3(4):e104 [FREE Full text] [doi: 10.2196/mhealth.4334] [Medline: 26678569]
3. Azar K, Lesser L, Laing B, Stephens J, Aurora M, Burke L, et al. Mobile applications for weight management: theory-based content analysis. Am J Prev Med 2013 Nov;45(5):583-589. [doi: 10.1016/j.amepre.2013.07.005] [Medline: 24139771]
4. Davis SF, Ellsworth MA, Payne HE, Hall SM, West JH, Nordhagen AL. Health behavior theory in popular calorie counting apps: a content analysis. JMIR Mhealth Uhealth 2016 Mar 02;4(1):e19 [FREE Full text] [doi: 10.2196/mhealth.4177] [Medline: 26935898]

5. Fakih El Khoury C, Karavetian M, Halfens R, Crutzen R, Khoja L, Schols J. The effects of dietary mobile apps on nutritional outcomes in adults with chronic diseases: a systematic review and meta-analysis. J Acad Nutr Diet 2019 Apr;119(4):626-651. [doi: 10.1016/j.jand.2018.11.010] [Medline: 30686742]
6. Plow M, Golding M. Using mHealth technology in a self-management intervention to promote physical activity among adults with chronic disabling conditions: randomized controlled trial. JMIR Mhealth Uhealth 2017 Dec 01;5(12):e185 [FREE Full text] [doi: 10.2196/mhealth.6394] [Medline: 29196279]
7. Bardus M, van Beurden SB, Smith JR, Abraham C. A review and content analysis of engagement, functionality, aesthetics, information quality, and change techniques in the most popular commercial apps for weight management. Int J Behav Nutr Phys Act 2016;13(1):35 [FREE Full text] [doi: 10.1186/s12966-016-0359-9] [Medline: 26964880]
8. DiFilippo KN, Huang W, Andrade JE, Chapman-Novakofski KM. The use of mobile apps to improve nutrition outcomes: a systematic literature review. J Telemed Telecare 2015 Jul;21(5):243-253. [doi: 10.1177/1357633X15572203] [Medline: 25680388]
9. McKay FH, Cheng C, Wright A, Shill J, Stephens H, Uccellini M. Evaluating mobile phone applications for health behaviour change: a systematic review. J Telemed Telecare 2018 Jan;24(1):22-30. [doi: 10.1177/1357633X16673538] [Medline: 27760883]
10. 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]
11. West J, Hall P, Arredondo V, Berrett B, Guerra B, Farrell J. Health behavior theories in diet apps. J Consum Health Internet 2013 Jan;17(1):10-24. [doi: 10.1080/15398285.2013.756343]
12. Ogden C, Carroll M, Kit B, Flegal K. Prevalence of childhood and adult obesity in the United States, 2011-2012. J Am Med Assoc 2014 Feb 26;311(8):806-814 [FREE Full text] [doi: 10.1001/jama.2014.732] [Medline: 24570244]
13. James W. WHO recognition of the global obesity epidemic. Int J Obes (Lond) 2008 Dec;32(Suppl 7):S120-S126. [doi: 10.1038/ijo.2008.247] [Medline: 19136980]
14. Popkin B, Adair L, Ng S. Global nutrition transition and the pandemic of obesity in developing countries. Nutr Rev 2012 Jan;70(1):3-21 [FREE Full text] [doi: 10.1111/j.1753-4887.2011.00456.x] [Medline: 22221213]
15. Young L, 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]
16. McKeown N, Meigs J, Liu S, Saltzman E, Wilson P, Jacques P. Carbohydrate nutrition, insulin resistance, and the prevalence of the metabolic syndrome in the Framingham Offspring Cohort. Diabetes Care 2004 Feb;27(2):538-546. [Medline: 14747241]
17. Elliott P. Observational studies of salt and blood pressure. Hypertension 1991 Jan;17(1 Suppl):I3-I8. [Medline: 1987009]
18. Drisko J, Bischoff B, Hall M, McCallum R. Treating irritable bowel syndrome with a food elimination diet followed by food challenge and probiotics. J Am Coll Nutr 2006 Dec;25(6):514-522. [Medline: 17229899]
19. Bergmann M, Caubet J, Boguniewicz M, Eigenmann P. Evaluation of food allergy in patients with atopic dermatitis. J Allergy Clin Immunol Pract 2013 Jan;1(1):22-28. [doi: 10.1016/j.jaip.2012.11.005] [Medline: 24229818]
20. Saccomani P. MobiLoud. 2016 Apr 03. What Factors Really Impact Your App Store Ranking? URL: https://www.mobiloud.com/blog/factors-really-impact-app-store-ranking/ [accessed 2017-10-04] [WebCite Cache ID 786oNvYfW]
21. Walz A. MOZ. 2015 May 27. Deconstructing the App Store Rankings Formula with a Little Mad Science URL: https://moz.com/blog/app-store-rankings-formula-deconstructed-in-5-mad-science-experiments [accessed 2019-05-03] [WebCite Cache ID 786ojvcsx]
22. Brooke J. Usability evaluation in industry. 1996. SUS-A quick and dirty usability scale URL: https://hell.meiert.org/core/pdf/sus.pdf [accessed 2019-05-07] [WebCite Cache ID 78BZEkej7]
23. Zapata B, Fernández-Alemán JL, Idri A, Toval A. Empirical studies on usability of mHealth apps: a systematic literature review. J Med Syst 2015 Feb;39(2):1. [doi: 10.1007/s10916-014-0182-2] [Medline: 25600193]
24. Brooke J. SUS: a retrospective. J Usability Stud 2013;8(2):29-40.
25. Liang J, Xian D, Liu X, Fu J, Zhang X, Tang B, et al. Usability study of mainstream wearable fitness devices: feature analysis and system usability scale evaluation. JMIR Mhealth Uhealth 2018 Nov 08;6(11):e11066 [FREE Full text] [doi: 10.2196/11066] [Medline: 30409767]
26. Peres SC, Pham T, Phillips R. Validation of the System Usability Scale (SUS): SUS in the Wild. Proc Hum Factors Ergon Soc Annu Meet 2013 Sep 30;57(1):192-196. [doi: 10.1177/1541931213571043]
27. System Usability Scale (SUS). URL: https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html [accessed 2019-05-03] [WebCite Cache ID 786pdFHzK]
28. Cane J, O'Connor D, Michie S. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci 2012 Apr 24;7:37 [FREE Full text] [doi: 10.1186/1748-5908-7-37] [Medline: 22530986]
29. Prochaska JO, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promot 1997;12(1):38-48. [Medline: 10170434]

30. United States Department of Agriculture. 2016 Oct 20. Food Consumption and Nutrient Intakes URL: https://www.ers.usda.gov/data-products/food-consumption-and-nutrient-intakes/food-consumption-and-nutrient-intakes/ [accessed 2019-05-03] [WebCite Cache ID 786q2yaB4]
31. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [Medline: 847061]
32. Cugelman B. Gamification: what it is and why it matters to digital health behavior change developers. JMIR Serious Games 2013 Dec 12;1(1):e3 [FREE Full text] [doi: 10.2196/games.3139] [Medline: 25658754]
33. Hale K, Capra S, Bauer J. A framework to assist health professionals in recommending high-quality apps for supporting chronic disease self-management: illustrative assessment of type 2 diabetes apps. JMIR Mhealth Uhealth 2015;3(3):e87 [FREE Full text] [doi: 10.2196/mhealth.4532] [Medline: 26369346]
34. Jospe MR, Fairbairn KA, Green P, Perry TL. Diet app use by sports dietitians: a survey in five countries. JMIR Mhealth Uhealth 2015 Jan 22;3(1):e7 [FREE Full text] [doi: 10.2196/mhealth.3345] [Medline: 25616274]
35. Chen J, Lieffers J, Bauman A, Hanning R, Allman-Farinelli M. The use of smartphone health apps and other mobile health (mHealth) technologies in dietetic practice: a three country study. J Hum Nutr Diet 2017 Dec;30(4):439-452. [doi: 10.1111/jhn.12446] [Medline: 28116773]
36. Academy of Nutrition and Dietetics. Eat Right Pro Trends and Reviews URL: http://www.eatrightpro.org/resources/media/trends-and-reviews/app-reviews?p=3 [accessed 2019-05-04] [WebCite Cache ID 786qWOuYw]
37. Bongers P, Jansen A, Havermans R, Roefs A, Nederkoorn C. Happy eating: the underestimated role of overeating in a positive mood. Appetite 2013 Aug;67:74-80. [doi: 10.1016/j.appet.2013.03.017] [Medline: 23583314]
38. Hou F, Xu S, Zhao Y, Lu Q, Zhang S, Zu P, et al. Effects of emotional symptoms and life stress on eating behaviors among adolescents. Appetite 2013 Sep;68:63-68. [doi: 10.1016/j.appet.2013.04.010] [Medline: 23603765]
39. Konttinen H, Männistö S, Sarlio-Lähteenkorva S, Silventoinen K, Haukkala A. Emotional eating, depressive symptoms and self-reported food consumption. A population-based study. Appetite 2010 Jun;54(3):473-479. [doi: 10.1016/j.appet.2010.01.014] [Medline: 20138944]
40. Loxton N, Dawe S, Cahill A. Does negative mood drive the urge to eat? The contribution of negative mood, exposure to food cues and eating style. Appetite 2011 Apr;56(2):368-374. [doi: 10.1016/j.appet.2011.01.011] [Medline: 21238524]
41. Tryon M, Carter C, Decant R, Laugero K. Chronic stress exposure may affect the brain's response to high calorie food cues and predispose to obesogenic eating habits. Physiol Behav 2013 Aug 15;120:233-242. [doi: 10.1016/j.physbeh.2013.08.010] [Medline: 23954410]
42. Lifesum. 2016 Oct 7. Lifesum's food database URL: https://lifesum.helpshift.com/a/lifesum---healthier-eating-better-living/ [accessed 2019-05-04] [WebCite Cache ID 786quYNRi]
43. MacManus R. Readwrite. 2015 Oct 7. How MyFitnessPal Became the King of Diet Trackers URL: https://readwrite.com/2015/02/23/trackers-myfitnesspal-excerpt/ [accessed 2019-05-04] [WebCite Cache ID 786r5cxfV]
44. Chen J, Berkman W, Bardouh M, Ng CY, Allman-Farinelli M. The use of a food logging app in the naturalistic setting fails to provide accurate measurements of nutrients and poses usability challenges. Nutrition 2018 May 25;57:208-216. [doi: 10.1016/j.nut.2018.05.003] [Medline: 30184514]
45. Griffiths C, Harnack L, Pereira M. Assessment of the accuracy of nutrient calculations of five popular nutrition tracking applications. Public Health Nutr 2018 Jan;21(8):1495-1502. [doi: 10.1017/S1368980018000393] [Medline: 29534771]
46. Teixeira V, Voci SM, Mendes-Netto RS, da Silva DG. The relative validity of a food record using the smartphone application MyFitnessPal. Nutr Diet 2018 Apr;75(2):219-225. [doi: 10.1111/1747-0080.12401] [Medline: 29280547]
47. Chen J, Berkman W, Bardouh M, Ng CY, Allman-Farinelli M. The use of a food logging app in the naturalistic setting fails to provide accurate measurements of nutrients and poses usability challenges. Nutrition 2018 May 25;57:208-216. [doi: 10.1016/j.nut.2018.05.003] [Medline: 30184514]
48. Tregarthen JP, Lock J, Darcy AM. Development of a smartphone application for eating disorder self-monitoring. Int J Eat Disord 2015 Nov;48(7):972-982. [doi: 10.1002/eat.22386] [Medline: 26213130]
49. Levinson C, Fewell L, Brosof L. My Fitness Pal calorie tracker usage in the eating disorders. Eat Behav 2017 Jan;27:14-16 [FREE Full text] [doi: 10.1016/j.eatbeh.2017.08.003] [Medline: 28843591]
50. Eikey EV, Reddy MC. "It's Definitely Been a Journey": A Qualitative Study on How Women with Eating Disorders Use Weight Loss Apps. In: Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. 2017 Presented at: CHI '17; May 06 - 11, 2017; Denver, Colorado p. 642-654. [doi: 10.1145/3025453.3025591]
51. Mansell R. From digital divides to digital entitlements in knowledge societies. Curr Sociol 2016 Jun 30;50(3):407-426. [doi: 10.1177/0011392102050003007]
52. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, et al. Addressing current criticism regarding the value of self-report dietary data. J Nutr 2015 Dec;145(12):2639-2645 [FREE Full text] [doi: 10.3945/jn.115.219634] [Medline: 26468491]

Abbreviations

iOS: iPhone operating system

mHealth: mobile health
SUS: System Usability Scale
TDF: Theoretical Domains Framework
USDA: United States Department of Agriculture

Edited by G Eysenbach; submitted 09.12.17; peer-reviewed by J Chen, K Davison, I Gabashvili, A Paglialonga; comments to author 15.03.18; revised version received 10.08.18; accepted 10.04.19; published 17.05.19.

Please cite as:
Ferrara G, Kim J, Lin S, Hua J, Seto E
A Focused Review of Smartphone Diet-Tracking Apps: Usability, Functionality, Coherence With Behavior Change Theory, and Comparative Validity of Nutrient Intake and Energy Estimates
JMIR Mhealth Uhealth 2019;7(5):e9232
URL: http://mhealth.jmir.org/2019/5/e9232/
doi:10.2196/mhealth.9232
PMID:31102369

Original Paper

A Qualitative Evaluation of the Acceptability of a Tailored Smartphone Alcohol Intervention for a Military Population: Information About Drinking for Ex-Serving Personnel (InDEx) App

Jo-Anne Puddephatt[1], BSc (Hons), MSc; Daniel Leightley[2], BSc (Hons), MSc, PhD; Laura Palmer[2], BA, MSc; Norman Jones[3], PhD; Toktam Mahmoodi[4], BSc (Hons), PhD; Colin Drummond[5,6], FRCPsych, MD; Roberto J Rona[2], PhD, FFPH; Nicola T Fear[2,3], BSc (Hons), MSc, DPhil (Oxon); Matt Field[7], BSc (Hons), PhD; Laura Goodwin[1], BSc (Hons), MSc, PhD

[1]Department of Psychological Sciences, Institute of Life and Human Sciences, University of Liverpool, Liverpool, United Kingdom

[2]King's Centre for Military Health Research, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

[3]Academic Department of Military Mental Health, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

[4]Department of Informatics, King's College London, London, United Kingdom

[5]Addictions Department, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, United Kingdom

[6]South London and Maudsley NHS Foundation Trust, London, United Kingdom

[7]Department of Psychology, University of Sheffield, Sheffield, United Kingdom

Corresponding Author:
Jo-Anne Puddephatt, BSc (Hons), MSc
Department of Psychological Sciences
Institute of Life and Human Sciences
University of Liverpool
Eleanor Rathbone Building
Bedford Street South
Liverpool, L69 7ZA
United Kingdom
Phone: 44 151 794 1401
Email: joannep@liverpool.ac.uk

Abstract

Background: Alcohol consumption in the UK Armed Forces is higher than in the general population, and this pattern continues after leaving the service. Smartphone apps may be useful to increase ex-serving personnel's awareness of their alcohol consumption, support self-monitoring, and prompt a change in behavior.

Objective: The study aimed to explore the acceptability of *Information about Drinking in Ex-serving personnel* (InDEx), a tailored smartphone app, combined with personalized short message service (SMS) text messaging designed to target ex-serving personnel who meet the criteria for hazardous alcohol use.

Methods: The InDEx intervention included 4 key modules: (1) assessment and normative feedback, (2) self-monitoring and feedback, (3) goal setting and review, and (4) personalized SMS text messaging. A semistructured telephone interview study was conducted with ex-serving personnel after using the app for a 28-day period. Interviews were used to explore the acceptability of app modules and its functionality and the perceived changes in participant's drinking. Interview transcripts were analyzed using inductive thematic analysis.

Results: Overall, 94% (29/31) participants who used InDEx agreed to take part in a telephone interview. Overall, 4 themes were identified: *Credibility*, *Meeting their needs*, *Simplicity*, and *Helpful for ex-serving personnel*. The importance of credibility, functionality, and meeting the individual needs of ex-serving personnel was emphasized. Acceptability and engagement with specific modules of the app and text messages were influenced by the following: (1) if they felt it provided credible information, (2) whether the content was appropriately personalized to them, (3) the ease of use, and (4) beliefs about their own drinking behaviors. Participants recommended that the app would be most suitable for personnel about to leave the Armed Forces.

Conclusions: InDEx was an acceptable smartphone app for ex-serving personnel for monitoring alcohol consumption and in providing meaningful feedback to the individual. Pages that met the participant's interests and provided real time personalized, credible feedback on their drinking and text messages tailored to participant's interactions with the app were particularly favored.

(JMIR Mhealth Uhealth 2019;7(5):e12267) doi:10.2196/12267

KEYWORDS

behavior; alcohol drinking; veterans; interview; smartphone

Introduction

Background

The prevalence of hazardous drinking within the UK Armed Forces is higher than in the general population [1], and this pattern continues after personnel leave service [2]. There is little research on drinking motivations in military populations, but some evidence indicates drinking as a maladaptive coping response to psychological distress [3] and for social and physical pleasure [4]. Research has shown that there is a perceived stigma around reporting alcohol problems [5], which may be a barrier to face-to-face help seeking [6]. There are also issues around individuals recognizing their drinking being a problem [7], specifically in the military where approximately one-third of those who meet the criteria for harmful alcohol use recognize a problem [8]. Given the high prevalence of hazardous drinking within the Armed Forces and the potential for additional pressures when transitioning to civilian life, it is important to identify a suitable intervention that can target alcohol use in this population.

Web- and smartphone app–based interventions (eg, Down Your Drink [9], A-CHESS [10], DrinkLess [11], and LBMI-A [12]), through formative methods, have been harnessed to increase the reach and provide real-time monitoring and personalized delivery. Self-monitoring and real-time information about one's drinking through the use of an app may help individuals identify if they have a problem.

Recent research indicates that smartphone alcohol app interventions may be effective for changing drinking behavior, for example, a systematic review of smartphone alcohol app interventions for adults found that 5 of the 6 studies reported significant changes in alcohol consumption [13]. Furthermore, digital alcohol interventions have been found to be as effective as face-to-face interventions [14]. Present alcohol apps have focused on self-monitoring, which is known to lead to behavior change [15,16], but a review of which behavior change techniques (BCTs) work in alcohol interventions has indicated that credible sources (communication from a credible source in favor of or against behaviors [17]) or suggesting alternative behaviors may also be effective techniques [16].

Although some alcohol apps have been shown to be effective, there are some issues to consider. One example is maintaining user engagement where research has shown that users engage less as the study progresses [18,19]. One possible way in which monitoring can be maintained is through the use of text messages encouraging adherence [20,21]. Another issue is the extent to which a theory is used to develop alcohol app interventions [22-24]; research has shown that more than half of digital alcohol interventions did not mention the use of theory [14,25] and only 38% used theory to inform intervention development [25]. Therefore, it is difficult to understand if it is effective in changing drinking behavior [14].

Given the prevalence of hazardous drinking in the Armed Forces and the potential for smartphone apps and text messages to target hard-to-reach populations and prompt changes in drinking behavior, we developed a theory-based intervention to target hazardous and harmful drinking in ex-serving military personnel which also incorporated BCTs: *Information about Drinking in Ex-serving personnel* (InDEx) [26,27]. InDEx is aimed at ex-serving personnel who are likely to be experiencing short-term consequences of their drinking but might not be aware of this being a problem. To the best of our knowledge, this is the first smartphone-based alcohol intervention that uses both an app and personalized and tailored text messaging.

We conducted a feasibility study with ex-serving personnel to explore adherence and engagement with InDEx over a 28-day period. Our usability findings have been published elsewhere [27], but briefly, we found good engagement (having at least 3 client-server interactions in a 7-day period) throughout this period, with 23 out of 31 participants using the app every week and 27 using the app in the final week. Self-monitoring and feedback modules were most frequently used compared with goal setting and normative feedback [27].

Although usability data from the app inform intervention developers of how participants engaged with certain features, it does not necessarily tell us why this occurred [28]. Using qualitative methods allows us to understand our usability findings further in addition to providing insight into the acceptability of the app for ex-serving personnel, particularly as this is the first of its kind to target alcohol use in this population. For the purpose of this study, we define acceptability in terms of how useful participants perceived InDEx for ex-serving personnel and whether it was useful for monitoring alcohol consumption in terms of positive and negative comments [29].

Objectives

The aim of this study, therefore, was to (1) explore the acceptability of InDEx for an ex-serving population based on their experience of using the app over a 28-day period and (2) explore participant's experience of using key modules to monitor their alcohol consumption using qualitative methods.

Methods

Ethical approval was obtained from the local research ethics committee at the University of Liverpool (reference: #0625).

Information About Drinking in Ex-Serving Personnel App

InDEx used agile development methodologies with each cycle, focusing firstly on the development and secondly on stakeholder and expert user testing [26] (see Figure 1).

InDEx was targeted to ex-serving personnel who had recently left the military. The app was informed by the Health Action Process Approach [30] and social norms theory [31], where delivery was split into 3 stages: (1) normative feedback, action self-efficacy, and self-monitoring; (2) maintenance self-efficacy and action planning; and (3) recovery self-efficacy and coping planning. The app was complemented by personalized and tailored text messaging to provide prompts to log into the app, suggest alternative behaviors, and provide feedback on goals set [26,27]. A full description of InDEx, including its source code, message bank, and rules-based approach for tailoring, is available in the study by Leightley et al [26,27]. A summary is provided hereafter.

Upon consenting to take part in the feasibility study, participants were sent a link via email to download InDEx. Participants were provided with a navigation page to explain how to use each of the following modules:

1. Assessment and normative feedback: participants were provided an infographic representing participants' self-reported alcohol consumption compared with the general population. This could be revisited at any time in the *How you compare* page under *My account*.
2. Self-monitoring and feedback: participants recorded their alcohol consumption using a drinks diary under the *Add drink* page from a list of predefined drinks; alcohol and calorie content and the price of a drink could be edited. Participants were provided feedback on their drinking under the *Dashboard* and *Drinks* pages.
3. Goal (setting and review): participants could set a goal based on implementation intentions [32] (ie, they identified the barrier and solution for each goal) using 1 of 4 options: (1) *I want to drink less on a night out*, (2) *I want to have more drink-free days during the week*, (3) *I want to spend less on alcohol this week*, or (4) *I will have a maximum of X drinks*. Goals set were viewed under the *Goals* page.
4. Personalized short message service (SMS) text messaging: a bank of 180 tailored and personalized text messages was developed and included BCTs found to be useful in alcohol interventions [16,22]. BCTs are defined as *active ingredients* of an intervention designed to change behavior [16,33,34]. The message bank and decision tree for sending text messages are available upon request from the corresponding author.

Participants

Participants in the feasibility study were recruited from the King's Centre for Military Health Research (KCMHR) Health and Wellbeing cohort [35-37], an ongoing study of UK Armed Forces personnel since 2003, including those who had left the military. Potential participants were selected from the cohort database based on meeting the following inclusion criteria: (1) were ex-serving personnel who left the Armed Forces within the last 2 years (at the time of sampling—May 2017); (2) aged 18 to 65 years; (3) who owned a smartphone; and (4) who had an Alcohol Use Disorder Identification Test (AUDIT) score of 8 to 19 at phase 2 of the cohort, which was conducted in 2009, indicating hazardous or harmful alcohol use according to the AUDIT. Those scoring above 20 on the AUDIT meet the criteria for probable alcohol dependency, and we felt that they may require more intensive treatment [27,38]. An additional criterion for the study was for participants to have taken part in the feasibility study [27].

For this study, purposive sampling was used to recruit all participants who had taken part in the feasibility study to ensure that we were able to capture all their experiences of using InDEx. Participants were invited to the interview using their email address after completing our feasibility study between June and August 2017. Participants were reimbursed £40 for using the app for a 28-day period.

Procedure

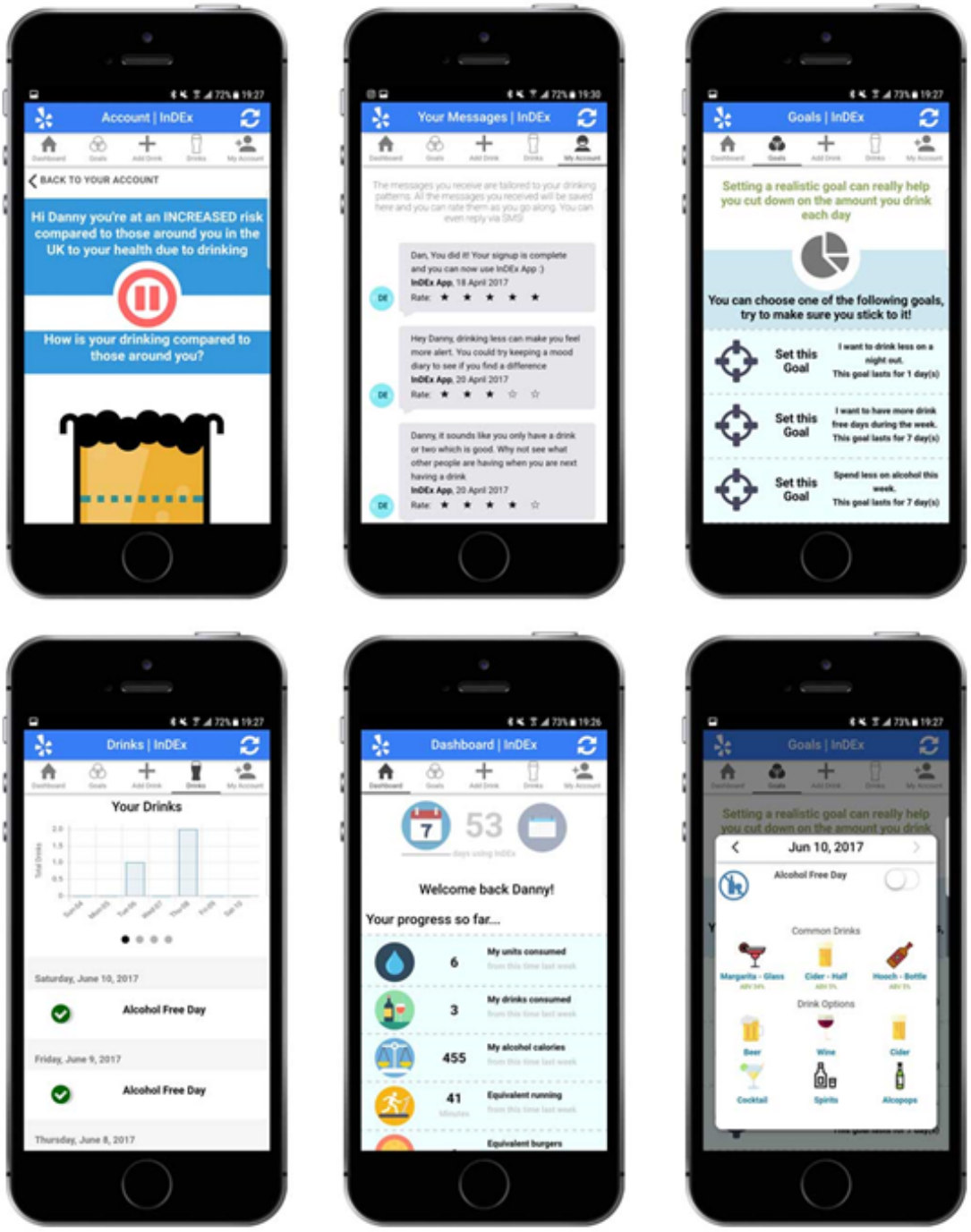
Participants were invited by JP, via email, to take part in a telephone interview up to 3 weeks after using InDEx for a 28-day period. If the participant did not respond to this initial email, a follow-up email and SMS text message was sent.

Semistructured telephone interviews were conducted by a trained postgraduate qualitative research assistant (JP); these were used to allow participants to discuss their experience of the app. No relationship was established between JP and participants before the study other than when inviting to the interview. Participants were aware that JP was a research assistant at the University of Liverpool and part of the InDEx app team. To minimize response bias, participants were briefed on the purpose of the study and were encouraged to be honest in their responses as it would be used to inform future intervention development.

The topic guide was developed by the authors and pilot-tested within the project team but did not influence interviews. Topics included the ease of use and accessibility of the app, thoughts about the content and frequency of the text messages, and perceptions about providing information about participants' drinking. The interview also included questions around the suitability of InDEx for ex-serving personnel. The topic guide can be found in Multimedia Appendix 1. Follow-up questions were allowed, and participants were able to add any additional comments that had not been covered. Field notes were also made during the interviews to inform potential follow-up questions.

Participants were informed before the interview that their responses would be audio-recorded using a digital Dictaphone and transcripts would be pseudoanonymized and stored securely. No other persons were present during the interview. Telephone interviews took place a median of 8 days after completing the feasibility study (range 2 to 23) and were conducted between June and August 2017. Repeat interviews did not take place. The mean length of the interview was 30 min and 22 seconds (range: 16 min and 38 seconds to 39 min and 37 seconds).

Figure 1. Example screenshots of interactions with the InDEx app. Top left to bottom right: normative feedback, personalized text message history, set a goal, drink diary, dashboard, and add a drink. InDEx: Information about Drinking in Ex-serving personnel.



Analysis

Interviews were audio-recorded and transcribed verbatim by a professional transcriptionist. Transcripts were checked with audio recordings to determine accuracy and were pseudoanonymized. Inductive thematic analysis was conducted to allow for the identification of patterns within data. This allows themes and codes to be strongly linked to the data, as opposed to being theory driven [39]. This involves a 5-phase process of familiarizing with the data, generating initial codes, searching for themes, reviewing themes, and finally defining and naming themes. This analytical method has been frequently used in previous studies aiming to explore the acceptability of smartphone health interventions [11,40,41].

Analysis followed an iterative process whereby themes were identified and merged or removed during the continual analysis of raw data. NVivo 10 was used to facilitate the coding process, and a subset of transcripts (N=3) were randomized and sent to a second coder (LP) along with a developed codebook to provide feedback on this codebook and code the transcripts and to establish reliability. JP reviewed the coded transcripts for coding consistency and any disagreements were reviewed by a third coder (LG). The codebook developed from our analysis is available upon request to the corresponding author.

We also triangulated our qualitative analysis with participants' AUDIT scores, interactions with InDEx, and self-reported units consumed from week 1 to week 4 of the study to provide context to their interview transcripts (see Multimedia Appendix 2).

Results

Overview

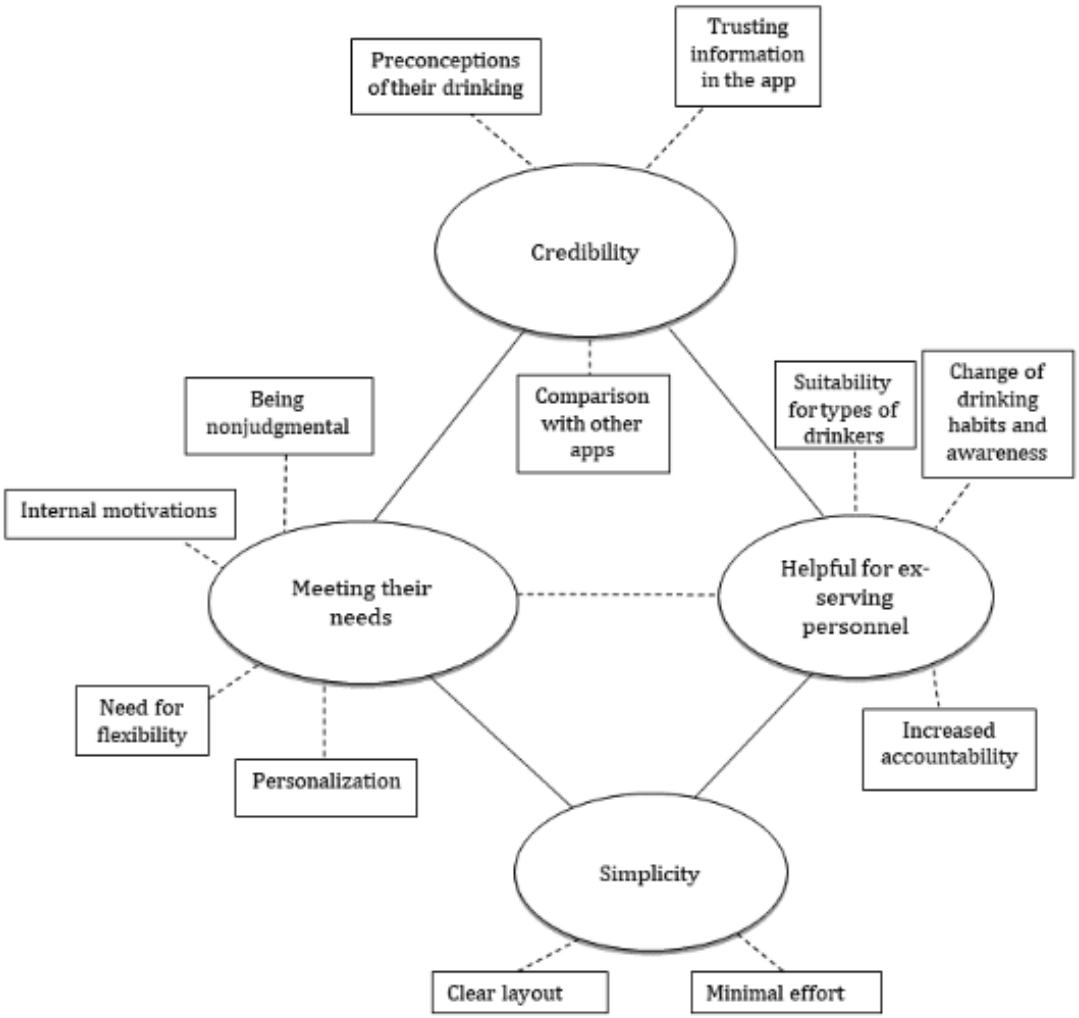
A total of 29 participants (94%, 29/31) agreed to be interviewed, of whom 25 (85%, 25/29) were male. Furthermore, 18 participants (62%, 18/29) served in the Army, 6 (21%, 6/29) in the Royal Air Force, and 5 (17%, 5/29) in the Navy. In addition, 23 participants (79%, 23/29) served as a noncommissioned officer or other rank, and 24 participants (83%, 24/29) served in the Armed Forces for more than 12 years. Moreover, 2 participants invited to take part in the interview did not respond to emails; therefore, it is not known why they did not wish to participate.

At baseline, participants reported a median AUDIT score of 11 (range 5 to 16), suggesting that some participants' AUDIT scores had decreased since taking part in the KCMHR Health and Wellbeing cohort (see Multimedia Appendix 2). Overall, 10 (35%, 10/29) participants reported using a health app previously, but no participants previously used an alcohol app. The median number of interactions (see Multimedia Appendix 2) reflects how frequently participants recorded a drink, completed questionnaires, and rated text messages.

Overall Themes

Overall, the InDEx app was acceptable for ex-serving personnel to monitor their alcohol consumption and understand their drinking habits based on positive participant feedback. A total of 4 overarching themes were developed: (1) *Credibility*, (2) *Meeting their needs*, (3) *Simplicity*, and (4) *Helpful for ex-serving personnel* (Figure 2).

Figure 2. Thematic map overview of themes and subthemes.



Theme 1: Credibility

The InDEx app was generally considered as a credible tool for participants; however, some features were deemed more credible than others, for example, the *Dashboard* page compared with the *How you compare* page. The theme of credibility hinged upon the 3 subthemes of preconceptions of their drinking, comparison with other apps, and trusting information in the app. These were interlinked as participants referenced their own drinking and experience of using other apps to determine how credible the app and its specific features were. These judgements influenced whether participants engaged with the app and

whether they believed the information provided, subsequently impacting perceived changes in accountability of their drinking behavior. Perceived credibility was therefore highly influential in determining the acceptability of InDEx and level of engagement.

Comparison With Other Apps

As participants had not previously used an alcohol app, they drew from their experiences using other health and fitness apps, for example, Fitbit, to develop recommendations of how to enhance the acceptability and credibility of InDEx:

> *I suppose it would be a bit more usable if there were (umm) a way to (umm) enter a...like a make or a (umm) a label scanner or something like that so that (umm) it was a bit more accurate.* [P3, male]

Some participants felt the app could include additional features, such as recording their diet, in accordance with other health apps, to provide an accurate reflection of the contexts of their drinking:

> *...a text message came through saying that...basically says eat whilst you're drinking...Yet on the App, ...you could add a drink (umm) you could say where you were, were you with anybody, but then you couldn't say whether you were eating...where I think that might be useful because...if you're promoting try eating when you drink as a means to reduce your consumption, it would be nice to know that you were actually out for a meal...I could then when I look at it...on the day-to-day drinking. So yeah I was out for a meal then, we had quite a lot to drink in effect, oh but we were out and we had a big meal or that at a restaurant. Just as an indication, because then it gives you...reasons why that...that consumption was higher that day than any other day sort of thing.* [P10, male]

Expectations derived from other app experiences therefore informed perceptions of InDEx's credibility and could impact participants' engagement with the app as a method for monitoring their alcohol consumption.

Preconceptions of Their Drinking

Not all participants perceived themselves as a problem drinker, neither did they use recommended drinking guidelines to inform their drinking. Instead, participants relied upon their own perceptions of both their personal drinking habits and those of others to determine the levels of risk. The below quote is from a participant with an AUDIT score of 14 who did not classify themselves as a problem drinker and seemed to attribute this to their drinking pattern:

> *Because again (sigh) I don't see myself as a...as a big drinker. So I...I'm not really doing more than twenty units in a week and I tend not to drink too much during the week...* [P21, male]

This belief appeared to influence the perceptions of whether the information participants received about their alcohol consumption was credible, in particular, the *How you compare* page, where participants assumed that they drank less than their peers and, therefore, questioned the credibility of this comparison information:

> *I class myself as a moderate drinker, even though I think the App reckons I do drink too much. (umm) But compared (laugh) to an awful lot of people I know, I don't.* [P2, male]
>
> *I think it (how you compare page) said that...I was higher than (sigh)...It said I was like (umm) the 60% of the population drinking and stuff.* [P21, male]
>
> *What did you think about that information that you were given?* [Interviewer]
>
> *(umm) I think that the population are good liars! (laugh)* [P21, male]

Trusting Information in the App

There were mixed opinions of how credible different pages of InDEx were and some pages appeared more credible than others:

> *I'm not too sure because (umm) it (how you compare page) said that 20% of the people in the country drink less than me! (laugh) And I find that quite hard! (laugh)...because I don't drink a lot! So (umm) in the time I've been using the App I've...I've probably only been drinking on about four occasions...in the month. So...(umm) to say that 20% of people in the country, that probably means people who don't drink!* [P20, male]
>
> *I've just got into the dashboard...seven days using it, (umm) bring up a new total, it knows my name and my progress. (umm) See because I've had a breather, everything is down to zero this week* [P6, male]

Participants also questioned where information about their drinking came from, particularly with the *How you compare* page, and the information was considered less credible when the source of information was not apparent, compared with the *Dashboard* page, which provides a summary of participants' drinking behaviors based upon participants' interactions with the app:

> *Well comparing you know forces drinking to others, I'm not sure where that information would have come from.* [P15, male]

The credibility of InDEx could be enhanced by the presentation of information, in addition to its content. It was suggested that long-term graphs of consumption could capture changes over an extended period, as currently only a 7-day summary graph was available:

> *I think that (month-long graph) would have been quite good because there are times when you have a heavier drinking week or weekend if you're away, a family party or whatever. And so the last week isn't necessarily representative. While at least over a month if you see at least...if you see you're drinking more heavily over a month, then that is far more than I think than a heavy day or two.* [P19, male]

Theme 2: Meeting Their Needs

Participants consistently emphasized that they used apps on a needs-only basis. In general, the app included a range of features, such as real-time feedback, comparisons with other health behaviors, and goal setting, which seemed to successfully accommodate participants' varying needs and interests.

Overall, 4 subthemes were developed to represent how InDEx helped to meet participants' needs, including personalization, internal motivations, being nonjudgmental, and the need for flexibility.

This theme was related to all other themes as the ease of use and credibility contributed to whether participants' needs were more likely to be fulfilled, which seemed to prompt some change in accountability of their drinking.

Personalization

Participants' preferences varied regarding the frequency of the text messages, the wording used within the app, and the need for InDEx to provide interactive feedback. Most participants reported using pages personalized to their interactions with the app, such as the *Dashboard* and *Add Drink* pages:

> *Just the comparisons really. (umm) Comparisons for what I've had the previous week...Yeah, just...just to see what I've had on a weekly basis really.* [P21, male]

This also extended to the text messages, with participants preferring texts that were personalized through their interactions with the app. Despite personalization being viewed positively, there were mixed responses about the level of personalization within the text messages:

> *Very informative. Yeah, they were good. (umm) They made...made you think of what you're putting in your body and what effect it can have and what other solutions you can use. So rather than have a drink, but still be sociable. Yeah, it was good.* [P27, male]

> *if you're filling in..., and you know you haven't put in your going round to see friends, you know friends places, you're putting in other information and the system spits out something...'We've noticed you drink mostly at friends houses' and that's incorrect. Straightaway it gets your back up.* [P22, male]

Although personalization was highly regarded, most participants did not want the app or text messages to be further personalized according to their former military occupation:

> *So I don't...I don't sort of look for (umm) I don't look to be communicated with from anybody in a different way because...what my job has been previously, if that makes sense?* [P1, male]

> *...I probably wouldn't go down the route of...of tying it in with loads of military style brandings and colours and everything. So even...even the language, I probably wouldn't use military terminology...* [P24, male]

Participants differed in their opinions of the frequency of text messages, which they felt could be tailored to usage over time or when participants were more likely to drink:

> *I suppose it just depends on the different...you know different people and what...what suits them. But (umm) but I mean to start with I think daily is fine. And...but then after that I don't know maybe once or twice a week I suppose.* [P4, male]

> *...if I started getting (umm) like Thursday to Sunday when I'm more likely to be drinking, it'll probably make me more mindful. Whereas if you give them all the time, you start to go yeah that's another message. Like if you're only getting them maybe on certain days of the week when you're likely to drink more...maybe that will make you a bit more mindful of it.* [P21, male]

The extent to which participants liked the frequency of the messages appeared to contribute to whether they felt InDEx met their needs. In this regard, some participants were dissuaded from engaging with the app on the basis that the frequency was not personalized and therefore applicable or relevant to them:

> *I was getting texts through the week and to be perfectly honest by the end I was thinking really? You know I'm not out...* [P13, male]

Internal Motivations

Most participants were content with their current drinking habits and therefore engaged with the general monitoring pages of InDEx only, that is, *Add drink* and *Dashboard*. The small proportion describing motivations to change were more likely to use the *Goals* page. Some participants reported becoming more motivated to change their drinking habits over the 28-day period:

> *Because essentially the App...how ready are you to give up alcohol? (umm) And that was always about an a...you know I've always sort of started off as a seven rather than an eight and then I headed towards a nine...* [P8, female]

Some participants reported that they had previously made significant changes to their alcohol consumption immediately after leaving the Armed Forces or had already set a goal to reduce their alcohol consumption and therefore found behavior change pages, such as the *Goals* page, less useful:

> *I've never thought my drinking was excessive, that I needed to set goals you know (umm) no. So maybe it hadn't really...it hadn't really applied to me as much as some others it might.* [P9, male]

Need for Flexibility

Participants expressed a preference for being able to record their alcohol consumption, set goals, and answer questions with little restriction to provide more context to their alcohol consumption. Overall, participants found InDEx relatively nonrestrictive and found some structured features particularly helpful:

> *Yeah there were enough (drink) options for (umm) which was good rather than having to put...input it*

> *yourself. So giving the options of what you had to drink (umm) was easy enough to use.* [P9, male]

However, some participants felt that drink categories could have been more specific, and drop-down menu options for weekly screening questions allow for open-ended responses so participants could add more contextual details, such as various drinking locations within a drinking session. Some participants suggested that this level of flexibility could improve the accuracy of personalization, particularly regarding the content of text messages:

> *I've gone through everything and for example (umm) it was 'location you're having a beer' (umm) sometimes I wasn't in any of the five option as in 'at a friend's/at a pub/service club/on my own...'* [P6, male]

> *I think (umm) one of the...one of the limitations then of actually you know giving your reasons or saying where you were and who you were with...(umm) Because there was a day when I did drink quite a lot, but actually I was at a wedding. But the App didn't know that.* [P26, female]

Being Nonjudgmental

Participants felt InDEx included positive content and language, and this was particularly important for meeting their needs. In this regard, participants seemed to become disengaged with pages they deemed as judgmental:

> *Yeah, they were good. Yes (umm) good (umm) advice and it wasn't too...well it wasn't patronising or anything, it was...it got back to...to a point.* [P24, male]

This subtheme relates to *Helpful for ex-serving personnel* as participants seemed more engaged and accepting of InDEx if they felt it was nonjudgmental.

Theme 3: Simplicity

Participants believed InDEx was intuitive and easy to use, and this complemented their desire to use apps on a needs-only basis. They preferred to engage with pages of InDEx that did not require too much effort, and its functionality appeared more important than its aesthetics. This also extended to their preference for short and direct text messages. This theme consists of 2 subthemes relating to the app's clear layout that required minimal effort to use.

Clear Layout

Most participants liked the clear layout of the app and felt the presentation of the text and infographics enabled easy interpretation and successful navigation. Participants preferred pages that did not have many hidden features, as these added unwanted complexity and time. Specifically, pages such as the *Dashboard* and *Add drink* were user-friendly and more acceptable to participants than the *Goals* page:

> *There wasn't too much information on there, it was all in the page that you needed, you didn't have to be scrolling around and searching for things and it was there.* [P7, female]

InDEx seemed to appeal to ex-serving personnel with a range of information technology abilities; however, participants felt they needed a step-by-step tutorial. Although there was an initial navigation page in InDEx, this was not apparent to participants.

> *So that's what I did and I think if you go back and look at my data you will see that at some point all of a sudden I went oh hang on, there's more stuff here! And it was asking me (umm) do I drink alone or with friends or with family, do I drink at home or in a pub? And I didn't see those...you know I don't know I didn't see them, but its...it took me a while to register and go oh I'm missing some information.* [P13, male]

Although participants generally found the *Add drink* page easy to use, some had difficulty understanding how to add drinks one at a time throughout a drinking session and, as a result, this changed the way they interacted with the app:

> *So I just remember being a bit confused by [how to add a drink] and thought right I'm not going to try that again.* [P1, male]

Minimal Effort

Participants seemed reluctant to spend too much time engaging with an app and appeared to engage with pages that were not time-consuming. They felt InDEx was user-friendly based on the minimal effort required to record information:

> *Yeah because...if you can't navigate your way around sometimes it...it makes it difficult so...you know the easier the better really. [InDEx] is user friendly so that's what you want or what I wanted anyway!* [P18, male]

Those who set a goal or recorded where and who they were drinking with found this process more complex and found difficulty in interpreting the goals set and amending where and who they were drinking with. This subtheme demonstrates the importance of simplicity and the subsequent impact on user engagement.

Theme 4: Helpful for Ex-Serving Personnel

Participants believed InDEx was particularly useful for heavier-drinking ex-serving personnel interested in reducing their alcohol consumption, despite the fact that most participants recruited in this study scored as hazardous drinkers themselves. If perceived as credible, participants felt that they became more accountable for their drinking habits by the end of the study, and this prompted them to make some changes to their alcohol consumption.

This theme therefore related to the helpfulness of InDEx for this population and consists of 3 subthemes, including participants' accountability for their alcohol consumption, changes in their alcohol consumption, and the app's usefulness for specific types of ex-serving drinkers.

Increased Accountability

Pages, such as the *Dashboard*, were useful for understanding the impact of participants' alcohol consumption upon aspects of their health and lifestyle by translating their drinks into alcoholic units, calories consumed, and the amount of money

spent on alcohol. Many participants reported a lack of knowledge about alcoholic units, and this facility therefore confronted them with the reality of their alcohol consumption:

> *Eye-opening, you know. All of a sudden it's in terms of...I've had (umm) nine units of alcohol and that's two burgers, or three burgers or...or that's an hours walking I've got to do! (laugh) Oh! Idiot! So it's sort of gives you a bit of time to sit back and reflect on.* [P10, male]

Graphs within the *Drinks* page further supported accountability by participants reviewing consumption over the past week:

> *I would look up you know how many...usually the drinks sort of tab to see how much I'd been drinking in the last week and think oh I ought...so it would make me look ahead and say well I'm going...if I'm going to need to fit in two or three drink-free days, then...I would use it to say well yeah how many days is it since I've had one?...* [P19, male]

These aspects facilitate participants' awareness of their alcohol consumption and related risks and increased accountability for their drinking habits.

Change of Drinking Habits and Awareness

Some participants reported reductions in their alcohol consumption, which they attributed partially to the information received from InDEx about the impact of their drinking upon their health. Where applicable, goals set by the participant contributed to this, even if they did not meet their goal:

> *The goal setting was good to get out of that habit. And once you break that habit which you know all is takes you know is that...that week. Then you know you're onto a good thing because you go what was all the fuss about?* [P28, male]
>
> *Yeah. And did you...how close were you to achieving your goals that you set on the App?* [Interviewer]
>
> *I think I...I folded on the six (laugh) I think I folded on the sixth day!* [P8, female]

If viewed as credible, the use of comparison information, in particular, appeared to prompt self-reported changes in behavior:

> *I'd drunk eight beef burgers and I was wondering if I could afford another beef burger!* [P28, male]

Suitability for Types of Drinkers

Participants believed that InDEx was acceptable for ex-serving personnel, particularly for those about to leave the Armed Forces, but questioned their motivations to download the app in the first instance:

> *(umm) Yeah, as I say I'd recommend [InDEx] to [ex-serving personnel]. It's whether they take it up.* [P17, male]

To support uptake, participants felt the app could be advertised within the resettlement package to provide additional support upon leaving service. A resettlement package is given to serving personnel before leaving service and is designed to help prepare them for entering the civilian job market [42].

> *I think it would be very useful as part of the resettlement package just as a...you know you get a housing brief, you get a finance brief...* [P25, male]

The app was also regarded as most suitable for heavier drinkers. As demonstrated by their engagement with the *Goals* page, specific features might be appropriate for drinkers incentivized to change their drinking. The following quote is from a participant with an AUDIT score of 10, which is toward the lower end of hazardous alcohol use:

> *The goals (umm) again it would say ask questions like 'Do you want to reduce your drinking?' and realistically no I don't because I don't drink that much.* [P3, male]

Despite drinking at hazardous or harmful levels, none of the participants identified as problem drinkers. The ability for the app to increase a sense of accountability and improve awareness, however, may support the recognition of problems through engagement with InDEx.

Discussion

Principal Findings

The main findings of this study indicated that InDEx was an acceptable and user-friendly app for ex-serving personnel to monitor their alcohol consumption. The successes of the app included its personalization, its abilities to meet participants' needs, the ease of use, and provision of real-time feedback on their drinking. Similarly, the text message facility, in providing more direct communication and personalization, prompted participants to log into InDEx and encouraged engagement. Credibility, relating to whether participants trusted the information provided by the app, was a central facet in determining participants' level of engagement. Many participants believed InDEx was suitable for heavier drinkers, and although the majority of participants in this study met the criteria for hazardous use, most did not perceive themselves as heavy drinkers.

During the interview, participants reported using self-monitoring pages (*Dashboard* and *Add drink*) the most compared with goal-setting pages (*Goals* and *How you compare)*; this indicates participants' needs, motivations for using InDEx, and our usability findings. Given that most participants reported that they were not motivated to change their drinking patterns, it is unsurprising that the majority did not set an intentional goal to reduce their drinking and instead engaged with self-monitoring pages of the app. Participants found the app most useful for identifying drinking patterns and understanding the potential impact of their drinking on other aspects of their health. The app may be further developed by incorporating comparison feedback from the *Dashboard* and *Drinks* pages to encourage goal setting. A relevant technique from a previous study used an algorithm that prompted a drinking limit goal via tailored text messaging based on an ecological momentary assessment. These suggested limits were set to be slightly lower than participants' self-reported alcohol consumption to ensure that the goal was relevant and realistic [43].

Usability findings from our feasibility study suggested a reduction in the number of units of alcohol consumed from week 1 to week 4 of the app study [27]. Our current findings on acceptability suggest self-monitoring and provision of information about participants' drinking facilitated this change, with many users reporting greater accountability for their own drinking. Self-monitoring, therefore, could be sufficient in prompting a change in drinking behavior without an active plan to change alcohol consumption.

Many participants did not depend on recommended drinking guidelines to identify their unit intake per week, and some participants did not believe the information presented on the *How you compare* page, in particular, comparisons of their drinking with the general population. We used the social norms approach to attempt to address the misperceptions of alcohol consumption in the general population [31]; yet, our sample appeared skeptical. It may be that future developments of InDEx could compare ex-serving personnel's alcohol consumption with other more relatable populations or include information on the impact of other health behaviors upon alcohol-attributable harms. The latter could be useful for this population as participants reported that the comparisons of alcohol with exercise and food intake were particularly enlightening.

Although personalized pages were deemed more acceptable, participants varied in how personalized they thought InDEx and text messages should be. Most participants preferred personalization based on their interactions with InDEx; for example, providing real-time feedback on their drinking, receiving text messages based upon their common drinks, and tailoring the frequency of the messages to reflect their drinking patterns. Surprisingly, participants indicated a preference for more generic, rather than military-specific, terminology, which may be due to a range of reasons. Previous research suggests that some ex-serving personnel experience problems adjusting to civilian life [5,44]; therefore, it was surprising that there was not a request for InDEx to be framed in the military context, which users may be more comfortable with. It could be that the suitability of military terms may depend on how long someone might have left service.

Although InDEx was acceptable, most participants felt that they would need an explicit incentive to download the app as they did not perceive themselves to be problem drinkers. It is a common perception that individuals believe others have drinking problems rather than themselves [45,46], even if they are drinking at the same level. This may be particularly relevant for the UK Armed Forces, where previous research indicates that those with alcohol problems are the least likely to recognize themselves as having a problem compared with other mental health issues [8,47].

Comparison With Previous Work

Some of our findings are consistent with previous work, for example, we found that credibility was a central facet of the acceptability of InDEx; Garnett et al's meta-regression [25] found that the credibility of sources was associated with a reduction in alcohol consumption. Furthermore, the persuasive system design model proposes that perceived system credibility is required for technologies to be believable and more persuasive [48,49]. Previous qualitative research found that information deemed untrustworthy undermined an app's credibility and users would be less inclined to engage with this information [50]. This may be particularly applicable for our target population as they may not identify themselves as problem drinkers, therefore requiring more persuasion.

We also found that functionality, personalization, and meeting user needs contribute to the acceptability of an app. This is consistent with previous research on the acceptability of smartphone apps in the general population [11,19,41,51]. Participants emphasized needing an app to be easy to use, not to require too much effort, and to contain features providing real-time feedback on their interactions. Contrary to Milward et al's findings [51], participants seemed to place more emphasis on the need of the app to meet their personal interests and its functionality instead of its aesthetics, preferring simplicity, which seemed to reflect engagement with specific modules [27]. Different preferences may be population specific, as Milward's study focused on young adults compared with our sample of ex-serving personnel, in which the majority were aged 40 years and older.

Our findings emphasized the need for personalizing both the app and text messages based upon the participant's interactions with InDEx in terms of usage and when they are likely to drink. This is consistent with previous research conducted in the general population in which the extent to which messages were personalized enhanced engagement and acceptability [23,52,53]. A meta-analysis on the efficacy of SMS text messaging–based interventions showed that message tailoring and personalization were associated with greater intervention efficacy [53]. This suggests that the need for personalization is a common theme across different populations.

One issue highlighted in our study was that participants did not identify themselves as problem drinkers, despite drinking at hazardous levels. This has been found in both military [54] and general populations [55], where individuals understate their drinking and overestimate the drinking of others. This could be explained by the cognitive dissonance theory, whereby ex-serving personnel considered other reasons for why their risk of drinking was higher compared with others, such as their comparators not being truthful about their alcohol consumption or the consumption of food mediating the effect of alcohol [56,57].

Another issue arising from this analysis related to participants' motivation to change, with many only using the self-monitoring modules, which may reflect that they did not identify having a drinking problem. Research has demonstrated that self-monitoring is associated with greater effect sizes from brief interventions [16] and larger decreases in alcohol consumption and AUDIT scores when using *active ingredients* of BCTs in an alcohol app [58]. Research has previously found motivation to be central to the utilization of smartphone apps and technology [40] and continued engagement for a sustained period of time [19]. More specifically, research on motivation and intention to change drinking behaviors, however, found changes in motivation and self-efficacy predicted drinking outcomes 8 weeks later in the general population [59], but this

may differ according to the level of alcohol risk [60]. This indicates that motivation to change may be one component of changing drinking behavior, with other components, such as engagement and self-monitoring, contributing toward this. Therefore, having a self-monitoring module may meet the needs of those who do not identify themselves as problem drinkers and are not motivated to change but may subsequently change their drinking as a result of this module.

There has been limited research on the use of smartphone apps for alcohol misuse for ex-serving personnel; however, 1 study compared these attitudes with smartphone apps for mental health for ex-serving personnel. They found positive overall attitudes to apps as a way of providing strategies to track symptoms [61]. These are similar to our findings where participants believed InDEx was useful for monitoring their consumption; therefore, smartphone apps may be an acceptable delivery method for monitoring drinking behavior in this population.

Limitations

Participants recruited to the study were selected from a cohort that consented to be contacted for future research and were compensated £40 for taking part; this may have increased engagement with the app. We tried to overcome this by not providing specific instructions on how to use InDEx other than to download it. This may also increase social validity as apps are typically downloaded without additional information [62].

Participants discussed their experience of the InDEx app with a researcher on the InDEx app team, which may have increased the risk of social desirability bias; however, previous research has found that conducting telephone interviews may reduce this risk because of feelings of anonymity [63]. JP also reminded participants throughout the interview that the purpose of the study was to inform intervention development and were encouraged to provide honest feedback. The analysis also involved an independent second coder who also reviewed the codebook and themes.

We also aimed to interview participants up to 1 week after testing the InDEx app, but, in some cases, there was a delay in interviewing participants of over 3 weeks. The researcher generally did not need to prompt participants throughout the interview; however, participants may have had difficulty recalling some pages of the app.

Conclusions

To the best of our knowledge, this is the first qualitative study to explore the acceptability of a smartphone-based alcohol app for ex-serving personnel. We found that InDEx was an acceptable app for ex-serving personnel to monitor their alcohol consumption. Credibility was a central facet to the acceptability of InDEx and its modules. Pages providing real-time, personalized, and easily interpretable feedback on their drinking and text messages reflecting participants' interactions with the app were most acceptable.

Acknowledgments

This study was funded by the Medical Research Council (MR/N028244/2). CD is part-funded by the National Institute for Health Research (NIHR) Biomedical Research Centre for Mental Health at South London and Maudsley National Health Service Foundation Trust and King's College London, and CD is part-funded by the NIHR Collaborations for Leadership in Applied Health Research and Care South London at King's College Hospital NHS Foundation Trust and is in receipt of an NIHR Senior Investigator award. The views expressed are those of the authors and do not necessarily reflect the views of the Department of Health, the NIHR, or the Medical Research Council.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Copy of topic guide.

[PDF File (Adobe PDF File), 57KB - mhealth_v7i5e12267_app1.pdf]

Multimedia Appendix 2

Demographic, alcohol use characteristics of participants and number of engagements with InDEx (Information about Drinking in Ex-serving personnel).

[PDF File (Adobe PDF File), 56KB - mhealth_v7i5e12267_app2.pdf]

References

1. Fear NT, Iversen A, Meltzer H, Workman L, Hull L, Greenberg N, et al. Patterns of drinking in the UK Armed Forces. Addiction 2007 Nov;102(11):1749-1759. [doi: 10.1111/j.1360-0443.2007.01978.x] [Medline: 17935583]
2. Goodwin L, Norton S, Fear NT, Jones M, Hull L, Wessely S, et al. Trajectories of alcohol use in the UK military and associations with mental health. Addict Behav 2017 Dec;75:130-137. [doi: 10.1016/j.addbeh.2017.07.010] [Medline: 28734152]

3. McDevitt-Murphy ME, Luciano MT, Tripp JC, Eddinger JE. Drinking motives and PTSD-related alcohol expectancies among combat veterans. Addict Behav 2017 Dec;64:217-222. [doi: 10.1016/j.addbeh.2016.08.029] [Medline: 27664565]
4. Marbac-Delon V, Créach-Juzan MA, Gillard J, Saintot M, Mayet A. Associations between alcohol use expectancies and drinking problems among French soldiers. Swiss Med Wkly 2015;145:w14101 [FREE Full text] [doi: 10.4414/smw.2015.14101] [Medline: 25665156]
5. Jones N, Campion B, Keeling M, Greenberg N. Cohesion, leadership, mental health stigmatisation and perceived barriers to care in UK military personnel. J Ment Health 2018 Feb;27(1):10-18. [doi: 10.3109/09638237.2016.1139063] [Medline: 27319607]
6. Keyes KM, Hatzenbuehler ML, McLaughlin KA, Link B, Olfson M, Grant BF, et al. Stigma and treatment for alcohol disorders in the United States. Am J Epidemiol 2010 Dec 15;172(12):1364-1372 [FREE Full text] [doi: 10.1093/aje/kwq304] [Medline: 21044992]
7. Kulesza M, Pedersen E, Corrigan P, Marshall G. Help-seeking stigma and mental health treatment seeking among young adult veterans. Mil Behav Health 2015 Jun;3(4):230-239 [FREE Full text] [doi: 10.1080/21635781.2015.1055866] [Medline: 26664795]
8. Hines LA, Goodwin L, Jones M, Hull L, Wessely S, Fear NT, et al. Factors affecting help seeking for mental health problems after deployment to Iraq and Afghanistan. Psychiatr Serv 2014 Jan 01;65(1):98-105. [doi: 10.1176/appi.ps.004972012] [Medline: 24037454]
9. Linke S, Brown A, Wallace P. Down your drink: a web-based intervention for people with excessive alcohol consumption. Alcohol Alcohol 2004;39(1):29-32 [FREE Full text] [Medline: 14691071]
10. Gustafson DH, McTavish FM, Chih M, Atwood AK, Johnson RA, Boyle MG, et al. A smartphone application to support recovery from alcoholism: a randomized clinical trial. JAMA Psychiatry 2014 May;71(5):566-572 [FREE Full text] [doi: 10.1001/jamapsychiatry.2013.4642] [Medline: 24671165]
11. Crane D, Garnett C, Brown J, West R, Michie S. Factors influencing usability of a smartphone app to reduce excessive alcohol consumption: think aloud and interview studies. Front Public Health 2017;5:39 [FREE Full text] [doi: 10.3389/fpubh.2017.00039] [Medline: 28421175]
12. Dulin PL, Gonzalez VM, Campbell K. Results of a pilot test of a self-administered smartphone-based treatment system for alcohol use disorders: usability and early outcomes. Subst Abus 2014;35(2):168-175 [FREE Full text] [doi: 10.1080/08897077.2013.821437] [Medline: 24821354]
13. Fowler LA, Holt SL, Joshi D. Mobile technology-based interventions for adult users of alcohol: a systematic review of the literature. Addict Behav 2016 Nov;62:25-34. [doi: 10.1016/j.addbeh.2016.06.008] [Medline: 27310031]
14. Kaner EF, Beyer FR, Garnett C, Crane D, Brown J, Muirhead C, et al. Personalised digital interventions for reducing hazardous and harmful alcohol consumption in community-dwelling populations. Cochrane Database Syst Rev 2017 Dec 25;9:CD011479. [doi: 10.1002/14651858.CD011479.pub2] [Medline: 28944453]
15. Garnett C, Crane D, Michie S, West R, Brown J. Evaluating the effectiveness of a smartphone app to reduce excessive alcohol consumption: protocol for a factorial randomised control trial. BMC Public Health 2016 Jul 08;16:536 [FREE Full text] [doi: 10.1186/s12889-016-3140-8] [Medline: 27392430]
16. 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]
17. 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]
18. Gajecki M, Berman AH, Sinadinovic K, Rosendahl I, Andersson C. Mobile phone brief intervention applications for risky alcohol use among university students: a randomized controlled study. Addict Sci Clin Pract 2014;9:11 [FREE Full text] [doi: 10.1186/1940-0640-9-11] [Medline: 24985342]
19. Attwood S, Parke H, Larsen J, Morton KL. Using a mobile health application to reduce alcohol consumption: a mixed-methods evaluation of the drinkaware track & calculate units application. BMC Public Health 2017 Dec 17;17(1):394 [FREE Full text] [doi: 10.1186/s12889-017-4358-9] [Medline: 28511698]
20. Suffoletto B, Kristan J, Chung T, Jeong K, Fabio A, Monti P, et al. An interactive text message intervention to reduce binge drinking in young adults: a randomized controlled trial with 9-month outcomes. PLoS One 2015;10(11):e0142877 [FREE Full text] [doi: 10.1371/journal.pone.0142877] [Medline: 26580802]
21. Muench F, van Stolk-Cooke K, Kuerbis A, Stadler G, Baumel A, Shao S, et al. A randomized controlled pilot trial of different mobile messaging interventions for problem drinking compared to weekly drink tracking. PLoS One 2017;12(2):e0167900 [FREE Full text] [doi: 10.1371/journal.pone.0167900] [Medline: 28146560]
22. Black N, Mullan B, Sharpe L. Computer-delivered interventions for reducing alcohol consumption: meta-analysis and meta-regression using behaviour change techniques and theory. Health Psychol Rev 2016 Sep;10(3):341-357. [doi: 10.1080/17437199.2016.1168268] [Medline: 26999311]
23. 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]

24. Garnett C, Crane D, Brown J, Kaner E, Beyer F, Muirhead C, et al. Reported theory use by digital interventions for hazardous and harmful alcohol consumption, and association with effectiveness: meta-regression. J Med Internet Res 2018 Feb 28;20(2):e69 [FREE Full text] [doi: 10.2196/jmir.8807] [Medline: 29490895]
25. Garnett CV, Crane D, Brown J, Kaner EF, Beyer FR, Muirhead CR, et al. Behavior change techniques used in digital behavior change interventions to reduce excessive alcohol consumption: a meta-regression. Ann Behav Med 2018 May 18;52(6):530-543. [doi: 10.1093/abm/kax029] [Medline: 29788261]
26. Leightley D, Puddephatt JA, Goodwin L, Rona R, Fear NT. InDEx: Open source iOS and Android software for self-reporting and monitoring of alcohol consumption. J Open Res Softw 2018 Mar 23;6:13 [FREE Full text] [doi: 10.5334/jors.207] [Medline: 29795769]
27. Leightley D, Puddephatt J, Jones N, Mahmoodi T, Chui Z, Field M, et al. A smartphone app and personalized text messaging framework (InDEx) to monitor and reduce alcohol use in ex-serving personnel: development and feasibility study. JMIR Mhealth Uhealth 2018 Sep 11;6(9):e10074 [FREE Full text] [doi: 10.2196/10074] [Medline: 30206054]
28. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. Br Med J 2015;350:h1258 [FREE Full text] [Medline: 25791983]
29. O'Cathain A, Hoddinott P, Lewin S, Thomas KJ, Young B, Adamson J, et al. Maximising the impact of qualitative research in feasibility studies for randomised controlled trials: guidance for researchers. Pilot Feasibility Stud 2015;1:32 [FREE Full text] [doi: 10.1186/s40814-015-0026-y] [Medline: 27965810]
30. Schwarzer R, Luszczynska A. How to overcome health-compromising behaviors. Eur Psychol 2008 Jan;13(2):141-151. [doi: 10.1027/1016-9040.13.2.141]
31. Berkowitz AD. An overview of the social norms approach. Changing the culture of college drinking: a socially situated health communication campaign. New York: Hampton Press; 2005:A-214.
32. Gollwitzer P, Oettingen G. Implementation intentions. In: Gellman M, Turner JR, editors. Encyclopedia of Behavioral Medicine. New York: Springer Verlag; 2013:1043-1048.
33. Michie S, Hyder N, Walia A, West R. Development of a taxonomy of behaviour change techniques used in individual behavioural support for smoking cessation. Addict Behav 2011 Apr;36(4):315-319. [doi: 10.1016/j.addbeh.2010.11.016] [Medline: 21215528]
34. Michie S, Wood CE, Johnston M, Abraham C, Francis JJ, Hardeman W. Behaviour change techniques: the development and evaluation of a taxonomic method for reporting and describing behaviour change interventions (a suite of five studies involving consensus methods, randomised controlled trials and analysis of qualitative data). Health Technol Assess 2015 Nov;19(99):1-188 [FREE Full text] [doi: 10.3310/hta19990] [Medline: 26616119]
35. Fear NT, Jones M, Murphy D, Hull L, Iversen AC, Coker B, et al. What are the consequences of deployment to Iraq and Afghanistan on the mental health of the UK armed forces? A cohort study. Lancet 2010 May 22;375(9728):1783-1797. [doi: 10.1016/S0140-6736(10)60672-1] [Medline: 20471076]
36. Hotopf M, Hull L, Fear NT, Browne T, Horn O, Iversen A, et al. The health of UK military personnel who deployed to the 2003 Iraq war: a cohort study. Lancet 2006 May 27;367(9524):1731-1741. [doi: 10.1016/S0140-6736(06)68662-5] [Medline: 16731268]
37. Stevelink S, Jones M, Hull L, Pernet D, MacCrimmon S, Goodwin L, et al. Mental health outcomes at the end of the British involvement in the Iraq and Afghanistan conflicts: a cohort study. Br J Psychiatry 2018 Dec;213(6):690-697 [FREE Full text] [doi: 10.1192/bjp.2018.175] [Medline: 30295216]
38. Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG. World Health Organizations. 2001. AUDIT: the alcohol use disorders identification test: guidelines for use in primary health care URL: https://apps.who.int/iris/bitstream/handle/10665/67205/WHO_MSD_MSB_01.6a.pdf;jsessionid=5AB840F7C411428D1275B45E6A830B95?sequence=1 [accessed 2019-04-29] [WebCite Cache ID 77zxYXYNt]
39. Braun V, Clarke V. Using thematic analysis in psychology. Qual Res Psychol 2006 Jan;3(2):77-101. [doi: 10.1191/1478088706qp063oa]
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. Peng W, Kanthawala S, Yuan S, Hussain SA. A qualitative study of user perceptions of mobile health apps. BMC Public Health 2016 Nov 14;16(1):1158 [FREE Full text] [doi: 10.1186/s12889-016-3808-0] [Medline: 27842533]
42. Ministry of Defence. Gov.UK. Information for Service Leavers URL: https://www.gov.uk/guidance/information-for-service-leavers [accessed 2019-04-29] [WebCite Cache ID 77zxqgkK8]
43. Suffoletto B, Chung T, Muench F, Monti P, Clark DB. A text message intervention with adaptive goal support to reduce alcohol consumption among non-treatment-seeking young adults: non-randomized clinical trial with voluntary length of enrollment. JMIR Mhealth Uhealth 2018 Feb 16;6(2):e35 [FREE Full text] [doi: 10.2196/mhealth.8530] [Medline: 29453191]
44. Blackburn D. Out of uniform: psychosocial issues experienced and coping mechanisms used by veterans during the military–civilian transition. J Mil Veteran Fam Health 2017 Apr;3(1):62-69. [doi: 10.3138/jmvfh.4160]

45. Keeling M, Knight T, Sharp D, Fertout M, Greenberg N, Chesnokov M, et al. Contrasting beliefs about screening for mental disorders among UK military personnel returning from deployment to Afghanistan. J Med Screen 2012 Dec;19(4):206-211. [doi: 10.1258/jms.2012.012054] [Medline: 23132954]
46. Garnett C, Crane D, West R, Michie S, Brown J, Winstock A. Normative misperceptions about alcohol use in the general population of drinkers: a cross-sectional survey. Addict Behav 2015 Mar;42:203-206 [FREE Full text] [doi: 10.1016/j.addbeh.2014.11.010] [Medline: 25482365]
47. Iversen A, van Staden L, Hughes JH, Browne T, Greenberg N, Hotopf M, et al. Help-seeking and receipt of treatment among UK service personnel. Br J Psychiatry 2010 Aug;197(2):149-155. [doi: 10.1192/bjp.bp.109.075762] [Medline: 20679269]
48. Oinas-Kukkonen H. Behavior change support systems: a research model and agenda. In: Ploug T, Hasle P, Oinas-Kukkonen H, editors. Persuasive Technology: 5th International Conference, Persuasive 2010, Copenhagen, Denmark, June 7-10, 2010, Proceedings (Lecture Notes In Computer Science). Berlin: Springer; 2010.
49. Oinas-Kukkonen H, Harjumaa M. Persuasive Systems Design: Key Issues, Process Model and System Features. London: Routledge; 2018:105-123.
50. Perski O, Blandford A, Ubhi HK, West R, Michie S. Smokers' and drinkers' choice of smartphone applications and expectations of engagement: a think aloud and interview study. BMC Med Inform Decis Mak 2017 Dec 28;17(1):25 [FREE Full text] [doi: 10.1186/s12911-017-0422-8] [Medline: 28241759]
51. Milward J, Deluca P, Drummond C, Watson R, Dunne J, Kimergård A. Usability testing of the BRANCH smartphone app designed to reduce harmful drinking in young adults. JMIR Mhealth Uhealth 2017 Aug 08;5(8):e109 [FREE Full text] [doi: 10.2196/mhealth.7836] [Medline: 28790022]
52. Kazemi DM, Borsari B, Levine MJ, Li S, Lamberson KA, Matta LA. A systematic review of the mHealth interventions to prevent alcohol and substance abuse. J Health Commun 2017 May;22(5):413-432. [doi: 10.1080/10810730.2017.1303556] [Medline: 28394729]
53. 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]
54. Stevelink SA, Jones N, Jones M, Dyball D, Khera CK, Pernet D, et al. Do serving and ex-serving personnel of the UK armed forces seek help for perceived stress, emotional or mental health problems? Eur J Psychotraumatol 2019;10(1):1556552 [FREE Full text] [doi: 10.1080/20008198.2018.1556552] [Medline: 30693074]
55. Emslie C, Hunt K, Lyons A. Older and wiser? Men's and women's accounts of drinking in early mid-life. Sociol Health Illn 2012 May;34(4):481-496 [FREE Full text] [doi: 10.1111/j.1467-9566.2011.01424.x] [Medline: 22034902]
56. Festinger L. A Theory Of Cognitive Dissonance. Stanford: Stanford University Press; 1962.
57. Harmon-Jones E, Mills J. Science Conference Series. Cognitive Dissonance: Progress on a Pivotal Theory in Social Psychology. In: An introduction to cognitive dissonance theory and an overview of current perspectives on the theory. Washington: American Psychological Association; 1999.
58. Crane D, Garnett C, Michie S, West R, Brown J. A smartphone app to reduce excessive alcohol consumption: identifying the effectiveness of intervention components in a factorial randomised control trial. Sci Rep 2018 Mar 12;8(1):4384 [FREE Full text] [doi: 10.1038/s41598-018-22420-8] [Medline: 29531280]
59. Kuerbis A, Armeli S, Muench F, Morgenstern J. Motivation and self-efficacy in the context of moderated drinking: global self-report and ecological momentary assessment. Psychol Addict Behav 2013 Dec;27(4):934-943 [FREE Full text] [doi: 10.1037/a0031194] [Medline: 23276318]
60. de Vocht F, Brown J, Beard E, West R, Michie S, Campbell R, et al. Motivation to reduce alcohol consumption and subsequent attempts at reduction and changes in consumption in increasing and higher-risk drinkers in England: a prospective population survey. Addiction 2018 May;113(5):817-827 [FREE Full text] [doi: 10.1111/add.14132] [Medline: 29368391]
61. Connolly SL, Miller CJ, Koenig CJ, Zamora KA, Wright PB, Stanley RL, et al. Veterans' attitudes toward smartphone app use for mental health care: qualitative study of rurality and age differences. JMIR Mhealth Uhealth 2018 Aug 22;6(8):e10748 [FREE Full text] [doi: 10.2196/10748] [Medline: 30135050]
62. Meredith SE, Alessi SM, Petry NM. Smartphone applications to reduce alcohol consumption and help patients with alcohol use disorder: a state-of-the-art review. Adv Health Care Technol 2015;1:47-54 [FREE Full text] [doi: 10.2147/AHCT.S65791] [Medline: 27478863]
63. Block ES, Erskine L. Interviewing by telephone: specific considerations, opportunities, and challenges. Int J Qual Methods 2012 Sep;11(4):428-445. [doi: 10.1177/160940691201100409]

Abbreviations

AUDIT: alcohol use disorder identification test
BCTs: behavior change techniques
InDEx: Information about Drinking in Ex-serving personnel
KCMHR: King’s Centre for Military Health Research
NIHR: National Institute for Health Research

SMS: short message service

Edited by G Eysenbach; submitted 27.09.18; peer-reviewed by J Milward, M Sewitch, C Jurberg; comments to author 08.01.19; revised version received 05.03.19; accepted 30.03.19; published 24.05.19.

Please cite as:
Puddephatt JA, Leightley D, Palmer L, Jones N, Mahmoodi T, Drummond C, Rona RJ, Fear NT, Field M, Goodwin L
A Qualitative Evaluation of the Acceptability of a Tailored Smartphone Alcohol Intervention for a Military Population: Information About Drinking for Ex-Serving Personnel (InDEx) App
JMIR Mhealth Uhealth 2019;7(5):e12267
URL: http://mhealth.jmir.org/2019/5/e12267/
doi:10.2196/12267
PMID:31127726

Original Paper

The Twazon Arabic Weight Loss App: App-Based Intervention for Saudi Women With Obesity

Aroub Alnasser[1,2], BSc (Hons), MSc; Janet Kyle[2], BSc (Hons), MSc, PhD; Najla Aloumi[1], BSc; Abdulrahman Al-Khalifa[1], MSc, PhD; Debbi Marais[3], MSc, PhD

[1]Food Science and Nutrition Dept, King Saud University, Riyadh, Saudi Arabia

[2]Institute of Applied Health Sciences, University of Aberdeen, Aberdeen, United Kingdom

[3]Warwick Medical School, University of Warwick, Coventry, United Kingdom

Corresponding Author:
Aroub Alnasser, BSc (Hons), MSc
Food Science and Nutrition Dept
King Saud University
Riyadh,
Saudi Arabia
Phone: 966 118056476
Email: aroub@ksu.edu.sa

Abstract

Background: By 2022, it is estimated that the rate of female obesity (78%) in Saudi Arabia will almost double that of males (41%). Despite being mainly attributed to poor diet, sedentary lifestyle, and a lack of health awareness, behavioral modification interventions are relatively new to the population; bariatric surgery continues to be the treatment of choice for comorbidities. However, neither pre nor postoperative diet and exercise are promoted. Evidence-informed mobile health (mHealth) weight loss apps and interventions may be an effective tool for delivering a culturally relevant intervention.

Objective: This study aimed to determine the feasibility of a weight loss intervention that tests the effectiveness of Twazon, an originally designed Arabic weight-loss app that promotes lifestyle modification specific to Arab populations.

Methods: A pre-post single - arm pilot study was carried out among a sample of 240 overweight volunteer Saudi women residing in Riyadh, Saudi Arabia who used the Twazon app over a 4-month period. Anthropometric, diet, and physical activity measures were assessed 3 times: baseline, 2-months and 4-months; frequency of app use and system usability were evaluated during the 2 latter data collection periods. Repeated measures analysis of variance was used to identify changes over time.

Results: A total of 40 participants completed the 4-month intervention with an attrition rate of 83%. An evaluation of the frequency of app use fostered 2 groups: engaged users (65%) and unengaged users (35%). At 4 months, the engaged users experienced more successful outcomes; body weight was lowered on average by 1.3 (SD 0.6) kg (P=.18), waist circumference (WC) was reduced by 4.9 (SD 1.1) cm (P<.001), and daily energy consumption was decreased by >600 calories (P=.002). Unengaged users experienced minor changes in body weight, WC, and reduced energy intake.

Conclusions: The findings have demonstrated that engagement with the Twazon app renders positive changes in body weight, WC, and energy intake. mHealth weight loss apps and interventions have the potential to be effective in promoting weight loss and healthy lifestyle modification in Saudi Arabia and similar populations.

(JMIR Mhealth Uhealth 2019;7(5):e10923) doi:10.2196/10923

KEYWORDS

obesity; weight loss; mobile applications; smartphone; obesity management; mHealth

Introduction

Background

The prevalence of obesity in the Gulf region has become a serious public health issue, with the Kingdom of Saudi Arabia (KSA) having the highest prevalence rates at almost 35% [1]. In line with the global trend, Saudi women are affected by obesity more than men with current rates at 42% and 31%, respectively [2]. Projection studies on the increasing rate of obesity have anticipated a steady rise for both genders but more so for women; by 2022, 78% of females compared with 41%

of males will be obese [3]. Studies have attributed the growing epidemic for Saudi women to a variety of factors, including low physical activity (PA), restrictive social norms based on gender inequality, and a transitioning diet [1]. The recent nutrition transition in Saudi Arabia is characterized by shifts in dietary patterns toward increased consumption of sugar, fat, and refined foods and can be linked to modernization, urbanization, economic development, and increased affluence within the population. Sedentary, inactive lifestyles are also more common under these conditions [4]. Despite this, research in the KSA has continued to focus on the descriptive analysis of age and gender differences rather than the effects of weight loss interventions [5].

Multicomponent lifestyle interventions aimed at modifying daily dietary and physical habits are effective in weight loss and weight management [6]. However, bariatric surgery is the preferred option in the KSA with 15,000 annual operations [7]; successful long-term outcomes are not enduring [8] and are associated with an absence of lifestyle and diet modification postsurgery [9]. A more sustainable and appropriate intervention design that includes behavior modification is needed for Saudis with obesity [6,10,11]. The worldwide popularity of mobile technology has provided a unique platform for delivering health information to the public via mobile apps, some of which have shown to positively promote weight loss and lifestyle modification [12]. In the KSA, the use of mobile phones that have an operating system capable of running downloaded apps (smartphones), has become ubiquitous [13]. There are many commercial Arabic weight loss apps available; however, an Arabic app screening [14] revealed that these apps show low adherence to evidence-informed practices (EIPs) for achieving weight loss and cultural insensitivity to diet and PA norms [14]. To fill the gap, the development of the Twazon Arabic weight loss app [15] was evidence informed and culturally adapted to meet the needs of its target audience.

Objectives

As the first app-based weight loss intervention in the region, this pilot study aimed to evaluate the effectiveness and feasibility of a complex weight loss intervention that utilized the Twazon app as its main component. Developed over several phases [15,16], the primary objective of this app-based intervention was to facilitate weight loss through diet and PA modification. Over a 4-month period, outcomes of change in nutritional status, including weight loss, body mass index (BMI) and waist circumference (WC), as well as lifestyle habits including energy intake, Mediterranean diet (MD) adherence, and PA, were monitored.

Methods

Study Design

A pre-post single - arm pilot study with 3 measurement points, this study was carried out to evaluate the effectiveness and feasibility of delivering a weight loss app intervention to overweight and obese Saudi women residing in Riyadh, Saudi Arabia.

Participants and Recruitment

Recruitment was carried out online via social networks such as Twitter and by placing posters at various locations in Riyadh, the KSA. Interested participants completed a Web-based screening questionnaire regarding background information, medical history, and current physical status. Eligibility requirements included the following: gender (female), residency in Riyadh, age (≥18 years), BMI (≥25), stability of weight (<2 kg variation in 3 months), no current weight loss program or weight-affecting medication, smartphone ownership, interest in losing weight, and consent to participate. Participants were excluded if they were ever diagnosed with medical conditions such as diabetes and cancer or if already pregnant, trying to get pregnant, or lactating. Participants who met the requirements were invited via short message service (SMS) text messaging to participate.

Measures

Anthropometric status, diet, and PA were assessed by trained researchers. Height and weight were measured to the nearest 0.1 cm and 0.01 kg, respectively, using the KERN electronic weighing scale (model MPB-P with stand). BMI was calculated as weight (kg)/height (m) using the World Health Organization (WHO) [17] cutoff points of 25.0 to 29.9 kg/m^2 for overweight and ≥30 kg/m^2 for obesity. WC was taken at the approximate midpoint between the lower margin of the last palpable rib and the top of the iliac crest [18] using a Seca 201 measuring tape to the nearest 0.1 cm; the WHO cutoff points define ≥80 cm as *increased risk of metabolic complications* and ≥88 cm as *substantially increased risk of metabolic complications* [18].

Total energy (kcal) average was calculated from two 24-hour dietary recalls on nonconsecutive days using the Automated Self-Administered 24-hour dietary recalls (ASA24) system [19]. To determine the nature of a participant's diet, a culturally adapted version of the validated 14-item MD questionnaire was used. Alcohol consumption is forbidden in the KSA; therefore, the question pertaining to wine was eliminated, and each participant received 1 point to properly calculate the MD score (MDS). The MDS is calculated by adding up each point given, with a maximum of 14. The level of adherence was defined as follows: score 0 to 5, lowest adherence; score 6 to 9, average adherence; and score ≥ 10, highest adherence [20].

A culturally-relevant self-administered PA questionnaire [21] was used to assess an individual's level of PA in terms of total metabolic equivalent task values (METs); METs levels are defined as *light* (<3 METs), *moderate* (3 to 6 METs), and *vigorous* (>6 METs) *intensity* [22]. A total of 60 min of moderate-intensity PA daily (eg, biking 10 to 12 miles per hour [mph] and walking 3 to 4 mph) can facilitate weight loss and maintenance for people with obesity or overweight [5]. The MET-min cutoff point of ≥1680 MET-min per week (60 min per day×7 days per week×4 METs) was used [23] to determine whether the recommendation was met.

The frequency of app use was heuristically determined through an algorithm (see below) created for Twazon that best calculated user activity or the frequency of user updates on certain measures (eg, diet, PA, and weight); time-stamped participant

data were continuously downloaded to a secure server using the following: *Number of Active Days/ Two Weeks Count=Frequency of Use.*

The *number of active days* was calculated by adding up the number of days users uploaded their personal updated data. The *2 weeks count* was calculated by dividing the number of days since registered by 7 and then again by 2. These 2 figures were then computed as shown above to determine *frequency of use*. A score of 14 or less was indicative of how many times users updated their info in the app; a score under 1 meant less than 1 update every 2 weeks, a score of 1 meant at least 1 update every 2 weeks, and a maximum score of 14 meant at least 1 update every day. Frequency of app use was calculated every 2 weeks and then averaged at the end of the intervention.

To determine the overall usability of the tool, the 10-item System Usability Scale (SUS) questionnaire [24] was given to those participants who were labeled as engaged. Final SUS scores are linked with adjective ratings and acceptability ranges with a score of >70 being considered as *acceptable* [25,26].

Intervention Implementation Procedure

A total of 3 assessments with a duration of 45 to 60 min were carried out at baseline, 2 months, and 4 months at 4 different locations in Riyadh; a voucher for 25 Saudi Arabian Riyals (or £5) or a small gift was given for participation. At the first assessment, individuals were interviewed and tutored on how to use the app. After completing the 2 ASA24 dietary recalls, participants were sent a link via SMS or WhatsApp to download the app, input personal information, and begin using the tool. Measures (refer to the subheading Measures) were evaluated at baseline, 2 months, and 4 months, with the exception of the SUS and frequency of app use scores, which were analyzed during the 2 latter assessments; height was measured at baseline. Reminders to attend the 2 month and 4 month assessments were also sent via SMS text messaging or WhatsApp.

Statistical Analysis

Changes in the frequency of app use were evaluated using a repeated measures analysis of variance. Assessing outcome measures for within-group differences was carried out by the time-by-group interaction effect, whereas between-group differences were attained through contrastive analysis. The *P* value significance was set at $P<.05$ and a Bonferroni correction was used. All analyses were performed using the SPSS version 24 (SPSS Inc). Power calculations for sample size indicated that 165 participants would provide 90% power to detect a difference in means of 7.1 kg. This translates to a clinically meaningful change of 10% weight loss using a paired t-test with a 0.05 2-tailed significance level. Assuming a 30% attrition rate, a sample of 215 participants was estimated.

Ethics

Ethical approval was obtained from the University of Aberdeen College of Life Sciences Ethics Review Board and the Scientific Research Ethics Committee in Riyadh, at King Saud University. This study complies with the WHO guidelines for the reporting of health interventions using mobile phones [27].

Results

Baseline Participants

Out of 1028 interested Saudi women, 240 were eligible and were invited to participate in the intervention. Figure 1 shows the flow of participants at baseline, 2 months, and 4 months.

Baseline participants were clinically obese with an average weight of 81.8 (SD 15.1) kg, WC of 89.0 (SD 11.2) cm, and BMI of 32.9 (SD 5.6) kg/m^2. In total, 49% (117/240) of all participants with a WC >88 cm were classified as *at high risk* of metabolic complications, 29% (69/240) were *at risk*, and 22% (52/240) had a healthy WC (Table 1).

Energy intake was reported at 1692 (SD 751) calories per day, and the average MDS was 7.3 (SD 1.9), indicating medium MD adherence 68% (162/240). Over 75% of baseline participants reported little to no consumption of *red meat* (187/240), *butter* (192/240), or *carbonated/sugar-sweetened beverages* (SSBs; 190/240) and low consumption scores for fruit 21% (50/240) and fish 17% (41/240), (Figure 2). PA was deemed insufficient with almost 62% of baseline participants (146/236) not meeting the recommendation for weight loss; 38% (90/236) met the recommendation, and 4 women did not complete the PA questionnaire.

Figure 1. Flow of study participants. *Most common reason for exclusion was having a medical condition such as hypothyroidism. **Two participants were excluded because of being on a new diet. ***Five participants' data were not included in the final analysis because of failure to attend the 2-month assessment. ****26= 21 users who stayed engaged +5 previously unengaged who became engaged at 4-months.

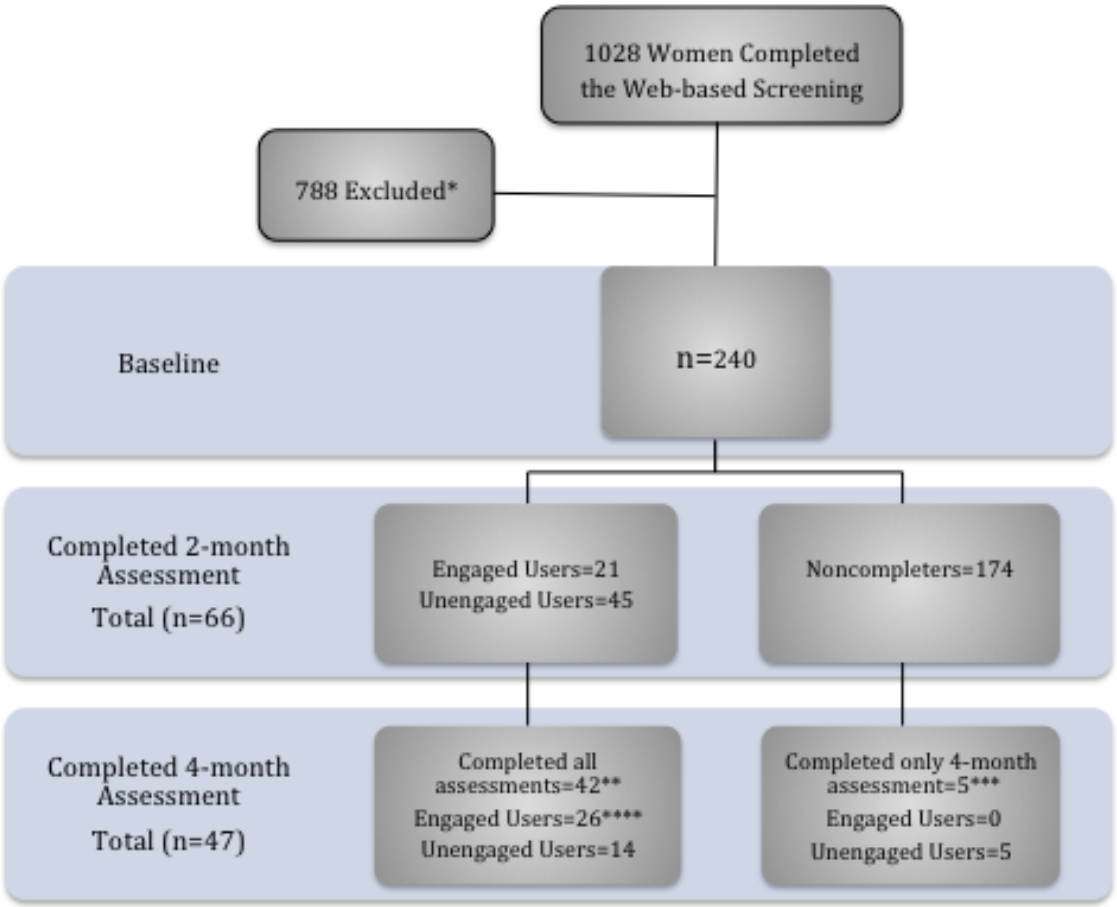
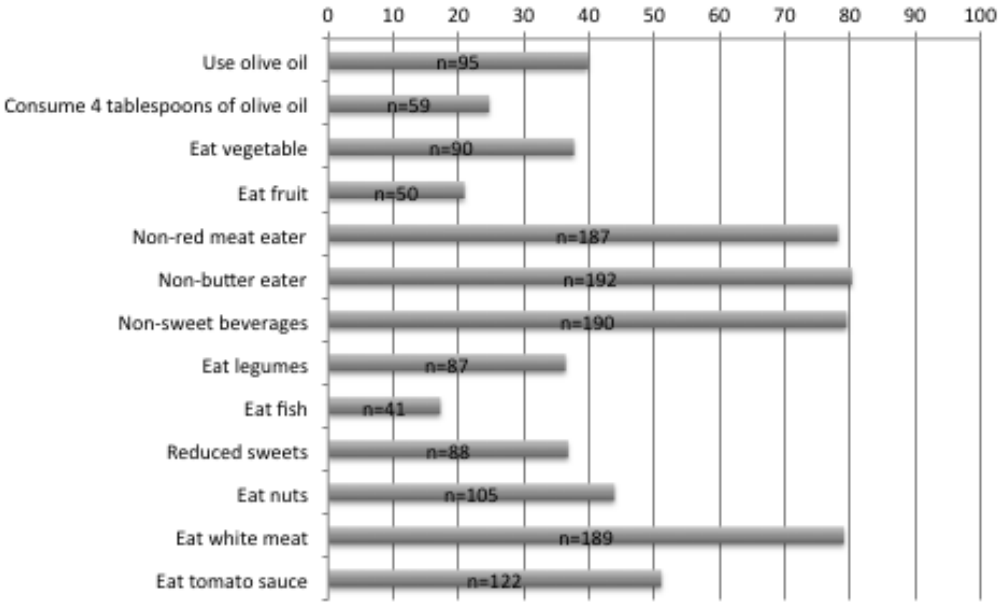


Table 1. Baseline characteristics of the 240 women.

Characteristics	Baseline values
Age (years), mean (SD)	31 (10)
Body weight (kg), mean (SD)	81.8 (15)
Body mass index (kg/m2), mean (SD)	**32.9 (5)**
Overweight (25.0-29.9), n (%)	84 (35)
Obese (≥30), n (%)	154 (65)
Waist circumference (cm), mean (SD)	**89.0 (11)**
Normal <80, n (%)	52 (22)
Risk ≥80, n (%)	69 (29)
High risk ≥88, n (%)	117 (49)
Energy intake (kcal/d), mean (SD)	1692 (751)
Total Mediterranean diet score (point), mean (SD)	7.26 (2)
Level of adherence, n (%)	
Low (≤5)	48 (20)
Medium (6-9)	162 (68)
High (≥10)	29 (12)
Physical activity (METs[a]-min/week), median (25^{th}-75^{th} percentile)	1109 (334-2570)
Met the Recommendation. ≥1680 METs-min/ week, n (%)	90 (38)
Less than 1680 METs-min/ week, n (%)	146 (62)

[a]METs: metabolic equivalent task values.

Figure 2. The number and percentage of the responses to Mediterranean diet questionnaire.



Frequency of App Use Scores

The results of app usage scores were heuristically determined to comprise 2 groups: *engaged* and *unengaged* users (See Measures in Methods). The more frequent the input, the higher the app use frequency score; ≥1 was labeled as engaged and <1 as unengaged. Users had to have updated their information in the app at least 8 times during the 16-week period to be considered engaged, which translates to a minimum of 1 update every 2 weeks. Of the 26 total engaged participants, 5 were

labeled as unengaged at the 2-month assessment, but they increased their usage in the last 2 months, which relabeled them as engaged at the 4-month assessment. No participant went from being labeled as engaged (at 2 months) to unengaged at 4 months.

Effects of 4-month Intervention

Anthropometric Changes

Engaged users lost 1.3 (SD 0.6) kg of body weight (*P*=.03), significantly reduced WC by 4.9 (SD 1.1) cm (*P*<.001), and lowered BMI by 0.6 (SD 0.2). Unengaged users lost 0.2 (SD 0.8) kg of body weight, reduced their WC by 3.0 (SD 0.8) cm, and exhibited no change in BMI (Table 2).

Diet and Lifestyle Changes

Energy Intake

Engaged users significantly reduced their energy intake by 672 calories per day (*P*=.002; see Multimedia Appendix 1), consuming 1470 (SD 535) calories per day, thereby meeting the reduction requirement [5] for decreasing weight-related risks. Unengaged users lowered their energy intake by 204 calories per day, consuming 1644 (SD 453) calories per day, but they failed to meet the requirement for weight loss (Table 2).

Mediterranean Diet

Engaged and unengaged participants had a mean MDS of 8.65 (SD 2.2) and 8.00 (SD 2.7), respectively, indicating a statistically significant improvement in overall MD adherence (*P*=.003). High MD adherence was achieved by 40% (16/40) of all participants at 4 months, compared with 12% (6/40) at baseline. Adherence to fruit (Q4; *P*=.005) and vegetable (Q3; *P*=.02) consumption was improved by all participants collectively. The engaged participants significantly lowered consumption of sweets and improved MD adherence (*P*<.001), whereas the unengaged participants increased their intake and lowered adherence. Table 3 shows mean MDS and adherence to MD questions. Multimedia Appendix 2 shows *P* values and adherence to MD questions (See Multimedia Appendix 2).

Physical Activity

At baseline, 69% (18/26) of engaged participants and 61% (8/14) of unengaged participants did not meet the PA recommendation, labeling them as inactive. The most common reason given for physical inactivity at any of the 3 assessments was *lack of time*. By 4 months, the engaged participants and unengaged participants increased their PA by 112 and 48 METs-mins per week, respectively (Table 2); however, no statistically significant changes were achieved by either of the 2 groups and neither met the recommendation (Table 4).

Table 2. Summary of effect results after the 4-month intervention.

Changes	Engaged		Unengaged		*P* values	
	Mean (SD)	%	Mean (SD)	%	Within - group difference	Between - group difference
Body weight (kg)	–1.3 (0.6)	–2	–0.2 (0.8)	0	.18	.34
Waist circumference (cm)	–4.9 (1.1)	–5	–3 (0.8)	–3	<.001	.35
Body mass index (kg/m^2)	–0.6 (0.2)	–2	0 (0.2)	0	.18	.27
Energy (kcal/d)	–672 (283)	–31	–204 (403)	–11	.002	.30
Adherence to Mediterranean diet score	1.3 (0.5)	18	0.54 (0.5)	7	.003	.39
Physical activity	112 (669)	8	48 (148)	3	.77	.49

Table 3. Adherence to Mediterranean diet score.

Level of adherence	Engaged (n=26)		Unengaged (n=14)	
	Baseline, n (%)	4 months, n (%)	Baseline, n (%)	4 months, n (%)
Low (≤5)	4 (15)	3 (12)	2 (15)	3 (23)
Medium (6-9)	19 (73)	13 (50)	8 (62)	4 (31)
High (≥10)	3 (12)	10 (39)	3 (23)	6 (46)

Table 4. Success rate of adherence to recommended daily physical activity.

Physical activity	Engaged (n=26)		Unengaged (n=14)	
	Baseline, n (%)	4 months, n (%)	Baseline, n (%)	4 months, n (%)
Met the recommendation ≥1680 METs-min/ week	8 (31)	7 (27)	5 (39)	5 (39)
Less than 1680 METs-min/ week	18 (69)	19 (73)	8 (62)	8 (62)

System Usability Scale

An average of the SUS scores calculated at 2 months and 4 months was determined for all participants. Engaged and unengaged users had an average SUS score of 69.3 (SD 10) and 64.3 (SD 8.6), respectively, indicating a marginally high acceptability rating (approximately >63) closer to an adjective rating of *good* (approximately >71.4) than of *ok* (approximately >50.9). A more in-depth analysis was carried out in a triangulated study on the relationship among actual user experiences, that is, the SUS score from this study, evidence-informed requirements [14], and potential user expectations [16], to explore the users' primary preferences and suggested improvements of the Twazon weight loss app [28].

Discussion

Principal Findings

This pilot study aimed to evaluate the effectiveness of the Twazon Arabic weight loss app and to determine the feasibility of implementing it in a weight loss intervention among overweight and obese Saudi women. Due to a limited sample size and some methodological concerns, no concrete conclusions can be made about the effectiveness of the Twazon app. This study also reveals that the intervention may not be feasible because of low retention, despite successful recruitment and no significant differences found in baseline data between noncompleters and completers. A detailed discussion of these issues and how they might be avoided follows in this section.

Effectiveness of the Twazon App

System usability was deemed to be of marginally high acceptability and those participants who engaged with the app more frequently, also experienced more successful outcomes. The engaged users lost an average of 0.2 (SD 0.1) kg by 2-months, and an additional 1.1 (SD 0.5) kg by the end of the intervention, with a total loss of 2% of baseline weight. This did not meet the 10% of clinically recommended weight loss for reducing risk factors associated with obesity, however, similar findings [29] in a smartphone-based weight-loss intervention showed an average loss of 1.8 kg after 6-months, indicating sustainable gradual weight loss over time. By 4-months, the engaged users significantly reduced WC by 4.9 (SD 1.1) cm (P<.001), going from a *substantially increased risk* to an *increased risk* for metabolic complications, and representing progress toward clinically significant weight-loss.

Energy intake and daily diet choices seem to have been positively influenced by increased interaction with the app. By the end of the intervention, engaged users reported an average intake of 1470 (SD 535) calories per day**,** representing the recommended daily intake of 1200-1600 calories per day for achieving weight loss [30]. Although overweight and obese women often underreport their energy intake [31], the ASA24 can be a useful tool for determining the dietary intake of this population; Widamen and colleagues [32] found only a 5% difference between energy reported and actual energy consumed. Despite the associations between obesity and the issue of underestimating, this study considered the process of self-monitoring to be more important than the accuracy of reported dietary intake [33].

MD adherence revealed that the rate of positive MD adjustment increased after 4-months for those who engaged with the app. Engaged participants reported consuming very little red meat, butter or carbonated and SSBs at baseline and 4-months, with small improvements made for the latter two. At 4-months the engaged participants were more successful at increasing or maintaining MD adherence with the most notable declines in adherence being for the consumption of sweets and SSBs. Changing trends in Saudi food choices that may contribute to the obesity problem should be considered for their role in health-related outcomes; for instance, some highly popular, foreign foods such as high-fat processed cheese products and spreads [34,35] may not have been reported because of their absence on the MD questionnaire. The low reporting of SSB consumption in this study is also questionable as research conducted in the KSA reported that at least 46% of adults are consuming one or two soft drinks, energy drinks and/or sweet beverages a day [36-38]. A preexisting knowledge of SSBs being associated with increased obesity [39] may have had an influence on low-reporting, however, it is also possible the question was not understood to include fruit drinks, juices or syrup-flavored coffee beverages which contain excessive amounts of sugar [40-42].

The unengaged users experienced some progress in diet by decreasing caloric intake by 204 (SD 403) calories per day and lost 0.2 (SD 0.8) kg of body weight. These results could be attributed to awareness-induced outcomes. A review study [43] indicated that participants were more conscious of certain diet and lifestyle habits after the intervention, despite not having used the app because of the information they received during the assessments and questionnaires. In terms of PA improvements, the lack of significant differences was not unexpected for 2 main reasons: first, a meta-analysis of weight-loss apps showed no significant changes in PA [12], and second, three-quarters of Saudi women are characterized as *insufficiently active* [44], ranking as the least physically active worldwide. Causes for the lack of PA in the KSA include an absence of locations designed for general PA such as sidewalks or green spaces and very limited options for female-oriented PA; only high-cost gyms or sports clubs exist which means that normal daily walking or aerobic movement is not a common activity for either gender [45].

Feasibility of the Twazon Intervention

Although high recruitment rates were achieved (n=240 participants, or 10% of the original recruitment target) only 17% of the sample was retained. It is difficult to identify the exact reasons for attrition; however, there are several possibilities. Seed and colleagues suggest that high recruitment rates have the potential to have a negative effect on retention rates, such as individuals initially participating but not committing to the full duration of an intervention [46]. Whether or not a participant had the intent to commit or was actually prepared to devote the necessary work to begin losing weight was not assessed before beginning the current study. Being interested in losing weight was a requisite but may not have

been an adequate indicator of what a participant was willing to invest in terms of behavior change and self-monitoring. A readiness questionnaire that also indicates what the commitment potentially could entail in terms of time and effort might assist in evaluating this. Participants who are new to lifestyle modification or weight loss may have felt unprepared, thereby affecting retention [47].

High attrition may be also explained by motivation; very little is known about what drives a Saudi man or woman to participate in a weight loss intervention. In line with this, an aversion to goals perceived as not being robust enough or disappointing [48,49] may have demotivated participants. During recruitment, many women reported losing interest in the intervention due to dissatisfaction with the goal of a 10% loss of total body weight [50,51]. Other app-based weight-loss interventions have found that many participants with obesity discontinued app use if the initial amount of weight lost was unsatisfactory [52,53]. Although Foster [48] and other studies [49,54] found that a weight-loss goal of 22-34% was expected and perceived as more motivating, attainment of 50% of said goal resulted in positive physical and psychosocial effects. If less-than-satisfactory weight loss goals can still be beneficial to obese individuals, then there may be a need for cognitive intervention aimed at improving personal body image and reducing the importance of weight as a number.

The amount of personal contact had with researchers could also have had an effect on a participant's adherence to the Twazon intervention. Weight loss interventions delivered via mobile phone that also included daily personal contact by those carrying out the intervention were found to be more effective than a stand-alone app [55]. Although there have not been any studies conducted in the KSA assessing the effect of personal contact, Aroian and colleagues' longitudinal study [56] in the United States states that explored the efficacy and inefficacy of strategies used to recruit and retain Arab Muslim immigrant women and their children for research. Their results indicated that personal contact with data collectors that they trusted and had good rapport with, was the most important motivator for adherence to the study. This should be explored by future research, as these findings may hold true in Arab nations like the KSA, and perhaps in other high-context communication cultures [57,58] that place substantial value on the community.

Strengths and Limitations

The Twazon weight loss app intervention is, to the best of our knowledge, the first of its kind in the Gulf region, and unique in its focus on adult Saudi women, in its use of an original Arabic weight loss app, and in its development being in compliance with EIPs and behavior change techniques for weight-loss. This study is not without its limitations, which include: (1) a lack of control group, (2) recruitment that combined traditional methods with social media, (3) discrepancies between changes in weight and WC, (4) normal fluctuations in body weight changes was not considered, (5) possible underreporting of energy intake, and finally, (6) no validated index of app use was available. Future research should be aware of these limitations when designing protocol for similar studies so that more conclusive evidence can support the findings.

Conclusions

The study set out to evaluate the effectiveness of the culturally-adapted and evidence-informed Arabic Twazon app which was developed to promote weight loss through lifestyle modification, and to determine the feasibility of using said app in an intervention conducted among obese and overweight Saudi women. Although some positive outcomes of change were observed, limited sample size did not allow for any concrete conclusions to be established from the current study's findings, and therefore, the effectiveness of Twazon cannot be confirmed. Whether the complexity of the Twazon app's multicomponent features was encouraging of weight loss or overly demanding on the participants is also still unknown. In regards to the feasibility of the Twazon app intervention, retention was challenging, and for this reason it is not certain if app-enabled mobile phone technology should be used in weight-loss interventions among the target population; researchers should conduct more in-depth exploratory analysis of potential issues before designing clinical trials. Regarding local policy in the KSA, a substantial increase in the availability of PA-promoting environments and the integration of health-and-diet education needs to be implemented for all citizens of all ages if the obesity epidemic is to be effectively managed. Future studies are sorely needed in and around the Middle East as well as among similar populations found outside the region.

Acknowledgments

The authors extend their appreciation to the Research and Development Office-Ministry of Education for funding this research work, through project number (HQI-001).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Repeated measure analysis of variance (ANOVA) analyses of 4-month weight loss intervention for the engaged users (n=26) andunengaged users (n=14).

[PDF File (Adobe PDF File), 78KB - mhealth_v7i5e10923_app1.pdf]

Multimedia Appendix 2

Results of repeated measure analysis of variance (ANOVA) of the Mediterranean diet Scores for the engaged users (n=26) and unengaged users (n=14).

[PDF File (Adobe PDF File), 71KB - mhealth_v7i5e10923_app2.pdf]

References

1. DeNicola E, Aburizaiza O, Siddique A. Doi.org. Obesity and public health in the Kingdom of Saudi Arabia URL: https://www.degruyter.com/view/j/reveh.2015.30.issue-3/reveh-2015-0008/reveh-2015-0008.xml [accessed 2019-05-14] [WebCite Cache ID 78NVRnuYf]
2. World Health Organization. Global health observatory data repository URL: http://apps.who.int/gho/data/node.main.A900A?lang=en [accessed 2018-05-02] [WebCite Cache ID 6z6dAy2bs]
3. Al-Quwaidhi AJ, Pearce MS, Critchley JA, Sobngwi E, O'Flaherty M. Trends and future projections of the prevalence of adult obesity in Saudi Arabia, 1992-2022. East Mediterr Health J 2014 Oct;20(10):589-595. [Medline: 25356689]
4. Popkin B. Global nutrition dynamics: the world is shifting rapidly toward a diet linked with noncommunicable diseases. Am J Clin Nutr 2006 Aug;84(2):289-298. [doi: 10.1093/ajcn/84.1.289] [Medline: 16895874]
5. Al-Khaldi Y. Obesity in Saudi Arabia……gaps in research. Saudi J Obesity 2014;2(1):1. [doi: 10.4103/2347-2618.137562]
6. Centre for Public Health Excellence at NICE (UK), National Collaborating Centre for Primary Care (UK). The National Institute for Health and Care Excellence. 2006 Jan 1. Obesity: the prevention, identification, assessment and management of overweight and obesity in adults and children [Internet] URL: https://www.nice.org.uk/guidance/cg189/evidence/obesity-update-appendix-m-pdf-6960327447 [accessed 2019-05-14] [WebCite Cache ID 78NWJXAru]
7. Al-Khaldi Y. Bariatric surgery in Saudi Arabia: the urgent need for standards. Saudi J Obesity 2016;4(1):1. [doi: 10.4103/2347-2618.184930]
8. Magro DO, Geloneze B, Delfini R, Pareja BC, Callejas F, Pareja JC. Long-term weight regain after gastric bypass: a 5-year prospective study. Obes Surg 2008 Jun;18(6):648-651 [FREE Full text] [doi: 10.1007/s11695-007-9265-1] [Medline: 18392907]
9. Karmali S, Brar B, Shi X, Sharma AM, de Gara C, Birch DW. Weight recidivism post-bariatric surgery: a systematic review. Obes Surg 2013 Nov;23(11):1922-1933. [doi: 10.1007/s11695-013-1070-4] [Medline: 23996349]
10. US Preventive Services Task Force, Bibbins-Domingo K, Grossman DC, Curry SJ, Davidson KW, Epling JW, et al. Screening for and management of obesity in adults: US Preventive Services Task Force recommendation statement. J Am Med Assoc 2017 Jan 24;317(4):407-414. [doi: 10.1001/jama.2016.20325] [Medline: 28118461]
11. Alfadda AA, Al-Dhwayan MM, Alharbi AA, Al Khudhair BK, Al Nozha OM, Al-Qahtani NM, et al. The Saudi clinical practice guideline for the management of overweight and obesity in adults. Saudi Med J 2016 Oct;37(10):1151-1162 [FREE Full text] [doi: 10.15537/smj.2016.10.14353] [Medline: 27652370]
12. Flores MG, Granado-Font E, Ferré-Grau C, Montaña-Carreras X. Mobile phone apps to promote weight loss and increase physical activity: a systematic review and meta-analysis. J Med Internet Res 2015;17(11):e253 [FREE Full text] [doi: 10.2196/jmir.4836] [Medline: 26554314]
13. Communications and Information Technology Commission. ICT Indicators End of 2017 Q2 Report URL: https://www.citc.gov.sa/en/indicators/Indicators%20of%20Communications%20and%20Information%20Techn/ICTIndicators-Q22017En.pdf#search=CITC%2C%20Q2%202017 [accessed 2017-02-05] [WebCite Cache ID 78L3OlBXZ]
14. Alnasser AA, Amalraj RE, Sathiaseelan A, Al-Khalifa AS, Marais D. Do Arabic weight-loss apps adhere to evidence-informed practices? Transl Behav Med 2016 Dec;6(3):396-402 [FREE Full text] [doi: 10.1007/s13142-015-0347-7] [Medline: 27528528]
15. Alnasser A, Sathiaseelan A, Al-Khalifa A, Marais D. Development of 'Twazon': an Arabic app for weight loss. JMIR Res Protoc 2016 May 16;5(2):e76 [FREE Full text] [doi: 10.2196/resprot.5497] [Medline: 27185568]
16. Alnasser AA, Alkhalifa AS, Sathiaseelan A, Marais D. What overweight women want from a weight loss app: a qualitative study on Arabic women. JMIR Mhealth Uhealth 2015;3(2):e41 [FREE Full text] [doi: 10.2196/mhealth.4409] [Medline: 25993907]
17. World Health Organization. BMI classification URL: http://apps.who.int/bmi/index.jsp?introPage=intro_3.html [WebCite Cache ID 6z6hWlsO1]
18. National Department of Health South Africa. Waist Circumference and Waist-Hip Ratio : Report of a WHO expert Consultation, Geneva, 8-11 December 2008. Geneva: World Health Organization; 2020.
19. Subar AF, Kirkpatrick SI, Mittl B, Zimmerman TP, Thompson FE, Bingley C, et al. The Automated Self-Administered 24-hour dietary recall (ASA24): a resource for researchers, clinicians, and educators from the National Cancer Institute. J Acad Nutr Diet 2012 Aug;112(8):1134-1137 [FREE Full text] [doi: 10.1016/j.jand.2012.04.016] [Medline: 22704899]
20. Martínez-González MA, García-Arellano A, Toledo E, Salas-Salvadó J, Buil-Cosiales P, Corella D, et al. A 14-item Mediterranean diet assessment tool and obesity indexes among high-risk subjects: the PREDIMED trial. PLoS One 2012;7(8):e43134 [FREE Full text] [doi: 10.1371/journal.pone.0043134] [Medline: 22905215]

21. Al-Hazzaa HM, Musaiger AO, ATLS Research Group. Arab Teens Lifestyle Study (ATLS): objectives, design, methodology and implications. Diabetes Metab Syndr Obes 2011;4:417-426 [FREE Full text] [doi: 10.2147/DMSO.S26676] [Medline: 22253540]
22. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Tudor-Locke C, et al. 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc 2011 Aug;43(8):1575-1581. [doi: 10.1249/MSS.0b013e31821ece12] [Medline: 21681120]
23. Al-Hazzaa HA, Abahussain NA, Al-Sobayel HI, Qahwaji DM, Musaiger AO. Physical activity, sedentary behaviors and dietary habits among Saudi adolescents relative to age, gender and region. Int J Behav Nutr Phys Act 2011 Dec 21;8(1):140 [FREE Full text] [doi: 10.1186/1479-5868-8-140] [Medline: 22188825]
24. Brooke J. SUS-a quick and dirty usability scale. In: Usability Evaluation In Industry. Florida, United States: CRC Press; Jun 1, 1996:4-7.
25. Bangor A, Kortum P, Miller J. Determining what individual SUS scores mean: adding an adjective rating scale. J Usability Stud 2009 Jan 1 [FREE Full text]
26. Bangor A, Kortum P, Miller J. An empirical evaluation of the System Usability Scale. Int J Hum Comput Interact. ? 2008;24(6):594. [doi: 10.1080/10447310802205776]
27. Agarwal S, LeFevre AE, Lee J, L'Engle K, Mehl G, Sinha C, et al. Guidelines for reporting of health interventions using mobile phones: mobile health (mHealth) evidence reporting and assessment (mERA) checklist. Br Med J 2016;352:i1174. [Medline: 26988021]
28. Alnasser A, Kyle J, Alkhalifah A, Marais D. Relationship between evidence requirements, user expectations, and actual experiences: usability evaluation of the Twazon Arabic weight loss app. JMIR Hum Factors 2018 Apr 17;5(2):e16 [FREE Full text] [doi: 10.2196/humanfactors.9765] [Medline: 29666042]
29. Allen JK, Stephens J, Dennison HC, Stewart KJ, Hauck S. Randomized controlled pilot study testing use of smartphone technology for obesity treatment. J Obes 2013;2013:151597 [FREE Full text] [doi: 10.1155/2013/151597] [Medline: 24392223]
30. Peterkin BB. Dietary guidelines for Americans, 1990 edition. J Am Diet Assoc 1990 Dec;90(12):1725-1727. [Medline: 2131342]
31. Wehling H, Lusher J. People with a body mass index 30 under-report their dietary intake: a systematic review. J Health Psychol 2017 Jul 1:1359105317714318. [doi: 10.1177/1359105317714318] [Medline: 28810493]
32. Widaman AM, Keim NL, Burnett DJ, Miller B, Witbracht MG, Widaman KF, et al. A potential tool for clinicians; evaluating a computer-led dietary assessment method in overweight and obese women during weight loss. Nutrients 2017 Mar 1;9(3) [FREE Full text] [doi: 10.3390/nu9030218] [Medline: 28257040]
33. 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;15(4):e32 [FREE Full text] [doi: 10.2196/jmir.2283] [Medline: 23587561]
34. Hussey M. Bord Bia - Irish Food Board. The Cheese market in Saudi Arabia reviewed URL: http://www.bordbia.ie/industry/manufacturers/insight/alerts/Pages/TheCheesemarketinSaudiArabiareviewed.aspx?year=2014&wk=29 [accessed 2017-02-05] [WebCite Cache ID 6z6fr0RW4]
35. Moradi-Lakeh M, El Bcheraoui C, Afshin A, Daoud F, AlMazroa MA, Al Saeedi M, et al. Diet in Saudi Arabia: findings from a nationally representative survey. Public Health Nutr 2017 Apr;20(6):1075-1081. [doi: 10.1017/S1368980016003141] [Medline: 27974061]
36. Alsunni AA, Badar A. Energy drinks consumption pattern, perceived benefits and associated adverse effects amongst students of University of Dammam, Saudi Arabia. J Ayub Med Coll Abbottabad 2011;23(3):3-9. [Medline: 23272423]
37. Faris MA. Patterns of caffeinated energy drinks consumption among adolescents and adults in Hail, Saudi Arabia. Food Nutr Sci 2014;05(02):158-168. [doi: 10.4236/fns.2014.52021]
38. Majeed F. Association of BMI with diet and physical activity of female medical students at the University of Dammam, Kingdom of Saudi Arabia. J Taibah Univ Sci 2015 Jun;10(2):188-196. [doi: 10.1016/j.jtumed.2014.11.004] [Medline: 27974061]
39. Epuru S, Shammary M. Nutrition knowledge and its impact on food choices among the students of Saudi Arabia. IOSR J Biotechnol Biochem 2014 Jun;13(1):68-74. [doi: 10.9790/0853-13156874]
40. Action on Sugar. 2016 Jan 1. Shocking amount of sugar found in many hot flavoured drinks URL: http://www.actiononsugar.org/news-centre/surveys/2016/shocking-amount-of-sugar-found-in-many-hot-flavoured-drinks.html [accessed 2019-05-15] [WebCite Cache ID 78NZoe7LO]
41. Wilson T, Dahl R, Temple N. Beverage tends affect future nutritional health impact. In: Beverage Impacts on Health and Nutrition. New York, NY: Springer International Publishing; Jan 1, 2016:351-358.
42. Gill JM, Sattar N. Fruit juice: just another sugary drink? Lancet Diabetes Endocrinol 2014 Jun;2(6):444-446. [doi: 10.1016/S2213-8587(14)70013-0] [Medline: 24731678]
43. Semper HM, Povey R, Clark-Carter D. A systematic review of the effectiveness of smartphone applications that encourage dietary self-regulatory strategies for weight loss in overweight and obese adults. Obes Rev 2016 Dec;17(9):895-906 [FREE Full text] [doi: 10.1111/obr.12428] [Medline: 27192162]

44. Al-Zalabani AH, Al-Hamdan NA, Saeed AA. The prevalence of physical activity and its socioeconomic correlates in Kingdom of Saudi Arabia: A cross-sectional population-based national survey. J Taibah Univ Sci 2015 Jun 14;10(2):208-215 [FREE Full text] [doi: 10.1016/j.jtumed.2014.11.001]
45. Sidawi B, Alhariri MT, Albaker WI. Creating a healthy built environment for diabetic patients: the case study of the eastern province of the Kingdom of Saudi Arabia. Glob J Health Sci 2014 Apr 14;6(4):136-147 [FREE Full text] [doi: 10.5539/gjhs.v6n4p136] [Medline: 24999135]
46. Seed M, Juarez M, Alnatour R. Improving recruitment and retention rates in preventive longitudinal research with adolescent mothers. J Child Adolesc Psychiatr Nurs 2009 Aug;22(3):150-153. [doi: 10.1111/j.1744-6171.2009.00193.x] [Medline: 19702968]
47. Laing BY, Mangione CM, Tseng CH, Leng M, Vaisberg E, Mahida M, et al. Effectiveness of a smartphone application for weight loss compared with usual care in overweight primary care patients: a randomized, controlled trial. Ann Intern Med 2014 Nov 18;161(10 Suppl):S5-12 [FREE Full text] [doi: 10.7326/M13-3005] [Medline: 25402403]
48. Foster GD, Wadden TA, Vogt RA, Brewer G. What is a reasonable weight loss? Patients' expectations and evaluations of obesity treatment outcomes. J Consult Clin Psychol 1997 Feb;65(1):79-85. [Medline: 9103737]
49. Pétré B, Scheen A, Ziegler O, Donneau AF, Dardenne N, Husson E, et al. Weight loss expectations and determinants in a large community-based sample. Prev Med Rep 2018 Dec;12:12-19 [FREE Full text] [doi: 10.1016/j.pmedr.2018.08.005] [Medline: 30116705]
50. National Heart, Lung, and Blood Institute NHLBI Obesity Education Initiative Expert Panel. Clinical guidelines on the identification, evaluation, and treatment of overweight and obesity in adults: executive summary. Expert Panel on the Identification, Evaluation, and Treatment of Overweight in Adults. Am J Clin Nutr 1998 Oct;68(4):899-917. [doi: 10.1093/ajcn/68.4.899] [Medline: 9771869]
51. McQuigg M, Brown JE, Broom J, Laws RA, Reckless JP, Noble PA, et al. The Counterweight programme: prevalence of CVD risk factors by body mass index and the impact of 10% weight change. Obes Res Clin Pract 2008 Mar;2(1):I-II. [doi: 10.1016/j.orcp.2008.01.002] [Medline: 24351674]
52. Chin SO, Keum C, Woo J, Park J, Choi HJ, Woo JT, et al. Successful weight reduction and maintenance by using a smartphone application in those with overweight and obesity. Sci Rep 2016 Dec 7;6(8):34563-31076 [FREE Full text] [doi: 10.1038/srep34563] [Medline: 27819345]
53. Kozak AT, Buscemi J, Hawkins MA, Wang ML, Breland JY, Ross KM, et al. Technology-based interventions for weight management: current randomized controlled trial evidence and future directions. J Behav Med 2017 Feb;40(1):99-111 [FREE Full text] [doi: 10.1007/s10865-016-9805-z] [Medline: 27783259]
54. Jeffery RW, Wing RR, Mayer RR. Are smaller weight losses or more achievable weight loss goals better in the long term for obese patients? J Consult Clin Psychol 1998 Aug;66(4):641-645. [doi: 10.1037//0022-006X.66.4.641] [Medline: 9735580]
55. Schippers M, Adam PC, Smolenski DJ, Wong HT, de Wit JB. A meta-analysis of overall effects of weight loss interventions delivered via mobile phones and effect size differences according to delivery mode, personal contact, and intervention intensity and duration. Obes Rev 2017 Dec;18(4):450-459. [doi: 10.1111/obr.12492] [Medline: 28187246]
56. Aroian KJ, Katz A, Kulwicki A. Recruiting and retaining Arab Muslim mothers and children for research. J Nurs Scholarsh 2006;38(3):255-261 [FREE Full text] [Medline: 17044343]
57. Zaharna R. Understanding cultural preferences of Arab communication patterns. Public Relat Rev 1995 Sep;21(3):241-255. [doi: 10.1016/0363-8111(95)90024-1]
58. Feghali E. Arab cultural communication patterns. Int J Intercult Relat 1997 Aug;21(3):345-378. [doi: 10.1016/S0147-1767(97)00005-9]

Abbreviations

ASA24: Automated Self-Administered 24-hour dietary recalls
BMI: body mass index
EIP: evidence-informed practice
KSA: Kingdom of Saudi Arabia
MD: Mediterranean diet
MDS: Mediterranean diet score
METs: metabolic equivalent task values
mHealth: mobile health
mph: miles per hour
PA: physical activity
SMS: short message service
SSB: sugar-sweetened beverage
SUS: System Usability Scale
WC: waist circumference
WHO: World Health Organization

Edited by G Eysenbach; submitted 02.05.18; peer-reviewed by JY Kim, PH Lin; comments to author 07.06.18; revised version received 02.11.18; accepted 10.02.19; published 28.05.19.

Please cite as:
Alnasser A, Kyle J, Aloumi N, Al-Khalifa A, Marais D
The Twazon Arabic Weight Loss App: App-Based Intervention for Saudi Women With Obesity
JMIR Mhealth Uhealth 2019;7(5):e10923
URL: http://mhealth.jmir.org/2019/5/e10923/
doi:10.2196/10923
PMID:31140444

Original Paper

Design and Implementation of a Novel System for Correcting Posture Through the Use of a Wearable Necklace Sensor

Hung-Yuan Chung[1], PhD; Yao-Liang Chung[2], PhD; Chih-Yen Liang[1], MS

[1]Department of Electrical Engineering, National Central University, Taoyuan, Taiwan

[2]Department of Communications, Navigation and Control Engineering, National Taiwan Ocean University, Keelung, Taiwan

Corresponding Author:
Yao-Liang Chung, PhD
Department of Communications, Navigation and Control Engineering
National Taiwan Ocean University
No. 2, Beining Rd, Jhongjheng District
Keelung, 20224
Taiwan
Phone: 886 928738102
Fax: 886 283693022
Email: ylchung@email.ntou.edu.tw

Abstract

Background: To our knowledge, few studies have examined the use of wearable sensing devices to effectively integrate information communication technologies and apply them to health care issues (particularly those pertaining to posture correction).

Objective: A novel system for posture correction involving the application of wearable sensing technology was developed in this study. The system was created with the aim of preventing the unconscious development of bad postures (as well as potential spinal diseases over the long term).

Methods: The newly developed system consists of a combination of 3 subsystems, namely, a smart necklace, notebook computer, and smartphone. The notebook computer is enabled to use a depth camera to read the relevant data, to identify the skeletal structure and joint reference points of a user, and to compute calculations relating to those reference points, after which the computer then sends signals to the smart necklace to enable calibration of the smart necklace's standard values (base values for posture assessment). The gravitational acceleration data of the user are collected and analyzed by a microprocessor unit-6050 sensor housed in the smart necklace when the smart necklace is worn, with those data being used by the smart necklace to determine the user's body posture. When poor posture is detected by the smart necklace, the smart necklace sends the user's smartphone a reminder to correct his or her posture; a mobile app that was also developed as part of the study allows the smart necklace to transmit such messages to the smartphone.

Results: The system effectively enables a user to monitor and correct his or her own posture, which in turn will assist the user in preventing spine-related diseases and, consequently, in living a healthier life.

Conclusions: The proposed system makes it possible for (1) the user to self-correct his or her posture without resorting to the use of heavy, thick, or uncomfortable corrective clothing; (2) the smart necklace's standard values to be quickly calibrated via the use of posture imaging; and (3) the need for complex wiring to be eliminated through the effective application of the Internet of Things as well as by implementing wireless communication between the smart necklace, notebook computer, and smartphone.

(JMIR Mhealth Uhealth 2019;7(5):e12293) doi:10.2196/12293

KEYWORDS

wearable sensing technology; necklace; posture correction; image recognition; internet of things

Introduction

Background

Apple released the iPhone 3G, the world's first 3G-capable phone, in 2008, thus ushering in the 3G revolution that has since significantly altered the lifestyles of phone users around the world, with countless users effectively becoming smartphone addicts. Unfortunately, such users may develop bone spurs in their cervical vertebrae because of the extended periods they spend with their backs in a slouched position while staring down

at their phones. This, in turn, has meant that those patients seeking treatment for cervical degeneration have been of younger and younger average ages in recent years. Moreover, the postural problems associated with smartphone use can also negatively affect the pelvis and the caudal, thoracic, and lumbar vertebrae, with these negative effects potentially leading, in turn, to functional leg length discrepancies, scoliosis, and distended stomachs. Typically, posture correction efforts have utilized 1 of 2 approaches, with the approach used being determined according to the severity of the given patient's condition. Specifically, kyphosis correction belts (which consist of highly elastic fabrics and are used to provide support to the back) and kyphosis correction exercises may be utilized for patients with only mild cases of kyphosis, whereas rigid back braces made with aluminum alloys are more likely to be used to support and maintain back straightness in patients with severe spinal diseases.

Information communication technologies relating to newly developed wearable sensing devices have been touted as key technologies that can be applied in the health care field, and the use of these advanced technologies to address the aforementioned issues will become a trend going forward.

Related Work

In recent years, the availability of increasingly smaller chips and greater computer power has accelerated the pace of development for wearable sensing devices. These developments have increased the applicability of such devices in the health monitoring and medical care field. Furthermore, wearable sensing devices also offer a considerably wide variety of application possibilities in other fields, such as force feedback devices, solutions for communication between people, environmental obstacle detection, and human-machine interface control.

In the health monitoring and medical care field, a study by Fallahzadeh et al [1] has examined the use of socks (embedded with accelerators and flexible stretch sensors) to detect ankle edema. In a study by Dauz et al [2], sleep quality was monitored by measuring skin potential activity, body temperature changes, and heart rates, and electrocutaneous stimulation was applied to the skin during the slow wave sleep stage to improve sleep quality. Sundaravadivel et al [3] and Boateng et al [4] explored the use of triple axis accelerometer readings to monitor the daily physical activity levels of individuals. In the study by Sundaravadivel et al [3], the data that were obtained were used to generate peak, mean, and standard deviation values, which were in turn used to differentiate between activities such as walking, stair climbing, and lying down. In the study by Boateng et al [4], a lightweight machine learning algorithm was used to analyze the obtained data and differentiate between the activities performed by users, so as to monitor and encourage physical activity among users. Liu et al [5] utilized accelerometers to measure electrocardiography signals and analyze user behavior. Martin and Voix [6] measured heart and respiratory rates by detecting sounds generated in the ear canal. Surrel et al [7] proposed a set of wearable devices capable of detecting sleep apnea. Takei et al [8] developed a wearable sensing device that can provide microelectric stimulation to the muscles and monitor muscle activity. Durbhaka [9] utilized shirts and pants embedded with triple axis accelerometers to analyze and assess human posture. Moreover, Gia et al [10] proposed a low-cost health monitoring system that involves the combined application of the Internet of Things with energy-saving sensor nodes and fog layers. This system utilized fog computing to automate services such as data sorting and channel management, allowing physicians to remotely monitor their patient's physical conditions.

When wearable sensing devices are used as force feedback devices, they can enhance the virtual reality experience and convey the data collected by the sensors of robots as tactile feedback to users. For example, Chinello et al [11] designed a wearable fingertip cutaneous device that utilized 3 servo and vibration motors to create a tactile sensation at the fingertips of users. Through the use of motor-driven belts, Meli et al [12] were able to design a wearable sensing device for the upper limb that was capable of controlling the movements of a robotic arm and of receiving movement resistance feedback from the robotic arm. Wearable sensing devices can also be used to address communication problems between people. Goncharenko et al [13] analyzed images to identify sign language and enable the instant translation of the said language into a textual or auditory output, such that a user is able to communicate with a deaf-mute individual. Furthermore, wearable sensing devices can be used to detect environmental obstacles. Through the use of shoes equipped with ultrasonic sensors and vibration motors, Patil et al [14] were able to develop a navigational aid for individuals suffering from amblyopia and blindness. When a user's foot comes too close to an obstacle, the shoes' vibration motors will alert the user to the situation. With respect to the use of human-machine interfaces to change traditional mouse-based controls for computers, Zhang et al [15] examined the use of pressure sensors to detect the distribution of wrist strength and control cursor movements.

However, although all of these studies [1-15] did make significant contributions in various respects, none examined the use of wearable sensing devices to effectively integrate information communication technologies and apply them to health care issues (particularly those pertaining to posture correction).

Study Objective

Motivated by the aforementioned lack of research, the objective of this paper was to design and implement a novel system for correcting posture that involved the use of a wearable necklace sensor and the integration of key emerging information communication technologies. The system was created with the aim of preventing the unconscious development of bad postures (as well as potential spinal diseases over the long term).

Methods

Overall Hardware Architecture

Figure 1 shows the overall hardware architecture proposed in this study. This architecture consists of 3 subsystems, namely, a smart necklace, notebook computer, and smartphone. The hardware specifications of the subsystems are described below:

1. Smart necklace: A Raspberry Pi 3 connected to an IT TRAINING [16] I/O board and, subsequently, to a Microprocessor Unit (MPU)-6050 sensor (the Raspberry Pi 3 requires a 5V power supply, which can be provided via a mobile power bank or transformer).
2. Notebook computer: A Micro-Star International GP62 notebook computer connected to a Kinect v1 camera [17].
3. Smartphone: A Sony Xperia Z with an Android operating system.

These 3 subsystems communicate through a Wi-Fi router. Data transfers between these 3 subsystems are carried out via the transmission control protocol/Internet Protocol (IP) network protocol. The mobile app developed in this study enables signal transmissions among the smartphone, the smart necklace, and the notebook computer. To facilitate later descriptions, the 2 primary transmission signals in this system are defined below:

1. Setting signals: These are signals (originating from the smartphone or notebook computer) used to adjust the smart necklace's internal parameters.
2. Reminder signals: These are signals sent from the smart necklace to the smartphone, for the purpose of reminding the user to adjust his or her posture.

Furthermore, it should be noted that the activation of the overall system does not imply that the subsystems are now connected to each other. A connection between the transmitting and receiving ends is only established when the former intends to transmit a setting signal, and the connection is broken off when the transmission is completed.

Overall System's Operations

Figure 2 is an outline flowchart that describes the overall system's operations. After a new user puts on the smart necklace, its internal parameters have to be recalibrated. The overall system is first activated, after which the mobile app is used to activate the notebook computer's image recognition function, which will then perform an automatic calibration of the smart necklace's internal parameters (primary calibration mechanism). Following this, the user may opt to perform fine manual adjustments to the smart necklace's internal parameters (secondary calibration mechanism), in which case he or she may do so via the smartphone. If the user opts not to make the above adjustments, the smart necklace's program for determining the user's posture will start. Subsequently, the mobile app is used to activate the function of receiving reminder signals from the smart necklace. The smart necklace will simultaneously be continuing its assessment of the user's posture. In the event that poor posture is detected, a reminder signal will be sent to the smartphone. Once the system reaches its set operating time, its operations will end (the overall system shuts down).

Figure 1. Hardware architecture of the overall system.

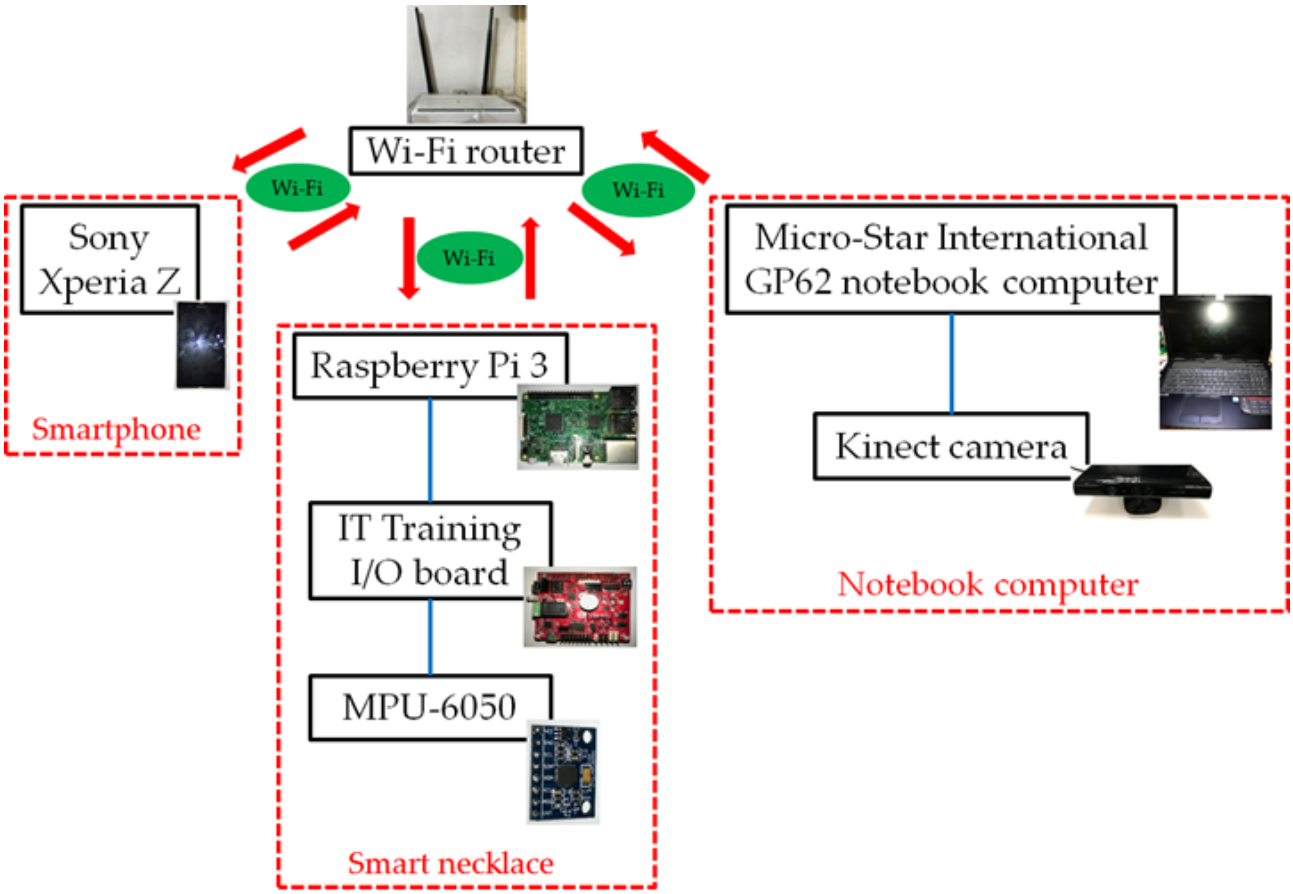
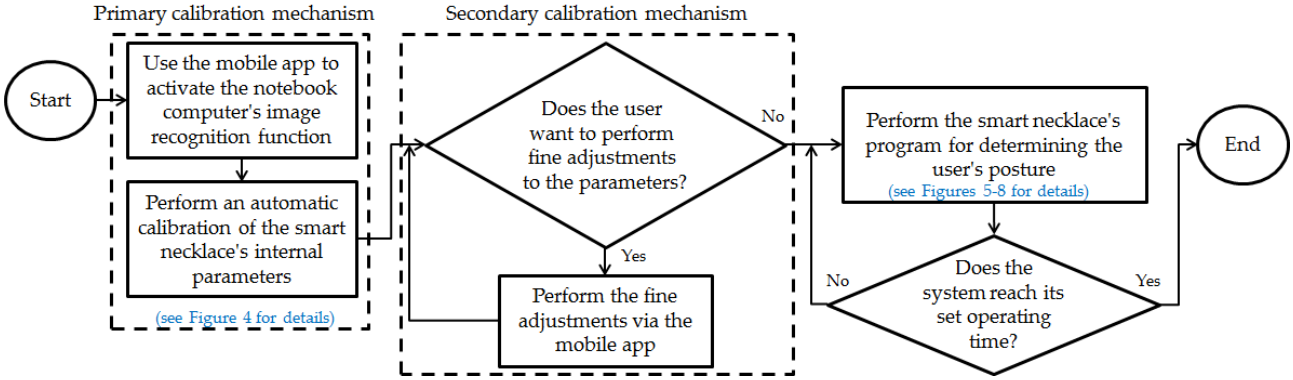


Figure 2. Outline flowchart of the overall system's operations.



When diagnosing spine-related diseases, an orthopedist first performs a visual diagnosis. To determine if a patient is suffering from kyphosis, the orthopedist will use the center points of the patient's head and neck as reference points and verify if the line connecting these 2 points is perpendicular to the ground (a kyphotic posture in Figure 3). As for the diagnosis of scoliosis, the orthopedist will use the left and right shoulder joints as reference points and verify if they are at an even height (a posture in which left shoulder is tilted downward in Figure 3). Thus, these criteria were used in the study as the basis for determining if the body posture of a person is poor. To facilitate later descriptions, we defined the variable S_i (unit: G) as the gravitational acceleration value (a reference value for determining posture) for the direction i, with $i= x$ representing the front-back direction and $i= y$ representing the left-right direction, measured by the MPU-6050 when a user has a standard posture. S_i will change according to the notebook computer's image recognition results (see Figure 4) or the fine manual adjustments to the parameters made via the smartphone (see the Smartphone section).

Figure 3. (a) Kyphosis and (b) left shoulder tilted downward.

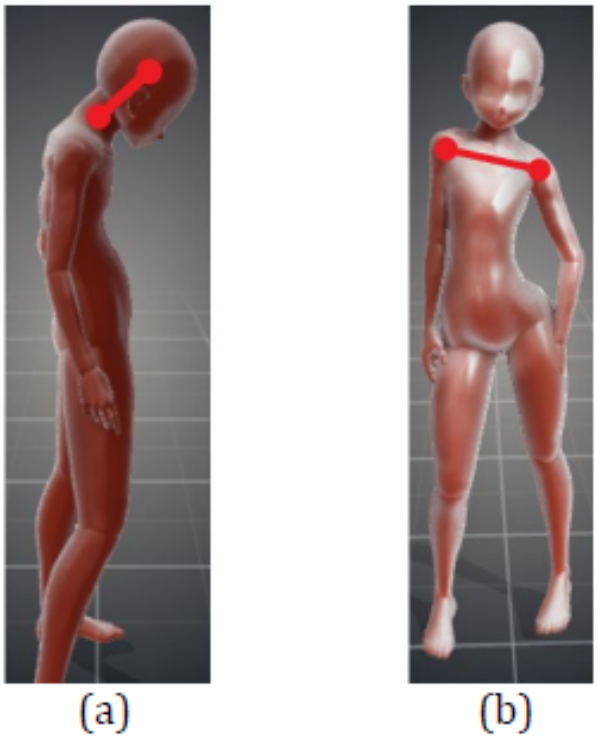


Figure 4. Interrelationship flowchart of the 3 subsystems when a new user utilizes the notebook computer's image recognition function to perform an automatic calibration of the smart necklace's internal parameters.

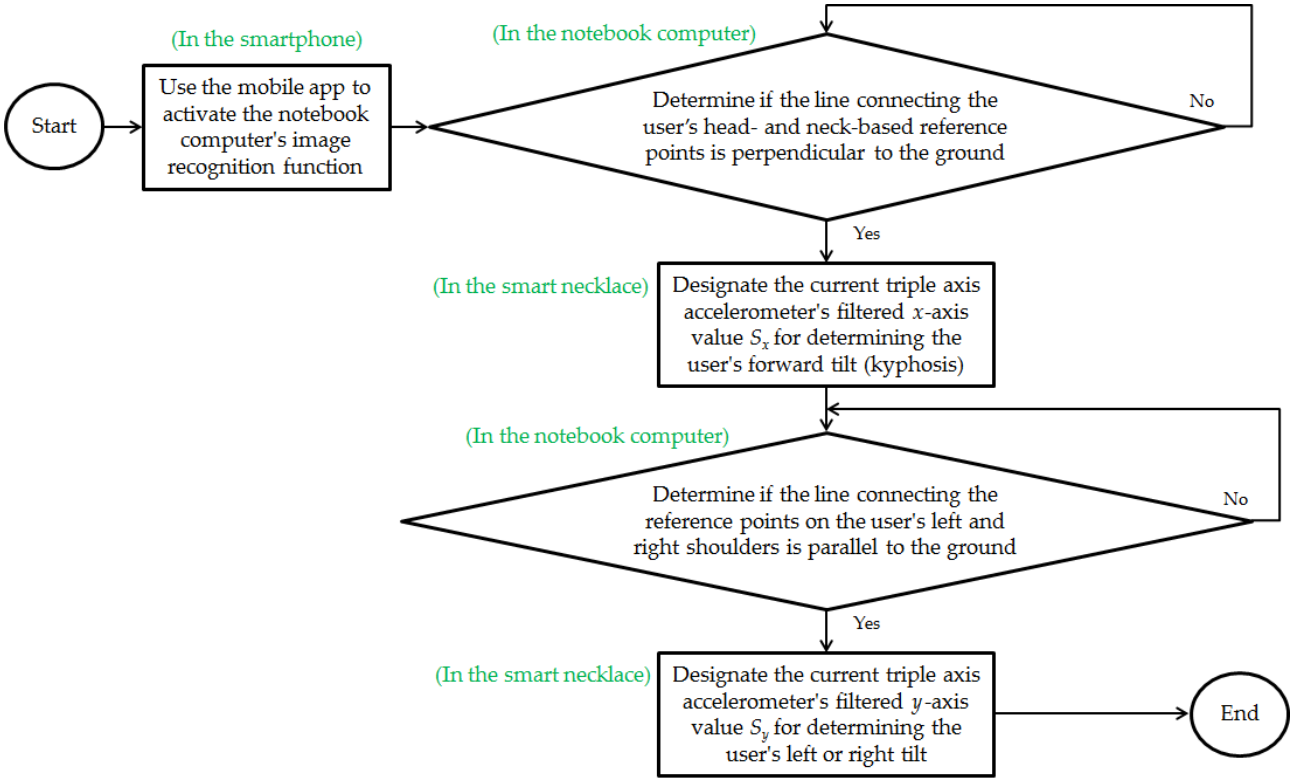


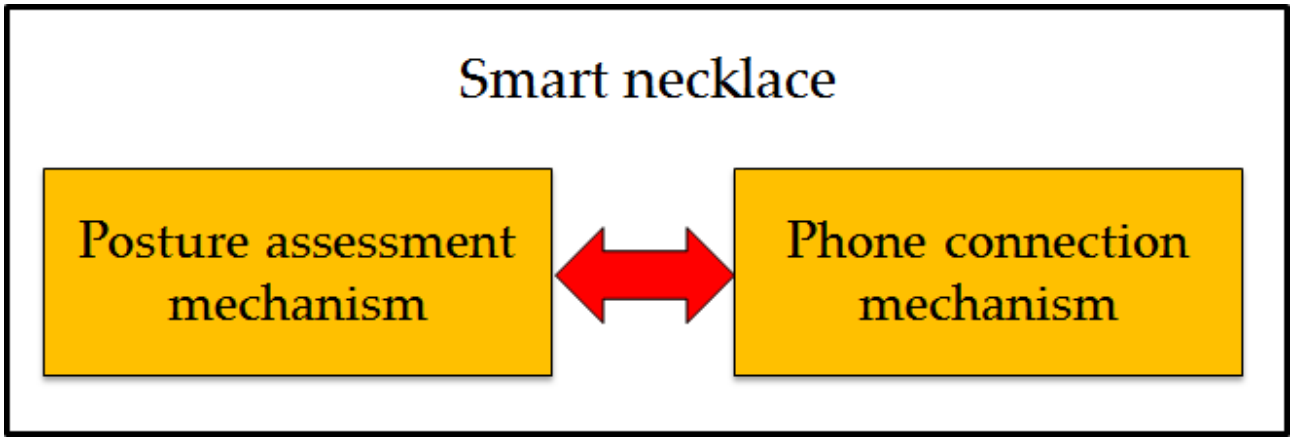
Figure 4 shows an interrelationship flowchart of the 3 subsystems when a new user utilizes the notebook computer's image recognition function to perform an automatic calibration of the smart necklace's internal parameters. The user first activates the smartphone's automatic calibration function, which then activates the notebook computer's skeletal structure image recognition function. Following this, the notebook computer assesses the user's side posture by determining if the line connecting his or her head- and neck-based reference points is perpendicular to the ground. When the line is not perpendicular to the ground, the notebook computer will continue with the image assessment process. When the line is perpendicular (indicating proper posture), the notebook computer will send a signal to the smart necklace, where the triple axis accelerometer's *x*-axis value will be digitally filtered (the current value and the previous values are averaged) and designated as the standard value S_x for determining the user's forward tilt (kyphosis). Next, the notebook computer will assess the user's frontal posture to determine if the line connecting the reference points on the user's left and right shoulders is parallel to the ground. If the line is not parallel to the ground, the image assessment process will continue. Conversely, if the line is parallel to the ground, the notebook computer will send a signal to the smart necklace, where the triple axis accelerometer's *y*-axis value will be digitally filtered and designated as the standard value S_y for determining the user's left or right tilt. After the internal parameters of the smart necklace have been successfully calibrated, the smartphone will receive a message informing the user of the successful calibration. At this point, the program running on the notebook computer will close.

In the following sections, the design of the smart necklace, operation of the notebook computer, and design and operation of the smartphone will be described.

Smart Necklace

Overall, 2 mechanisms were incorporated into the design of the smart necklace, namely, the posture assessment and phone connection mechanisms, which operate concurrently (see Figure 5). The posture assessment mechanism allows for the posture analysis of the gravitational acceleration data recorded by the MPU-6050, whereas the phone connection mechanism allows for reminder signals to be sent to the smartphone.

Figure 5. Two mechanisms of the smart necklace, namely, the posture assessment and phone connection mechanisms, responsible for performing posture assessments and sending reminder signals to the smart phone, respectively.



Posture Assessment Mechanism

Figure 6 is a flowchart describing the posture assessment performed by the smart necklace. The core of the assessment involves the use of a digital hysteresis comparator (see Figure 7), which contains 4 key elements, namely, the input (*in*), output (*out*), and 2 thresholds ($Lref_i$ and $Href_i$). It should be noted that Figure 6 implicitly shows 3 comparators, which are used for the frontal, left, and right tilt values. Here, the left and right tilt indexes pertaining to the thresholds are expressed as *yl* and *yr*, respectively. We defined A_i as the value obtained after digitally filtering the triple axis accelerometer's continuously changing *i*-axis value (note that S_i is the A_i when the user's posture is correct). The Raspberry Pi 3 utilizes an I2C transfer protocol to read the data and compare the digitally filtered values (ie, A_i) with the thresholds ($Lref_i$ and $Href_i$), so as to determine if the user's posture is correct. In other words, A_i is substituted into *in*. When *out* is equal to 0, no reminder signals will be sent. However, when *in* becomes higher than $Href_i$, the value of *out* then becomes 1 and a reminder signal (indicating that the smart necklace has detected poor posture) will be sent via the phone connection mechanism. When *out* is initially equal to 1, a reminder signal will be sent. However, when in is lower than $Lref_i$, the value of *out* then becomes 0 and no further reminder signals will be sent (indicating that the smart necklace has determined the user's posture to be correct).

Figure 6. Flowchart describing the posture assessment performed by the smart necklace.

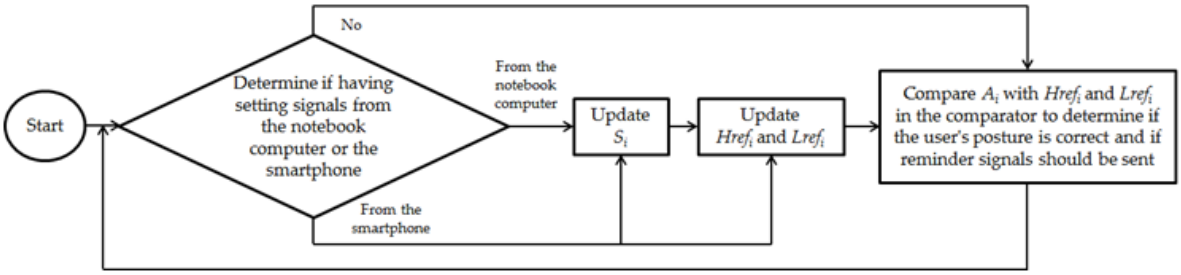
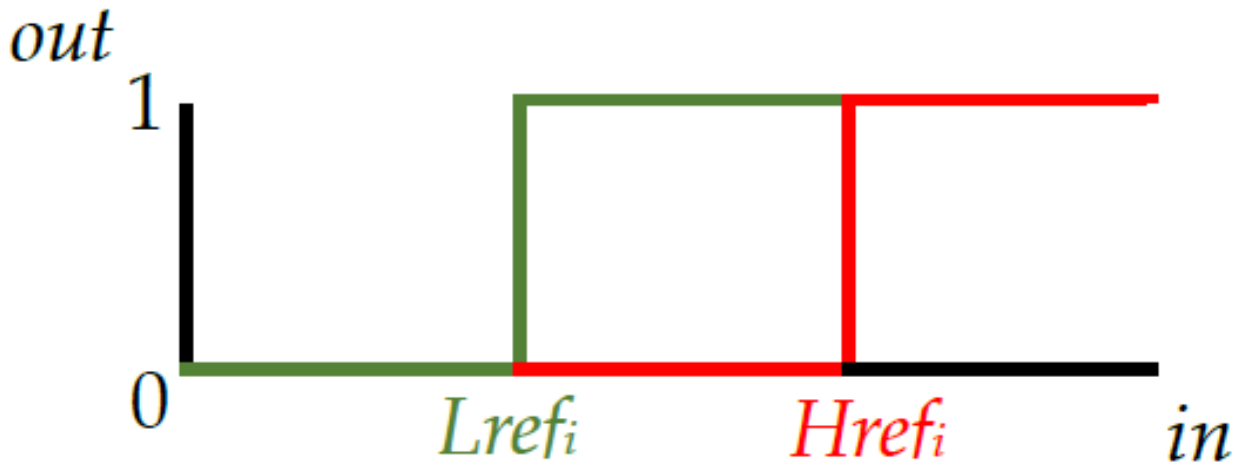


Figure 7. Hysteresis comparator.



Each individual's skeletal structure is different, and shoulder width is an important indicator that helps to define the characteristics of a skeletal structure. On the basis of the shoulder width characteristics of a normal individual, 3 users (whose shoulder widths were generally considered to be short, moderate, and long) were selected for the tests. To some degree, this selection method allowed us to cover the entire range of normal skeletal structures. Thus, this design approach should be valid for the general population (however, the system's threshold values must be set separately for skeletal structures with abnormal characteristics, and it is recommended that future studies could focus on this aspect of the research).

The threshold values were set based on posture assessments of the 3 users (indexed by *j*, with *j*=1, 2, 3) that were performed by 3 orthopedists. First, a user would tilt forward until the orthopedist determines that a poor posture has been achieved, at which point the triple axis accelerometer's digitally filtered *x*-axis value (ie, the value) would be recorded. Next, the user would tilt leftward until the orthopedist determines that a poor posture has been achieved, at which point the triple axis accelerometer's digitally filtered *y*-axis value (ie, the value) would be recorded. Finally, the user would tilt rightward until the orthopedist determines that a poor posture has been achieved, at which point the triple axis accelerometer's digitally filtered *y*-axis value (ie, the value) would be recorded. On the basis of the standards set by the 3 orthopedists, the values (ie, , *i*=*x*, *yl*, *yr*) recorded for the 3 users are presented in Tables 1-3, respectively, with | – | representing the change in user *j*'s tilt toward direction *i* at the point when the value was recorded.

To implement stricter posture assessment standards for the system, the value ($\min_{j=1, 2, 3}$ | – |), which represents the minimal degree of change, was used as the basis for the assessment. In other words, even a slight tendency to tilt will be identified as poor posture by the smart necklace, which will send a reminder to the user to prompt a correction.

More specifically, Tables 1-3 show that the minimum value of the change in the 3 users' forward tilt was 0.01159G. Thus, $Href_x$ was defined as:

$$Href_x=S_x-0.01159 \quad [1]$$

To prevent the smart necklace from sending reminders at an erratic frequency, numerous rounds of data evaluation were performed, and the value 0.00854G (which is approximately equal to the value of 0.01159G minus 0.003G) was selected for use to set $Lref_x$. Thus, $Lref_x$ was defined as:

$$Lref_x=S_x-0.00854 \quad [2]$$

Table 1. Measurement results of user number 1 based on orthopedists' standards.

Model	Value (G)	Change (G) | – |
	0.99288	—[a]
	0.93354	0.05934
	−0.06689	—
	0.00896	0.07585
	−0.17777	0.11087

[a]Not applicable.

Table 2. Measurement results of user number 2 based on orthopedists' standards.

Model	Value (G)	Change (G) | – |
	0.98999	—[a]
	0.94897	0.04102
	−0.16921	—
	−0.14087	0.02834
	−0.21952	0.05031

[a]Not applicable.

Table 3. Measurement results of user number 3 based on orthopedists' standards.

Model	Value (G)	Change (G) | – |
	0.99846	—[a]
	0.98687	0.01159
	−0.06850	—
	−0.01312	0.05538
	−0.12673	0.05823

[a]Not applicable.

Furthermore, the minimum value of the change in the 3 users' left/right tilt was 0.02834G, hence the use of this value to set $Href_{yl}$ and $Href_{yr}$ and the selection of the value 0.02441 (which is approximately equal to the value of 0.02834G minus 0.004G) to set $Lref_{yl}$ and $Lref_{yr}$. Thus, $Href_{yl}$, $Lref_{yl}$, $Href_{yr}$, and $Lref_{yr}$ were set as:

$$Href_{yl}=S_y+0.02834 \quad [3]$$

$$Lref_{yl}=S_y+0.02441 \quad [4]$$

$$Href_{yr}=S_y-0.02834 \quad [5]$$

and

$$Lref_{yr}=S_y-0.02441 \quad [6]$$

Phone Connection Mechanism

Figure 8 shows the process whereby a reminder signal is sent from the smart necklace to the smartphone. The smart necklace first waits for the smartphone to establish a connection with it. If a connection is not established, the smart necklace will continue to wait. If a connection is established, the posture assessment mechanism will be used to determine if a reminder signal should be sent (ie, the outcome described in the *Posture Assessment Mechanism* section). If the outcome involves the sending of a reminder, the smart necklace will send a reminder signal to the smartphone. The aforementioned process is repeated continuously to enable the continuous assessment of the user's posture, that is, until the set operating time is reached and the system's operation ends (the overall system shuts down).

Notebook Computer

Figure 9 shows the operating process whereby the notebook computer is used to perform posture recognition. Among the considerable number of methods used to detect human skeletal structures, the method proposed by Microsoft [18] was selected for this study. This method involves the use of the Kinect device to obtain depth image data, which are then analyzed using random decision trees and forests [19] to identify joint points. Once a user's joint points are identified, he or she will turn to allow the camera to capture his or her side profile. Calculations are then performed to determine if the line connecting his or her head- and neck-based reference points is perpendicular to the ground, so as to determine the presence or absence of kyphosis. Next, the user will turn again to allow the camera to capture his or her front profile. Calculations are then performed to determine if the line connecting the user's left and right shoulders is parallel to the ground, so as to determine if the user's standing posture is correct. The use of this method makes it possible to perform posture assessments and, thus, calibrate the smart necklace's internal parameters.

Figure 8. Flowchart describing the process whereby a reminder signal is sent from the smart necklace to the smartphone.

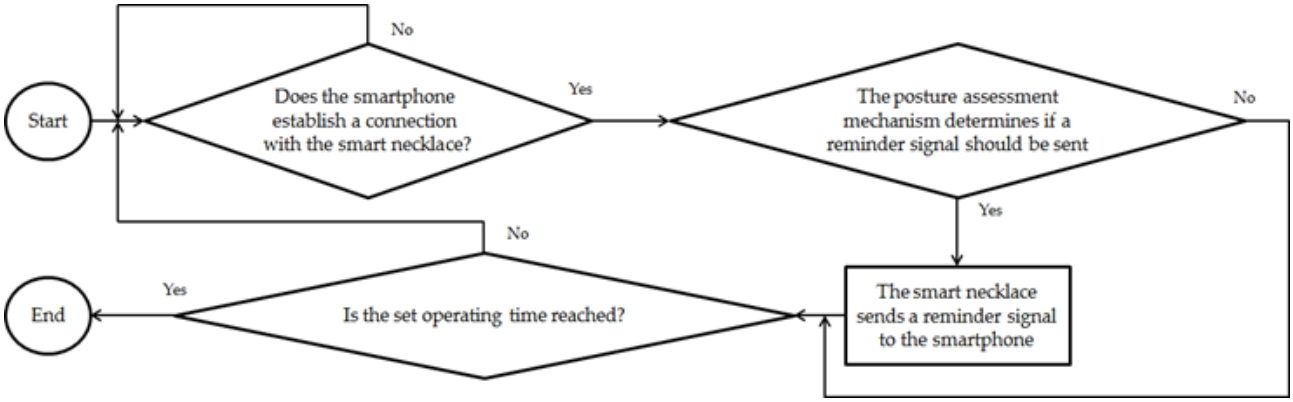
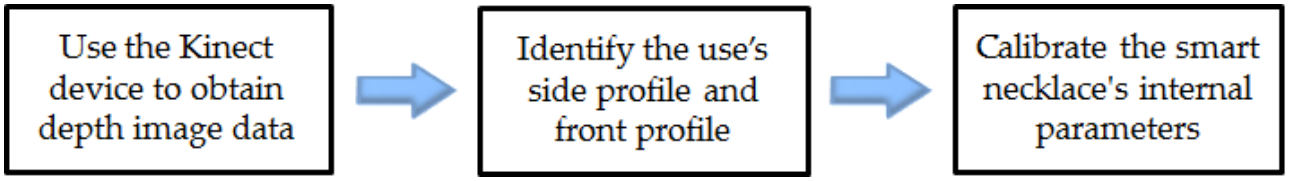


Figure 9. Operating process whereby the notebook computer is used to perform posture recognition.

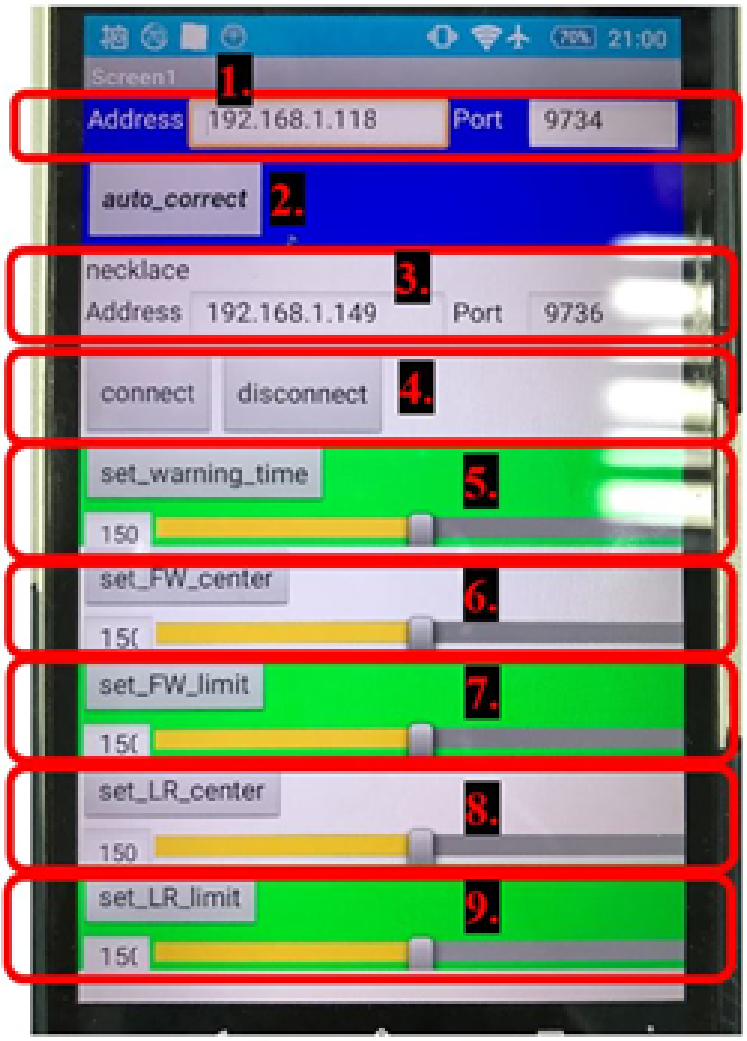


Smartphone

The App Inventor 2 [20], which was developed by the Massachusetts Institute of Technology Center for Mobile Learning, is used in this study as it does away with the need to use Java to open up Blocks Editor (which is integrated into the website and can be used immediately) and makes it easy for developers to write Android apps. For the human-machine interface, 2 IPv4 addresses are required for signal transmission, with one serving as the IP address for the camera linked to the notebook computer and the other serving as the IP address for the smart necklace (see items 1 and 3 of Figure 10). The *auto_correct* button (ie, item 2 of Figure 10) is used to "activate the notebook computer's image recognition function and the automatic calibration of the smart necklace's internal parameters." The *connect* and *disconnect* buttons (item 4 of Figure 10) are used to "enable the smartphone to connect to the smart necklace and receive reminder signals from the smart necklace" and "disconnect the smartphone and smart necklace (no reminder signals will be received)," respectively. Furthermore, it is possible to manually set the smartphone's internal parameters using items 5 to 9 of Figure 10. The *set_warning_time* button (ie, item 5) is used to "adjust the frequency at which the smart necklace sends reminder signals when poor posture is detected (the higher the value, the higher the frequency)." The *set_FW_center* button (ie, item 6) is used to "set parameter *a* (adjustable range of 1-300 and central value of 150) so as to make minor adjustments to the S_x value of the smart necklace," with the adjustment formula being:

$$S_x=S_x\times(a/150) \text{ [7]}$$

Figure 10. Mobile app interface.



The *set_FW_limit* button (ie, item 7) is used to "set parameter *b* (adjustable range of 1-300 and central value of 150) so as to make minor adjustments to the $Href_x$ and $Lref_x$ values of the smart necklace," with the adjustment formulas being:

$$Href_x=S_x–(0.01159×[b/150]) \quad [8]$$

and

$$Lref_x=S_x–(0.00854×[b/150]) \quad [9]$$

The *set_LR_center* button (ie, item 8) is used to "set parameter *c* (adjustable range of 1-300 and central value of 150) so as to make minor adjustments to the S_y value of the smart necklace," with the adjustment formula being:

$$S_y=S_y×(c/150) \quad [10]$$

The *set_LR_limit* button (ie, item 9) is used to "set parameter *d* (adjustable range of 1-300 and central value of 150) so as to make minor adjustments to the $Href_{yl}$, $Lref_{yl}$, $Href_{yr}$, and $Lref_{yr}$ values of the smart necklace," with the adjustment formulas being:

$$Href_{yl}=S_y+(0.02834×[d/150]) \quad [11]$$

$$Lref_{yl}=S_y+(0.02441×[d/150]) \quad [12]$$

$$Href_{yr}=S_y–(0.02834×[d/150]) \quad [13]$$

and

$$Lref_{yr}=S_y–(0.02441×[d/150]) \quad [14]$$

It should be noted that the 4 parameters *a*, *b*, *c*, and *d* are independent of each other when the adjustments are being made.

Recall that formulas (1) to (6) involve the utilization of the notebook computer's image recognition function to perform an automatic calibration of the smart necklace's S_i value (primary calibration mechanism), whereas formulas (7) to (14) correspond to the fine manual adjustments to the smart necklace's internal parameters that a user may want to further perform (secondary calibration mechanism). Thus, when formulas (7) to (14) are used, the S_i, $Href_i$, and $Lref_i$ values that are already present in formulas (1) to (6) will be updated by formulas (7) to (14).

Results

This section describes the practical tests performed on the proposed posture correction system (see Figure 1 for hardware specifications) to demonstrate its effectiveness. During the tests, the smart necklace was worn around the back of the user's neck. The position and appearance of the smart necklace are shown in Figure 11. In addition, a demo video has been provided to further demonstrate the advantages of our system (see Multimedia Appendix 1).

The Raspberry Pi 3 reads the MPU-6050 value once every 0.1 seconds and repeats this process 99 times to record the relationship between the user's body movement and the readings. Figure 12 shows the values recorded when the user's body was tilted forward, including both the prefiltered and filtered values. The horizontal and vertical axes represent time (measured at 0.1-second intervals) and gravitational acceleration G, respectively, and ax represents the triple axis accelerometer's *x*-axis value, which decreased from 1G to approximately 0.73G. Figure 13 shows the values recorded when the user's body was tilted leftward, with ay representing the triple axis accelerometer's *y*-axis value, which is shown to have increased from 0.1G to approximately 0.9G. When the user's body was tilted rightward, the values changed in the opposite direction. As shown in Figure 14, the ay value decreased from 0.2G to around −0.6G. Furthermore, ax was affected to a greater degree than ay under walking conditions, which led to a considerable level of noise interference. To address this issue, both the prefiltered and filtered ax values were recorded to demonstrate the changes that occurred (see Figure 15).

Figure 11. The position and appearance of the smart necklace when worn around the back of the user's neck, with the blue and red arrows indicating the directions of the triple axis accelerometer's *x*-axis and *y*-axis, respectively.

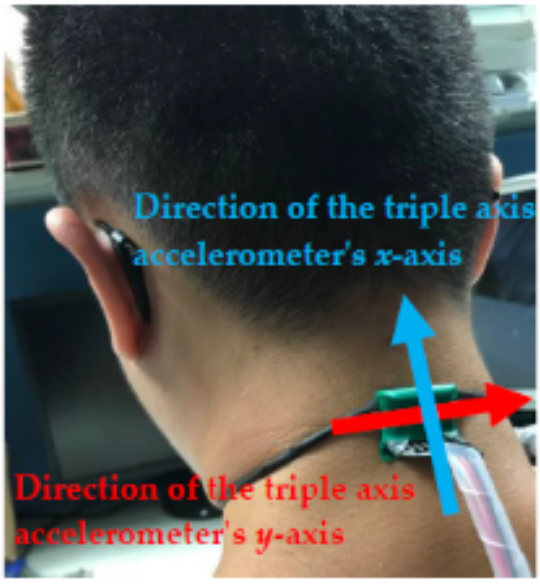


Figure 12. Triple axis accelerometer's pre-filtered *x*-axis values (ax) and filtered *x*-axis values (ax being filtered) when the user's body was tilted forward.

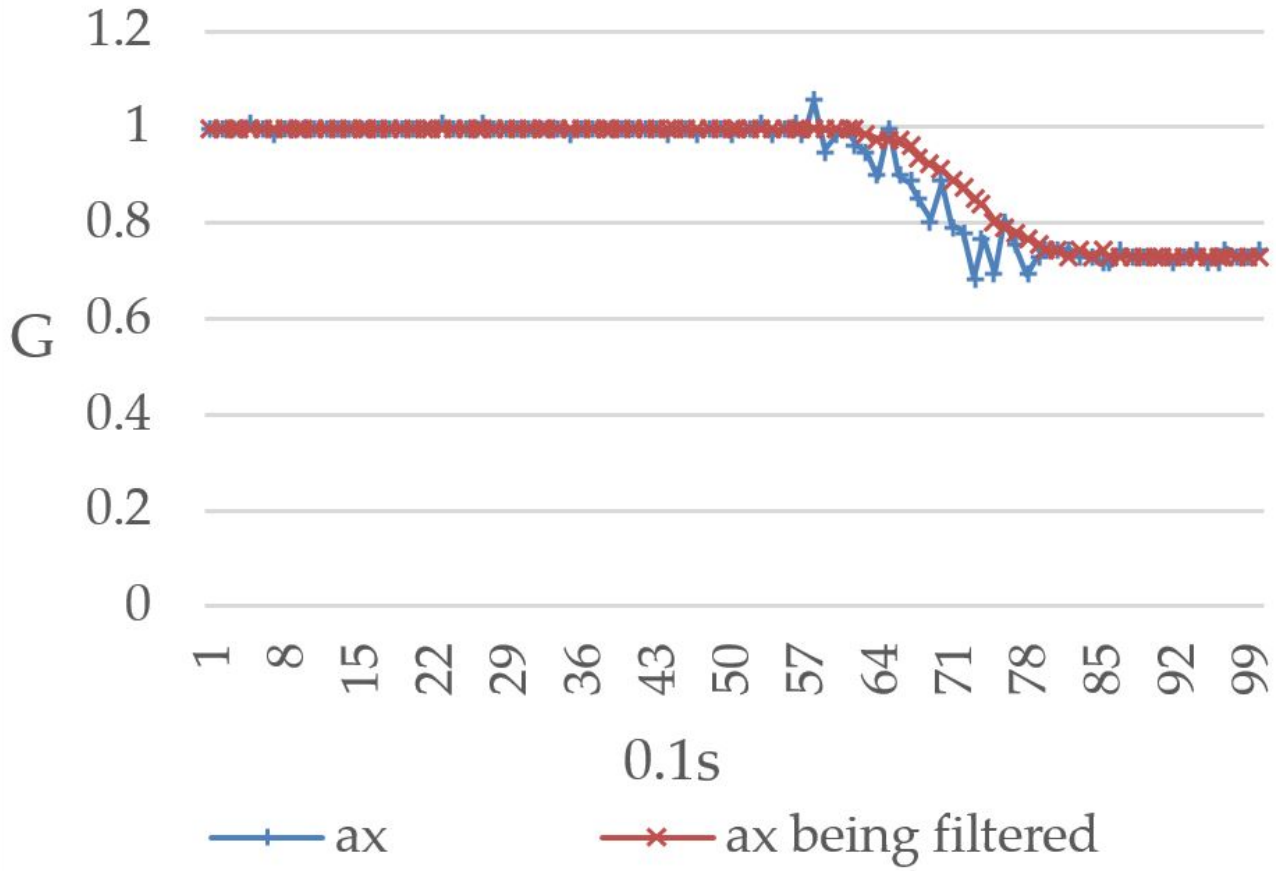


Figure 13. Triple axis accelerometer's pre-filtered *y*-axis values (ay) and filtered *y*-axis values (ay being filtered) when the user's body was tilted leftward.

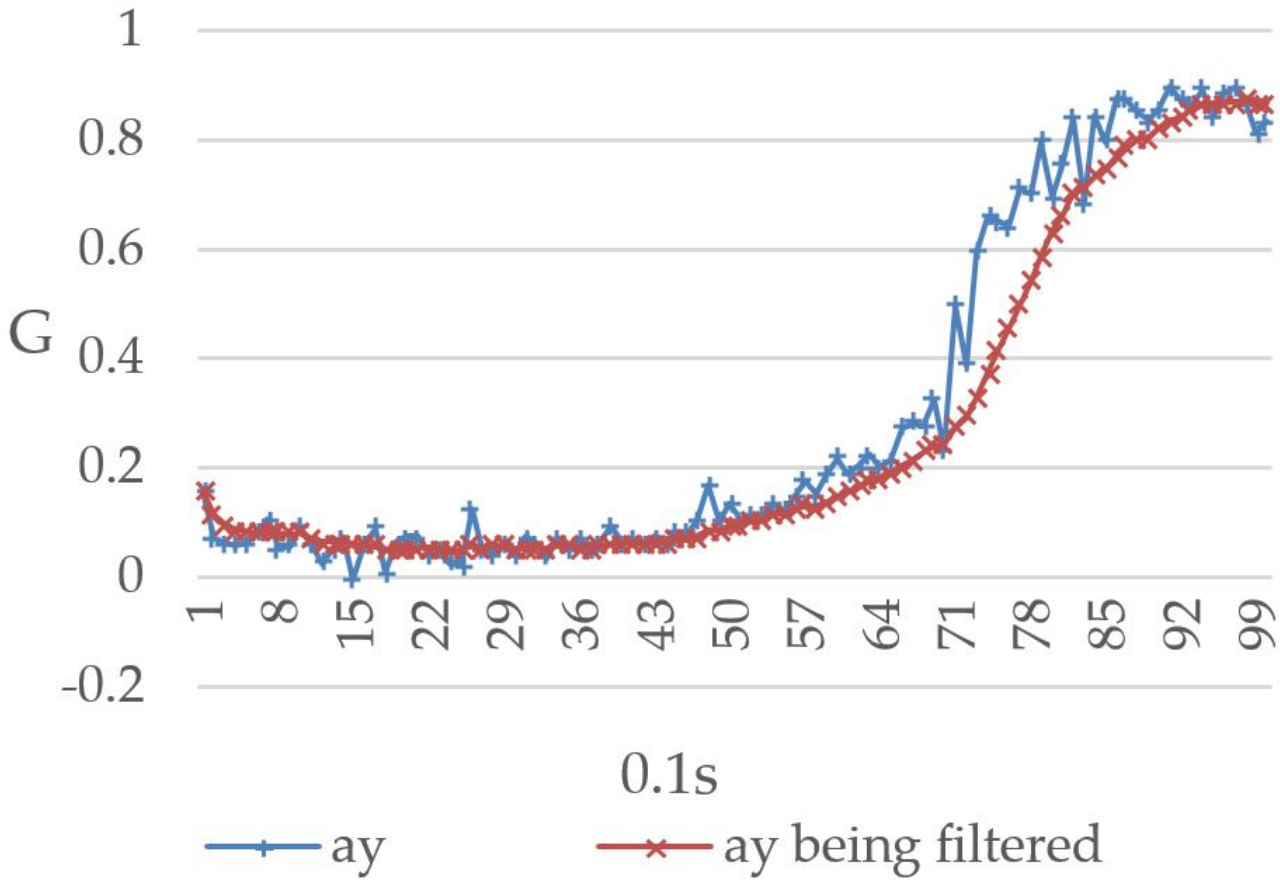


Figure 14. Triple axis accelerometer's pre-filtered *y*-axis values (ay) and filtered *y*-axis values (ay being filtered) when the user's body was tilted rightward.

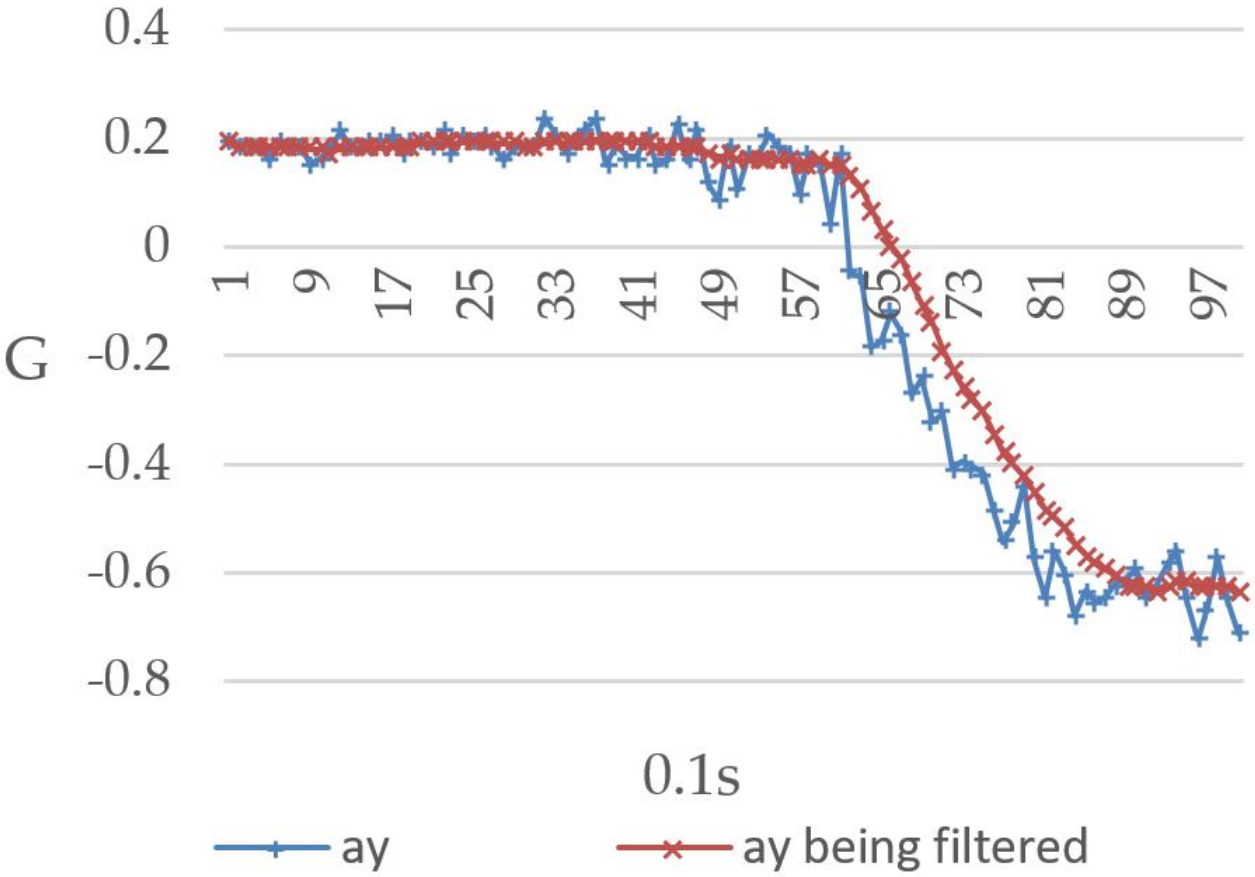
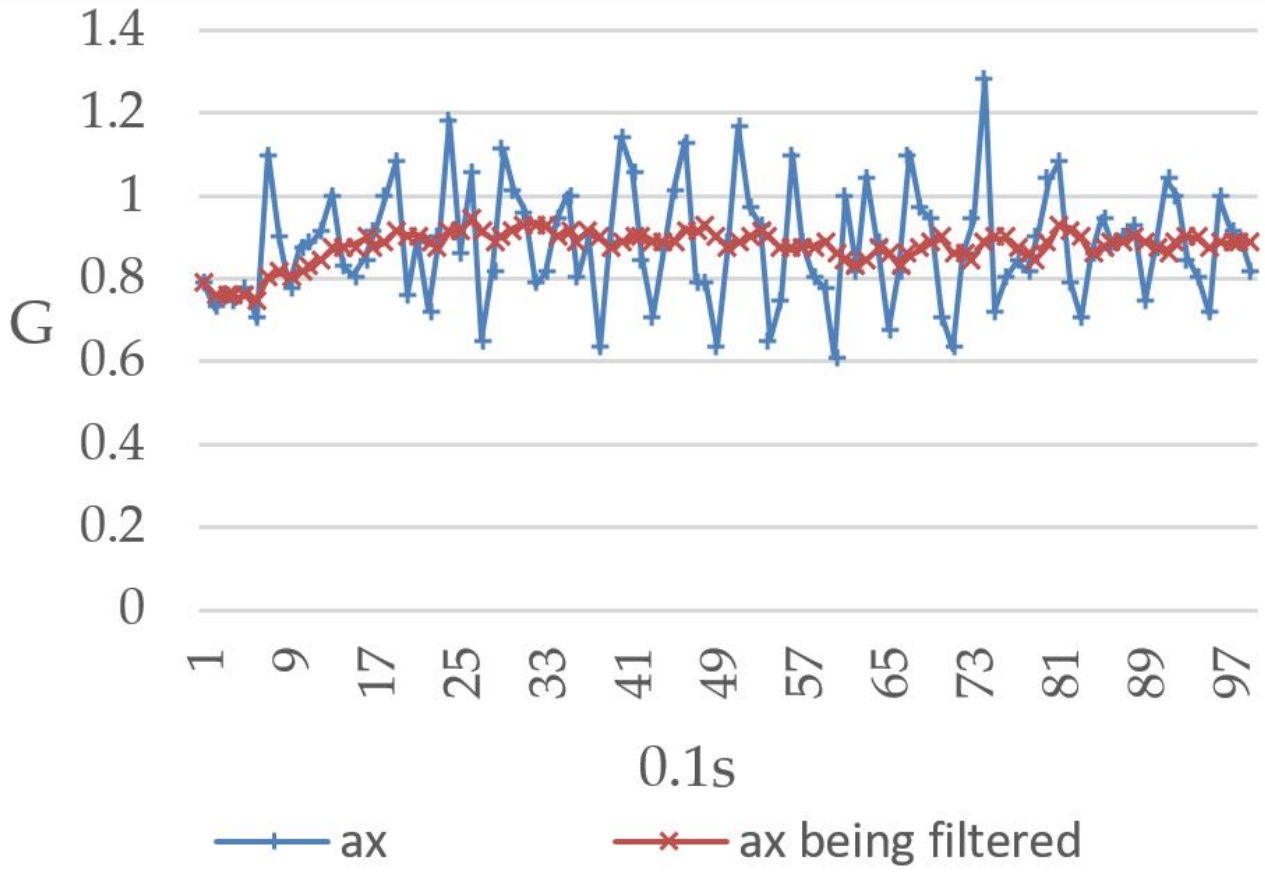


Figure 15. Triple axis accelerometer's pre-filtered *x*-axis values (ax) and filtered *x*-axis values (ax being filtered) under walking conditions.



Given that the values will become somewhat unstable when there is noise interference, the mean value of the current reading and the previous 9 readings was thus used in this study for the purpose of eliminating noise interference and preventing the system from making errors. Looking at Figures 12-15, it is clear that the filtering has caused the values to become more stable and reduced signal fluctuations.

Automatic Calibration of Smart Necklace's Internal Parameters Through the Notebook Computer's Image Recognition Function

Measurements Performed on a Single User

The user first activates the notebook computer (which is linked to a camera). Next, the mobile app is activated and the *auto_correct* button (see Figure 16) is clicked to activate the notebook computer's image recognition function and perform the automatic calibration of the smart necklace's internal parameters, after which the *Connected* status will be displayed on the smartphone. At this point, the notebook computer will attempt to detect the user's skeletal structure and mark out the head, neck, left shoulder, and right shoulder positions using 4 pink points. The user first turns to allow the camera to capture his side profile, and when the line connecting his head- and neck-based reference points is perpendicular to the ground, the computer will transmit a setting signal to the smart necklace to calibrate S_x and set it as the standard value for determining the presence of kyphosis. Upon the calibration's completion, the head- and neck-based points will disappear. Following this, the user then turns to face the camera frontally, and when the line connecting the reference points on the user's left and right shoulders is parallel to the ground, the computer will transmit a setting signal to the smart necklace to calibrate S_y and set it as the standard value for determining the presence of a left/right tilt. Upon the calibration's completion, the image program closes automatically and the *Correct Finish!!* status is displayed on the smartphone to indicate that the calibration has been completed.

Measurement Data of and Individual Differences Between 3 Users

Define and , respectively, as the forward tilt and left/right tilt standard values of the 3 users (indexed by j, with j=1, 2, 3) who were tested. These are the values obtained after the notebook computer's image recognition function was utilized to perform an automatic calibration of the smart necklace's internal parameters as shown in Figure 2 (to facilitate the description of the core concepts, the mobile app was not used to perform fine manual adjustments in this scenario).

Figure 17 is a 2-dimensional scatter plot showing the respective and values of the 3 users. As shown in the figure, the coordinate points (,) for each user's body are different, primarily because skeletal structures differ from individual to individual.

Thus, a user new to the system would have to recalibrate his or her and values to ensure that accuracy of his or her posture assessment.

Manipulation of the Smart Necklace's Internal Parameters via the Smartphone

Adjustment of the Smart Necklace's Internal Parameters

This section describes the manner in which a user can make further fine adjustments to the smart necklace's internal parameters. For example, the S_x value may be adjusted by dragging the corresponding horizontal slider to the desired value (see Figure 18) based on formula (7). The bigger this value is, the greater the extent to which the S_x value is being corrected toward the back (see the red arc in Figure 19 that is pointing toward "Backward"). Conversely, the smaller this value is, the greater the extent to which the S_x value is being corrected toward the front (see the red arc in Figure 19 that is pointing toward "Forward"). Subsequently, the *set_FW_center* button is clicked to transmit a setting signal to the smart necklace. This results in the appearance of the *Set forward center succeed* message on the screen of the smartphone, which indicates that the adjustment of S_x is complete.

Figure 16. Program that utilizes the notebook computer's image recognition function to perform an automatic calibration of the smart necklace's internal parameters.

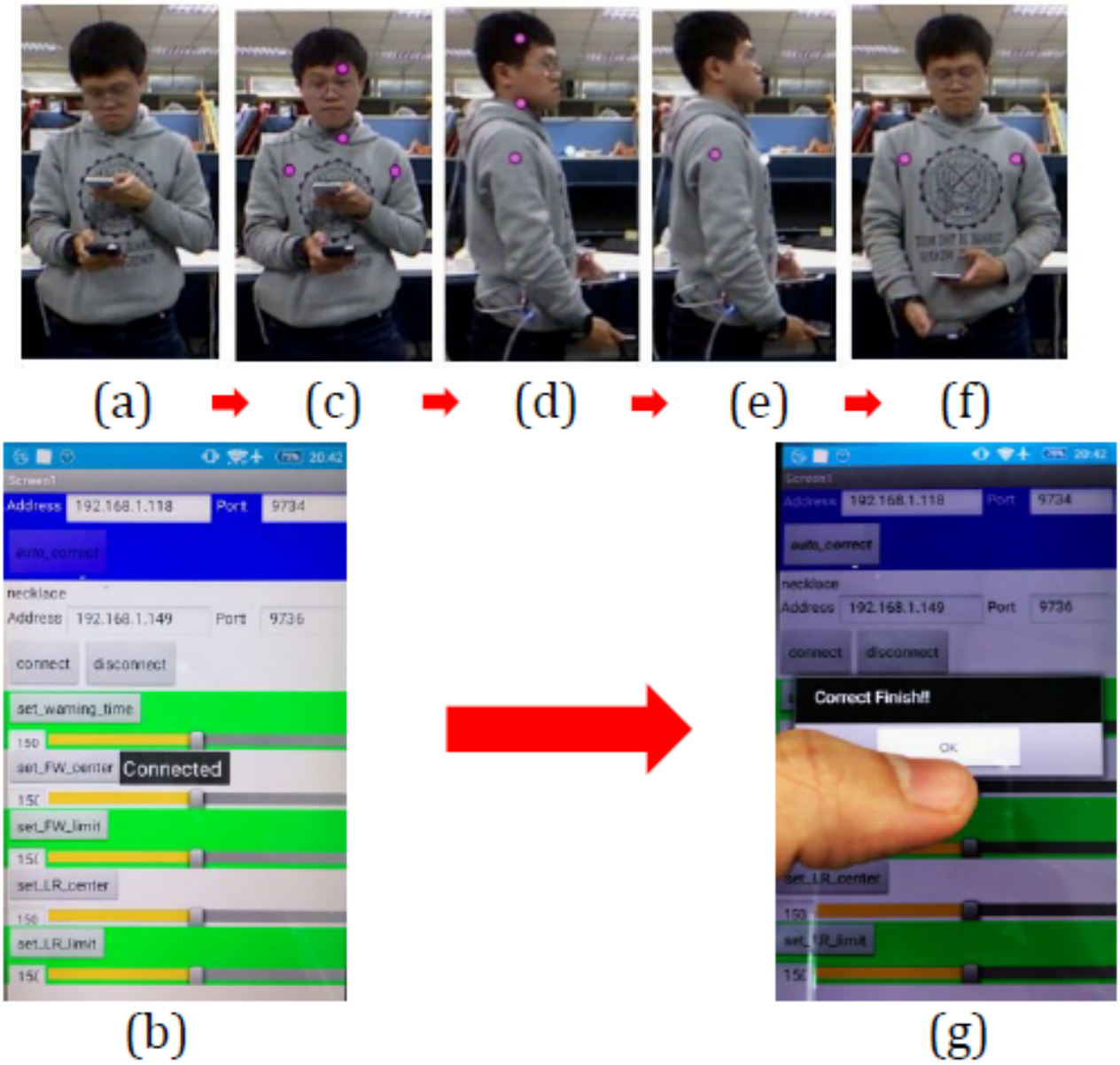


Figure 17. Distribution of the standard values of the 3 users (who were tested) across a 2-dimensional scatter plot.

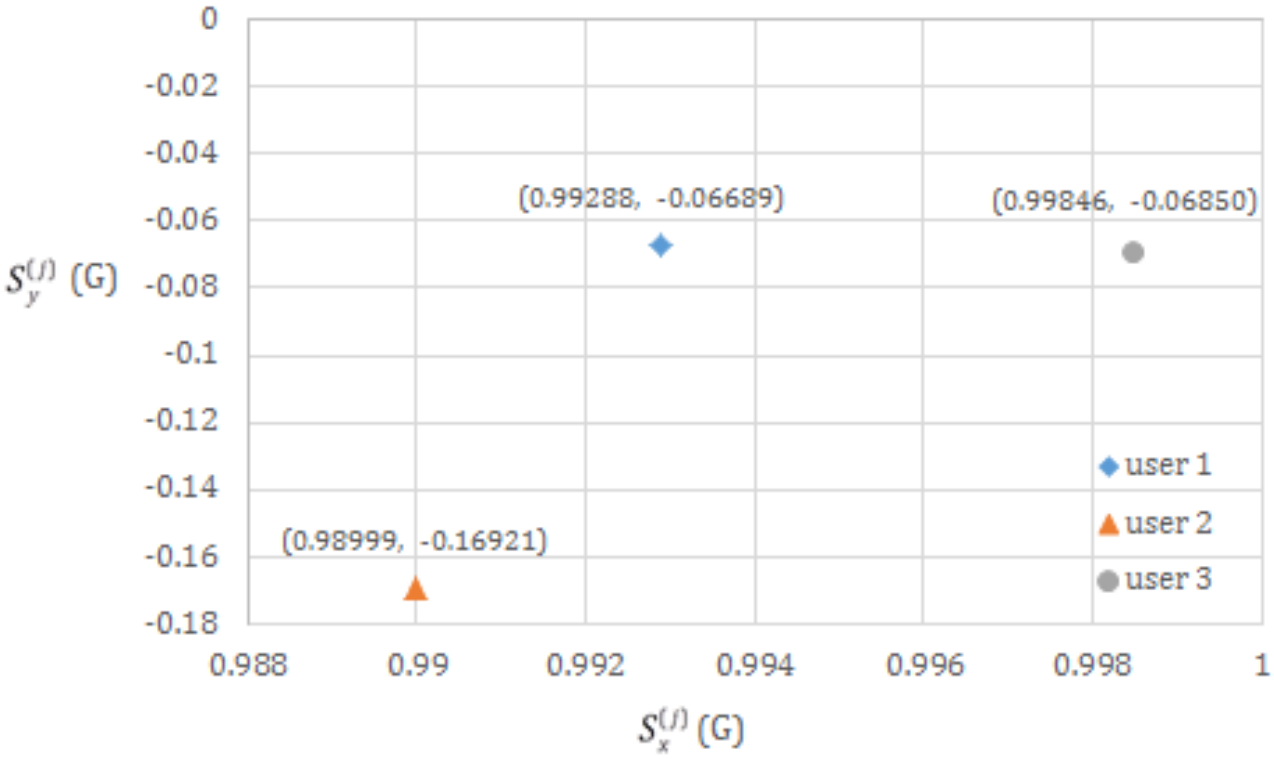


Figure 18. Setting of the smart necklace's internal parameters manually.

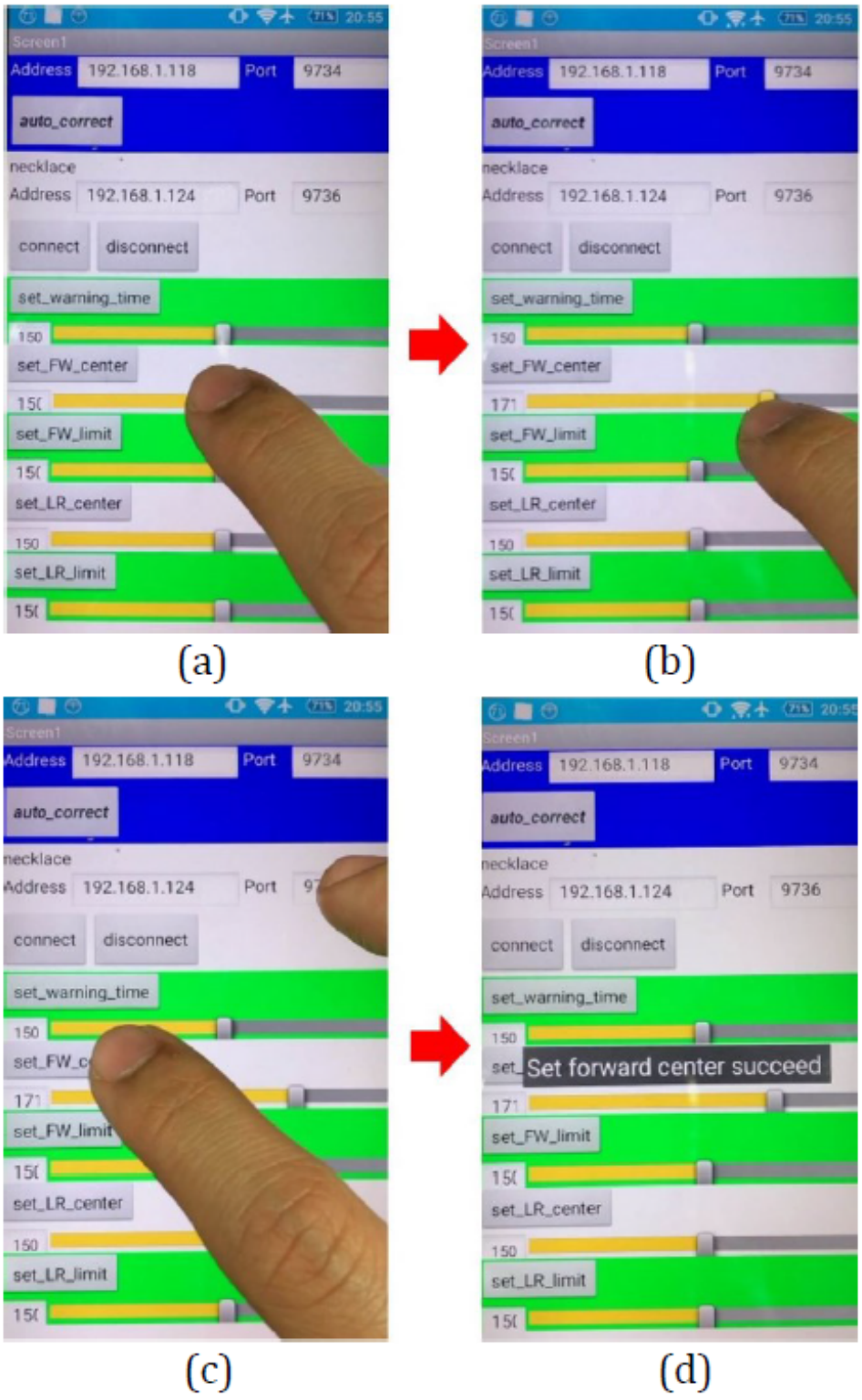
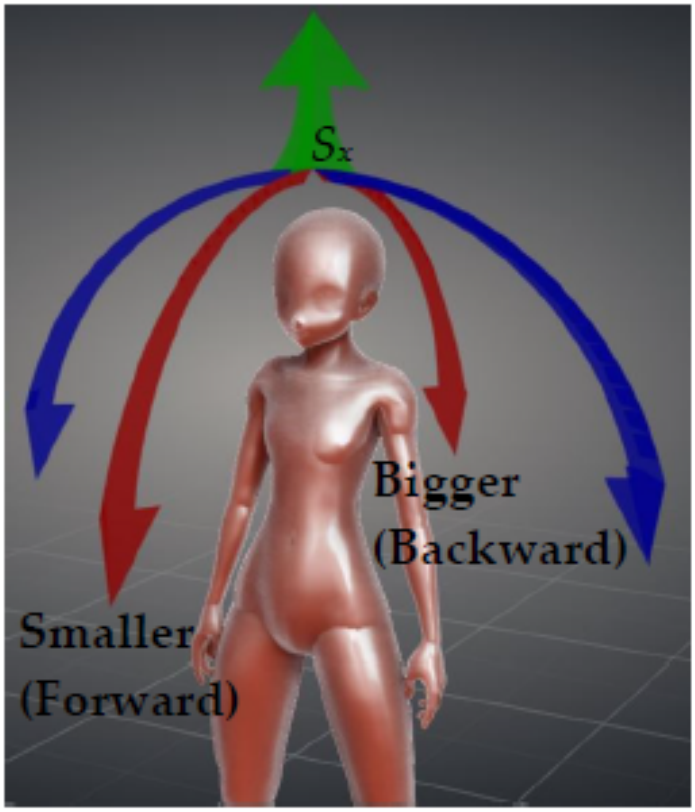


Figure 19. Diagram showing the correction of posture standard values.



Activation to Receive the Smart Necklace's Reminders

The mobile app is activated, and the *connect* button is clicked (see Figure 20) to link the smartphone and smart necklace. At this point, the *Connected* message will be displayed on the screen of the smartphone, which has now activated its function for receiving reminders from the smart necklace.

When a user displays the tendency to shift from a correct posture to an incorrect one, the corresponding yellow Light-Emitting Diode (LED) on the IT TRAINING expansion board will light up. As shown in Figure 21, the IT TRAINING expansion board has 4 LEDs. The activation of the middle LED indicates correct posture, the activation of the top LED indicates a forward tilt, the activation of the left LED indicates a left tilt, and the activation of the right LED indicates a right tilt. For instance, when a user shifts from a correct posture to one with a left and forward tilt, the left and top LEDs on the IT TRAINING expansion board will light up simultaneously, whereas the middle LED will turn off. At the same time, the smart necklace will send a reminder signal to the smartphone, causing the *Too Forward & Left!!* reminder to appear on the screen of the smartphone. This serves to remind the user that he is displaying a forward-left tilt that requires correction.

For the sake of completeness, another example is provided (see Figure 22). In this case, the user shifts from a correct posture to one with a right tilt; this change will be detected by the smart necklace and reflected on the IT TRAINING expansion board's LEDs. In other words, the middle LED will turn off while the right LED will light up. At the same time, the smart necklace will send a reminder signal to the smartphone, causing the *Too Right!!* reminder to appear on the screen of the smartphone. This serves to remind the user that he is displaying a right tilt that requires correction.

If a user does not want to receive any reminders regarding poor posture, he or she may click the *disconnect* button (see Figure 23) to cut off the connection between the smartphone and the smart necklace and cause the *Not Connected!* message to be displayed on the smartphone's screen. At this point, the smartphone has already deactivated its function for receiving reminders from the smart necklace.

Figure 20. Activation of the smartphone's function for receiving reminders from the smart necklace.

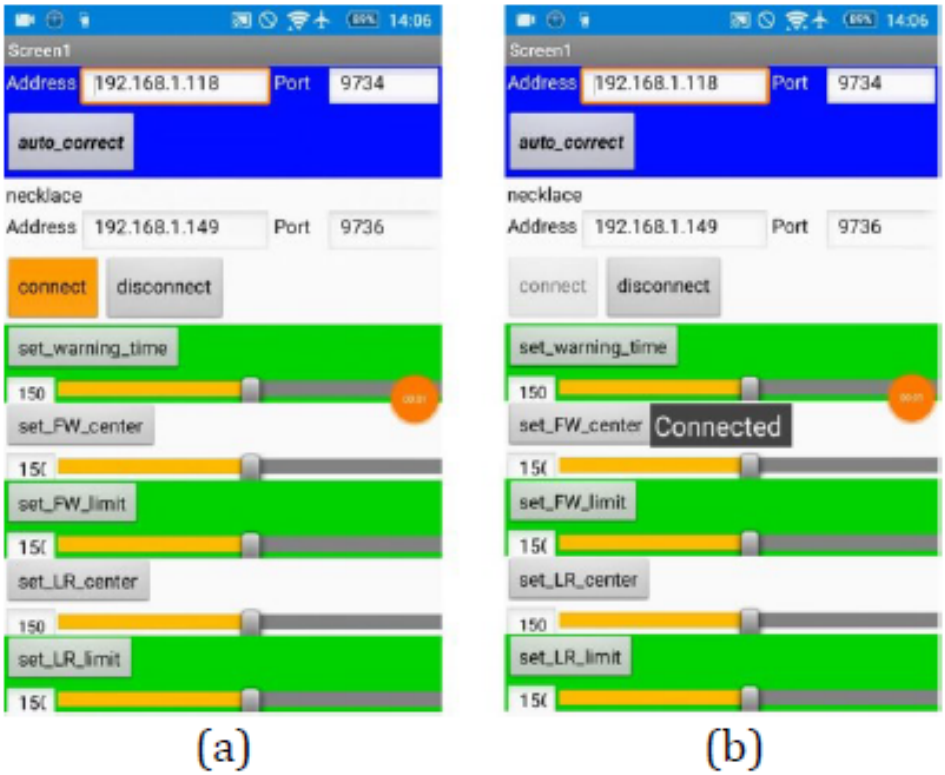


Figure 21. Posture with forward-left tilt.

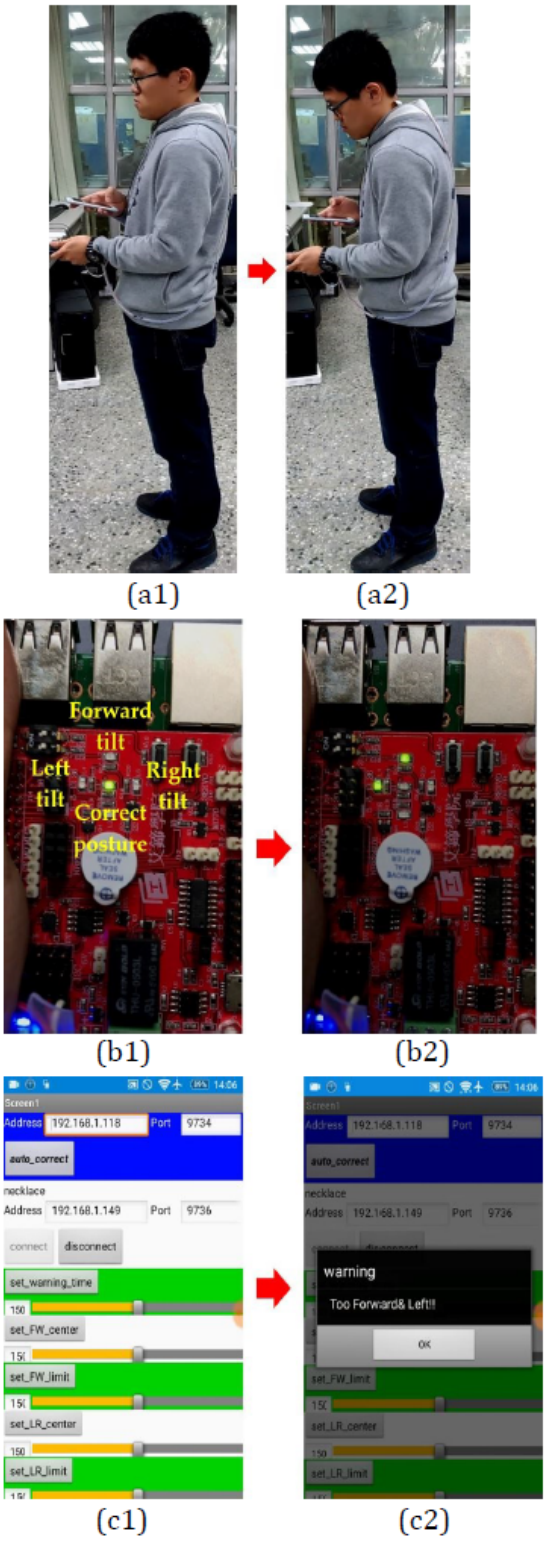


Figure 22. Posture with right tilt.

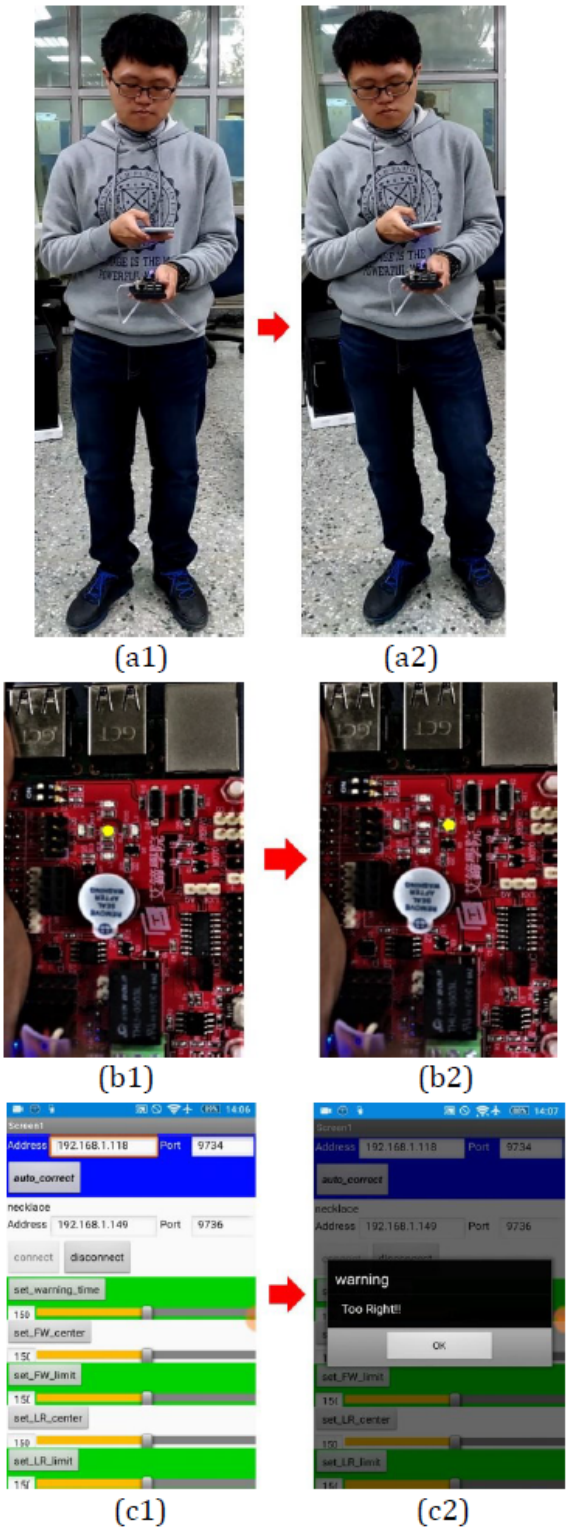
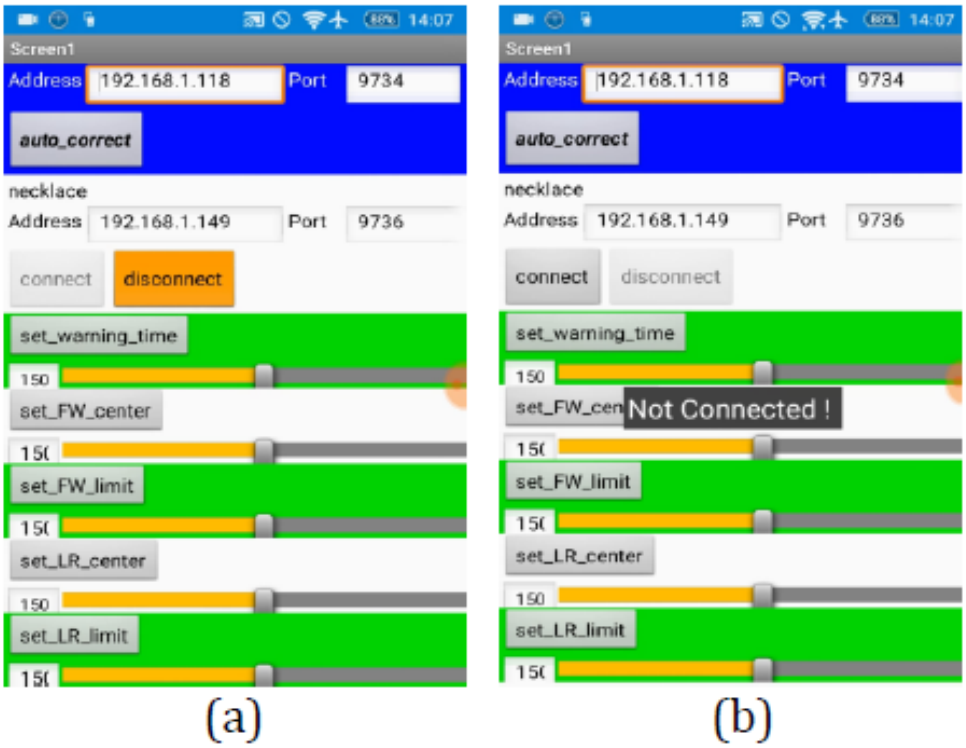


Figure 23. Deactivation of the smartphone's function for receiving reminders from the smart necklace.



Discussion

Principal Findings

In this study, a novel posture correction system was developed through the application and integration of image recognition, wearable sensing, and Internet of Things technologies, with this new system consisting of 3 subsystems: a smart necklace, notebook computer, and smartphone. The system constitutes a pioneering achievement with respect to research on the application of wearable sensing devices in posture correction.

The notebook computer is equipped with a Kinect camera and can be used to calibrate the smart necklace and ensure that it does not lose its effectiveness in the event that it is improperly used. The posture image recognition function of the notebook computer gives the user the ability to quickly calibrate the internal parameters of the smart necklace subsystem. By wearing the lightweight smart necklace and carrying the smartphone with the mobile reminder system, the user is able to self-correct his or her posture and prevent spine-related diseases without the use of corrective clothing that is heavy, thick, and uncomfortable. Moreover, because the system utilizes wireless communication among the interfaces of its different subsystems, the user does not have to deal with complex wiring, thus allowing him or her to utilize the new system in a convenient manner.

Limitations

The posture correction system's 3 subsystems communicate with each other via a Wi-Fi router, making it suitable for office or home use. However, this design also requires the 3 subsystems to be located in close proximity to each other, so as to enable the near-instant transmission of signals between them. Using the system used in this study as a foundation, future studies could explore ways to enable users to perform instant posture correction regardless of their location (ie, via a system that can be used in any environment).

Comparison With Prior Work

The developmental trends in information and communications technology are pointing toward the substantial application of technologies such as wearable sensors, image recognition, and the Internet of Things in smart medical systems. However, to our knowledge, few studies have looked into the effective integration and application of these technologies for posture correction. In addition, physicians typically do not recommend that patients suffering from only mild or temporary kyphosis symptoms make use of heavy, thick, or uncomfortable corrective clothing. At the same time, it is often difficult for patients who attempt to rely only on their own willpower to achieve good results. The system introduced herein was thus developed with these issues in mind to assist users in detecting poor posture, correcting their postures as appropriate, and, in turn, protecting themselves from spinal diseases resulting from the long-term effects of bad posture.

Conclusions

The spine is the most important group of bones in the human body, as it supports not only the weight of the body but also enables it to perform the full range of torsional motions. When poor posture is maintained on a frequent basis or over a long period, issues including spine-related problems or even the onset of premature aging and geriatric diseases among young adults may occur. Thus, the most basic approach to protecting our spines consists of fixing our bad postures. With that in mind, the proposed system can effectively enable a user to monitor and correct his or her own posture, which in turn will assist the

user in preventing spine-related diseases and, consequently, in living a healthier life. We believe that this system can contribute to our health and serve as a pioneering system with respect to the application of integrated forward-looking technologies to posture correction. Furthermore, this system is expected to meet current and future needs and possess commercial potential.

Acknowledgments

The authors wish to acknowledge the financial support of the Ministry of Science and Technology (MOST), Republic of China, under Contract MOST 105-2628-E-019-002-MY3.

Authors' Contributions

HYC, YLC, and CYL conceived the idea of the study; HYC and YLC designed the methods and experiments; CYL performed the experiments; and YLC wrote the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Demo video. A demo video for further demonstrating the effectiveness of the developed posture correction system.

[MP4 File (MP4 Video), 129MB - mhealth_v7i5e12293_app1.mp4]

References

1. Fallahzadeh R, Pedram M, Ghasemzadeh H. SmartSock: a wearable platform for context-aware assessment of ankle edema. Conf Proc IEEE Eng Med Biol Soc 2016 Dec;2016:6302-6306. [doi: 10.1109/EMBC.2016.7592169] [Medline: 28269690]
2. Dauz F, Strumban E, Maev RG. Wearable device for increasing the slow wave sleep stage by electrocutaneous stimulation. 2017 Presented at: 2017 IEEE 30th Canadian Conference on Electrical and Computer Engineering (CCECE); April 30-May 3, 2017; Windsor, ON, Canada.
3. Sundaravadivel P, Mohanty SP, Kougianos E, Yanambaka VP, Ganapathiraju MK. Smart-walk: An intelligent physiological monitoring system for smart families. 2018 Presented at: 2018 IEEE International Conference on Consumer Electronics (ICCE); January 12-14, 2018; Las Vegas, NV, USA.
4. Boateng G, Batsis JA, Proctor P, Halter R, Kotz D. GeriActive: Wearable app for monitoring and encouraging physical activity among older adults. 2018 Presented at: 2018 IEEE 15th International Conference on Wearable and Implantable Body Sensor Networks (BSN); March 4-7, 2018; Las Vegas, NV, USA.
5. Liu J, Chen J, Jiang H, Jia W, Lin Q, Wang Z. Activity recognition in wearable ECG monitoring aided by accelerometer data. 2018 Presented at: 2018 IEEE International Symposium on Circuits and Systems (ISCAS); May 27-30, 2018; Florence, Italy.
6. Martin A, Voix J. In-ear audio wearable: measurement of heart and breathing rates for health and safety monitoring. IEEE Trans Biomed Eng 2018 Dec;65(6):1256-1263. [doi: 10.1109/TBME.2017.2720463] [Medline: 28678697]
7. Surrel G, Aminifar A, Rincón F, Murali S, Atienza D. Online obstructive sleep apnea detection on medical wearable sensors. IEEE Trans Biomed Circuits Syst 2018;12(4):762-773. [doi: 10.1109/TBCAS.2018.2824659]
8. Takei Y, Yoshida M, Takeshita T, Kobayashi T. Wearable muscle training and monitoring device. 2018 Presented at: 2018 IEEE Micro Electro Mechanical Systems (MEMS); January 21-25, 2018; Belfast, UK.
9. Durbhaka GK. Adaptive wearable smart fabric based on body posture and temperature. 2016 Presented at: 2016 2nd International Conference on Advances in Computing, Communication, & Automation (ICACCA); September 30-October 1, 2016; Bareilly, India.
10. Gia TN, Jiang M, Sarker VK, Rahmani AM, Westerlund T, Liljeberg P, et al. Low-cost fog-assisted health-care IoT system with energy-efficient sensor nodes. 2017 Presented at: 2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC); June 26-30, 2017; Valencia, Spain.
11. Chinello F, Pacchierotti C, Malvezzi M, Prattichizzo D. A three revolute-revolute-spherical wearable fingertip cutaneous device for stiffness rendering. IEEE Trans Haptics 2018;11(1):39-50. [doi: 10.1109/TOH.2017.2755015]
12. Meli L, Hussain I, Aurilio M, Malvezzi M, O'Malley MK, Prattichizzo D. The hBracelet: a wearable haptic device for the distributed mechanotactile stimulation of the upper limb. IEEE Robot Autom Lett 2018;3(3):2198-2205. [doi: 10.1109/LRA.2018.2810958]
13. Goncharenko AA, Voronova LI, Voronov VI, Ezhov AA, Goryachev DV. Automated support system designing for people with limited communication. 2018 Presented at: 2018 Systems of Signals Generating and Processing in the Field of on Board Communications; March 14-15, 2018; Moscow, Russia.

14. Patil K, Jawadwala Q, Shu FC. Design and construction of electronic aid for visually impaired people. IEEE Trans Human-Mach Syst 2018;48(2):172-182. [doi: 10.1109/THMS.2018.2799588]
15. Zhang Y, Liu B, Liu Z. WristMouse: Wearable mouse controller based on pressure sensors. 2018 Presented at: 2018 IEEE 15th International Conference on Wearable and Implantable Body Sensor Networks (BSN); March 4-7, 2018; Las Vegas, NV, USA.
16. Ai Wei College. IT Training URL: https://www.ittraining.com.tw/ittraining/ [accessed 2019-03-23] [WebCite Cache ID 775uf06Nq]
17. Wikipedia. Kinect URL: https://en.wikipedia.org/wiki/Kinect [accessed 2019-03-23] [WebCite Cache ID 775ul0Hdu]
18. Shotton J, Fitzgibbon A, Cook M, Sharp T, Finocchio M, Moore R, et al. Real-time human pose recognition in parts from single depth images. 2011 Presented at: CVPR 2011; June 20-25, 2011; Colorado Springs, CO, USA URL: https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/BodyPartRecognition.pdf
19. Amit Y, Geman D. Shape quantization and recognition with randomized trees. Neural Comput 1997;9(7):1545-1588. [doi: 10.1162/neco.1997.9.7.1545]
20. MIT App Inventor. URL: http://appinventor.mit.edu/explore/ [accessed 2019-03-23] [WebCite Cache ID 775vCvpgB]

Abbreviations

IP: Internet Protocol
LED: Light-Emitting Diode
MOST: Ministry of Science and Technology
MPU: Microprocessor Unit

Edited by G Eysenbach; submitted 21.09.18; peer-reviewed by J Zhou, S Prior; comments to author 02.02.19; revised version received 22.02.19; accepted 25.02.19; published 28.05.19.

Please cite as:
Chung HY, Chung YL, Liang CY
Design and Implementation of a Novel System for Correcting Posture Through the Use of a Wearable Necklace Sensor
JMIR Mhealth Uhealth 2019;7(5):e12293
URL: https://mhealth.jmir.org/2019/5/e12293/
doi:10.2196/12293
PMID:31140439

Original Paper

The Use of Wireless, Smartphone App–Assisted Home Blood Pressure Monitoring Among Hypertensive Patients in Singapore: Pilot Randomized Controlled Trial

Eui Whan Moon[1], MD; Ngiap Chuan Tan[2], MBBS, MMED; John Carson Allen[3], PhD; Tazeen Hasan Jafar[1], MD, MPH

[1]Programme in Health Services and Systems Research, Duke-NUS Medical School, Singapore, Singapore

[2]SingHealth Polyclinics, Singapore, Singapore

[3]Centre for Quantitative Medicine, Office of Clinical Sciences, Duke-NUS Medical School, Singapore, Singapore

Corresponding Author:
Tazeen Hasan Jafar, MD, MPH
Programme in Health Services and Systems Research
Duke-NUS Medical School
8 College Road
Singapore, 169857
Singapore
Phone: 65 6601 2582
Fax: 65 6534 8632
Email: tazeen.jafar@duke-nus.edu.sg

Abstract

Background: Reliable home blood pressure monitoring (HBPM) is essential to effective hypertension management; however, manual recording is subject to underreporting and inaccuracies. Mobile health technologies hold great potential as HBPM tools, but the fidelity of a smartphone app in HBPM has not been adequately assessed.

Objective: The primary aim of the trial was to compare the fidelity of a smartphone app to that of a handwritten logbook in making HBPM data available to clinicians at follow-up visits. Fidelity was defined as the percentage of scheduled blood pressure (BP) recordings over a 3-week period that were properly recorded and reported to the clinic. The secondary aims were to investigate patient factors associated with HBPM fidelity and to explore the effect of time on the fidelity.

Methods: A 2-arm, parallel, unblinded, randomized controlled pilot trial was conducted in a government polyclinic in Singapore. Hypertensive adults, aged 40 to 70 years, who were on antihypertensive medication and owned a smartphone were recruited and randomized by a computer-generated randomization schedule to 3 weeks of either semiautomated HBPM utilizing a Bluetooth-enabled BP monitor and a smartphone app or a fully manual process utilizing a conventional handwritten logbook. The primary outcome was home BP recording fidelity.

Results: Of the 80 patients randomized, 79 (smartphone app: 38 and logbook: 41) were included in the final analysis. Although fidelity was higher among the app users, it did not differ significantly between study arms (smartphone app: 66.7% and logbook: 52.4%; *P*=.21). Chinese and Indian ethnicities were associated with higher fidelity (absolute percent and 95% CI) by 35.6% (4.27 to 66.9) and 45.0% (8.69 to 81.3), respectively, in comparison with other ethnicities (*P*=.03); longer smartphone ownership increased fidelity on an average of 10.5% (0.83 to 20.2) per year (*P*=.03); the number of apps on the smartphone decreased fidelity at a rate of −0.32% (−0.58 to −0.05) per app (*P*=.02); years of hypertension morbidity increased fidelity at a rate of 1.56% (0.03 to 3.09) per year (*P*=.046); and the number of people working in the household decreased fidelity at a rate of −8.18% (−16.3 to −0.08) per additional working person (*P*=.048). The fidelity of the app was significantly higher in the first week (64.4%) than the second (55.1%, *P*=.001) and third (58.2%, *P*=.03) weeks of monitoring.

Conclusions: Amid the increasing integration of health technologies into clinical practice, our study demonstrates the feasibility of smartphone app–assisted HBPM in hypertensive adults of Singapore. Our pilot study found no statistically significant difference in mean BP recording fidelity between a smartphone app and conventional handwritten logbook. However, the small sample size precludes definitive conclusions and highlights the need for a larger, adequately powered trial.

Trial Registration: ClinicalTrials.gov NCT03209024; https://clinicaltrials.gov/ct2/show/NCT03209024 (Archived by WebCite at http://www.webcitation.org/78EVWBg0T)

(JMIR Mhealth Uhealth 2019;7(5):e13153) doi:10.2196/13153

KEYWORDS

mHealth; smartphone; hypertension; home blood pressure monitoring; self blood pressure monitoring; health informatics; data collection methods; personal health records

Introduction

Background

Hypertension is the leading attributable risk factor for cardiovascular disease and death, globally [1]. The 2 blood pressure (BP) monitoring methods commonly used in primary care settings are home BP monitoring (HBPM) and office-based monitoring. HBPM is superior to office-based measurements as a predictor of cardiovascular disease and all-cause mortality [2]. Furthermore, the prognostic value of HBPM improves with the number of home BP measurements that patients are able to provide to their health care providers [3], thus emphasizing the importance of having a reliable means of collecting and reporting home BP data at each office visit. A systematic review and meta-analysis of 37 randomized controlled trials (RCTs) comparing HBPM with office BP measurements for effectiveness in reducing BP showed HBPM to be more effective in achieving BP control [4]. Nonetheless, at least one-third of known hypertensive patients in Singapore do not have adequate BP control per conventionally recommended levels [5]. This is especially problematic among the elderly in whom the prevalence of hypertension (73.9%) is higher, as is the rate of uncontrolled BP (75.9%) [6]. The lack of reliability in the conventional HBPM method could be an important contributing factor in the failure to achieve effective BP control and cardiovascular risk reduction in these patients. The shortcomings of conventional HBPM using handwritten logbooks are well known and include inaccuracies, underreporting of data [7], and failure to bring logbooks to clinic visits [8]. The purpose of HBPM is undermined and the value of reported measurements is diminished without an effective means of making accurate home BP readings available to clinicians. It is these considerations that motivate and necessitate the exploration of more reliable methods of communicating home BP values to health care providers.

Mobile Health

Mobile health (mHealth) technology has been increasingly evaluated for chronic disease management. mHealth is broadly defined as any health care practice supported by mobile devices and their functionalities [9]. With the widespread use of smartphones in recent years, mobile apps have gained attention as an mHealth modality [10]. A content analysis of the top 107 apps for hypertension management showed that 72% include a tracking function for BP values [11]. This simple feature, when coupled with the wireless data transfer capabilities of Bluetooth-enabled BP monitors, would allow a mobile app to function as an electronic logbook that is more convenient to use and more readily accessible to clinicians than a handwritten log. Singapore, with a 91% smartphone penetration rate, holds favorable conditions to utilize a wireless platform in a clinical setting [12]. However, there are relatively few studies in the literature that present a quantitative comparison between a smartphone app and manual logbook in terms of their respective reliability as a recording tool for HBPM by patients. In addition, because operating technological devices is highly user-dependent, specific patient factors associated with the effectiveness of smartphone app–assisted HBPM need to be explored.

Aims of This Study

Our pilot study aimed to begin addressing the abovementioned knowledge gaps by assessing whether there is any benefit in using mHealth technology (smartphone app) to store home BP values compared with using handwritten logbooks in terms of making these records available at clinic visits. The primary aim of our RCT was to compare the home BP recording fidelity over a 3-week period using a smartphone app versus a handwritten logbook in the Singaporean hypertensive patient population. Fidelity was defined as the percentage of scheduled home BP readings that are compliant with the HBPM regimen and are successfully reported at the follow-up visit. The null hypothesis postulated no difference in BP recording fidelity between a smartphone app and handwritten logbook. As there was no a priori basis for postulating greater fidelity with the app, the null hypothesis was tested against a 2-sided alternative, although the desired outcome was higher recording fidelity with the smartphone app. Secondary aims were (1) to explore associations between participant characteristics and the recording fidelity within each study arm and (2) to explore the effect of time on the weekly recording fidelity in each study arm. With detailed participant demographic data as well as HBPM records obtained via a logbook and the app, we compared the fidelity of the 2 home BP recording modalities, identified patient characteristics associated with higher fidelity for each monitoring method, described the attenuation pattern of fidelity in each study arm, and compared the fidelity of the 2 recording modalities on a weekly interval.

Methods

Trial Design

This study was an open-label, parallel group RCT of 2 study arms with a 1:1 allocation ratio.

Ethics Approval and Trial Registration

The study protocol was approved by the SingHealth Centralised Institutional Review Board (reference #2017/2014) and registered under ClinicalTrials.gov (NCT03209024).

Participants

Setting

Participants were recruited from Pasir Ris Polyclinic, a public primary care clinic serving the multi-ethnic population of a district in Singapore composed of approximately 143,000 residents.

Inclusion and Exclusion Criteria

Singaporean citizens or permanent residents aged between 40 and 70 years, visiting Pasir Ris Polyclinic for at least 1 year, diagnosed with essential hypertension and taking at least 1 antihypertensive medication, owning a compatible smartphone, and able to communicate in English were eligible.

Patients with cardiac arrhythmia, end-stage renal disease, cancer, history of stroke or myocardial infarction, or any other physical or mental disability that would prevent self-monitoring of BP at home were excluded. Other exclusions were arm circumference exceeding the cuff size, extensive travel overseas during the study period, working night shifts, or participation in another clinical trial.

Screening and Recruitment

Patients were enrolled in the study via convenience sampling. Patients in the polyclinic waiting area were approached, a brief explanation of the study was provided, and prescreening questions were administered. Interested patients were subsequently screened for eligibility based on the study inclusion/exclusion criteria, and written informed consent was obtained from eligible patients. The informed consent procedure included an explanation of the purpose of the study, the study procedures and visit schedule, participants' responsibilities and rights, and confidentiality of medical records.

Randomization and Allocation Concealment

A computer-generated treatment allocation sequence accommodating 80 patients was generated by the study statistician (JCA) using permuted block randomization with blocks of size 6 and one block of size 8. Sequentially numbered, opaque, sealed envelopes (SNOSEs) were prepared, with each envelope containing a treatment group assignment. The primary investigator enrolled all patients, whereas allocation concealment and sequential dispensing of envelopes were enforced by an on-site research coordinator. Patients were randomly assigned to either the smartphone app or the handwritten logbook study groups.

Home Blood Pressure Monitoring

Home Blood Pressure Monitoring Methods

Smartphone app–assisted HBPM was performed using the Bluetooth-enabled Omron HEM7280T BP monitor (Textbox 1) to wirelessly record BP values onto the Omron Connect app (Textbox 2), which was available free of charge on both Google Play store and Apple App Store and did not undergo major updates during the evaluation process. In brief, this was a semiautomated process that required patients to refresh the app's home screen upon completion of a BP measurement to initiate the transfer of the reading from the HEM7280T BP monitor into the app's electronic log. Logbook HBPM was performed by reading the BP values displayed on the BP monitor and recording them into a physical logbook. Sample screenshots of the Omron Connect app on iOS and Android platforms, as well as a sample image of the manual HBPM logbook, can be found in Multimedia Appendices 1,2, and 3, respectively. At the baseline visit, all the participants were instructed on how to properly record BP values using the HBPM method to which they were assigned, and they were given the opportunity to practice this process under supervision. All participants were provided with a phone number for technical support should they require any help troubleshooting errors with the app or BP monitor.

Textbox 1. Study device (Omron HEM7280T) specifications.

Device Equivalent

- M6 AC (HEM-7322-E)—validated by the European Society of Hypertension protocol

Mode

- Fully automatic, oscillometric blood pressure (BP) monitor

Wireless feature

- Bluetooth-enabled device
- BP data (ie, systolic pressure, diastolic pressure, pulse rate, and time of measurement) transfer onto a smartphone app

BP measurement range

- 0-299 mmHg

Accuracy

- +/−3 mmHg

Internal memory capacity

- 2 user settings; 100 readings per user

Upper arm cuff circumference

- 22-42 cm

Textbox 2. Study device (Omron Connect Smartphone App) specifications.

Wireless feature

- Synchronization and storage of the user's blood pressure (BP) data from Omron HEM7280T via Bluetooth

Data export

- Able to export the log of BP data (date, time, time zone, systolic BP, diastolic BP, pulse, irregular heartbeat, cuff wrap guide, and BP device model) as a comma-separated value file using email or other apps (eg, WhatsApp and iMessage)

Compatible devices

- Compatible with both iOS and Android operating systems

Home Blood Pressure Monitoring Regimen and Blood Pressure Measurement Technique

All participants received the same instructions on the home BP recording regimen and the correct BP measurement technique. The home BP recording regimen was based on guidelines and recommendations of the European Society of Hypertension (ESH) [13,14] and consisted of consecutive duplicate readings in the morning (06:00 to 09:00 hours) and evening (18:00 to 21:00 hours). The patients were asked to follow the recording regimen over a 3-week (21 days) study period for a total of 84 measurements. The instructions for the BP measurement technique were adapted from recommendations by the authoritative sources [14-17].

Data Collection

The Baseline Data Questionnaire was administered to obtain information on the participants' sociodemographics, economic status, clinical characteristics, and exposure to technology. Baseline BP was recorded as the average of the last 2 of 3 consecutive BP measurements taken at 1-min intervals after 5 min of initial rest. The Hypertension Self-Care Profile questionnaire [18] is a validated tool to assess self-care behaviors in the domains of Behavior, Motivation, and Self-Efficacy. The version adapted to the Singaporean population was administered with permission from the authors [19].

At the end of the 3-week HBPM period, participants returned to the clinic for a single (final) follow-up visit. During the visit, the electronic log on the app used by the smartphone app group was exported to the study database for further analysis; similarly, a copy of the logbooks from the patients in the logbook group was uploaded into the study database. For security precautions, all participant data were anonymized and stored in password-protected computers or in locked cabinets.

Upon completion of the study, each participant received a Singapore $30 grocery store voucher.

Statistical Analysis

Outcome Measures

The primary outcome measure was home BP recording fidelity, defined a priori as the percentage of scheduled home BP readings over the 3-week HBPM regimen which was recorded, regimen compliant, and successfully reported at the final clinic visit.

A secondary outcome of this study was *time-independent fidelity* which loosens the definition of fidelity by modifying the strict recording time criteria for determining which readings will be considered *HBPM regimen compliant* to include in the calculation of fidelity: the timeframe for a measurement to be considered regimen compliant was expanded to 01:01 to 13:00 hours for morning readings and 13:01 to 01:00 hours for evening readings. This study's home BP recording regimen and recording time criteria (specified above) is useful to standardize outcome measures in the research setting, but these timeframes are largely arbitrary and are not followed strictly in the clinical setting. The secondary outcome of *time-independent fidelity* would allow this study to explore whether the strict recording time criteria of the HBPM regimen has any influence on the fidelity for each study arm. Weekly fidelity calculated as the percentage of scheduled weekly readings (28), which was HBPM regimen compliant and reported at the final clinic visit, was also assessed as a secondary outcome.

Description of Analytic Models

All data analysis was performed using SAS University Edition. Statistical significance was at $P<.05$.

Our study was an RCT. In consideration of the modest departure of home BP recording fidelity from a Gaussian distribution, the primary comparison on fidelity between the study groups was tested using the Wilcoxon rank-sum test appropriate for 2 independent samples. The median difference was estimated by the Hodges-Lehmann shift parameter estimate. In the sensitivity analyses, mean fidelity was also compared using (1) the 2-sample *t* test and (2) a general linear mixed-model, repeated-measures analysis to adjust the treatment group comparison for possible confounders of age, gender, and baseline systolic and diastolic BP, as well as to assess the effects of follow-up time (week) and time × treatment group interaction. In the latter statistical analysis, participants were included as random effects and time as a repeated-measures fixed effect within participants. The same sensitivity analyses were also performed on the outcomes defined under the *time-independent fidelity* between the 2 study arms. The data from the sole patient who withdrew from the study were omitted from the analyses.

The same general linear mixed-model was also used to assess the effect of time on weekly home BP recording fidelity in each study arm. Comparisons among weekly fidelities in the context of the mixed-model are analogous to paired *t* tests based on within-participant differences. A subgroup analysis on elderly

patients (60≤age<70) was also performed using the same statistical method.

To investigate associations between participant baseline characteristics and home BP recording fidelity and to assess the predictive potential, a univariate and multivariate analysis of covariance was performed within each study arm. Baseline variables exhibiting significance at $P<.20$ in the univariate analysis were included in a stepwise multiple linear regression analysis to identify possible predictors of fidelity. Variables specific to exposure to technology were included in the analysis of the smartphone app arm but were excluded from the analysis of the logbook arm. From the variables selected in the stepwise analysis, only those significant at $P \leq .10$ were retained in the final model as potential predictors of home BP recording fidelity.

Sample Size Calculation

On the basis of a 2-sided 2-sample *t* test at alpha=.05, a sample size of n=35 patients per study arm was calculated to provide 80% power to detect an effect size (ES, Cohen *d*) of 0.6, where 0.5 is generally considered a *medium* ES. Overall, 80 participants were recruited anticipating a 10% withdrawal rate.

Results

Participant Flow and Baseline Characteristics

A total of 928 patients were approached during the recruitment and follow-up period (03/15/2017–07/15/2017). As shown in the Consolidated Standards of Reporting Trials flow diagram (Figure 1), of the 102 patients undergoing screening, 83 were eligible for enrollment. Of those eligible, 80 were randomized. One patient randomized to the smartphone app arm was reallocated to the logbook arm at the baseline visit before commencing HBPM owing to unexplained smartphone incompatibility with the study app during the initial set-up process. One participant in the smartphone app arm withdrew from the study. In the smartphone app arm, 38 participants were included in the final analysis and 11 had home BP recording fidelity >80% at the end of the 3-week follow-up period. For the logbook arm, 41 were included in the analysis and 7 had home BP recording fidelity >80% at the end of 3 weeks.

The smartphone app and logbook arms were comparable in all baseline characteristics with the exception of systolic BP (SBP; Table 1). There was no evidence of SBP as a confounder.

Home Blood Pressure Monitoring Fidelity

In the primary analysis on 79 participants (smartphone app: 38 and logbook: 41), higher median fidelity was achieved with the use of the smartphone app (66.7%) compared with the logbook (52.4%), although the difference was not statistically significant (Wilcoxon rank-sum, $P=.21$; Table 2). Similar results were obtained from comparisons of mean fidelity using the *t* test ($P=.21$) and the general linear mixed-model ($P=.14$).

In comparison with the primary outcome, the assessment of *time-independent fidelity* exhibited higher fidelity in both study arms, although the difference between arms was smaller and not statistically significant ($P=.30$; Table 2).

Participant Characteristics Associated With Fidelity

For smartphone app–assisted HBPM, a multivariate analysis identified 5 independent baseline variables that exhibited statistically significant associations with fidelity: (1) The number of people working in the household was associated with decreased fidelity (absolute percent and 95% CI) by –8.18% (−16.3 to −0.08) per additional working person ($P=.048$); (2) Years of hypertension increased fidelity by 1.56% (0.03 to 3.09) per year ($P=.046$); (3) Years of current smartphone use increased fidelity on an average of 10.5% (0.83 to 20.2) per year of use ($P=.03$); (4) The number of apps on the current smartphone decreased fidelity by –0.32% (−0.58 to −0.05) per app ($P=.02$); (5) Mean fidelity had a significant association ($P=.03$) with ethnicity, with Chinese (n=24) and Indians (n=4) exhibiting higher fidelity by 35.6% (4.27 to 66.9) and 45.0% (8.69 to 81.3), respectively, in comparison with other ethnicities (n=3) composed of Pakistani, Ceylonese, and Eurasian ethnicities.

In the logbook arm, a multivariate analysis identified 3 independent baseline variables that exhibited significant associations with fidelity: (1) Working participants had significantly lower fidelity by 18.6% (1.51 to 35.8; $P=.03$) compared with nonworking participants; (2) Age was a positive factor with an average increase in fidelity by 1.21% (0.05 to 2.36) per year increase in age ($P=.04$); (3) For every additional child, fidelity declined by an average of –9.20% (–16.4 to –2.04; $P=.01$).

Self-care capacity, both in individual domains and the cumulative score of Hypertension Self-Care Profile, exhibited no evidence of association with fidelity.

Figure 1. Consolidated Standards of Reporting Trials flow diagram.

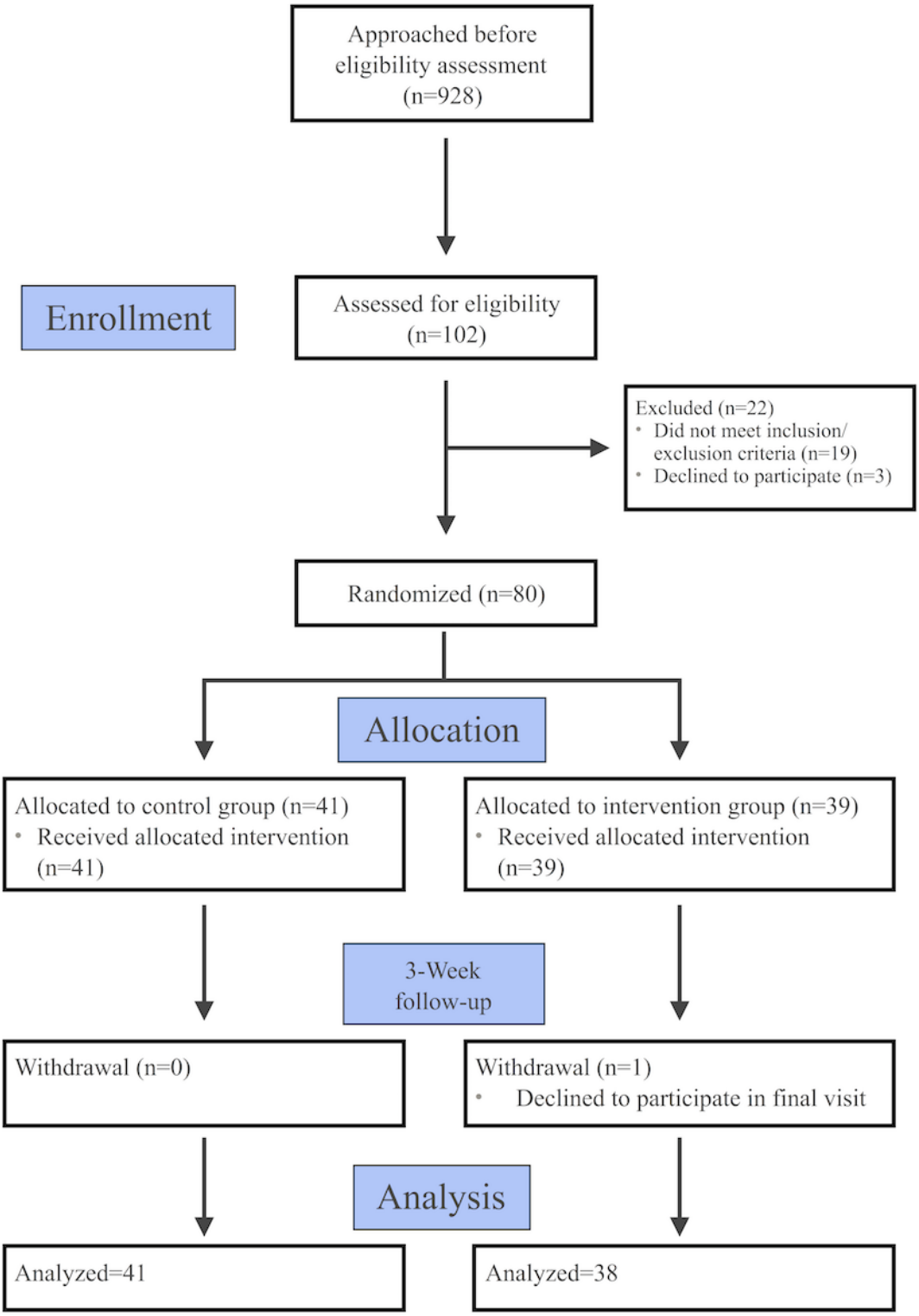


Table 1. Baseline characteristics of study participants.

Variable	Total (N=79)	Logbook (n=41)	Smartphone app (n=38)	*P* value
Sociodemographic characteristics				
Age, mean (SD)	56.3 (7.2)	56.1 (6.8)	56.5 (7.7)	.83
Gender, n (%)				
Male	37 (46.8)	19 (46.3)	18 (47.4)	>.99
Ethnicity, n (%)				
Chinese	52 (65.8)	28 (68.3)	5 (12.2)	.92
Malay	13 (16.5)	6 (14.6)	7 (18.4)	.92
Indian	9 (11.4)	5 (12.2)	4 (10.5)	.92
Other	5 (6.3)	2 (4.9)	3 (7.9)	.92
Highest level of education completed, n (%)				
Secondary or lower	34 (43.0)	14 (34.2)	20 (52.6)	.27
Postsecondary or Polytechnic diploma	28 (35.4)	17 (41.5)	11 (29.0)	.27
Degree or above	17 (21.5)	10 (24.4)	7 (18.4)	.27
Marital status, n (%)				
Married	68 (86.1)	36 (87.8)	32 (84.2)	.75
Not married	11 (13.9)	5 (12.2)	6 (15.8)	.75
Number of children, mean (SD)	2.1 (1.1)	2.2 (1.1)	2.0 (1.1)	.49
Economic status				
Currently working, n (%)	58 (73.4)	30 (73.2)	28 (73.7)	>.99
Ownership of house, n (%)				
Yes	75 (94.9)	40 (97.6)	35 (80.0)	.61
On rent	1 (1.3)	0 (0)	1 (2.6)	.61
Office accommodation	1 (1.3)	0 (0)	1 (2.6)	.61
Owned by relative or parents	2 (2.5)	1 (2.4)	1 (2.6)	.61
Type of housing, n (%)				
Public housing (<5 rooms)	20 (25.3)	10 (24.4)	10 (26.3)	.77
Public housing (5+ rooms or HUDC[a] or executive flat or studio)	48 (60.8)	24 (58.5)	24 (63.2)	.77
Private property	11 (13.9)	7 (17.1)	4 (10.5)	.77
Number of people in the house, mean (SD)	3.9 (1.6)	3.9 (1.6)	4.1 (1.4)	.56
Number of people working in the household, mean (SD)	2.2 (1.1)	2.2 (1.1)	2.1 (1.0)	.79
Gross average monthly income of household, n (%)				
Below Singapore $8000	45 (57.0)	24 (58.5)	21 (55.3)	.82
Singapore $8000 and above	34 (43.0)	17 (41.5)	17 (44.7)	.82
Exposure to technology				
Number of personal electronic devices, mean (SD)	3.5 (1.8)	3.5 (1.8)	3.1 (1.6)	.27
Regular use of computer, n (%)	59 (74.7)	29 (70.7)	30 (79.0)	.45
Phone OS[b], Apple, n (%)	44 (55.7)	23 (56.1)	21 (55.3)	>.99
Years of smartphone use, mean (SD)	8.7 (3.3)	8.7 (3.3)	7.8 (3.2)	.23
Years of current smartphone use, mean (SD)	1.8 (1.3)	1.8 (1.3)	1.8 (0.9)	.95
Number of apps on current smartphone, mean (SD)	84.8 (54.9)	84.8 (54.9)	68.7 (34.1)	.12
Hours of smartphone use per day, mean (SD)	4.7 (3.4)	4.7 (3.7)	4.4 (3.7)	.77

Variable	Total (N=79)	Logbook (n=41)	Smartphone app (n=38)	*P* value
Clinical characteristics				
Use of personal BP[c] monitor at home, n (%)	72 (91.1)	37 (90.2)	35 (92.1)	>.99
Years of personal BP monitor use, mean (SD)	4.8 (4.0)	4.8 (4.0)	4.6 (3.4)	.82
Baseline systolic BP, mean (SD)	131.4 (15.5)	127.2 (17.1)	135.9 (12.1)	.01
Baseline diastolic BP, mean (SD)	82.8 (9.9)	81.8 (11.1)	84.0 (8.5)	.33
Years of hypertension, mean (SD)	8.0 (6.1)	8.3 (6.9)	7.7 (5.1)	.67
Number of antihypertensive drugs, mean (SD)	1.6 (1.2)	1.6 (1.2)	1.6 (0.6)	.88
Diabetes, n (%)	39 (49.4)	18 (43.9)	21 (55.3)	.37
Hypertension Self-Care Profile—Behavior, mean (SD)	46.2 (8.5)	45.5 (9.14)	47.0 (7.79)	.43
Hypertension Self-Care Profile—Motivation, mean (SD)	34.5 (9.2)	35.3 (9.24)	33.5 (9.11)	.38
Hypertension Self-Care Profile—Self-Efficacy, mean (SD)	38.0 (9.4)	37.9 (10.3)	38.1 (8.49)	.91
Hypertension Self Care Profile—Cumulative, mean (SD)	119.0 (22.9)	119.3 (25.6)	118.7 (20.1)	.90

[a]HUDC: Housing and Urban Development Company.

[b]OS: operating system.

[c]BP: blood pressure.

Table 2. Comparison of home blood pressure recording fidelity in all participants

Outcome measures and parameter	Smartphone app (n=38)	Logbook (n=41)	Difference (95% CI)	*P* value
Fidelity[a]				
Median (IQR[b])	66.7 (32.1 to 83.3)	52.4 (29.8 to 72.6)	8.33[c] (−4.76 to 22.6)	.21[d]
Mean (SD)	58.6 (29.2)	50.7 (27.2)	7.95 (−4.67 to 20.6)	.21[e]
LS mean[f] (RMSE[g])	59.2 (28.3)	49.8 (28.3)	9.43 (−3.15 to 22.0)	.14[h]
Time-independent fidelity				
Median (IQR)	88.7 (57.1 to 97.6)	84.5 (56.0 to 96.4)	2.38[c] (−3.57 to 11.9)	.30[d]
Mean (SD)	78.4 (22.4)	73.2 (27.4)	5.21 (−6.07 to 16.5)	.36[e]
LS mean (RMSE)	78.8 (26.2)	72.8 (26.2)	6.01 (−5.36 to 17.4)	.30[h]

[a]Fidelity is defined as the percentage of specified home blood pressure readings over the 3-week home blood pressure monitoring regimen which was recorded, regimen compliant, and successfully reported at the final clinic visit.

[b]IQR: interquartile range.

[c]Hodges-Lehmann shift parameter estimate (median difference).

[d]Wilcoxon rank-sum test.

[e]2-sample *t* test.

[f]LS mean: least-squares mean.

[g]RMSE: root mean square error.

[h]Mixed-model longitudinal analysis of variance (ANOVA) adjusting for age, study arm, age by study arm interaction, and baseline systolic and diastolic blood pressures.

Fidelity by Study Week

In assessing the weekly fidelity trend over the 3-week period within each study arm, fidelity was highest in the first week and lower in subsequent weeks (Figure 2). This difference was most pronounced among the app users, whose fidelity in the first week (64.4%) was significantly higher than that of the second (55.1%; *P*=.001) and third (58.2%; *P*=.03) weeks of monitoring. Though a similar trend was seen in the logbook arm, no change in fidelity among weeks was significant.

The participants using the smartphone app had higher weekly fidelity than their counterparts using the logbook in the first week, with the least-squares mean difference approaching significance (*P*=.06). Although mean fidelity for the smartphone app arm continued to outperform the logbook arm in the second

and third weeks, these differences were less pronounced (Table 3).

Figure 2. Weekly fidelity trend over 3 weeks in the smartphone app and logbook arms. HBPM: home blood pressure monitoring.

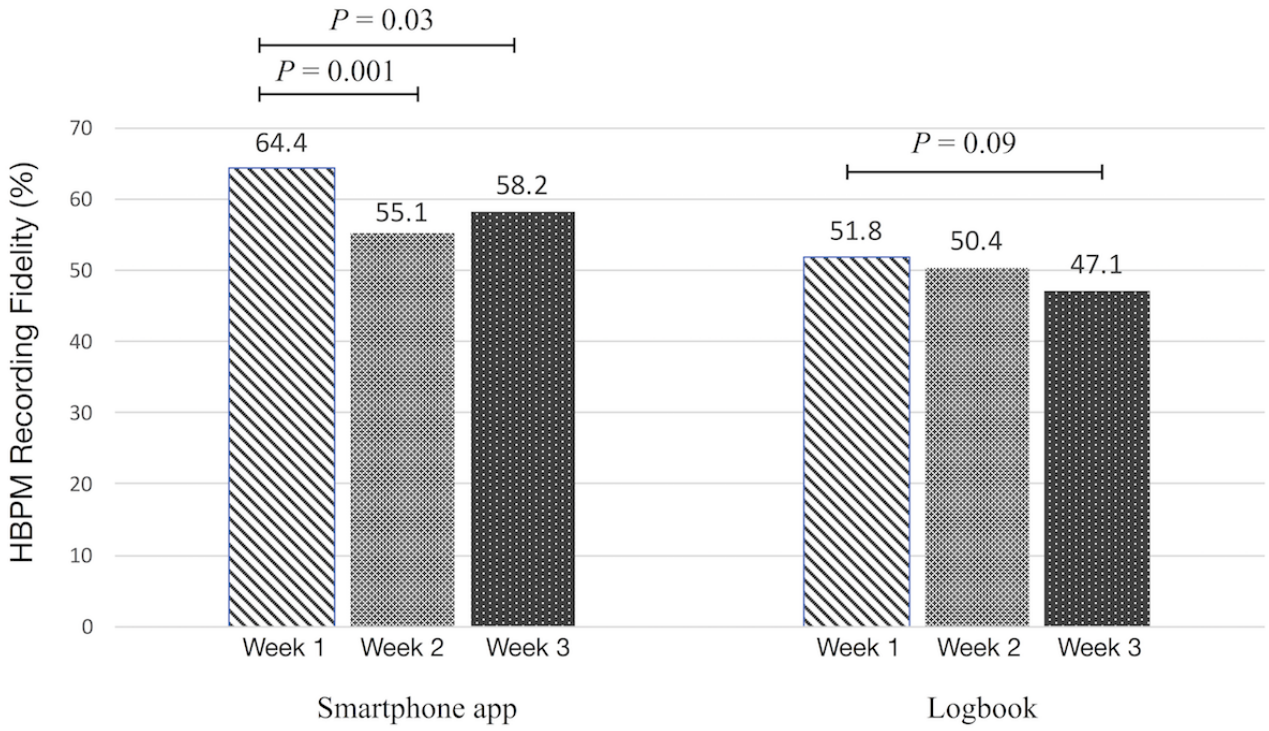


Table 3. Comparison of home blood pressure recording fidelity by study weeks.

Parameter and week	Fidelity[a]			
	Smartphone app	Logbook	Difference (95% CI)	*P* value
Week 1; smartphone app (n=38) and logbook (n=41)				
LS mean[b] (RMSE[c])	64.4 (28.3)	51.8 (28.3)	12.6 (−0.63 to 25.8)	.06
Median (IQR[d])	64.3 (42.9 to 92.9)	53.6 (28.6 to 75.0)	14.3 (−3.57 to 25.0)	.12
Week 2; smartphone app (n=38) and logbook (n=41)				
LS mean (RMSE)	55.1 (28.3)	50.4 (28.3)	4.68 (−8.54 to 17.9)	.49
Median (IQR)	53.6 (21.4 to 85.7)	50.0 (32.1 to71.4)	3.57 (−10.7 to 17.9)	.62
Week 3; smartphone app (n=38) and logbook (n=41)				
LS mean (RMSE)	58.2 (28.3)	47.1 (28.3)	11.0 (−2.22 to 24.2)	.10
Median (IQR)	64.3 (35.7 to 78.6)	50.0 (28.6 to 71.4)	10.7 (−3.57 to 25.0)	.16

[a]Fidelity is defined as the percentage of specified weekly home blood pressure readings which was recorded, regimen compliant, and successfully reported at the final clinic visit.

[b]LS mean: least-squares mean.

[c]RMSE: root mean square error.

[d]IQR: interquartile range.

Subgroup Analysis in the Elderly

In a posthoc assessment of the elderly participants (60≤age<70), there was no significant difference in overall mean fidelity between app and logbook users (*P*=.10; Table 4). However, in the assessment of weekly recording fidelities, app users exhibited significantly higher fidelity (79.0%) than logbook users (58.8%) during the first week of monitoring (*P*=.048; Table 5).

Table 4. Comparison of home blood pressure recording fidelity by age.

Parameter and age	Fidelity[a]			
	Smartphone app	Logbook	Difference (95% CI)	*P* value
40<age<50; smartphone app (n=8); logbook (n=6)				
LS mean[b] (RMSE[c])	49.5 (35.8)	51.7 (35.8)	−2.26 (−46.7 to 42.2)	.91
Median (IQR[d])	58.9 (31.5 to 79.8)	51.8 (20.2 to 86.9)	0.595 (−46.4 to 50.0)	>.99
50≤age<60; smartphone app (n=17); logbook (n=21)				
LS mean (RMSE)	48.3 (30.2)	44.8 (30.2)	3.49 (−16.5 to 23.5)	.72
Median (IQR)	40.5 (25.0 to 72.6)	40.5 (20.2 to 66.7)	2.38 (−15.5 to 23.8)	.68
60≤age<70; smartphone app (n=13) logbook (n=14)				
LS mean (RMSE)	74.9 (22.2)	59.0 (22.2)	15.9 (−3.13 to 35.0)	.10
Median (IQR)	79.8 (71.4 to 92.9)	58.3 (45.2 to 72.6)	18.4 (−1.19 to 36.9)	.07

[a]Fidelity is defined as the percentage of specified home blood pressure readings over the 3-week home blood pressure monitoring regimen which was recorded, regimen compliant, and successfully reported at the final clinic visit.

[b]LS mean: least-squares mean.

[c]RMSE: root mean square error.

[d]IQR: interquartile range.

Table 5. Comparison of home blood pressure recording fidelity in the elderly subgroup (60≤age<70).

Parameter and week	Fidelity[a]			
	Smartphone app	Logbook	Difference (95% CI)	*P* value
Week 1; smartphone app (n=13) and logbook (n=14)				
LS mean[b] (RMSE[c])	79.0 (22.2)	58.8 (22.2)	20.1 (0.22 to 40.0)	.048
Median (IQR[d])	92.9 (64.3 to 96.4)	62.5 (50.0 to 78.6)	21.4 (3.57 to 39.3)	.04
Week 2; smartphone app (n=13) and logbook (n=14)				
LS mean (RMSE)	73.2 (22.2)	62.4 (22.2)	10.8 (–9.13 to 30.7)	.28
Median (IQR)	82.1 (50.0 to 92.9)	64.3 (50.0 to 78.6)	14.3 (–7.14 to 32.1)	.18
Week 3; smartphone app (n=13) and logbook (n=14)				
LS mean (RMSE)	72.7 (22.2)	55.8 (22.2)	16.9 (–3.04 to 36.8)	.10
Median (IQR)	71.4 (60.7 to 85.7)	53.6 (42.9 to 71.4)	17.9 (0.00 to 35.7)	.08

[a]Fidelity is defined as the percentage of specified home blood pressure readings over the 3-week home blood pressure monitoring regimen which was recorded, regimen compliant, and successfully reported at the final clinic visit.

[b]LS mean: least-squares mean.

[c]RMSE: root mean square error.

[d]IQR: interquartile range.

Discussion

Principal Findings

Our study compared the recording fidelity of an app-mediated electronic record versus a handwritten logbook in performing HBPM. Although higher fidelity was observed in the smartphone app arm compared with the logbook arm, indicating the potential for improved fidelity with the use of the app, statistical significance was not achieved. Meanwhile, the findings of our secondary aims have provided valuable information on the fidelity of the app-assisted method by identifying a unique set of potential predictors and describing an attenuation pattern over time. Finally, the post hoc subgroup analysis by age group showed significantly higher recording fidelity with the app versus the handwritten logbook in the elderly during week 1 of monitoring, suggesting that app-assisted HBPM is feasible across a wide range of ages. These findings have promising implications for the expanding use of mHealth technology in hypertension management and warrant further investigation in future studies.

There remains a scarcity of the literature addressing a direct comparison of home BP recording fidelity between app-mediated and handwritten methods. Although some studies do report outcome measures similar to fidelity as we define it

(eg, level of adherence to a given HBPM regimen), heterogeneity in regimen frequency and duration make comparisons of study results difficult [3]. In addition to reliable recording fidelity, adequate patient support may also be important to achieving effective clinical outcomes. The findings from a recent systematic review and meta-analysis showed that HBPM alone did not have a higher association with BP lowering and BP control compared with no-home monitoring (usual care), whereas HBPM with additional patient support via feedback, education, and counselling demonstrated a significantly higher association [20]. This emphasizes the importance of future studies to further explore smartphone app use, not only as a high-fidelity recording tool for HBPM but also as a platform for delivering effective patient adherence support such as health information and reminders [21].

From the secondary aim of our study, we found that the participant factors associated with higher fidelity were not the same for the 2 HBPM modalities, suggesting that determinants of fidelity in manual recording of home BP are different from those in smartphone use. To our knowledge, this is the first study to identify potential predictors of fidelity for app-assisted HBPM, and this novel information could be utilized in future studies to develop a clinically useful predictive model that provides decision support for identifying patients most suitable for mHealth-based monitoring. Although determining the mechanism of how certain patient factors lead to higher HBPM fidelity is beyond the scope of this study, we do offer plausible explanations for these associations. Among app users, longer ownership of a smartphone may mean greater proficiency with one's own device, thus improving the performance of app-mediated HBPM. The number of apps on the phone exhibited a negative association with fidelity, although we find the clinical relevance of this finding questionable as the impact of an additional app on fidelity was negligible. A longer duration of hypertension and fewer working household members were also associated with higher fidelity, perhaps because patients who have been hypertensive for many years may be more accustomed to following an HBPM regimen, whereas those who have more employed family members may have fewer people available at home to provide family support for hypertension management. For handwritten logbook-mediated HBPM, age was associated with fidelity whereas employment and number of children exhibited negative associations. This may reflect the difficulty in manually recording home BP values on a rigorous regimen while simultaneously being occupied with commitments at work or home. With advancing age and eventual retirement, adherence may improve as demands at home and work decrease.

The significant attenuation of fidelity from weeks 1 to 2 in the smartphone app arm may be explained in part by the duration of BP monitoring. The ESH guidelines specify a 7-day period of monitoring for its HBPM regimen [13,14], and prolonging this rigorous regimen beyond the recommended period of 1 week may have led to study fatigue among participants in the smartphone app arm.

The post hoc subgroup analysis by age suggested that the elderly perform HBPM better with the app than the manual logbook, especially during week 1 of monitoring. This finding cannot be regarded as conclusive but suggests that the elderly do not lack in the smartphone proficiency needed for recording home BP, which challenges the common preconception that mHealth-based HBPM may be too complicated for older patients.

The use of a 2-sample test in the primary analysis with no adjustment for potential confounders was appropriate. Of the 30 covariates measured at baseline, only one (SBP) resulted in a significant difference between study groups, which is consistent with the expected false positive error rate for 30 individual tests performed at the 5% significance level. Moreover, there was no evidence that SBP was associated with fidelity.

Limitations and Strengths

This study has a number of limitations. First, as a pilot trial, the sample size was limited by time and available resources. As we had no knowledge of either the expected difference in mean fidelity or the population SD, the sample size calculation was based on a targeted Cohen ES of 0.6, which was the smallest feasible and realistically achievable ES, given the available resources and timeframe of the study. The observed ES for fidelity was 0.33, hence the study was obviously underpowered. Second, recruitment was limited to one polyclinic and the participants were exclusively current smartphone users, which could preclude the generalizability of the results to the greater hypertensive population in Singapore or around the world. Third, this was an unblinded study because participants were required to learn the procedures for their respective HBPM methods; investigators were not masked, but we attempted to reduce bias by defining our primary outcome measure and the primary statistical method a priori. Finally, the Hawthorne effect could not be precluded, and the magnitude of the effect on the 2 study arms may have differed. A future full-scale trial of longer duration may mitigate the Hawthorne effect as patients become accustomed to home BP recording over the course of several months or years.

With regard to the strengths of our study, to our knowledge, this is the first RCT to directly compare the recording fidelity of mHealth and the conventional logbook methods of HBPM within the multi-ethnic Southeast Asian population. In addition, although the HBPM regimen used in research settings can vary widely, the regimen used in our study was adopted from an established clinical guideline and recommendations by ESH, which contributes to the reproducibility of this study and standardizes its outcome to allow comparison with other studies that also use this well-established HBPM regimen.

Future Directions

For the future full-scale trial to be sufficiently powered, the sample SD of approximately 28% and mean difference in fidelity of about 8% observed in this study can be used to ensure that a 2-arm study with n=175 participants per group would provide 80% power to detect an ES of 8/28 (approximately 0.30) at alpha=.05.

Although beyond the scope of this study, there remains a need to categorically quantify and compare the influence of over-reporting, underreporting, and falsifying data on the fidelity of logbook- and app-mediated HBPM to guide future app design

to promote accuracy in data transfer from home to clinics. Moreover, other important clinical outcomes in hypertension management, such as the magnitude of BP lowering and proportion of BP control, must be assessed.

As with any novel health care implementation, the sustainability of app-mediated HBPM is another area that deserves the focus of future research. As intuition would suggest, health apps are not immune to attenuation of usage over time, which has direct implications on the home BP recording fidelity as was observed in our results as well. One study on the effect of the usage pattern of a health care app on user retention suggests the frequency of utilizing the self-monitoring function of the app—as opposed to other functions such as outpatient support service or medication functions—significantly increases the probability of the app to stay in use [22]. The authors of that study suggest that the benefit of achieving better health outcomes by modifying health behavior based on tracking one's own health parameters is the mechanism behind how frequent utilization of the self-monitoring function of a health app promotes its sustained use. The self-monitoring function is at the core of app-mediated HBPM, therefore it is vital for future studies not only to describe attenuation patterns of app use over time but also to explore modifiable factors (including BP recording regimens and the ease of app user interface) that promote frequent self-monitoring, thereby improving the app's retention rate.

Conclusions

As health technologies continue to find an increasingly profound integration into clinical practice, our study makes a contribution to the sparsity of available knowledge on the fidelity of smartphone app–assisted HBPM. Although our pilot study did not find a significant difference in BP recording fidelity between app-assisted and logbook-mediated HBPM, our results have identified unique determinants of fidelity in the 2 recording methods of home BP, which has not been reported before, and have characterized the attenuation of fidelity in time. Our findings also suggest that app-assisted HBPM is viable across a wide age spectrum that includes the elderly.

Acknowledgments

This study was funded by the author's (THJ) block grant from the Singapore Ministry of Education. Study devices (HEM-7280T) were sponsored by OMRON Healthcare Singapore. The funding bodies had no role in the design of the study; collection, analysis, and interpretation of data; and in writing the manuscript. The full trial protocol is available from the corresponding author upon request.

The authors would like to thank OMRON Healthcare Singapore for sponsoring the BP monitors for this study, Reena Subramanian Chandhini for on-site support with dispensing the SNOSEs, Liyana Ghazali for assembling the SNOSEs, Patricia Kin and Caris Tan for administrative oversight, Marissa Iskandar for logistical support, the staff at Pasir Ris Polyclinic for accommodating our study, and all the participants for their generous contribution in time and effort. The authors also wish to thank Dr Han Hae-Ra and her team for permission to use the Hypertension Self-Care Profile Tool, which was validated in SingHealth Polyclinics.

Authors' Contributions

EM designed the trial as part of his Doctor of Medicine (MD) thesis project under the mentorship of THJ, and co-mentorship of NCT. EM performed data analysis under the supervision of JCA, and all authors contributed to the interpretation of results. EM wrote the first draft of the paper, and all the authors reviewed and proofread the manuscript to its final version.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshot of Omron Connect app on iOS platform.

[PNG File, 4MB - mhealth_v7i5e13153_app1.png]

Multimedia Appendix 2

Screenshot of Omron Connect app on the Android platform.

[PNG File, 4MB - mhealth_v7i5e13153_app2.png]

Multimedia Appendix 3

Image of manual home blood pressure monitoring logbook.

[PNG File, 4MB - mhealth_v7i5e13153_app3.png]

Multimedia Appendix 4

CONSORT - EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 662KB - mhealth_v7i5e13153_app4.pdf]

References

1. Lawes CM, Vander HS, Rodgers A, International Society of Hypertension. Global burden of blood-pressure-related disease, 2001. Lancet 2008 May 03;371(9623):1513-1518. [doi: 10.1016/S0140-6736(08)60655-8] [Medline: 18456100]
2. Niiranen TJ, Hänninen M, Johansson J, Reunanen A, Jula AM. Home-measured blood pressure is a stronger predictor of cardiovascular risk than office blood pressure: the Finn-Home study. Hypertension 2010 Jun;55(6):1346-1351. [doi: 10.1161/HYPERTENSIONAHA.109.149336] [Medline: 20385970]
3. Ohkubo T, Asayama K, Kikuya M, Metoki H, Hoshi H, Hashimoto J, Ohasama Study. How many times should blood pressure be measured at home for better prediction of stroke risk? Ten-year follow-up results from the Ohasama study. J Hypertens 2004 Jun;22(6):1099-1104. [Medline: 15167443]
4. Agarwal R, Bills JE, Hecht TJ, Light RP. Role of home blood pressure monitoring in overcoming therapeutic inertia and improving hypertension control: a systematic review and meta-analysis. Hypertension 2011 Jan;57(1):29-38 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.110.160911] [Medline: 21115879]
5. Epidemiology & Disease Control Division, Ministry of Health, Republic of Singapore. National Health Survey. 2010. 2010 URL: https://www.moh.gov.sg/resources-statistics/reports/national-health-survey-2010 [accessed 2018-10-09] [WebCite Cache ID 732JHLvNP]
6. Malhotra R, Chan A, Malhotra C, Østbye T. Prevalence, awareness, treatment and control of hypertension in the elderly population of Singapore. Hypertens Res 2010 Dec;33(12):1223-1231. [doi: 10.1038/hr.2010.177] [Medline: 20882026]
7. Mengden T, Hernandez RM, Beltran B, Alvarez E, Kraft K, Vetter H. Reliability of reporting self-measured blood pressure values by hypertensive patients. Am J Hypertens 1998 Dec;11(12):1413-1417. [Medline: 9880121]
8. Tamaki S, Nakamura Y, Teramura M, Sakai H, Takayama T, Okabayashi T, et al. The factors contributing to whether or not hypertensive patients bring their home blood pressure record to the outpatient clinic. Intern Med 2008;47(18):1561-1565 [FREE Full text] [Medline: 18797113]
9. World Health Organization. 2011. mHealth: new horizons for health through mobile technologies URL: http://www.who.int/goe/publications/goe_mhealth_web.pdf [WebCite Cache ID 732H5R0IR]
10. Rehman H, Kamal AK, Morris PB, Sayani S, Merchant AT, Virani SS. Mobile health (mHealth) technology for the management of hypertension and hyperlipidemia: slow start but loads of potential. Curr Atheroscler Rep 2017 Mar;19(3):12. [doi: 10.1007/s11883-017-0649-y] [Medline: 28210974]
11. Kumar N, Khunger M, Gupta A, Garg N. A content analysis of smartphone-based applications for hypertension management. J Am Soc Hypertens 2015 Feb;9(2):130-136. [doi: 10.1016/j.jash.2014.12.001] [Medline: 25660364]
12. Consumer Barometer with Google. Trended Data: Singapore URL: https://www.consumerbarometer.com/en/trending/?countryCode=SG&category=TRN-NOFILTER-ALL [WebCite Cache ID 732GpeGXt]
13. O'Brien E, Asmar R, Beilin L, Imai Y, Mallion J, Mancia G, European Society of Hypertension Working Group on Blood Pressure Monitoring. European Society of Hypertension recommendations for conventional, ambulatory and home blood pressure measurement. J Hypertens 2003 May;21(5):821-848. [doi: 10.1097/01.hjh.0000059016.82022.ca] [Medline: 12714851]
14. Parati G, Stergiou GS, Asmar R, Bilo G, de Leeuw P, Imai Y, et al. European Society of Hypertension guidelines for blood pressure monitoring at home: a summary report of the Second International Consensus Conference on Home Blood Pressure Monitoring. J Hypertens 2008 Aug;26(8):1505-1526. [doi: 10.1097/HJH.0b013e328308da66] [Medline: 18622223]
15. Mancia G, Fagard R, Narkiewicz K, Redón J, Zanchetti A, Böhm M, et al. 2013 ESH/ESC Guidelines for the management of arterial hypertension: the Task Force for the management of arterial hypertension of the European Society of Hypertension (ESH) and of the European Society of Cardiology (ESC). J Hypertens 2013 Jul;31(7):1281-1357. [doi: 10.1097/01.hjh.0000431740.32696.cc] [Medline: 23817082]
16. Parati G, Stergiou GS, Asmar R, Bilo G, de Leeuw P, Imai Y, ESH Working Group on Blood Pressure Monitoring. European Society of Hypertension practice guidelines for home blood pressure monitoring. J Hum Hypertens 2010 Dec;24(12):779-785. [doi: 10.1038/jhh.2010.54] [Medline: 20520631]
17. Pickering TG, Miller NH, Ogedegbe G, Krakoff LR, Artinian NT, Goff D, American Heart Association, American Society of Hypertension, Preventive Cardiovascular Nurses Association. Call to action on use and reimbursement for home blood pressure monitoring: executive summary: a joint scientific statement from the American Heart Association, American Society Of Hypertension, and Preventive Cardiovascular Nurses Association. Hypertension 2008 Jul;52(1):1-9 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.107.189011] [Medline: 18497371]
18. Han HR, Lee H, Commodore-Mensah Y, Kim M. Development and validation of the Hypertension Self-care Profile: a practical tool to measure hypertension self-care. J Cardiovasc Nurs 2014;29(3):E11-E20 [FREE Full text] [doi: 10.1097/JCN.0b013e3182a3fd46] [Medline: 24088621]

19. Koh YL, Lua YH, Hong L, Bong HS, Yeo LS, Tsang LP, et al. Using a Web-Based Approach to Assess Test-Retest Reliability of the "Hypertension Self-Care Profile" Tool in an Asian Population: A Validation Study. Medicine (Baltimore) 2016 Mar;95(9):e2955 [FREE Full text] [doi: 10.1097/MD.0000000000002955] [Medline: 26945410]
20. Tucker KL, Sheppard JP, Stevens R, Bosworth HB, Bove A, Bray EP, et al. Self-monitoring of blood pressure in hypertension: a systematic review and individual patient data meta-analysis. PLoS Med 2017 Sep;14(9):e1002389 [FREE Full text] [doi: 10.1371/journal.pmed.1002389] [Medline: 28926573]
21. Bobrow K, Farmer AJ, Springer D, Shanyinde M, Yu L, Brennan T, et al. Mobile phone text messages to support treatment adherence in adults with high blood pressure (SMS-Text Adherence Support [StAR]): a single-blind, randomized trial. Circulation 2016 Feb 09;133(6):592-600 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.017530] [Medline: 26769742]
22. Lee K, Kwon H, Lee B, Lee G, Lee JH, Park YR, et al. Effect of self-monitoring on long-term patient engagement with mobile health applications. PLoS One 2018;13(7):e0201166 [FREE Full text] [doi: 10.1371/journal.pone.0201166] [Medline: 30048546]

Abbreviations

BP: blood pressure
ES: effect size
ESH: European Society of Hypertension
HBPM: home blood pressure monitoring
mHealth: mobile health
RCT: randomized controlled trial
SBP: systolic blood pressure
SNOSE: sequentially numbered, opaque, sealed envelope

Edited by G Eysenbach; submitted 15.12.18; peer-reviewed by M Lee, A Farmer, M Ponum, MS Aslam; comments to author 20.02.19; revised version received 04.03.19; accepted 24.03.19; published 28.05.19.

Please cite as:
Moon EW, Tan NC, Allen JC, Jafar TH
The Use of Wireless, Smartphone App–Assisted Home Blood Pressure Monitoring Among Hypertensive Patients in Singapore: Pilot Randomized Controlled Trial
JMIR Mhealth Uhealth 2019;7(5):e13153
URL: http://mhealth.jmir.org/2019/5/e13153/
doi:10.2196/13153
PMID:30905872

Original Paper

Self-Administered Auricular Acupressure Integrated With a Smartphone App for Weight Reduction: Randomized Feasibility Trial

Lorna Suen[1], BN, MPH, PhD; Wenru Wang[2], BScMed, MSN, PhD; Kenneth King Yip Cheng[3], BSc, PhD; Matthew Chin Heng Chua[4], BEng, PhD, CSM, PMP; Jerry Wing Fai Yeung[1], BCM, BSc, PhD; Wai Kin Koh[4], BEng; Simon Kai Wang Yeung[1], BScPych; Janice Yuen Shan Ho[1], BCM, PhD

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

[2]Alice Lee Centre for Nursing Studies, Yoog Loo Lin School of Medicine, National University of Singapore, Singapore

[3]Department of Health Technology and Informatics, The Hong Kong Polytechnic University, Hong Kong, China (Hong Kong)

[4]Smart Health Leadership Centre, Institute of Systems Science, National University of Singapore, Singapore

Corresponding Author:
Lorna Suen, BN, MPH, PhD
School of Nursing
The Hong Kong Polytechnic University
Hong Kong
China (Hong Kong)
Phone: 852 27667475
Fax: 852 23649663
Email: lorna.suen@polyu.edu.hk

Abstract

Background: Obesity is a common global health problem and increases the risk of many chronic illnesses. Given the adverse effects of antiobesity agents and bariatric surgeries, the exploration of noninvasive and nonpharmacological complementary methods for weight reduction is warranted.

Objective: The study aimed to determine whether self-administered auricular acupressure (AA) integrated with a smartphone app was more effective than using AA alone or the controls for weight reduction.

Methods: This study is a 3-arm randomized waitlist-controlled feasibility trial. A total of 59 eligible participants were randomly divided into either group 1 (AA group, n=19), group 2 (AA plus smartphone app, n=19), or group 3 (waitlist control, n=21). A total of 6 reflective zones or acupoints for weight reduction were chosen. The smartphone app could send out daily messages to the subjects to remind them to perform self-pressing on the 6 ear acupoints. A “date picker” of the 8-week treatment course was used to enable the users to input the compliance of pressing and the number of bowel movement daily instead of using the booklet for recordings. The app also served as a reminder for the subjects regarding the dates for returning to the center for acupoint changing and assessments. Treatment was delivered 2 times a week, for 8 weeks. Generalized estimating equations were used to examine the interactions among the groups before and after intervention.

Results: Subjects in group 2 expressed that the smartphone app was useful (7.41 out of 10). The most popular features were the daily reminders for performing self-pressing (88%), the ear diagram indicating the locations and functions of the 6 ear points (71%), and ear pressing method demonstrated in the video scripts (47%). Nearly 90% of the participants completed the 8-week intervention, with a high satisfaction toward the overall arrangement (8.37 out of 10). The subjects in group 1 and 2 achieved better therapeutic effects in terms of body weight, body mass index (BMI), waist circumference, and hip circumference and perceived more fullness before meals than the waitlist controls. Although no significant differences in the pairwise comparisons between the 2 groups were detected (P>.05), the decrease in body weight, BMI, body fat, visceral fat rating and leptin level, and increase in adiponectin level were notable in group 2 before and after the intervention.

Conclusions: The high compliance rate and high satisfaction toward the trial arrangement indicate that AA can be used to achieve weight reduction and applied in future large-scale studies. AA integrated with the smartphone app has a more notable effect than using AA alone for weight reduction. Larger sample size should be considered in future trials to determine the causal relationship between treatment and effect.

Trial Registration: ClinicalTrials.gov NCT03442712; https://clinicaltrials.gov/ct2/show/NCT03442712 (Archived by WebCite at http://www.webcitation.org/78L2tO8Ql)

(JMIR Mhealth Uhealth 2019;7(5):e14386) doi:10.2196/14386

KEYWORDS

acupressure; auriculotherapy; overweight; obesity; smartphone; leptin; adiponectin; randomized controlled trial

Introduction

Background

Obesity is a common global health problem caused by different factors such as endocrine disorder, metabolic syndrome, improper diet, drugs, or heredity [1,2]. It increases the risk of many chronic illnesses, including but not limited to hypertension, type 2 diabetes mellitus, cardiovascular diseases, musculoskeletal disorders, sleep apnea, and certain types of cancer [2-5].

Conventional approaches to alleviate obesity include medications [4], exercise [6], dietary control [7], behavior modification therapy [8,9], or bariatric surgeries [10]. However, the safety of antiobesity agents is a concern because they may induce depressed mood disorders, anxiety, or even increased risk of suicide during treatment [4]. Gastric bypass surgery and other bariatric surgeries also pose potential risks such as excessive bleeding, bowel obstruction, dumping syndrome, hernias, or stomach perforation [11].

Auriculotherapy or auricular treatment involves stimulating points on the ear with sterile acupuncture needles or acupressure with magnetized pellets or seeds [12]. Auriculotherapy offers a more effective and economical option to treat obesity than conventional approaches [13]. Auricular acupuncture has been frequently used to treat obesity in many countries [14-17]. However, the use of needles for auricular acupuncture may not be acceptable to people with needle phobia and could pose a risk of blood-borne transmission through needle prick injuries. Other methods of auricular acupoint stimulation include the application of electroacupuncture [18,19] or auricular acupressure (AA) using magnetic pellets or seeds [17-20]. Some researchers [21,22] attempted to adopt a combined approach by integrating auriculotherapy with diet restriction. Although auriculotherapy was found to be effective for weight reduction and management of dyslipidemia, the effect solely attributed to auriculotherapy could not be determined.

AA is a safe, noninvasive, inexpensive, and easily self-administered approach that causes few adverse effects [13,15,17]. A number of studies have attempted to examine the efficacy of using different materials for AA in weight reduction. The authors [13,15] concluded that using *semen vaccariae* is more effective than magnetic pellets for lowering body weight. During AA treatment, self-administered pressing on the seeds by the patients is necessary to achieve adequate acupoint stimulation. However, previous therapists may have difficulties in monitoring the compliance of subjects to perform seed pressing [17,21,23,24] that might affect the intervention dosages. Some controversial findings have indicated that AA does not significantly change the anthropometric parameters between experimental and control groups [25] probably because of a low compliance rate of subjects in conducting seed pressing, which the researchers fail to monitor.

Smartphones are currently the most popular communication tools, with approximately 70% of the global population using them. Accelerometer-based tracking devices and smartphone apps have been increasingly used to promote health because of their potential to influence self-regulation of a person's behavior [26-28].

Given that weight reduction increases the serum adiponectin concentration, resulting in a decrease in leptin levels [29], outcome measures should not only include anthropometric indices but also hormonal changes (leptin concentration and adiponectin level) associated with weight reduction. Determining the changes in plasma leptin and adiponectin levels could facilitate our understanding of the association of these biomarkers and the underlying mechanisms of the treatment protocol on weight reduction. A smartphone app (namely 'Auricular Acupressure for Weight Reduction, V1') was developed in this trial to monitor and enhance subjects' compliance on performing pressings to the acupoints.

Objectives

Our work aims to determine whether self-administered AA integrated with smartphone app was more effective than using AA alone or the controls for weight reduction. Subjects' satisfaction level toward the treatment protocol and the smartphone app, the recruitment and attrition rate of the subjects were evaluated. The preliminary effects of the treatment protocols, the hormonal changes associated with weight reduction, and the effect sizes of the treatment protocols were also determined. The findings of this feasibility study could provide valuable information for future large-scale studies.

Methods

Settings and Participants

This study is a 3-arm randomized waitlist-controlled feasibility trial. On the basis of the previous studies [20,30] of the effectiveness of AA on weight reduction, a medium effect size is considered for sample size calculation. According to Whitehead et al [31], a sample size of 15 per treatment arm is adequate. Considering an attrition rate of 20%, 19 to 20 subjects per arm were recruited for this feasibility study.

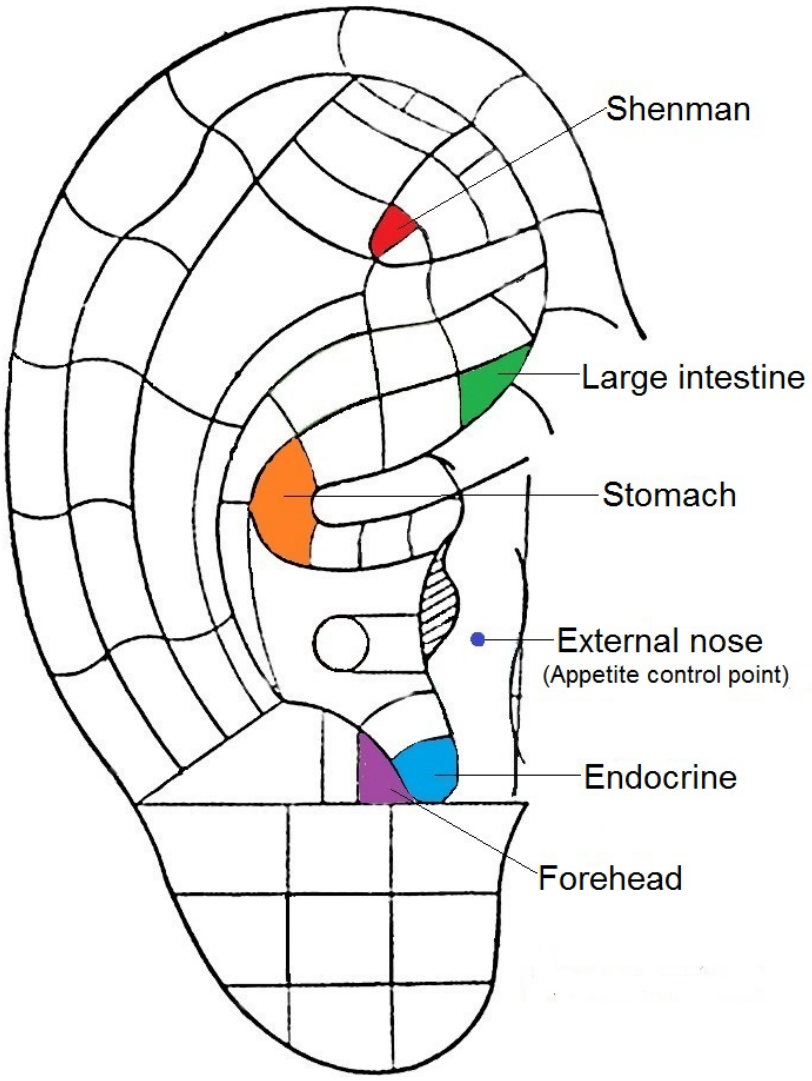
Subjects who met the following inclusion criteria were recruited from the community through snowball sampling via social network platforms (WhatsApp and Facebook): (1) age 18 years or older; (2) overweightness, with body mass index (BMI) ≥25.0 kg/m^2 in accordance with the BMI classification of the World

Health Organization (WHO) [2]; (3) neither received other weight control measures nor experienced medical or drug history within the last 3 months; (4) no ear injuries, such as inflammation or lesions, and no medical history of ear surgery within the last 6 months; and (5) smartphone user. Exclusion criteria were (1) diabetes, severe hypertension, heart disease or endocrine abnormalities; (2) pregnancy; (3) SCOFF (a questionnaire utilizing an acronym in 5 simple questions on "Sick, Control, One stone (6.5 kg), Fat, and Food") score ≥2 out of 5 items, which indicates eating disorders [14]; and/or (4) psychiatric and mental disorders.

Groupings and Procedure

Eligible subjects were randomly allocated to 1 of the 3 groups by using a computer-generated randomized table. The random allocation sequence was placed in an opaque, sequentially numbered, sealed envelope to guarantee adequate allocation concealment. The therapy was administered by a researcher (SY) who had received intensive coaching by the research team (LS, JY), and reliability on the accuracy of ear point identification was established. A total of 6 reflective zones or acupoints for weight reduction were chosen. These acupoints included "shenmen" (TF_4), "stomach" (CO_4), "endocrine" (CO_{18}), "external nose" ($TG_{1,\ 2i}$), "large intestine" (CO_7), and "forehead" (AT_1). The Chinese Standard Ear Acupoints system [32,33] and the nomenclature of the Nogier auricular acupoints (European system) [34,35] were taken as reference for acupoint selection and location identification (Figure 1). The principles of acupoint selection are to reduce excessive calorie intake and promote waste excretion to achieve weight reduction. The research team, which included an academic with over 20 years of research experience on auriculotherapy (LS) and 2 team members registered as traditional Chinese medicine practitioners of Hong Kong (JY and JH), selected the acupoints to be used.

Figure 1. Ear acupoints for weight reduction.



Group 1: Auricular Acupressure Only

AA treatment using *semen vaccariae* seeds [36] was performed by the researcher. The seeds were kept in place by a piece of adhesive tape (Figure 2) and were applied on 1 ear only. The researcher met with the subjects twice weekly to change the tapes every 3 to 4 days to the opposite ear to prevent skin irritation. Subjects were requested to apply pressure 20 times using a constant rhythm to each point thrice per day, preferably within 30 min before meals. Coaching on how to self-administer AA on the acupoints was given to the subjects, and a return demonstration was required to ensure that the treatment was

performed properly. An information booklet containing acupoint location and functions and the self-pressing methods was given to the subjects. The subjects were requested to record the frequency of daily pressing and bowel movement per day and show their records to the researcher in every visit for checking purposes.

Figure 2. Semen vaccariae for auricular acupressure.

Group 2: Auricular Acupressure Plus Smartphone App

A smartphone app, namely "Auricular Acupressure for Weight Reduction, V1," which is applicable for iPhone operating system (iOS) and android users and was written in Chinese, was developed (Figures 3-6). To obtain better performance and improve user experience, the smartphone app was developed natively by using Java (Version 8.0, Oracle Corporation) for Android phone user, and Swift (Version 4.0, Apple Inc) for iPhone user. Apart from providing the AA treatment and information booklet for the subjects, the developed app was installed in the smartphone of the subjects. The app contains an ear diagram indicating the locations and functions of the 6 ear points, video scripts that demonstrated proper ear pressing method, and precautions for performing AA. To minimize the

size of the app, the video scripts were stored in the cloud, and the video was streamed on the fly. During the first launch of the app, a "date picker" prompted the user to choose the starting date of the treatment. Once the user selected the starting date, it was verified in the background, and the 8-week treatment schedule was populated automatically as shown in Figure 5. Once user clicked on the schedule, it directed the user to the weekly schedule list where the user would be able to input the compliance of pressing and the number of bowel movement daily (Figure 6) instead of using the booklet for recordings. The data were then cached in the phone. The app sent out daily notifications to the subjects to remind them to perform self-pressing on the 6 ear acupoints. An algorithm based on the decision tree method was embedded in the app to determine the number of reminders and content of the notification. The frequency of sending reminders to the subjects by the app was accorded with the subjects' compliance of self-pressing. Normally, only 1 reminder/day was sent to the subject if they indicated good compliance (ie, 2-3 pressings/day), and an additional reminder would be sent to them the next day if the compliance is poor (ie, 0-1 time/day). In response to subjects' compliance on acupoint pressing, some messages such as "Excellent job" for those with good compliance and "You are almost there, keep going" for those with poor compliance would "pop up" upon data entry as positive reinforcement. The app also served as a reminder for the subjects regarding the dates for returning to the center for acupoint changing and assessments. In case the subjects encountered any difficulties during the process, they could send out a message to the researcher for receiving timely advice via the app. More details about the screenshots (in Chinese) of the "Auricular Acupressure for Weight Reduction, V1" app are shown in Multimedia Appendix 1.

Figure 3. An ear diagram indicating the locations and functions of the 6 ear points in the app.

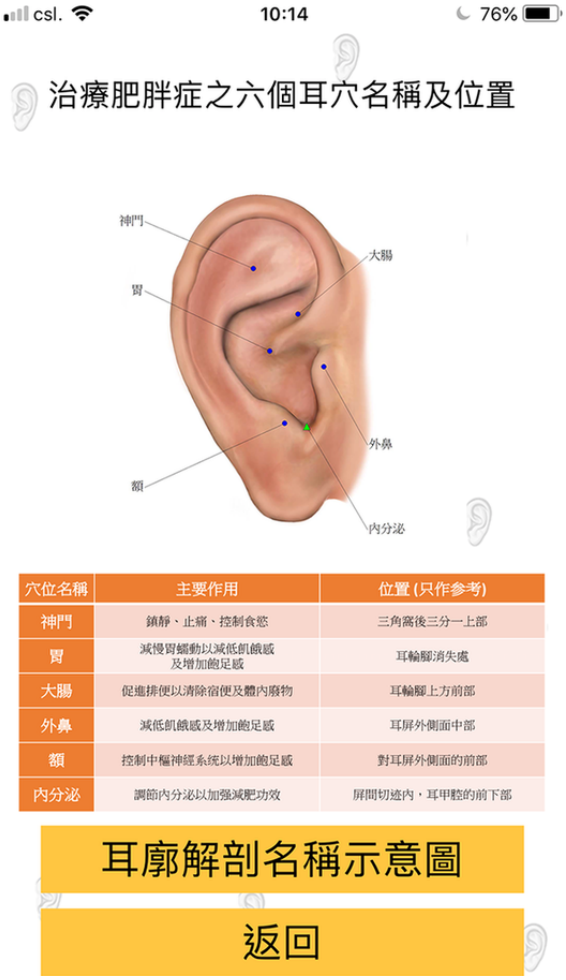


Figure 4. Video script that demonstrated proper ear pressing method in the app.

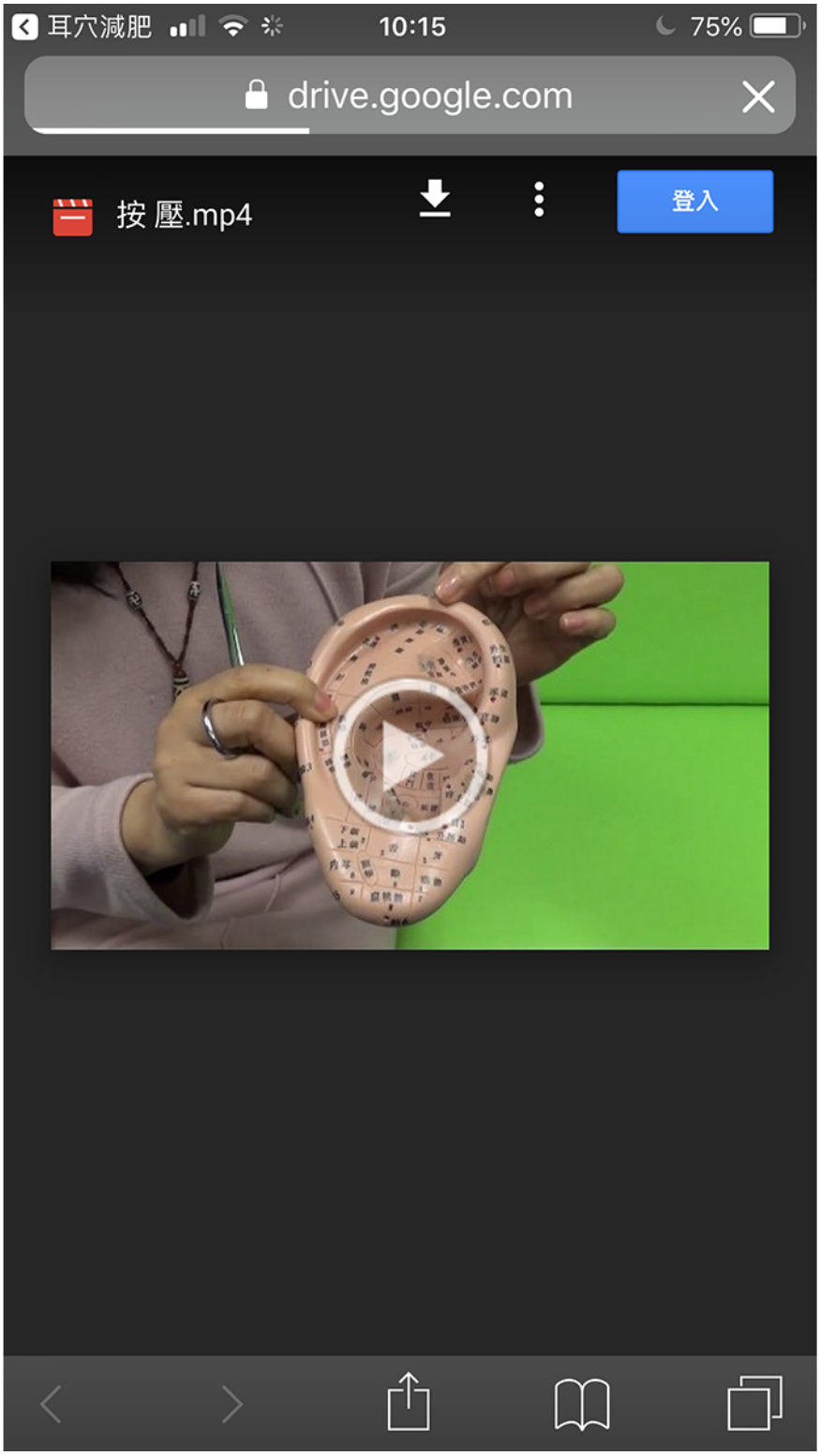


Figure 5. The eight-week treatment schedule was populated automatically once the user selected the starting date in the app.

Figure 6. The weekly schedule list which allows the user to input the compliance of pressing and the number of bowel movement daily in the app.

Group C: Waitlist Control Group

The subjects in the waitlist control group were told to maintain their usual dietary and exercising patterns during the waiting period. They were also required to receive the assessments similar to subjects in the other 2 groups. AA treatment plus smartphone app was given to these subjects after the 8-week intervention of the 2 experimental groups.

The following procedures were standardized for the subjects in groups 1 and 2. The auricle of the subject was cleaned using 75% isopropyl alcohol before therapy administration. Only 1 ear received treatment at a time. The experimental objects (seeds) were applied to the reactive region of these acupoints indicated by an acupoint finder (Pointer Plus) [37]. Treatment was firstly applied to the right ear during the first visit, followed by the left ear during the second visit, and so on. The experimental objects were replaced twice weekly to avoid local irritation of the auricular points under treatment. The total treatment period lasted for 8 weeks.

Ethical approval was obtained from the Human Research Ethics Review Committee of the Hong Kong Polytechnic University. The study was conducted in accordance with the Declaration of Helsinki. Participation in the study was voluntary. Written informed consent was obtained from each subject upon explanation of the risks and benefits of their participation. Given the multiple visits to the centers for receiving the protocol, a travel subsidy in the form of supermarket coupons (approximately US $25) was given to each subject upon completion of the study.

Outcome Measures

Another assistant who was unaware of the type of treatment modality received by the subjects evaluated the effect of treatment to achieve evaluator blinding. All outcome measurements were conducted at baseline and post intervention at 8 weeks. The BMI (kg/m^2) was taken as the primary outcome of the study. Secondary outcomes included other anthropometric indices such as body weight (kg), body fat (%), body water (%), and visceral fat rating, which were determined by a body composition analyzer (Model: Tanita BC-545N) [38]. Waist (cm) and hip circumferences (cm) were taken twice using standard method [39] to ensure accuracy.

Laboratory tests for leptin and adiponectin testing were conducted via standardized methods. Leptin concentration was measured according to the manufacturer's instructions using a commercial sandwich ELISA kit [40] comprising ready-to-use components, which were either concentrated or lyophilized. The dilution factor was considered during the calculation of leptin concentrations. The limit of detection of the leptin assay was 0.2 ng/mL. The research team would repeat the tests if results exceeded a leptin concentration of 50 ng/mL for diluted samples. Precision intra-assay (within-run), coefficient of variability (CV)=5.9%, and inter-assay (run-to-run) CV=5.6% were reported to express the precision or repeatability of the immunoassay test result [5]. Adiponectin level was measured with a commercially available sandwich ELISA kit [41]. The assay has a sensitivity of 1.5 ng/mL^{-1}. The serum concentrations of leptin and adiponectin were calculated based on standard curves plotted according to the manufacturer's instructions [29].

Subjects were asked to rate the fullness level before lunch and dinner using a visual analog scale of 0 to 10 (adapted from Rock et al) [42] for 3 consecutive days at baseline and post intervention. The question was "How full do you feel?," with anchor values ranging from 0 ("Not at all full") to 10 ("Totally full"). Before the therapy, the subjects were asked regarding their confidence and perceived usefulness of the treatment to manage their overweight problem. Upon completion of the 8-week protocol, the satisfaction of the subjects toward the therapy, the information booklet, and the smartphone app (if applicable) was also evaluated using a 10-point scale, with higher scores indicating greater satisfaction.

Data Analyses

Descriptive statistics for sociodemographic characteristics of the subjects were presented. The estimated mean and SE of the outcome variables before and after intervention were computed. The association between categorical variables was examined using χ^2 test or Fisher exact test. The Mann–Whitney *U* test or Kruskal–Wallis test was used for detecting group differences where appropriate.

Primary analysis was conducted using the generalized estimating equations with an auto-regression correlation structure to examine the interactions among the groups before and after intervention on the primary and secondary outcome variables, including anthropometric indices, perceived hunger level before meals, leptin concentration (ng/mL), and adiponectin level (ng/mL). Missing data were addressed using the GEE model and assumed to be missed at random [43]. Apart from conducting analyses on all participants, we repeated the main analysis of the completers who had finished the treatment protocol for sensitivity analysis. Correlation analyses among the anthropometric indices and the biomarkers were conducted. The reported adverse effects, expectation, and satisfaction toward the therapy were evaluated. We used SPSS version 25.0 (IBM Corporation) for all statistical analyses. All statistical tests were 2-sided, with the significance level set to 0.05.

Results

The data were collected from April to November 2018. The recruitment rate was fairly high (72%). Within 2 weeks of promotion via the social platforms (WhatsApp and Facebook), 82 enquiries regarding the project were received. A total of 59 eligible participants were randomly divided into 3 groups (group 1=19, group 2=19, and group 3=21). Participants who were excluded were mainly because of having metabolic syndrome with uncontrolled diabetic or hypertensive conditions. After 8 weeks of intervention, the attrition rate was only 10% (n=6). The flow diagram of the participants of this trial is illustrated in Multimedia Appendix 2. The trial is reported in accordance with Consolidated Standards of Reporting Trials of Electronic and Mobile HEalth Applications and onLine TeleHealth in Multimedia Appendix 3.

Subject Characteristics

The recruited subjects had an average age of 49.15 years (SD 10.54), with a mean BMI of 30.35 (SD 4.53 kg/m^2). The waist–hip ratio was 0.95 (SD 0.63) and 0.90 (SD 0.05) for males and females, respectively. Males only accounted for 15% (n=9) of the subjects. The groups were essentially comparable and well balanced in age, gender distribution, BMI, education level, marital status, comorbid illnesses, exercising level, daily fluid intake, smoking status, and alcohol consumption. Majority of the participants had eating-out habits at least once per day (83%), perceived themselves as a gluttonous person (48%), ever attempted other means of weight reduction (58%), and worried about their overweight problem. Women significantly felt unhappier regarding their overweight problem than men (P<.05). None of the participants had eating disorders as verified by the SCOFF questionnaire [14]. Among the subjects, there were more android users (37/59, 63%) than iOS users (22/59, 37%; Table 1).

Table 1. Sociodemographic and baseline characteristics of the subjects (N=59).

Characteristics	All	Group 1: AA[a] (n=19)	Group 2: AA + app (n=19)	Group 3: Waitlist control (n=21)	*P* value
Age (years), mean (SD)	49.15 (10.54)	47.58 (11.59)	49.21 (9.70)	50.52 (10.58)	.87[b]
Gender, n (%)					
Male	9 (15)	4 (21)	3 (16)	2 (10)	.60[c]
Female	50 (85)	15 (79)	16 (84)	19 (90)	—[d]
Education level, n (%)					
Primary or below	1 (2)	0 (0)	1 (5)	0 (0)	.36[c]
Secondary	32 (54)	13 (68)	8 (42)	11 (52)	—
Tertiary or above	26 (44)	6 (32)	10 (53)	10 (48)	—
Marital status, n (%)					
Single, divorced/widowed	19 (32)	10 (53)	5 (26)	4 (19)	.08[c]
Married	40 (68)	9 (47)	14 (74)	17 (81)	—
Body mass index (kg/m^2), mean (SD)	30.35 (4.53)	30.32 (4.65)	30.65 (5.41)	30.11 (3.70)	.97[b]
Waist-hip ratio, mean (SD)					
Male	0.95 (0.63)	0.97 (0.59)	0.91 (0.08)	0.97 (0.06)	.04[b]
Female	0.90 (0.05)	0.91 (0.05)	0.89 (0.05)	0.89 (0.05)	.25[b]
Comorbid illness, n (%)					
No	37 (63)	10 (53)	12 (63)	15 (71)	.48[e]
Yes	22 (37)	9 (47)	7 (37)	6 (29)	—
Regular drugs taken, n (%)					
No	39 (66)	11 (58)	13 (68)	15 (71)	.70[e]
Yes	20 (34)	8 (42)	6 (32)	6 (29)	—
Regular exercise, n (%)					
No	30 (51)	10 (53)	8 (42)	12 (57)	.72[e]
Yes	29 (49)	9 (47)	11 (58)	9 (43)	—
Smoking habit, n (%)					
Never	56 (95)	16 (84)	19 (100)	21 (100)	.06[c]
Ex-smoker	2 (3)	2 (11)	0 (0)	0 (0)	—
Current smoker	1 (2)	1 (5)	0 (0)	0 (0)	—
Drinker, n (%)					
Never	21 (36)	7 (37)	7 (37)	7 (33)	>.99[e]
Social drinker	38 (64)	12 (63)	12 (63)	14 (67)	—
Cups of fluid intake/day, mean (SD)	6.84 (2.04)	6.63 (2.41)	6.84 (1.73)	7.05 (2.02)	.82[b]
Number of times/day eat-out, n (%)					
0	10 (18)	3 (16)	4 (21)	3 (16)	.42[c]
1-2	37 (64)	13 (68)	13 (68)	11 (58)	—
3-4	10 (18)	3 (16)	2 (11)	5 (26)	—
Claim to be a gluttonous person, n (%)					
Sometimes	31 (53)	14 (74)	7 (37)	10 (48)	.05[e]

Characteristics	All	Group 1: AA[a] (n=19)	Group 2: AA + app (n=19)	Group 3: Waitlist control (n=21)	*P* value
Always	28 (47)	5 (26)	12 (63)	11 (52)	—
Bowel habit, n (%)					
Regular bowel open	43 (73)	12 (63)	13 (68)	18 (86)	.36[c]
Occasion constipation	11 (19)	4 (21)	4 (21)	3 (14)	—
Frequent constipation	5 (8)	3 (16)	2 (11)	0 (0)	—
Ever try other means for weight reduction, n (%)					
No	25 (42)	7 (37)	8 (42)	10 (48)	.76[e]
Yes	34 (58)	12 (63)	11 (58)	11 (52)	—
Worry about overweight, n (%)					
Always	17 (29)	6 (32)	7 (37)	4 (19)	.23[c]
Sometimes	34 (58)	8 (42)	11 (58)	15 (71)	—
Never	8 (13)	5 (26)	1 (5)	2 (10)	—
Unhappy owing to overweight, n (%)					
Always	10 (17)	5 (26)	4 (21)	1 (5)	.29[c]
Sometimes	33 (56)	8 (42)	10 (53)	15 (71)	—
Never	16 (27)	6 (32)	5 (26)	5 (24)	—
Phone model, n (%)					
Android	37 (63)	9 (47)	13 (68)	15 (71)	.29[e]
iPhone operating system	22 (37)	10 (53)	6 (32)	6 (29)	—
Sick, Control, One stone (6.5 kg), Fat, and Food score, mean (SD)	0.32 (0.80)	0.47 (1.17)	0.37 (0.68)	0.14 (0.36)	.52[b]

[a]AA: auricular acupressure.

[b]Kruskal-Wallis test.

[c]Fisher exact test.

[d]Not applicable.

[e]Chi-square test.

Compliance, Expectation, and Satisfaction Toward Treatment

The compliance to the intervention protocol was high, with an average of 90% (n=53) having completed the 8-week intervention. Even though over 76% of the subjects did not experience receiving complementary and alternative treatment, they generally exhibited a positive attitude toward AA (6.53 out of 10) before the trial. After the intervention, their satisfaction with the overall arrangement was high (8.37 out of 10). The compliance for performing self-pressing was satisfactory, with 79% of subjects in groups 1 and 2 performing self-pressing 3 times or more before meals throughout the protocol. No significant difference in the pressing compliance and frequency of bowel movement between groups was noted.

Subjects in group 2 expressed that the smartphone app was useful (7.41 out of 10). The most popular features were the daily reminders for performing self-pressing (88%), the ear diagram indicating the locations and functions of the 6 ear points (71%), and ear pressing method demonstrated in the video scripts (47%). Majority of the subjects in groups 1 and 2 (n=34, 97%) indicated that they would consider recommending this therapy to others. No specific adverse effects arising from the therapy were observed, except 2 participants (5%) who reported having mild skin irritation on the ears because of the adhesive tapes for holding the experimental objects in place. The most tender acupoint felt by the subjects was “stomach” (74%), followed by “forehead” (34%), “endocrine” (32%), “external nose” (18%), “shenmen” (16%), and “large intestine” (16%). Only 2 male subjects (6%) felt embarrassed because of the adhesive tapes put on the auricles (Table 2).

Table 2. Reported adverse effects, expectations, and satisfaction toward the therapy (N=59).

Variables	All	Group 1: AA[a] (n=19)	Group 2: AA + app (n=19)	Group 3: Waitlist control (n=21)	*P* value
Have you used complementary therapies in the past?[b]					
No	45 (76%)	16	14	15	.65[c]
Yes	14 (24%)	3	5	6	—[d]
How much confidence do you have in complementary therapies in general (0-10)[b]	6.05 (1.91)	5.84 (1.89)	6.05 (2.17)	6.24 (1.73)	.93[e]
Perceived usefulness of the treatment being received(0-10)[b]	6.53 (1.52)	6.16 (1.54)	6.68 (1.34)	6.71 (1.68)	.37[e]
What will you expect about the overweight problem after 8 weeks (post intervention)?[b]					
Totally resolve	0	0	0	0	.28[c]
Improve greatly	9	3	2	4	—
Moderate improvement	22	10	6	6	—
Little improvement	20	3	10	7	—
Same as before	8	3	1	4	—
Ear itchiness[f]	2 (5%)	2	0	—	—
Tenderness on acupoints[f]					
Shenmen	6 (16%)	2	4	—	—
Stomach	28 (74%)	14	14	—	—
External nose	7 (18%)	4	3	—	—
Forehead	13 (34%)	7	6	—	—
Large intestine	6 (16%)	5	1	—	—
Endocrine	12 (32%)	7	5	—	—
Change in satiety level[f]					
Decrease appetite	23 (62%)	12	10	—	.20[c]
No change	12 (32%)	5	7	—	—
Increase appetite	2 (5%)	1	0	—	—
Compliance on performing pressing on acupoints[f]					
<3 times/day	7	4	3	—	>.99[c]
3 times/day	21	11	10	—	—
>3 times/day	9	4	5	—	—
Satisfaction toward the overall arrangement (0- 10)[f]	8.37 (1.68)	8.61 (1.85)	8.12 (1.50)	—	.18[e]
Satisfaction toward the treatment effect (0-10)[f]	5.63 (2.46)	5.89 (2.25)	5.35 (2.71)	—	.65[e]
Usefulness of the information booklet[f], mean (SD)	6.03 (2.53)	7.00 (2.30)	5.00 (2.40)	—	.02[e]
Usefulness of the mobile app[f], mean (SD)	—	—	7.41 (2.53)	—	—
Usefulness of the features in the app[f]					
Reminder	—	—	15 (88%)	—	—

Variables	All	Group 1: AA[a] (n=19)	Group 2: AA + app (n=19)	Group 3: Waitlist control (n=21)	*P* value
Video script	—	—	6 (36%)	—	—
Contact the researcher	—	—	4 (24%)	—	—
Ear diagram	—	—	12 (71%)	—	—
Ear pressing method demonstration	—	—	8 (47%)	—	—
Will recommend this therapy to others[f]					>.99[e]
Definitely will	10 (29%)	5	5	—	—
Maybe	24 (69%)	12	12	—	—
No	1 (3%)	1	0	—	—
Embarrass owing to ear plaster[f]					
No	33	18	15	—	.23[e]
Yes	2	0	2	—	—

[a]AA: auricular acupressure.

[b]Evaluated before the intervention.

[c]Fisher exact test.

[d]Not applicable.

[e]Mann-Whitney *U* test or Kruskal-Wallis test as appropriate.

[f]Evaluated after the intervention has been completed.

Treatment Effect

In general, the subjects in groups 1 and 2 achieved better therapeutic effects in terms of body weight (kg), BMI (kg/m^2), waist circumference (cm), and hip circumference (cm), and perceived more fullness before meals than the waitlist controls. Although no significant differences in the pairwise comparisons between the 2 groups were detected, the decrease in body weight, BMI, body fat, visceral fat rating and leptin level, and increase in adiponectin level were notable in group 2 before and after the intervention (Table 3). When the effect size was estimated using the primary outcome (BMI), a medium effect size (d=0.4928) was determined in group 1 (AA), whereas a large effect size (d=0.7798) was detected in group 2 (AA plus app), taking the waitlist control group as reference. Correlation analyses indicate that significant correlations were present among the anthropometric indices and the leptin concentration (ng/mL) after the intervention (Multimedia Appendix 4). A completer's analysis showed consistent findings on the outcome variables of the trial.

Table 3. Outcome variables across the 3 groups before and after intervention.

Measures	Grouping						Pairwise comparisons between groupsBeta (95% CI)		
	Group 1: AA[a] (n=19)		Group 2: AA + app (n=19)		Group 3: Waitlist control (n=21)		Group 1 versus group 2	Group 1 versus group 3	Group 2 versus group 3
	Estimated mean (SE)	Mean difference (*P* value)	Estimated mean (SE)	Mean difference (*P* value)	Estimated mean (SE)	Mean difference (*P* value)			
Body weight (kg)									
Baseline	78.87 (2.83)	1.33 (0.005)	78.00 (3.15)	1.56 (0.000)	75.68 (2.18)	0.23 (0.515)	–0.23 (–1.38 to 0.92)	1.10 –0.07 to 2.26)	1.33 (–0.36 to 2.29)
Postintervention	77.54 (2.63)	__[b]	76.44 (3.09)	—	75.44 (2.12)	—	—	—	—
Body mass index (kg/m^2)									
Baseline	30.32 (1.04)	0.50 (0.013)	30.65 (1.21)	0.60 (0.000)	30.11 (0.79)	0.12 (0.435)	–0.10 (–0.56 to 0.36)	0.38 (–0.12 to 0.87)	0.48 (0.09 to 0.87)
Postintervention	29.82 (0.98)	—	30.05 (2.21)	—	29.99 (0.75)	—	—	—	—
Body fat (%)									
Baseline	39.58 (1.72)	0.17 (0.640)	40.73 (2.18)	0.39 (0.012)[c]	41.34 (1.44)	0.66 (0.274)	–0.23 (–1.00 to 0.54)	–0.49 (–1.88 to 0.89)	–0.27 (–1.49 to 0.96)
Postintervention	39.42 (1.71)	—	40.33 (2.19)	—	40.68 (1.51)	—	—	—	—
Body water (%)									
Baseline	45.21 (1.04)	–0.33 (0.190)	44.66 (1.33)	0.23 (0.144)	44.83 (0.89)	0.17 (0.641)	–0.56 (–1.15 to 0.02)	–0.51 (–1.38 to 0.37)	0.06 (–073 to 0.84)
Postintervention	45.54 (1.06)	—	44.43 (1.31)	—	44.66 (0.84)	—	—	—	—
Visceral fat rating									
Baseline	11.05 (0.85)	0.28 (0.021)	11.13 (0.64)	0.62 (0.066)	10.45 (0.57)	–0.28 (0.477)	–0.34 (–1.04 to 0.36)	0.56 (–0.25 to 1.36)	0.90 (–0.12 to 1.92)
Postintervention	10.78 (0.80)	—	10.51 (0.61)	—	10.73 (0.62)	—	—	—	—
Waist circumference (cm)									
Baseline	101.08 (2.61)	2.81 (0.001)	98.85 (2.55)	1.60 (0.006)	97.65 (1.54)	–0.76 (0.185)	1.21 (–0.79 to 3.21)	3.57 (1.59 to 5.55)	2.36 (0.75 to 3.96)
Postintervention	98.27 (2.31)	—	97.25 (2.52)	—	98.41 (1.51)	—	—	—	—
Hip circumference (cm)									
Baseline	109.66 (2.29)	1.53 (0.003)	110.74 (2.36)	1.41 (0.002)	109.07 (1.57)	–0.26 (0.612)	0.12 (–1.23 to 1.46)	1.78 (–2.55 to –0.51)	1.67 (0.35 to 2.99)
Postintervention	108.14 (2.19)	—	109.33 (2.42)	—	109.32 (1.48)	—	—	—	—
Waist-to-hip ratio									
Baseline	0.92 (0.13)	0.01 (0.084)	0.89 (0.01)	0.00 (0.694)	0.90 (0.01)	0.00 (0.259)	0.01 (–0.01 to 0.03)	0.02 (0.00 to 0.03)	0.01 (–0.01 to 0.02)
Postintervention	0.91 (0.01)	—	0.89 (0.01)	—	0.90 (0.01)	—	—	—	—

Measures	Grouping						Pairwise comparisons between groupsBeta (95% CI)		
	Group 1: AA[a] (n=19)		Group 2: AA + app (n=19)		Group 3: Waitlist control (n=21)		Group 1 versus group 2	Group 1 versus group 3	Group 2 versus group 3
	Estimated mean (SE)	Mean difference (*P* value)	Estimated mean (SE)	Mean difference (*P* value)	Estimated mean (SE)	Mean difference (*P* value)			
Leptin (ng/ml)									
Baseline	63.87 (9.97)	1.98 (0.850)	77.81 (14.23)	12.11 (0.243)	70.05 (9.50)	14.24 (0.071)	–10.13 (–38.98 to 18.71)	–12.27 (–37.93 to 13.39)	–2.14 (–27.67 to 23.40)
Postintervention	61.90 (9.14)	—	65.70 (11.19)	—	55.80 (6.32)	—	—	—	—
Adiponectin (ng/ml)									
Baseline	14896.48 (2190.72)	–370.60 (0.671)	16163.83 (2671.44)	–1468.54 (0.717)	14117.40 (2063.87)	–239.12 (0.876)	1097.94 (–7036.93 to 9232.82)	–131.47 (0.36 to 3321.20)	–1229.42 (–9730.36 to 7271.53)
Postintervention	15267.08 (2309.31)	—	17632.37 (3384.36)	—	14356.52 (1846.75)	—	—	—	—
Fullness (lunch)									
Baseline	4.54 (0.58)	–0.60 (0.333)	4.51 (0.47)	–1.06 (0.047)[c]	6.02 (0.50)	–0.28 (0.399)	0.82 (–1.15 to 2.07)	–0.32 (–1.70 to 1.06)	1.06 (0.01 to 2.12)[c]
Postintervention	5.15 (0.44)	—	5.57 (0.43)	—	6.30 (0.53)	—	—	—	—
Fullness (dinner)									
Baseline	4.60 (0.62)	–0.71 (0.321)	4.51 (0.51)	–1.40 (0.010)	6.76 (0.51)	0.26 (–0.328)	–0.69 (–1.06 to 2.45)	–0.97 (–2.45 to 0.52)	1.40 (033 to 2.47)
Postintervention	5.30 (0.47)	—	5.91 (040)	—	6.50 (0,59)	—	—	—	—

[a]AA: auricular acupressure.

[b]Not applicable.

[c]Estimated mean and standard error (SE) from generalized estimating equations.

Discussion

Principal Findings

The findings of this 3-arm randomized waitlist-controlled study on 59 eligible participants demonstrates that the trial arrangement was feasible. The high compliance rate and high satisfaction toward the study indicate that AA can be used to achieve weight reduction and applied in future large-scale studies. Although no significant differences in the pairwise comparisons between AA integrated with the smartphone app (group 2) and AA only (group 1) were detected, the decrease in body weight, BMI, body fat, visceral fat rating, and leptin level and increase in adiponectin level were notable in group 2 before and after the intervention. The most popular features of the smartphone app were the daily reminders for performing self-pressing, the ear diagram indicating the locations and functions of the 6 ear points, and ear pressing method demonstrated in the video scripts. Majority of the subjects receiving AA treatment indicated that they would consider recommending this therapy to others. No specific adverse effects arising from the therapy were observed.

The high recruitment rate of this study indicates that overweight and obesity are common problems among the population. According to the 2016 statistics of the World Health Organization (WHO), nearly 2.0 billion adults around the world were overweight, of which 34.0% were obese [2]. The eligible participants had a BMI of 30.35, which is classified as obese according to the WHO criteria [2]. Only a small number of males (n=9) participated in this study possibly because of the gender disparities in attitude toward the overweight problem. Although the waist–hip ratio was above the normal range in both genders (males vs females: 0.95 vs 0.90), women significantly felt unhappier regarding the overweight problem than men. This finding may possibly be attributed to the active participation of women in weight reduction programs to establish a better image and appearance after becoming slim. However, android obesity (apple-shaped) is becoming more common in males, and the visceral fat, which lies deep inside the abdomen surrounding the internal organs, affects vital organs such as the heart, liver, kidney, and the lungs [44], thus increasing the risk of cardiovascular disease, hypertension, diabetes, sleep apnea, colorectal cancer, and premature death

[44-46]. Therefore, additional strategies are needed to encourage males to deal with their overweight problem actively.

The high compliance rate, high satisfaction toward the trial arrangement, and the low attrition rate indicate the feasibility of the trial. These indicators also reflected a high motivation among the subjects (groups 1 and 2) who participated in the trial. Nearly all participants (97%) said that they would consider recommending this therapy to others.

Majority of the subjects had eating-out habits at least once per day (83%), perceived themselves as a gluttonous person, and ate even if they were not hungry (48%). Acupoint selection was mainly guided by the principles to reduce excessive calorie intake and to promote waste excretion to achieve weight reduction. For example, the use of "stomach" could reduce the motility of the stomach, diminish the sense of hunger, and raise serotonin levels, thereby suppressing appetite [13,15,35]. "External nose" is also called "hunger point" and "appetite control point," which could heighten satiety and curb appetite [34]. "Forehead" could inhibit the central nervous system and increase satiety [33]. "Large intestine" can promote excretion [33], "endocrine" regulates endocrine function [33], and "shenmen" can reduce anxiety associated with weight loss and suppress the appetite of the subjects [13,15,35].

In general, the subjects in groups 1 and 2 achieved better therapeutic effect in certain anthropometric indices before meals than the waitlist controls. These findings are in accordance with the study conducted by Kim et al [17] who also demonstrated that AA is an effective intervention to reduce weight and BMI, decrease the sensation of hunger, and increase satiety of a person. Auricular stimulation modulates the hypothalamic neuronal activities associated with feeding, thereby curbing appetite [18,22]. Seeds for the administration of AA induce pressure to the skin according to the principle of neural reflexes similar to that of acupuncture [15] and influence appetite-related hormone peptides [16,47].

Although no significant differences in the pairwise comparisons between the 2 groups could be observed ($P>.05$), the decrease in body weight, BMI, body fat, visceral fat rating, and leptin level and increase in adiponectin level were more notable in group 2 before and after the intervention. The effect size of group 2 was also larger than that of group 1, indicating that the effect is substantial according to Cohen terminology [48]. Taking a conservative approach, a medium effect size with a sample size of 32 subjects/arm could be considered in future studies to detect 5% level of significance with 95% power.

Smartphones and wireless devices as mobile health have been widely used in recent decades to improve health outcomes. Previous studies have demonstrated the effectiveness of using smartphone apps as a useful tool to self-regulate diet for weight loss in patients with overweightness or obesity [49-53]. A meta-analysis on 11 studies indicated that reminder systems could significantly increase patient adherence to treatment [54]. Therefore, the daily reminders sent out to the subjects were hypothesized to increase their compliance to perform self-pressing, thereby achieving adequate acupoint stimulation. However, the absence of significant difference in pressing compliance between the 2 groups might indicate that the subjects in group 1 were independent and self-disciplined enough to follow the instructions and perform self-pressing as instructed even though no reminders were sent to them.

In general, the subjects in group 1 had higher perceived usefulness of the information booklet than those in group 2 mainly because the booklet was the only source of written information related to the therapy delivered to them. Subjects in group 2 expressed that the smartphone app was useful and were especially impressed by the daily reminders for performing self-pressing, the ear diagram indicating the locations and functions of the 6 ear points, and ear pressing method demonstrated in the video scripts. The stronger therapeutic effect observed in the protocol of group 2 might be because of the additional support, encouragement, and subjects' interactions when using the app.

Adiponectin and leptin are 2 major hormones derived from adipose tissue (fat) [55]. Adiponectin controls body weight by regulating food intake and sensitizing insulin action in body tissues [22,56]. The adiponectin level is inversely correlated with body mass, especially visceral fat and metabolic health [22,57]. Weight reduction increases the serum adiponectin concentration, resulting in a decrease in leptin levels [29]. The adiponectin/leptin ratio in patients with obesity and diabetes is significantly lower than that in patients without obesity [58]. The elevated level of adiponectin also has a positive health impact because of its anti-inflammatory, antioxidant and antiatherosclerotic properties [59].

By contrast, leptin secretion may be an important mechanism, in which adipose tissue sends signals to the hypothalamic nuclei that the body has stored enough fat and thus no longer requires food intake [5]. Previous evidence demonstrates that AA is associated with decreased leptin levels, which are associated with weight loss [5,16,22,60]. The measurement of leptin and adiponectin can serve as a surrogate marker of metabolic heath and indicate whether the treatment has effects on adipose tissue function and systemic metabolism. In this trial, correlation analyses indicate that a significant correlation exists among the anthropometric indices and the leptin concentration (ng/mL) after the intervention. However, the changes of these biomarkers (leptin and adiponectin) among the subjects in group 2 are notable but not significant. Therefore, further investigations should be conducted before conclusive results of the pathway can be drawn.

Only 2 male subjects felt embarrassed because of the adhesive tapes placed on the auricles. No specific adverse effects arising from the therapy were observed, except 2 participants who reported having mild skin irritation on the ears because of the adhesive tapes used for positioning the experimental objects in place. The most tender acupoint felt by the subjects was "stomach," followed by "forehead," "endocrine," "external nose," "shenmen," and "large intestine." According to the auricular diagnosis system, the areas of the auricle with heightened tenderness upon touching correspond to specific areas of the body where some pathological conditions exist [61,62]. Applying seeds may induce physical pressure on the ear acupoints and cause tenderness, especially in cases with disequilibrium of the bodily functions (in this case, overweight

problem) corresponding to specific acupoints. The tenderness on the reflective acupoints experienced by the subjects could be taken as a part of the treatment process rather than as adverse effects of auriculotherapy.

Limitations and Recommendations of the Study

As this is a feasibility study enrolled with a small number of participants, therefore generalization of results is limited. A larger sample size with more male subjects should be considered in future trials to evaluate whether any gender disparities may have an impact on the treatment effect. Moreover, longer follow-up (a minimum of 3 months) after the therapy may be considered in evaluating the sustained treatment effect. The effectiveness of AA should also be evaluated in other populations, such as those with metabolic syndrome or children with obesity, so that their overweight problem can be tackled using a self-manipulated and nontraumatic approach.

Conclusions

The high compliance rate, high satisfaction toward the trial arrangement, and low attrition rate indicate that AA can be used to achieve weight reduction and applied in future large-scale studies. AA integrated with the smartphone app has a more notable effect than using AA alone in terms of the decrease in body weight, BMI, body fat, visceral fat rating, and leptin level and increase in adiponectin level. Larger sample size should be considered in future trials to determine the causal relationship between treatment and effect.

Acknowledgments

This project was supported by the Internal Research Fund, School of Nursing, The Hong Kong Polytechnic University, Hong Kong (G-UADZ). The authors extend their appreciation to Dr HC Wong and his team at the University Health Services, The Hong Kong Polytechnic University for providing the support in the blood sampling. Special thanks to Mr Ego Ho and Mr Cato To for the technical support and advice in the video filming and the smartphone app installation. The authors are also grateful to Ms Kiki Lo who offered her assistance in conducting the assessments; and to the subjects for their participation in this study.

Authors' Contributions

LS conceived the study and study design, analysis, and manuscript preparation. WW participated in the design of the study. KC and JH offered assistance in biomarker measurements and analyses. MC and WKK were involved in smartphone app development. SY was involved in recruitment of subjects and therapy administration. All authors have read and approved the final version of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Screenshots of the Auricular Acupressure for Weight Reduction, V1 app.

[PPTX File, 2MB - mhealth_v7i5e14386_app1.pptx]

Multimedia Appendix 2

Flowchart of the trial.

[DOCX File, 94KB - mhealth_v7i5e14386_app2.docx]

Multimedia Appendix 3

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 7MB - mhealth_v7i5e14386_app3.pdf]

Multimedia Appendix 4

Correlational analyses among anthropometric indices (postintervention).

[DOCX File, 39KB - mhealth_v7i5e14386_app4.docx]

Multimedia Appendix 5

CONSORT-EHEALTH checklist (V 1.6.1).

[PDF File (Adobe PDF File), 7MB - mhealth_v7i5e14386_app5.pdf]

References

1. El Kabbaoui M, Chda A, Bousfiha A, Aarab L, Bencheikh R, Tazi A. Prevalence of and risk factors for overweight and obesity among adolescents in Morocco. East Mediterr Health J 2018 Jul 29;24(6):512-521 [FREE Full text] [doi: 10.26719/2018.24.6.512] [Medline: 30079946]
2. World Health Organisation. News: Obesity and overweight URL: https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight [accessed 2019-05-08] [WebCite Cache ID 78Dc5mCqM]
3. Batsis JA, Gill LE, Masutani RK, Adachi-Mejia AM, Blunt HB, Bagley PJ, et al. Weight loss interventions in older adults with obesity: a systematic review of randomized controlled trials since 2005. J Am Geriatr Soc 2017 Feb;65(2):257-268 [FREE Full text] [doi: 10.1111/jgs.14514] [Medline: 27641543]
4. Christensen R, Kristensen PK, Bartels EM, Bliddal H, Astrup A. Efficacy and safety of the weight-loss drug rimonabant: a meta-analysis of randomised trials. Lancet 2007 Nov 17;370(9600):1706-1713. [doi: 10.1016/S0140-6736(07)61721-8] [Medline: 18022033]
5. Darbandi S, Darbandi M, Mokarram P, Owji A, Zhao B, Ghayor-Mobarhan M, et al. Effects of body electroacupuncture on plasma leptin concentrations in obese and overweight people in Iran: a randomized controlled trial. Altern Ther Health Med 2013;19(2):24-31. [Medline: 23594450]
6. Beavers KM, Ambrosius WT, Rejeski WJ, Burdette JH, Walkup MP, Sheedy JL, et al. Effect of exercise type during intentional weight loss on body composition in older adults with obesity. Obesity (Silver Spring) 2017 Dec;25(11):1823-1829. [doi: 10.1002/oby.21977] [Medline: 29086504]
7. Perreault L. UpToDate. Obesity in adults: Dietary therapy URL: http://www.uptodate.com/contents/dietary-therapy-for-obesity?view [accessed 2019-01-06] [WebCite Cache ID 78DcfyNRm]
8. Golan M, Fainaru M, Weizman A. Role of behaviour modification in the treatment of childhood obesity with the parents as the exclusive agents of change. Int J Obes Relat Metab Disord 1998 Dec;22(12):1217-1224. [Medline: 9877257]
9. Kaukua J, Pekkarinen T, Sane T, Mustajoki P. Health-related quality of life in obese outpatients losing weight with very-low-energy diet and behaviour modification--a 2-y follow-up study. Int J Obes Relat Metab Disord 2003 Oct;27(10):1233-1241. [doi: 10.1038/sj.ijo.0802379] [Medline: 14513072]
10. Sjöström L, Narbro K, Sjöström CD, Karason K, Larsson B, Wedel H, Swedish Obese Subjects Study. Effects of bariatric surgery on mortality in Swedish obese subjects. N Engl J Med 2007 Aug 23;357(8):741-752. [doi: 10.1056/NEJMoa066254] [Medline: 17715408]
11. Mayo Clinic. Bariatric surgery URL: https://www.mayoclinic.org/tests-procedures/bariatric-surgery/about/pac-20394258 [accessed 2019-05-08] [WebCite Cache ID 78DcpR2eB]
12. Suen LK, Wong TK, Leung AW. Is there a place for auricular therapy in the realm of nursing? Complement Ther Nurs Midwifery 2001 Aug;7(3):132-139. [doi: 10.1054/ctnm.2001.0565] [Medline: 11855508]
13. Hsieh C, Su TJ, Fang YW, Chou PH. Effects of auricular acupressure on weight reduction and abdominal obesity in Asian young adults: a randomized controlled trial. Am J Chin Med 2011;39(3):433-440. [doi: 10.1142/S0192415X11008932] [Medline: 21598412]
14. Bradford A, Tuffrey V, Polley M. Effect of auricular acupuncture on appetite: a pilot study. Med Acupunct 2010;22(3):1-7. [doi: 10.1089/acu.2010.0751]
15. Hsieh C, Su T, Fang Y, Chou P. Efficacy of two different materials used in auricular acupressure on weight reduction and abdominal obesity. Am J Chin Med 2012;40(4):713-720. [doi: 10.1142/S0192415X1250053X] [Medline: 22809026]
16. Hsu C, Wang C, Hwang K, Lee T, Chou P, Chang H. The effect of auricular acupuncture in obese women: a randomized controlled trial. J Womens Health 2009;18(6):813-818. [Medline: 19445642]
17. Kim D, Ham O, Kang C, Jun E. Effects of auricular acupressure using Sinapsis alba seeds on obesity and self-efficacy in female college students. J Altern Complement Med 2014 Apr;20(4):258-264. [doi: 10.1089/acm.2012.0283] [Medline: 24070326]
18. Schukro R, Heiserer C, Michalek-Sauberer A, Gleiss A, Sator-Katzenschlager S. The effects of auricular electroacupuncture on obesity in female patients--a prospective randomized placebo-controlled pilot study. Complement Ther Med 2014 Feb;22(1):21-25. [doi: 10.1016/j.ctim.2013.10.002] [Medline: 24559812]
19. Weng C, Hung Y, Shyu L. A study of electrical conductance of meridian in the obese during weight reduction. Am J Chin Med 2004;32(3):417-425. [Medline: 15344425]
20. Ceng H, Nie B. Electroacupuncture, threads embedding on acupoints, integrated with auricular acupressure for simple obesity. J Clin Acupunct 2006;22(7):13-14 [FREE Full text]
21. Abdi H, Abbasi-Parizad P, Zhao B, Ghayour-Mobarhan M, Tavallaie S, Rahsepar A, et al. Effects of auricular acupuncture on anthropometric, lipid profile, inflammatory, and immunologic markers: a randomized controlled trial study. J Altern Complement Med 2012 Jul;18(7):668-677. [doi: 10.1089/acm.2011.0244] [Medline: 22788576]
22. Yeh M, Chu N, Hsu M, Hsu C, Chung Y. Acupoint stimulation on weight reduction for obesity: a randomized sham-controlled study. West J Nurs Res 2015 Dec;37(12):1517-1530. [doi: 10.1177/0193945914548707] [Medline: 25183702]
23. Chien L, Chen F, Hu H, Liu C. Correlation of electrical conductance in meridian and autonomic nervous activity after auricular acupressure in middle-aged women. J Altern Complement Med 2014 Aug;20(8):635-641. [doi: 10.1089/acm.2012.0900] [Medline: 24865945]

24. Tang L. Auricular seed pressing for simple obesity [Chinese]. Health Newsletter 2007;12(7):2 [FREE Full text]
25. Yeh C, Yeh S. Effects of ear points' pressing on parameters related to obesity in non-obese healthy and obese volunteers. J Altern Complement Med 2008 Apr;14(3):309-314. [doi: 10.1089/acm.2007.0678] [Medline: 18377231]
26. Eisenhauer C, Hageman P, Rowland S, Becker B, Barnason S, Pullen C. Acceptability of mHealth technology for self-monitoring eating and activity among rural men. Public Health Nurs 2016;34(2):138-146. [Medline: 27757986]
27. Simons D, De Bourdeaudhuij I, Clarys P, De Cocker K, Vandelanotte C, Deforche B. A smartphone app to promote an active lifestyle in lower-educated working young adults: development, usability, acceptability, and feasibility study. JMIR Mhealth Uhealth 2018 Feb 20;6(2):e44. [doi: 10.2196/mhealth.8287]
28. Hides L, Dingle G, Quinn C, Stoyanov SR, Zelenko O, Tjondronegoro D, et al. Efficacy and outcomes of a music-based emotion regulation mobile app in distressed young people: randomized controlled trial. JMIR Mhealth Uhealth 2019 Jan 16;7(1):e11482. [doi: 10.2196/11482]
29. Abbenhardt C, McTiernan A, Alfano CM, Wener MH, Campbell KL, Duggan C, et al. Effects of individual and combined dietary weight loss and exercise interventions in postmenopausal women on adiponectin and leptin levels. J Intern Med 2013 Aug;274(2):163-175 [FREE Full text] [doi: 10.1111/joim.12062] [Medline: 23432360]
30. Kim D, Ham OK, Kang C, Jun E. Effects of auricular acupressure using Sinapsis alba seeds on obesity and self-efficacy in female college students. J Altern Complement Med 2014 Apr;20(4):258-264. [doi: 10.1089/acm.2012.0283] [Medline: 24070326]
31. Whitehead AL, Julious SA, Cooper CL, Campbell MJ. Estimating the sample size for a pilot randomised trial to minimise the overall trial sample size for the external pilot and main trial for a continuous outcome variable. Stat Methods Med Res 2016 Jun;25(3):1057-1073 [FREE Full text] [doi: 10.1177/0962280215588241] [Medline: 26092476]
32. Nomenclature and location of auricular points (GB/T 13734-2008). Beijing: China Zhijian Publishing House; 2008.
33. Feng C, Bai X, Du Y. Chinese auricular therapy. Beijing: Scientific and Technical Documents Publishing House; 1994.
34. Oleson T. Auriculotherapy Manual. 4th edition. Edinburgh: Churchill Livingstone Elsevier; 2014.
35. Abbate S. Chinese Auricular Acupuncture. Boca Raton: CRC Press; 2004.
36. CARBO: ACP Medical Supplies Inc. Ear Seeds (Processed Vaccaria seeds on surgical tape, ES-01) URL: https://www.carbo.ca/store/do/category/Ear+Seeds.html [accessed 2019-05-08] [WebCite Cache ID 78DdluoF8]
37. Tens Plus. Pointer Plus (EC Certificate Registration Number: DD600751270001) URL: http://www.tensplus.com/pointerplus.asp [accessed 2019-05-08] [WebCite Cache ID 78DdzG9Z8]
38. Tanita Corporation of American, Inc. Illinois, USA Body Composition Monitor (Model: BC-545N) URL: https://tanita.eu/bc-545n [accessed 2019-05-08] [WebCite Cache ID 78DeFmYQS]
39. Wikipedia. 2019 Jan 05. Waist-hip ratio measurement URL: https://en.wikipedia.org/wiki/Waist%E2%80%93hip_ratio [accessed 2019-05-08] [WebCite Cache ID 78DeMDr5E]
40. BioVendor. Leptin Human ELISA, Clinical Range (Cat. No. RD191001100) URL: https://www.biovendor.com/leptin-human-elisa-clinical-range [accessed 2019-05-08] [WebCite Cache ID 78DeWp8cR]
41. Antibody and Immunoassay Service, The University of Hong Kong. Human adiponectin, rapid ELISA kit (Cat. No. 31012) URL: https://www.antibody.hku.hk/ELISA.php [accessed 2019-05-08] [WebCite Cache ID 78Dewi2SP]
42. Rock CL, Flatt SW, Barkai H, Pakiz B, Heath DD. Walnut consumption in a weight reduction intervention: effects on body weight, biological measures, blood pressure and satiety. Nutr J 2017 Dec 04;16(1):76 [FREE Full text] [doi: 10.1186/s12937-017-0304-z] [Medline: 29202751]
43. Bell M, Horton N, Dhillon H, Bray VJ, Vardy J. Using generalized estimating equations and extensions in randomized trials with missing longitudinal patient reported outcome data. Psychooncology 2018 Sep;27(9):2125-2131. [doi: 10.1002/pon.4777] [Medline: 29802657]
44. Peate I. Male obesity: a gender-specific approach to nurse management. Br J Nurs 2005;14(3):134-138. [doi: 10.12968/bjon.2005.14.3.17517] [Medline: 15788931]
45. Mayo Clinic. Belly fat in men: Why weight loss matters URL: https://www.mayoclinic.org/healthy-lifestyle/mens-health/in-depth/belly-fat/art-20045685 [accessed 2019-05-08] [WebCite Cache ID 78Dez4JLt]
46. Goh V, Hart W. Association of general and abdominal obesity with age, endocrine and metabolic factors in Asian men. Aging Male 2016;19(1):27-33. [doi: 10.3109/13685538.2015.1088825] [Medline: 26444311]
47. Ito H, Yamada O, Kira Y, Tanaka T, Matsuoka R. The effects of auricular acupuncture on weight reduction and feeding-related cytokines: a pilot study. BMJ Open Gastroenterol 2015 Dec 31;2(1):e000013. [doi: 10.1136/bmjgast-2014-000013]
48. Portney L, Watkins M. Foundations of clinical research: applications to practice (3rd ed). Upper Saddle River, NJ: Pearson/Prentice Hall; 2009.
49. Semper H, Povey R, Clark-Carter D. A systematic review of the effectiveness of smartphone applications that encourage dietary self-regulatory strategies for weight loss in overweight and obese adults. Obes Rev 2016 Dec;17(9):895-906. [doi: 10.1111/obr.12428] [Medline: 27192162]
50. Granado-Font E, Flores-Mateo G, Sorlí-Aguilar M, Montaña-Carreras X, Ferre-Grau C, Barrera-Uriarte M, et al. Effectiveness of a smartphone application and wearable device for weight loss in overweight or obese primary care patients: protocol for a randomised controlled trial. BMC Public Health 2015 Jun 04;15:531 [FREE Full text] [doi: 10.1186/s12889-015-1845-8] [Medline: 26041131]

51. Pellegrini C, Duncan J, Moller A, Buscemi J, Sularz A, DeMott A, et al. A smartphone-supported weight loss program: design of the ENGAGED randomized controlled trial. BMC Public Health 2012 Nov 30;12:1041 [FREE Full text] [doi: 10.1186/1471-2458-12-1041] [Medline: 23194256]
52. Patel M, Hopkins C, Brooks T, Bennett G. Comparing self-monitoring strategies for weight loss in a smartphone app: randomized controlled trial. JMIR Mhealth Uhealth. . 3081 2019;7(2):6851. [doi: 10.2196/12209.Medline:]
53. Wang J, Cai C, Padhye N, Orlander P, Zare M. A behavioral lifestyle intervention enhanced with multiple-behavior self-monitoring using mobile and connected tools for underserved individuals with type 2 diabetes and comorbid overweight or obesity: pilot comparative effectiveness trial. JMIR Mhealth Uhealth 2018 Apr 10;6(4):e92 [FREE Full text] [doi: 10.2196/mhealth.4478] [Medline: 29636320]
54. Fenerty SD, West C, Davis SA, Kaplan SG, Feldman SR. The effect of reminder systems on patients' adherence to treatment. Patient Prefer Adherence 2012;6:127-135 [FREE Full text] [doi: 10.2147/PPA.S26314] [Medline: 22379363]
55. Deng Y, Scherer P. Adipokines as novel biomarkers and regulators of the metabolic syndrome. Ann N Y Acad Sci 2010;1212:E1-E19. [Medline: 21276002]
56. Wang P, Mariman E, Renes J, Keijer J. The secretory function of adipocytes in the physiology of white adipose tissue. J Cell Physiol 2008 Jul;216(1):3-13. [doi: 10.1002/jcp.21386] [Medline: 18264975]
57. Ma W, Huang T, Zheng Y, Wang M, Bray G, Sacks F. Weight-loss diets, adiponectin, and changes in cardiometabolic risk in the 2-year POUNDS Lost Trial. J Clin Endocrinol Metab 2016;101(6):2415-2422. [Medline: 27055193]
58. Musil F, Blaha V, Ticha A, Hyspler R, Haluzik M, Lesna J, et al. Effects of body weight reduction on plasma leptin and adiponectin/leptin ratio in obese patients with type 1 diabetes mellitus. Physiol Res 2015;64(2):221-228 [FREE Full text] [Medline: 25317681]
59. Sarlak H, Akhan M, Cakar M, Kurt O, Arslan E, Balta S. A larger weight reduction is necessary to elicit an increase in adiponectin and a decrease in leptin levels. J Intern Med 2013 Dec;274(6):614 [FREE Full text] [doi: 10.1111/joim.12122] [Medline: 23957881]
60. Trayhurn P, Bing C, Wood IS. Adipose tissue and adipokines--energy regulation from the human perspective. J Nutr 2006 Dec;136(7 Suppl):1935S-1939S [FREE Full text] [doi: 10.1093/jn/136.7.1935S] [Medline: 16772463]
61. Suen L, Yeh C, Kwan J, Lee PH, Yeung GS, Wong EC, et al. Association of auricular reflective points and status of type 2 diabetes mellitus: a matched case-control study. Evid Based Complement Alternat Med 2015;2015:981563 [FREE Full text] [doi: 10.1155/2015/981563] [Medline: 26089953]
62. Suen L, Yeh C, Yeung S. Medicines. Association between auricular signals and the risk factors of metabolic syndrome URL: http://www.mdpi.com/2305-6320/4/3/45_doi:10.3390/medicines4030045 [accessed 2019-05-08] [WebCite Cache ID 78Df9DDIH]

Abbreviations

AA: auricular acupressure
BMI: body mass index
CV: coefficient of variability
iOS: iPhone operating system
WHO: World Health Organization

Edited by G Eysenbach; submitted 15.04.19; peer-reviewed by A Leung, V Lopez, R Ciptaningtyas; comments to author 27.04.19; revised version received 29.04.19; accepted 29.04.19; published 29.05.19.

Please cite as:
Suen L, Wang W, Cheng KKY, Chua MCH, Yeung JWF, Koh WK, Yeung SKW, Ho JYS
Self-Administered Auricular Acupressure Integrated With a Smartphone App for Weight Reduction: Randomized Feasibility Trial
JMIR Mhealth Uhealth 2019;7(5):e14386
URL: http://mhealth.jmir.org/2019/5/e14386/
doi:10.2196/14386
PMID:31144666

Original Paper

A Two-Way Interactive Text Messaging Application for Low-Income Patients with Chronic Medical Conditions: Design-Thinking Development Approach

Monika Marko-Holguin[1], BS, MS; Stephanie Luz Cordel[1,2], BS, MSW; Benjamin William Van Voorhees[1], MD, MPH; Joshua Fogel[3], PhD; Emily Sykes[1], BA, MA; Marian Fitzgibbon[1], PhD; Anne Elizabeth Glassgow[1], PhD

[1]Department of Pediatrics, College of Medicine, University of Illinois at Chicago, Chicago, IL, United States

[2]All Voices Consulting, LLC, Phoenix, AZ, United States

[3]Department of Business Management, Brooklyn College, Brooklyn, NY, United States

Corresponding Author:
Monika Marko-Holguin, BS, MS
Department of Pediatrics
College of Medicine
University of Illinois at Chicago
840 S Wood Street
Chicago, IL, 60612
United States
Phone: 1 3124131165
Email: mmarko@uic.edu

Abstract

Background: Two-way interactive text messaging between patient and community health workers (CHWs) through mobile phone SMS (short message service) text messaging is a form of digital health that can potentially enhance patient engagement in young adults and families that have a child with chronic medical conditions such as diabetes mellitus, sickle cell disease, and asthma. These patients have complex needs, and a user-centered way can be useful for designing a tool to address their needs.

Objective: The aim of this study was to utilize the user-centered approach of design thinking to develop a two-way interactive communication SMS text messaging tool for communication between patients or caregivers and CHWs.

Methods: We applied a design thinking methodology for development of the SMS text messaging tool. We collected qualitative data from 127 patients/caregivers and 13 CHWs, health care professionals, and experts. In total, 4 iterative phases were used to design the final prototype.

Results: The design thinking process led to the final SMS text messaging tool that was transformed from a one-dimensional, template-driven prototype (phases 1 and 2) into a dynamic, interactive, and individually tailored tool (phases 3 and 4). The individualized components consider social factors that influence patients' ability to engage such as transportation issues and appointment reminders. SMS text messaging components also include operational factors to support staff such as patient contact lists, SMS text messaging templates, and technology chat support.

Conclusions: Design thinking can develop a tool to meet the engagement needs of patients with complex health care needs and be user-friendly for health care staff.

(JMIR Mhealth Uhealth 2019;7(5):e11833) doi:10.2196/11833

KEYWORDS

mobile applications; telemedicine; patient participation; patient acceptance of health care; delivery of health care; family health; community health services; healthcare disparities; information technology; cell phone use

Introduction

Patients with Chronic Medical Conditions

Low-income Medicaid beneficiaries, in particular children and young adults with chronic medical conditions (CMCs), are at greater risk for poor health outcomes as compared with the general population [1-4]. Children and young adults with CMCs typically have one or more chronic health conditions, have extensive health service needs, and often have functional limitations [5,6]. Managing their condition requires strict adherence to their treatment plan, which can be impacted by psychosocial factors including economic barriers. Children and youth with CMCs have a greater reliance on multiple caregivers in their lives and must face a transition to self-management as they grow older and more independent, which can be difficult for many because of moving from a familiar pediatric setting to an unfamiliar adult one [7]. Factors such as fragmented health care services, lack of access to mental health, and other social support services exacerbate CMCs and lead to high utilization of health care resources [8] and overall poor engagement with health care services [9]. Children with CMCs account for 15% to 33% of the total pediatric health care utilization and costs, which is estimated to be about US $50 to US $110 billion annually [5,6,10-12].

Patient Engagement Using Short Message Service Technology

Patients actively engaged in their health and health care have better health outcomes and lower health care costs [13,14]. For patients with CMCs, high engagement supports greater adherence to their often complex treatment plan, which can prevent additional hospital visits and lower their risk for increased morbidity and mortality [7]. In recent years, digital health, namely electronic health and mobile health (mHealth), has increased in usage and has the potential to provide an enhanced approach to patient engagement in a cost-efficient way through reminders and patient-provider communication supporting treatment adherence and overall health [15,16]. Text messaging through mobile phone short message service (SMS) is a form of digital health that might be used in a low-cost manner to enhance engagement of those with CMCs who have resource-intensive health needs. Currently, texting technology is widely used by those from many cultures, socioeconomic backgrounds, and age groups [17]. Texting technology can reach people in an effortless and low-cost manner. Many low-income families do not have computers, but at least 96% of them have cell phones and 81% of them have unlimited text plans. Texting technology offers an opportunity to engage with these patients in a way that connects to their everyday life [18]. Compared with app software, texting technology does not require a data plan nor the latest cell phone software, which may be difficult for low-income patients to obtain [16]. SMS text messaging has the highest reach with 98% read rates compared with any communication form. Some studies have reported successful implementation of SMS text messaging as a health intervention tool [17,19-21]. Other studies have reported texting as an adjunct tool to enhance health interventions such as in reminders for medication adherence [22-24] and mental health treatment [25] as well as an add-on strategy to reduce the number of severe mental illness episodes in adults [26] and suicide prevention [27]. Texting technology can be characterized as human-based (human-to-human interaction), computer-based (computer-to-human interaction), or hybrid texting (computer sends out bulk messages to participants and a human addresses the replies in a tailored manner) [28].

Coordinated Health Care for Complex Kids Health Initiative

In an effort to address the health needs of children and young adults with CMCs, the Coordinated Health Care for Complex Kids (CHECK) program began in 2014 [8]. As part of the Medicare and Medicaid Innovation initiative from the Centers of Medicare and Medicaid Services, CHECK focuses on identifying pathways for reducing health care costs, school absenteeism, and promoting engagement of low-income children and young adults with CMCs and their families with their health and care. Participants, ranged from newborn to 25 years, are enrolled in Medicaid (fee-for-service or managed care) and have a diagnosis of asthma, diabetes mellitus, sickle cell disease, or prematurity. A majority of CHECK participants reside in zip codes identified as having a high community need index, a designation based on income, culture, education, health insurance, and housing [29].

To increase engagement and reduce health care costs for low-income patients with CMCs, we developed a hybrid (computer interacting with humans and vice versa) two-way interactive (patient to staff and vice versa) SMS text messaging technology tool in the context of the larger CHECK health initiative utilizing community health workers (CHWs). The combination of mHealth and CHWs aims to address high health care costs by increasing patient engagement and self-efficacy with their health and health care and reducing the number of preventable emergency room visits and hospitalizations [8,30-33]. It also offers an opportunity to identify barriers to patient engagement and support patients in accessing health-related resources that can enhance their overall well-being and ability to tend to their health care needs [34-36]. As patient engagement is a central factor in health care outcomes, it is vital that any tool used to connect with patients be developed with the patient experience in mind [37-40]. Staff operational needs also need to be considered for maximum utilization [39-41]. To ensure that the tool was user-friendly and effective, we utilized design thinking theoretical guidelines [42] for our approach in designing this texting app. The design thinking model is a user-informed design process that takes into account a large amount of user feedback for a bottom-up approach. The focus is on pragmatic, context-dependent, need-based solutions that improve the likelihood of achieving the full potential of an outcome or event. In doing so, design thinking affords a stable method of alternative solutions to redefine the outcome and seamlessly improve upon previous iterations [42].

The purpose of this study was to describe the development of a hybrid texting app to promote patient engagement among low-income Medicaid children and young adults with CMCs. We utilized design thinking, an iterative development process that incorporates feedback at multiple stages, to ensure an

effective, relevant prototype. This study has provided a descriptive overview of how to tailor a digital health tool such as SMS text messaging to patients with CMC needs in the context of a CHW-based initiative.

Methods

Design Thinking Process

We utilized the design thinking process to guide the development of the SMS text messaging tool [42]. Design thinking is an organic, bottom-up process, which allows the audience to guide and help design the product, hence increasing the likelihood of a successful outcome. There are 5 steps within the design thinking process that can be repeated as the developer gains more information through prototyping and testing [42].

Step 1 focuses on understanding and empathizing with the audience in a holistic way. This allows the designer to immerse with the audience while trying to objectively assess their needs. The designer can then create a product that greatly benefits the users and encourages their commitment beyond product development into successful implementation and ongoing use. Step 2 is to identify the need and define the need-based outcome through a shared viewpoint across multiple stakeholders. To achieve this, the design thinking approach accepts that human behavior is context-dependent and recognizes the importance of developing solutions that match the user's self-efficacy [43]. Step 3 uses ideation through a multidisciplinary team approach bringing together professional and *lived-experience* experts who are end-users who have real-life experience with the problem that is being targeted by the proposed solution/product [27] to brainstorm on innovative design solutions that meet the need-based outcome. Step 4 is to create one or more prototypes representing the proposed solution. Step 5 is to test the prototype in an attempt to gather end-user feedback for further molding of the prototype toward its full potential to meet the need-based outcome.

The development of the SMS text messaging tool was spread out across 4 phases over a period of 2 years, starting in March 2015 (see Table 1 for details). The design thinking process was adopted for each phase of the development, including a variety of approaches described below, to collect the information necessary to define the users' needs [44].

Design thinking encourages the development of rapid and basic prototypes because this permits a more efficient process toward gathering prompt feedback, quickly analyzing it, and only investing in the prototype that would best match the solution under study [42]. We adopted this process across all 4 phases of our tool development and started with an existing texting app (Figure 1).

Table 1. Phases of the design thinking process.

Phases and activities	Design thinking steps
Phase 1 (March to May 2015): Understand the audience, context, and need requirements	
Identify our audience as several end-user[a] groups comprised patients, parents, health care professionals, clinicians, analysts, software engineers, and research designers; create a codesign team of end-users composed of *lived experience*[b] experts, professionals, and researchers; literature review on SMS[c] use in health care to understand previous findings about end-users/target population; and literature review on the US population's broad use of technology-based tools	Understand
Identify targeted health conditions to determine the set of need requirements with a technology-based tool; and consult with *lived experience* experts, health care, and SMS expert researchers to confirm the need requirements for useful technology-based tools, such as SMS, within the context of health care	Define
Phase 2 (June to July/August 2015): Ideate a context-based solution, build a prototype, and test	
Create an SMS use case[d] tailored to CHECK[e] patients with CMCs[f] and build prototype mockups[g]; ideate with *lived experience* and other expert researchers in a roundtable discussion to evaluate the use case mockups; and build prototype 1 using the outcome results from phase 1 and feedback from the roundtable discussion	Ideate/Prototype I
Test prototype 1 via a pilot study with CHECK patients to gather patient end-users' need requirements feedback; and consult with SMS experts and licensed health care professionals on the outcome results	Test
Phase 3 (August to November 2015): Redesign and test	
Modify the SMS prototype solution according to feedback on prototype 1 received by end-users during phase 2 testing; and develop SMS prototype 2 that incorporates the feedback on improving patient end-user engagement from phase 2	Ideate/Prototype II
Test prototype 2 via a pilot study with CHECK patient end-users to measure their engagement with the tool; and test prototype 2 with a team of health care professional end-users to gather their feedback	Test
Phase 4 (February 2016 to March 2017): Full reiteration	
Expand the codesign team to include software engineers and health data professionals; and conduct individual interviews with CHW[h] staff to gather need requirements for an SMS tool in the context of health care delivery by CHWs	Understand
Create prototype 3 to program staff need requirements	Ideate/Prototype III
Conduct 6 individual interviews to gather requirements from staff as end-users	Test
Create prototype 4, integrating the patient and program staff need requirements	Ideate/Prototype IV
Gather feedback through an anonymous survey and 4 individual interviews with staff	Test

[a]Refers to a person who uses or is intended to use the final product.

[b]Refers to end-users who have real-life experience with the problem that is being targeted by the proposed solution/product.

[c]SMS: short message service.

[d]A use case acts as a software-modeling technique that defines the features to be implemented and the resolution of any errors that may be encountered [45].

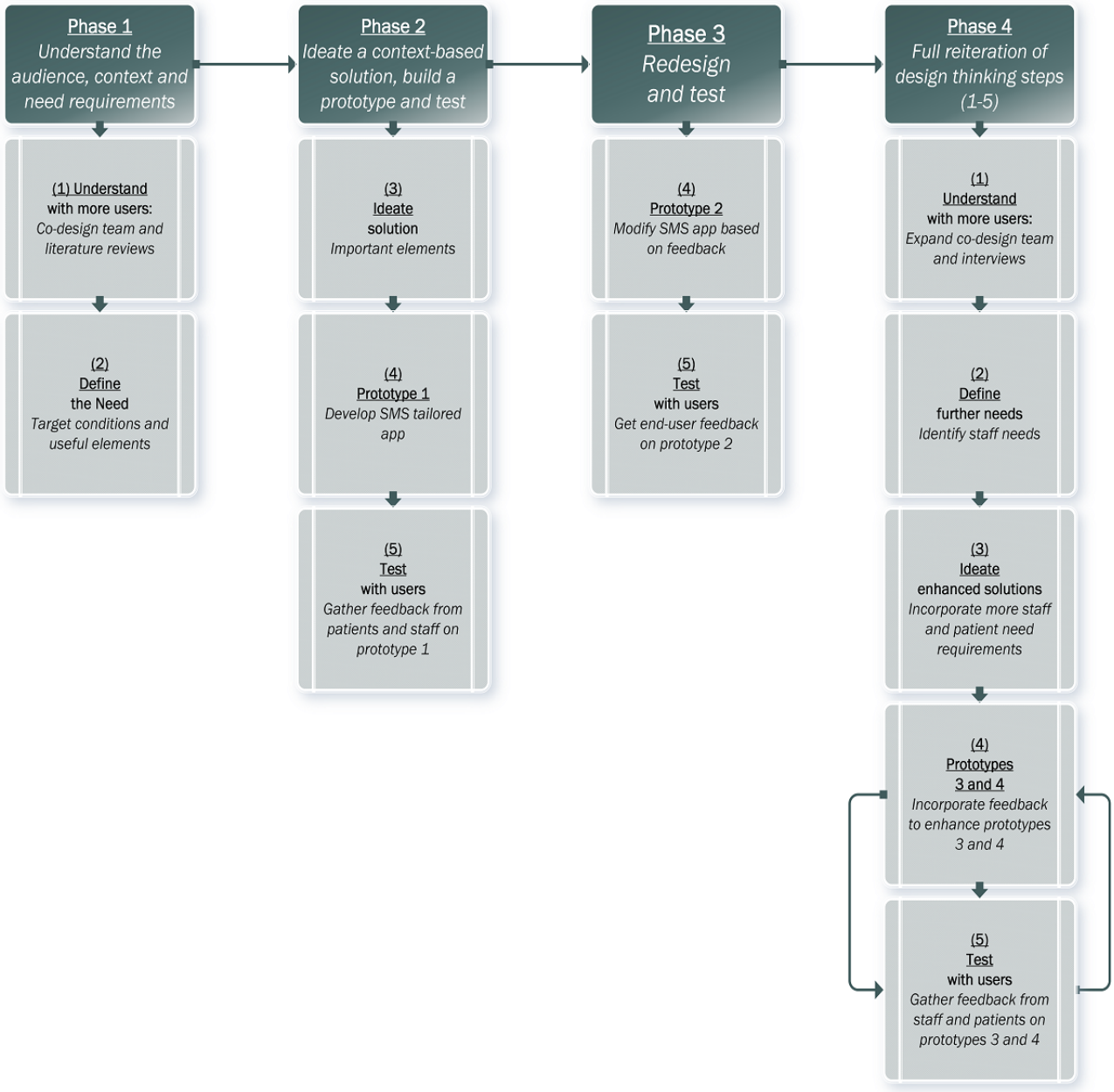
[e]CHECK: Coordinated Health Care for Complex Kids.

[f]CMC: Chronic medical condition.

[g]Refers to a model or replica of a software model used to provide a visual representation of the model's app in real life.

[h]CHW: community health worker.

Figure 1. Project phases and design thinking process.



Phase 1: Understand the Audience, Context, and Need Requirements

Phase 1, which started in March 2015, was guided by the principles of understand and define in the design thinking process. The objectives of this phase were to identify the end-users; understand their needs for a technology-based solution within the health care context; discover end-users' need-based requirements; and define the technology-based solution that addresses the needs within the context of health care. We operationalized these objectives into a number of activities, which included instituting a codesign team of end-users, conducting several literature reviews to confirm use of texting in the context of health care for the Medicaid population living with CMCs, as well as determine a set of need requirements through direct feedback from the various groups of end-users.

Understanding the End-Users

We identified our audience (end-users) to be comprised several end-user groups (Table 2). The pediatric patients and their parents who formed the Parent Advisory Board (PAB) and their community advocates/leaders, which formed the Community Advisory Board (CAB), were identified as the *lived experience* expert end-users, given their personal experience with living or caring for someone with CMCs [46]. The other end-user groups included the CHECK program staff comprised health care professionals, such as CHWs, licensed clinicians, and health data analysts, and the design team comprised a social scientist, software engineers, a texting research expert, and research designers.

Table 2. Summary of end-user groups by phase of the design thinking process.

End-user group	Phase 1, N	Phase 2, N	Phase 3, N	Phase 4, N
Patients/parents	—[a]	95	52	800
Lived experience experts	5	5	—	—
Health care professionals (CHWs)[b]	—	4	3	12
Other health care professionals (operation managers/ supervisors)	—	—	—	2
Licensed clinicians	5	—	—	4
Health data analysts and scientists	—	—	—	6
Software engineers	2	—	—	3
SMS research expert	1	1	1	—
Research Designers	5	5	5	3

[a]Not applicable.

[b]CHWs: community health workers.

We assembled a codesign team of end-users composed of *lived experience* experts, licensed clinicians, software engineers, and researchers. This was informed by previous research highlighting that to address the design requirements of a health care system that aims to meet the goals of both patients and clinicians [47], one needs to have in place an integrated, multidisciplinary design team of professionals [42,48] and *lived experience* experts [27] to successfully execute the design [46].

We also conducted several literature reviews on SMS text messaging use in health care to understand previous findings about end-users/target populations that are difficult to reach (eg, patients aged under 18 years). We used published empirical evidence to inform our understanding of the patient stakeholder group's needs with their engagement in health and health care [27,46]. This was subsequently followed by a generic literature search and review to understand the US population's broad uses for technology-based tools [49].

Defining the Need Requirements

To begin developing a technology-based solution that satisfied the end-users' needs, we started by first defining the end-user's needs. The process of closely refining the definition to meet the end-user's needs within a given context (ie, health care) is an approach that has led to designing successful outcomes [42,44].

The literature findings on the patient end-users' need requirements were discussed with the *lived experience* experts, the licensed clinicians, and the research professionals. The discussion resulted in a revised and narrowed definition of the patient end-users' need requirements for better engagement in their health and care assessed by the use of any technology-based tools. A final targeted literature review was conducted to determine the role and patient use of the technology-based tools within the context of health care [49].

Phase 2: Ideate a Context-Based Solution, Build a Prototype, and Test

Phase 2, which started in June 2015, was guided by the principles outlining the steps of ideate, prototype, and test in the design thinking process. The objective of this phase was to ideate a use case of the SMS text messaging model app, to demonstrate the texting prototype to the end-users and collect feedback. In this phase, we operationalized the objectives into the activities that included the development of use case mockups that targeted the engagement of CHECK patient end-users with their health and care via a texting app, engaging in a roundtable discussion with the end-user expert groups to gather feedback for the prototype development and test the texting prototype 1 with a group of end-users.

Ideate, Develop, and Test Prototype 1

The information gathered through the literature reviews and numerous discussions during Phase 1 converged into the SMS text tool as the technology-based solution to address CHECK patient end-users' engagement within the context of health care (Table 3). The proposed solution, the texting use case mockups tailored to CHECK patients with CMCs, was presented to the end-user team comprised the *lived experience* and texting research expert, along with other health care professionals. After reading the narrative version of the texting use case templates and reviewing the SMS text messaging app mockups, the end-users engaged in a roundtable discussion with the researchers. The sample questions used during the discussion included "How frequently do you think text messages should be sent out from the provider to the patient?"; "What time of day should a text message be sent out?"; "Is the timing even important and why?"; and "Are text messages that request a response helpful or annoying to you?" The feedback from this roundtable discussion was used to inform the development of the texting prototype app.

Table 3. Summary of iterations and prototype outcomes.

Iteration	Ideation	Prototype (patients)	Prototype (staff)	Example
Prototype 1	A two-way SMS[a] app, non–Health Insurance Portability and Accountability Act (HIPAA) secure, low to no cost, no data or Wi-Fi required, live communication through preset logic to gather insight into patients' common health-related needs and language that would encourage engagement with their health and care	Explore patients' language of engagement using 3 methods of engagement: Informative SMS using encouraging and self-motivating text; informative SMS using direct text; generic SMS using neutral text. Explore patients' needs through 5 common preidentified themes: (1) General information (2) Goal tracking (3) General support (4) Social reinforcement (5) Humor/religious	—[b]	Goal tracking: Do you think you get enough sleep at night? Reply 1=Yes or 2=No; Preset logic for *No* choice: Sleep is important. It is recommended that the average adult gets 7 to 9 hours of sleep each night. Try to make sleep a priority! Preset logic for *Yes* choice: Great! It is recommended that the average adult gets 7 to 9 hours of sleep each night. Keep it up.
Prototype 2	A two-way SMS non–HIPAA secure, low to no cost, no data or Wi-Fi required, live communication through preset logic on 3 themes identified as the patients' top needs and delayed interactive support by CHW[c] staff	Explore patients' level of engagement using text (English and Spanish) to meet their 3 top needs in a delayed two-way interaction with staff: (1) Transportation coordination (2) Social support service delivery (3) Appointment scheduling and reminders	Explore CHWs' operational needs in addressing and meeting their patients' top 3 needs in a timely fashion and executed per protocol	Transportation Scheduling: Dear Ms. Doe, the transportation service for your appointment has been scheduled. Your transportation provider is [Fake Transport]. It will pick you up at the address we have on file, on 01/01/01 at 01:00 AM/PM. Please reply *1 or Y* to confirm and keep this service; please reply *2 or N* to cancel it. Thank you, Your CHW; Preset logic for *Yes*: Thank you Ms. Doe; you are all set! We will send you a reminder closer to the date. Your CHW; Preset logic for *No*: Thank you, I will call you shortly to reschedule. Your CHW
Prototype 3	A two-way SMS HIPAA secure, low to no cost, no data or Wi-Fi required, real live communication, and interactive support by CHW staff	To meet patients' 3 top needs using a live, two-way interaction with staff: (1) Transportation coordination (2) Social support service delivery (3) Appointment scheduling and reminders	To meet on-the-field staff operational needs for live two-way interaction (English and Spanish) with patients, using an editable and smart library of SMS templates to address patients' top 3 needs; A self-maintaining secure password reset system for staff	See Figure 2
Prototype 4	A two-way SMS HIPAA secure, low to no cost, no data or Wi-Fi required, real live communication, and interactive support by CHW staff	To meet patients' 3 top needs using a live, two-way interaction with staff: (1) Transportation coordination (2) Social support service delivery (3) Appointment scheduling and reminders; An opt-out and HIPAA disclaimer feature for patients; A self-managed stop and start enrollment feature for patients	To meet patient-centered care support operational needs for live two-way interaction (English and Spanish): CHWs to patients (1:1 and 1:many) and vice versa; staff-to-staff and global texting. An automated grouping feature (based on any combination of patients' chronic condition, assigned risk, and zip code); An automated scheduling and reminder feature; A tech-support group chat feature triaging staff issues in real time; A self-learning training manual with 24-hour Web access	See Figures 3 and 4

[a]SMS: short message service.

[b]—Not applicable.

[c]CHWs: community health workers.

In July 2015, we developed the first prototype of the texting app. The goal of prototype 1 was to examine the health-related needs of Medicaid patients with CMCs around 5 common themes, as well as to evaluate patient end-users' engagement via the use of texting. The SMS text messaging contained language based on the 5 common themes resulting from phase 1 consultations (Table 3). The engagement archetypes were adopted from evidence-based techniques used in motivational interviewing [50] and behavioral activation [51].

The 5 themes were framed around the 3 engagement approaches in the following manner: (1) SMS text messages containing motivational language (eg, "Feeling stressed? Exercise is a great way to help! Grab a friend and plan a walk or bike ride today! Reply 1=great idea! Or 2=no thanks"); (2) SMS text messages containing direct language (eg, "Your doctor wants to know if you have problems or side effects from your medicine. Give him or her a call if that happens. Reply 1=Thanks, I'm all set or 2=I need help with that"), and (3) SMS text messages containing neutral/generic language (eg, "Are you eating mindfully? Look, smell, TASTE your food. Savor each bite. Take your time to allow your MIND to take in what you are eating and you will eat less.") A total of 50 different texting templates were generated for testing with patient end-users.

In August 2015, we launched the testing of the texting prototype 1. The prototype was tested with a group of 127 enrolled CHECK patient end-users randomly selected and comprising 35% of those who met the Medicaid eligibility criteria and were automatically enrolled into CHECK. The final sample size was 95 patients. Those excluded were 14 patients who asked to stop receiving text messages by replying *Stop*, *I don't want to participate*, or *No thanks* and 18 patients with nonworking phone numbers. The sample was representative of CHECK premature infant patients and children and young adult patients with asthma, diabetes, and sickle cell disease. Each patient participant was randomly assigned to 1 of the 3 patient engagement archetypes and received approximately 3 SMS text messages (each up to 140 characters long) per week, for a period of 4 weeks, for a total of approximately 12 SMS text messages. Data were collected on patient end-user's engagement level, as measured by the number of SMS replies received per end-user and patient's health-related needs as measured by the number of SMS replies per end-user on each of the 5 themes.

Figure 2. Tool illustration built on an open-source short message service platform (Heroku/Twilio). SMS: short message service.

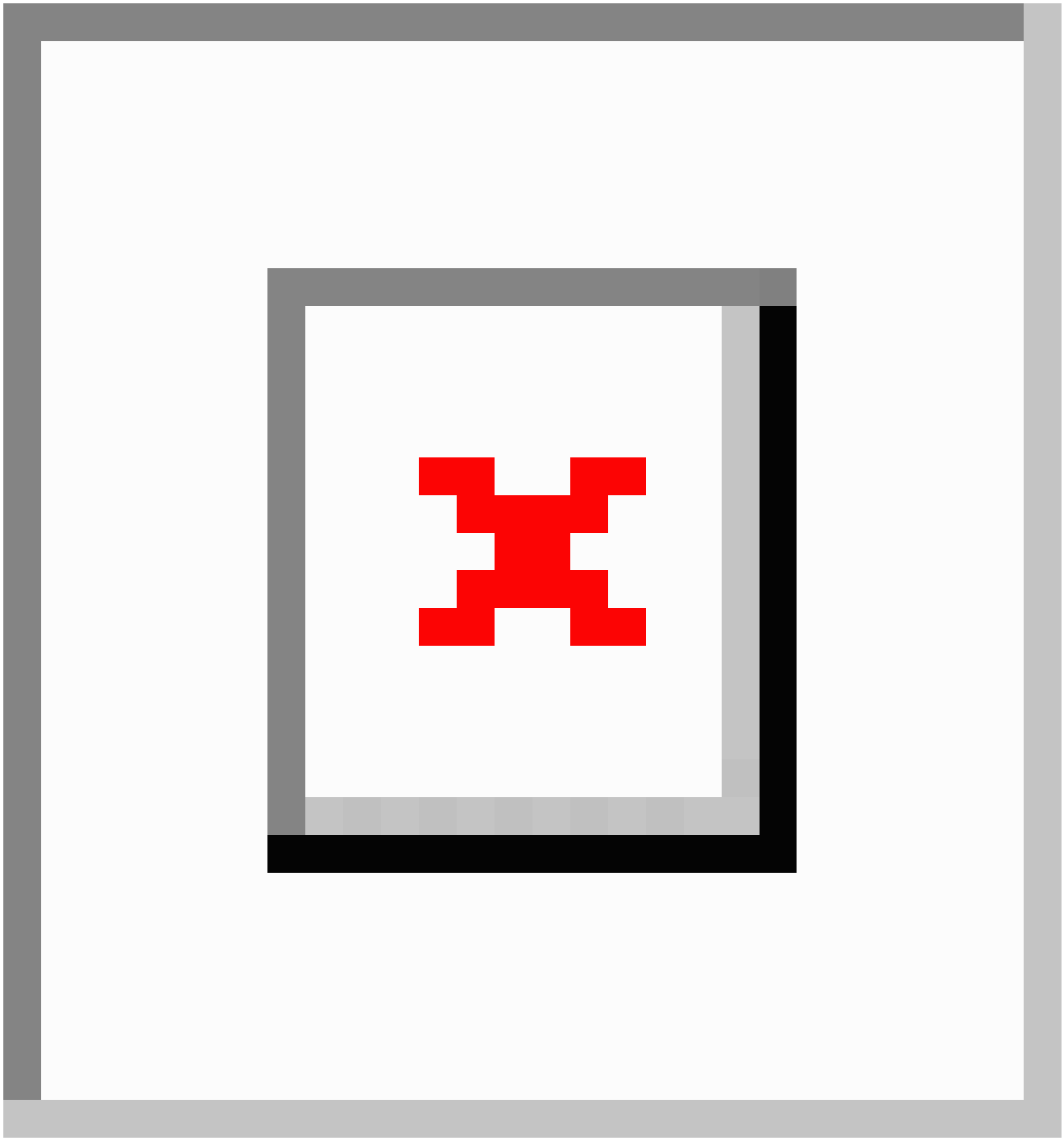


Figure 3. Technology support group chat feature.

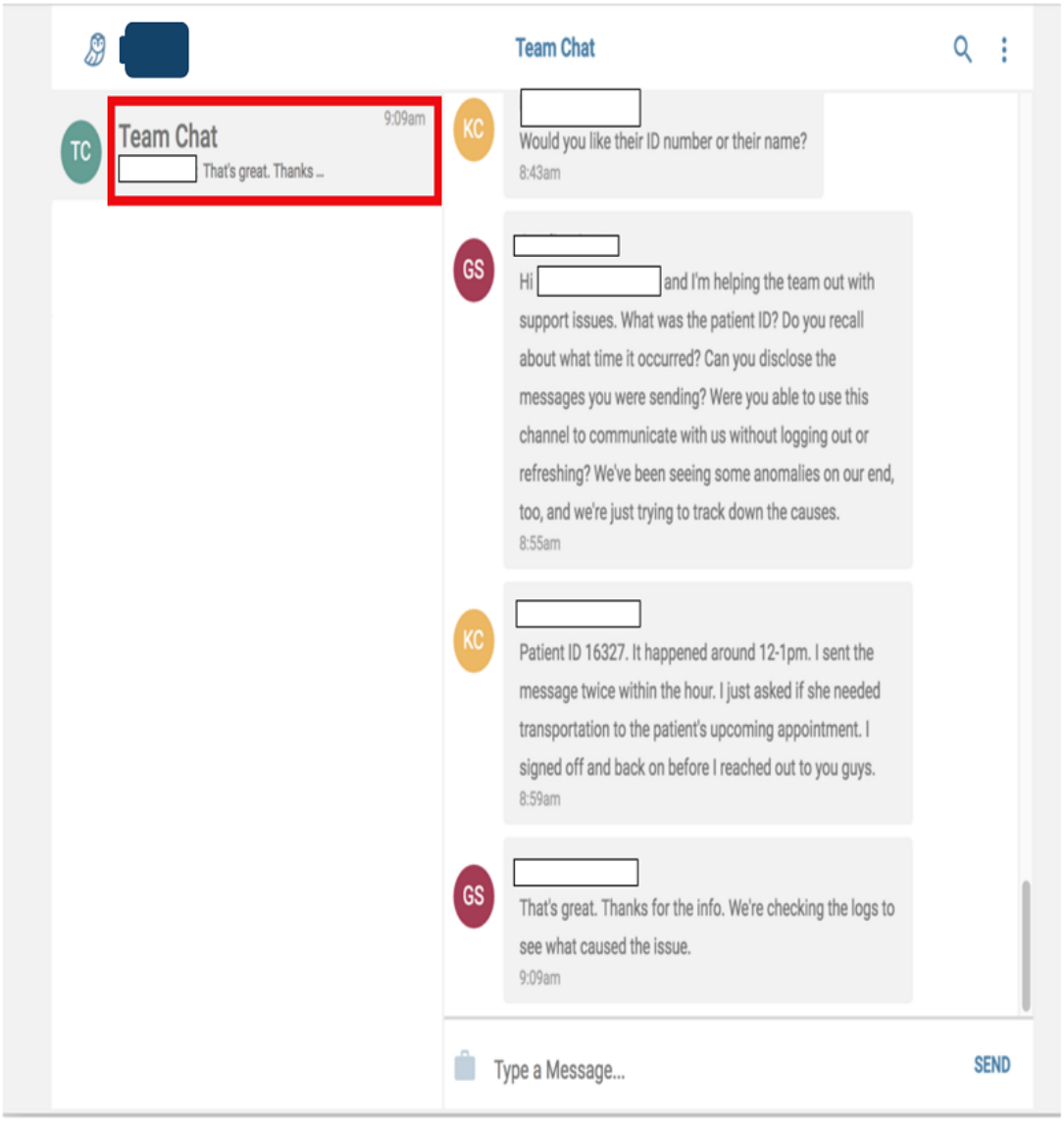
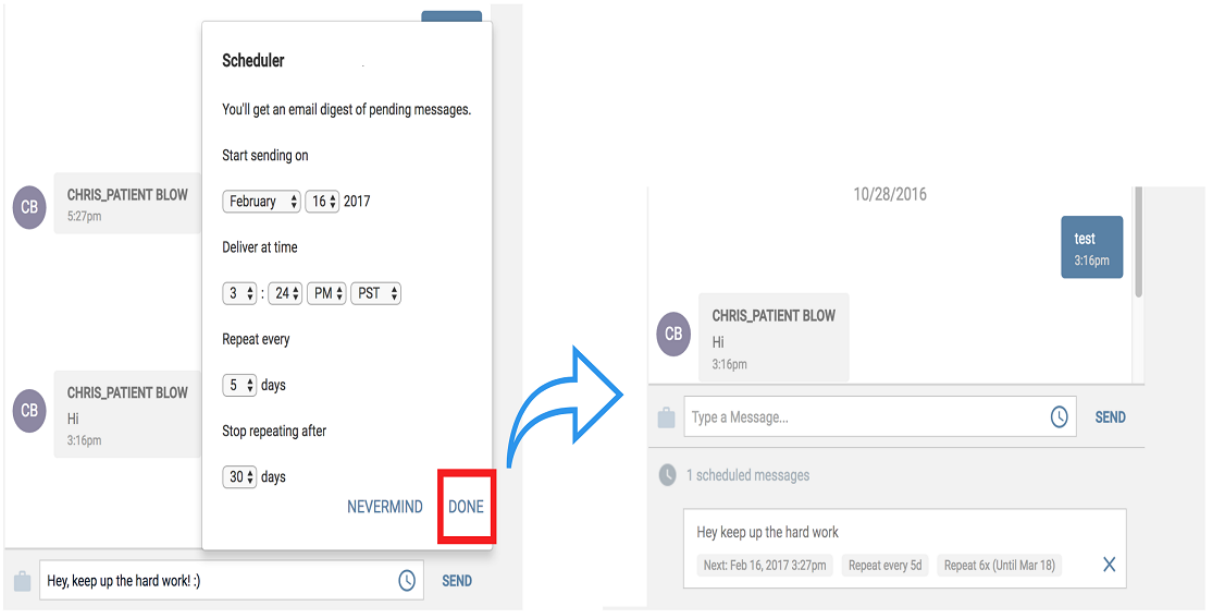


Figure 4. Automated scheduling and reminder feature.



Phase 3: Redesign Prototype and Test

Phase 3, which started in August 2015, was guided by the principles of redefining the need requirements, prototyping, and testing steps of the design thinking process. The objective of this phase was to ideate, develop, and test iterative versions of the texting app based on newly discovered need requirements from the health care professional (CHW) end-user group. In this phase, we operationalized the objective into activities that included developing the SMS app prototype 2, targeting need requirements from 2 different end-user groups, testing the prototype, and collecting feedback from both end-user groups.

Redefining the Need Requirements

The findings from prototype 1 testing were discussed with the CHWs in the form of semistructured interviews (n=4), each lasting approximately 60 min [52]. As the patient's health care navigators, the CHWs' feedback exposed a set of operational considerations in addressing their patients' specific needs in a timely fashion. Additional discussions were held with the texting research expert and the research design team.

Ideate, Develop, and Test Prototype 2

The proposed solution for prototype 2 was informed by the patient data collected and analyzed at the end of the second phase as well as the staff feedback. The goal of the second prototype was to address CHW end-user's need requirements for better patient engagement with the texting tool by targeting specific patient needs in a timely manner. This approach aligned with the self-efficacy principle that emphasizes that individuals' engagement in certain activities is strongly correlated with their level of efficacy [53] as measured by the outcome of the activity [54]. Prototype 2 included (1) preloaded SMS text messaging reply templates narrowed down to 3 main themes identified by the CHWs as most engaging for patients and (2) preloaded SMS text messaging reply templates that were reconstructed to meet the timing requirements in addressing the patient's needs.

The second prototype was piloted on 52 CHECK enrolled patient end-users and their respective CHWs (n=3) in October 2015. All 4 CHW staff were instructed to use the tool for a period of 8 weeks with their respective patients needing (1) transportation coordination, (2) social service referral delivery, and (3) appointment scheduling and reminder (all 3 themes identified as most engaging for patients during phase 2). The CHWs were trained to use the Web-based version of the texting prototype tool to send and reply to patient SMS text messages using preloaded SMS text messaging templates centered on these 3 services (Table 3). Patients received at least 1 SMS text message (up to 140 characters long) matching any of the 3 service needs. None of the 52 participants opted out during the testing period. Data were collected on the patient end-user's engagement level, as measured by the number of SMS replies received per end-user on each of the 3 themes.

Staff interviews conducted during this phase (n=9) were unstructured and were 40 min each [47]. We used unstructured creative interviewing [48] to collect oral reports from our CHW respondents in a more conversational and organic way. Sample questions used during these interviews include "Does the idea of text messaging with patients/caregivers makes sense?"; "Does it make sense for the text message to be integrated into a patient relation management software or electronic health record, or both?"; and "What do you find to be the benefits to connecting with patients via a text message?"

Phase 4: Full Reiteration

Phase 4, which started in February 2016, involved a full reiteration of the 5 design thinking steps: empathizing, defining, ideating, prototyping, and testing. The objective of this phase was to connect and understand multiple audiences, ideate integrated solutions that addressed the needs of the multiple end-users (patients, staff, and clinicians), and develop and test iterative versions of the texting app prototype that satisfied the final outcome of successfully engaging all end-user groups. In this phase, we operationalized the objective into activities that

included identifying and understanding additional new end-user groups; redefining the need requirements; ideating an integrated solution that incorporated the newly defined need requirements as well as developing and testing the texting app prototypes 3 and 4; and collecting feedback from the end-user groups.

Identifying and Understanding New End-User Groups

The first 3 phases became instrumental in recognizing the gap of end-user perspectives that led us to expand the codesign team to include not only software engineers and health data professionals but also health care professional support staff (operation managers and supervisors), which we identified as being key leaders to successful implementation and adaptation of the texting app tool by CHWs.

Redefining the Need Requirements

The findings from phases 1 to 3 were discussed with the newly added end-user groups through unstructured creative interviews [55], and it was discovered that the first 2 prototypes lacked implementation consideration. This included the delivery of services within the broader health care system and addressed issues such as privacy and security for patient data information exchange and additional reporting needs (see the Results section for specific details). These considerations were used to redefine the need requirements that addressed patient and staff engagement with the tool from this vantage point.

Ideate, Develop, and Test Prototype 3

The findings from the first 3 phases and the newly redefined need requirements were operationalized into an integrated solution of the texting tool prototype 3, which encompassed better security, live two-way interaction, and smart support for better triaging (triaging in health care is a concept that deals with the sorting and allocation of services/treatments to patients based on the severity level of their health condition).

A pilot of 10 staff members tested prototype 3 to ensure that its technical aspects met the requirements of the larger health care system for delivery of health care services. The patients' SMS text messaging activities were not tracked during this testing period. The staff interviews throughout this phase modeled those of phase 3 with the exception of 6 interviews that were preceded by a wireframe SMS text messaging mockup presentation, which was comprised a set of texting prototype tool images that illustrated the working elements of the tool. The interviews during this phase included 4 CHWs, 3 other health care professionals, 1 health data analyst, 1 software engineer, and 1 research designer.

Ideate, Develop, and Test Prototype 4

The findings from the prototype 3 testing highlighted the limitation of the patient-centered approach to care, which is a care delivery model whereby patient services/treatments are triaged through a primary care provider ensuring appropriate care at the right time [8]. Prototype 4 attempted to address this need by expanding on key additional features of the SMS tool.

A 23-item survey containing questions about the usability and feasibility of the texting prototype was administered to CHW staff (n=12). The purpose of the survey was to assess the staff's perception of the texting app final prototype 4. The survey was built and delivered via Qualtrics software that enables researchers to build and deliver Web-based surveys for the purpose of data collection [56]. A sample *yes/no* question was "In the past two weeks have you been using the SMS app?" A sample Likert-type (1=strongly agree to 5=strongly disagree) question was "Thinking back about your experience with the SMS app, please rate the following statement: 'SMS app is easy to use'." A sample open-ended question was "What is your least favorite thing about using the SMS app?" The patients' text messaging activities were not tracked during this testing period.

Data Analysis

All the interviews and discussions across the 4 phases were conducted by at least 2 members of the codesign team (1 leading the discussion and the other recording notes). In total, 2 independent researchers, with master's degrees, transcribed and open-coded the text separately and then openly compared the themes to reach a consensus on the final coding themes [57]. The emerging themes guided the content of the texting use cases and the foundational blocks for the several prototypes built and tested during the third and fourth phase (see Table 1).

The data collected during the pilot studies and the self-administered Web-based staff questionnaire were deidentified, coded, and analyzed using STATA SE, a statistical software package used by social science researchers [58]. The information gathered through all 4 phases was ethically conducted and was approved by the Institutional Review Board at the University of Illinois, Chicago.

Results

Phase 1: Understand the Audience, Context, and Need Requirements

Identifying and Understanding the End-Users

We identified our end-user groups to be pediatric patients and their parents as well as their community advocates/leaders comprising the *lived experience* expert group; the CHECK program staff comprised CHWs, licensed clinicians, other health care professionals (eg, supervisors) and health data analysts, and a scientist; and the design team comprised software engineers, a texting research expert, and research and social scientists. Further details are shown in Table 2.

To inform our understanding of parents/caregivers and their communities' needs for better engagement and useful technologies, we used recommendations from members of the CHECK PAB and CAB who served as *lived-experience* advisors and evaluators of the CHECK program components. The importance of recognizing the feedback of those with *lived experience* to guide the design and implementation process of a technology-enhanced tool is in alignment with the patient-clinician-designer framework [44] and has been successfully implemented in one other study [27]. CHWs are trusted members of the community serving as public health workers with a deep understanding of the issues and struggles of their communities [59]. The CHWs' unique and unusual position as both public health workers and patients of the same health care system allows them to provide innovative solutions

to health care–related problems. This approach to problem solving made CHWs an integral part of the multidisciplinary design team. Finally, the CHECK health care professionals whose expertise included clinical social workers, clinical providers, public health, and social scientists helped inform our understanding of the structural and systemic blocks on the benefits and barriers facing technology-based tools in health care [60].

The literature review matching the patients as end-users within the context of health care delivery using SMS text messaging technology was conducted in March 2015. A general search retrieved 911 citations, and 60 relevant studies were reviewed. Patient groups most similar to CHECK's target population, such as patients with HIV/AIDS (9/60, 15%) and diabetes (8/60, 13%), were identified as the most commonly targeted groups for SMS text messaging studies. The majority of the studies (46/60, 77%) reported improved outcomes. The most commonly reported improved outcomes by SMS text messaging studies focused on improving adherence to medication or treatment (24/46, 52%), increasing appointment attendance (11/46, 24%), and decreasing no-show appointment rates (11/46, 24%). Limitations to using texting within the context of health care centered around privacy concerns with 93% (56/60) of the studies reporting to have remedied this by applying safeguards such as omitting the patient's name or other Patient Health Information–related information from the body of the SMS text messages. Other limitations cited wrong or invalid mobile numbers because of patients changing them and not reporting the new phone number to the study staff [61] or because of data entry errors [62]. Recommendations for using texting in the context of health care delivery focused on the frequency and themes of the text message content tailored to fit the study's program goals and the target population.

Defining the Need Requirements

The results from the interviews with the *lived experience* experts, professionals, and researchers revealed certain common elements that helped frame the development for a useful SMS text messaging approach applicable to an already existing texting tool. The feedback highlighted the need for a real-time two-way interaction with a feedback loop between patients and providers that was secure but did not require data or a constant internet connection to function properly. Another important point obtained was the need for a tool that highlighted human interaction and noncomputerized intervention programs and placed the focus of communication on the language of engagement and non-disease specific topics, such as goal tracking or social support (see Table 3 prototypes).

Phase 2: Ideate a Context-Based Solution, Build a Prototype, and Test

Ideate, Develop, and Test Prototype 1

The first prototype that was ideated, developed, and tested was a two-way texting tool, non–Health Insurance Portability and Accountability Act (HIPAA) secure, low to no cost, no data or Wi-Fi required, and with live communication through preset logic to gather insight into patients' common needs and language that would encourage engagement and self-efficacy, particularly with goal setting.

A total of 835 SMS text messages were delivered successfully to 95 respondents over a 4-week period. Of all the sent SMS text messages, 7% (58/835) received responses. The engagement language did not appear to influence whether the patient chose to respond to the SMS text message. The majority of the patients who engaged in SMS replies (14/23, 63%) preferred setting goals and then tracking the goals (7/14, 51%). One patient followed up with the following text reply when prompted to create a goal implying enhanced self-efficacy: "...yeah [that] be so slick like let's take care of this problem first, I be on the roll." An overwhelming majority of the respondents (22/23, 99%) preferred SMS text message reminders to keep them on track with their goals.

Phase 3: Redesign Prototype and Test

Ideate, Develop, and Test Prototype 2

The second prototype was designed and developed in response to the patient data collected during Phase 2. Prototype 2 featured a similar app as the first prototype, namely, two-way texting, non-HIPAA secure, low to no cost, no data or Wi-Fi required, and with live communication through preset logic on 3 themes identified as the patients' top needs. In this prototype, we added a delayed interactive support feature by the CHW staff.

For the testing of prototype 2, over the 8-week period, a total of 64 SMS text messages (both computer-to-human and human-to-human) were sent out to 52 unique patients and averaging 6 to 7 patients per week, who exchanged communication regarding their service needs via the texting tool. The overall average response rate (SMS text messages sent-to-received) was 88%. This rate increased to 100% for the categories of transportation coordination and appointment scheduling and reminders.

All 3 CHWs found the prototype compatible with their patients' needs. Positive feedback included that it "...allows [additional] freedom to operate and customize text as needed" and "The [SMS] web tool is easy to use...with up to 10 patients per week." The use of the template library was seen as a major benefit to "...save [me] time..." and "... stay in compliance [with HIPAA guidelines]...". The CHWs also reported operational limitations presented by the use of the texting tool prototype. One such limitation was the lack of a universal SMS text messaging phone number similar to a toll-free telephone number. The CHWs found the lack of front-end user interface inconvenient but preferred it as a more comprehensive way of using the tool. A major operational barrier that was reported was having to manually reinsert the patient's name and phone number for every SMS text messaging interaction, even though this approach ensured that the phone number used to texting the patient was the most recent on file. The unilateral Web-based only platform for use was reported as a major barrier by the CHWs, because it restricted them to manual operations of varying frequency. They felt the need for a more instant approach to monitor and track SMS replies, especially when dealing with at-risk patients.

Phase 4: Full Reiteration

Identifying and Understanding New End-User Groups

As previously described for the texting tool, we identified our stakeholder groups as end-users of various systems within the context of health care. The pediatric patients, their parents, and their community advocates/leaders helped us define the SMS text messaging as a live, two-way interaction with staff for transportation coordination, social support service delivery, and appointment scheduling plus reminders. CHECK program staff comprised CHWs and care coordinators who identified operational needs as an editable and smart library of SMS text messaging templates to address patients' needs as well as a self-maintaining secure password reset system for staff. During this phase, we expanded our stakeholder audience to include members of yet another important system within the health care context. CHECK professional staff (health care operation managers and health data analysts) helped inform our understanding of the structural and systemic blocks on the benefits and barriers facing technology-based tools in health care.

Redefining the Need Requirements

Data collected through informal interviews with these professional staff revealed needs for technology-enhanced staff training, a built-in application programming interface for better integration with existing software cloud-based apps (eg, care management software), built-in dashboards with automated data queries for data exchange, and summary reports with varying levels of security for staff access and supervision. The results pointed to a need for automated feedback. The lack of automatic notification to notify staff of an SMS reply and the delayed SMS text messaging interaction was described as "...very restrictive...not enough [front-end] choices." Finally, the testing period of the previous 2 prototypes also highlighted the need for diversifying the texting platform to include (1) collecting quality improvement data on the CHECK program as a whole and (2) sending mass notifications to patients about program-funded/organized events with the ability to interact beyond RSVP capabilities.

Ideate, Develop, and Test Prototype 3

In February 2016, we redesigned the texting prototype. This redesign addressed the technology-enhanced operational needs and automated feedback feature identified across the stakeholder groups and from the testing phases of the first 2 prototypes. This approach identified a HIPAA disclaimer issue, which was resolved in real time.

The redesigned prototype 3 offered a two-way SMS tool that was HIPAA secure, low to no cost, no data or Wi-Fi required, real live communication, and interactive support by CHW staff. To address the operational needs for better integration with existing software, a troubleshooting feature was implemented. This allowed the CHWs to report their feedback or request technical support via a technology support group chat that would triage the issues to 3 different technology support teams (development, data, and content-based teams).

In response to the automated feedback issues, the texting app included a group feature for staff-to-staff communication and also a built-in scheduling reminder texting feature that allowed staff to customize it in terms of frequency and language used to fit the specific needs of the patient (Figure 5).

Furthermore, prototype 3 offered 3 core tools: (1) A secure mode of communicating about the CHECK program and its services; (2) An easy-to-navigate and accessible-from-anywhere information exchange tool that is not restricted by data or Wi-Fi–only devices (eg, smartphone, tablet, or computers); and (3) An instant and interactive mode of information exchange (eg, similar to chat messages; see Table 3 for a full list of features).

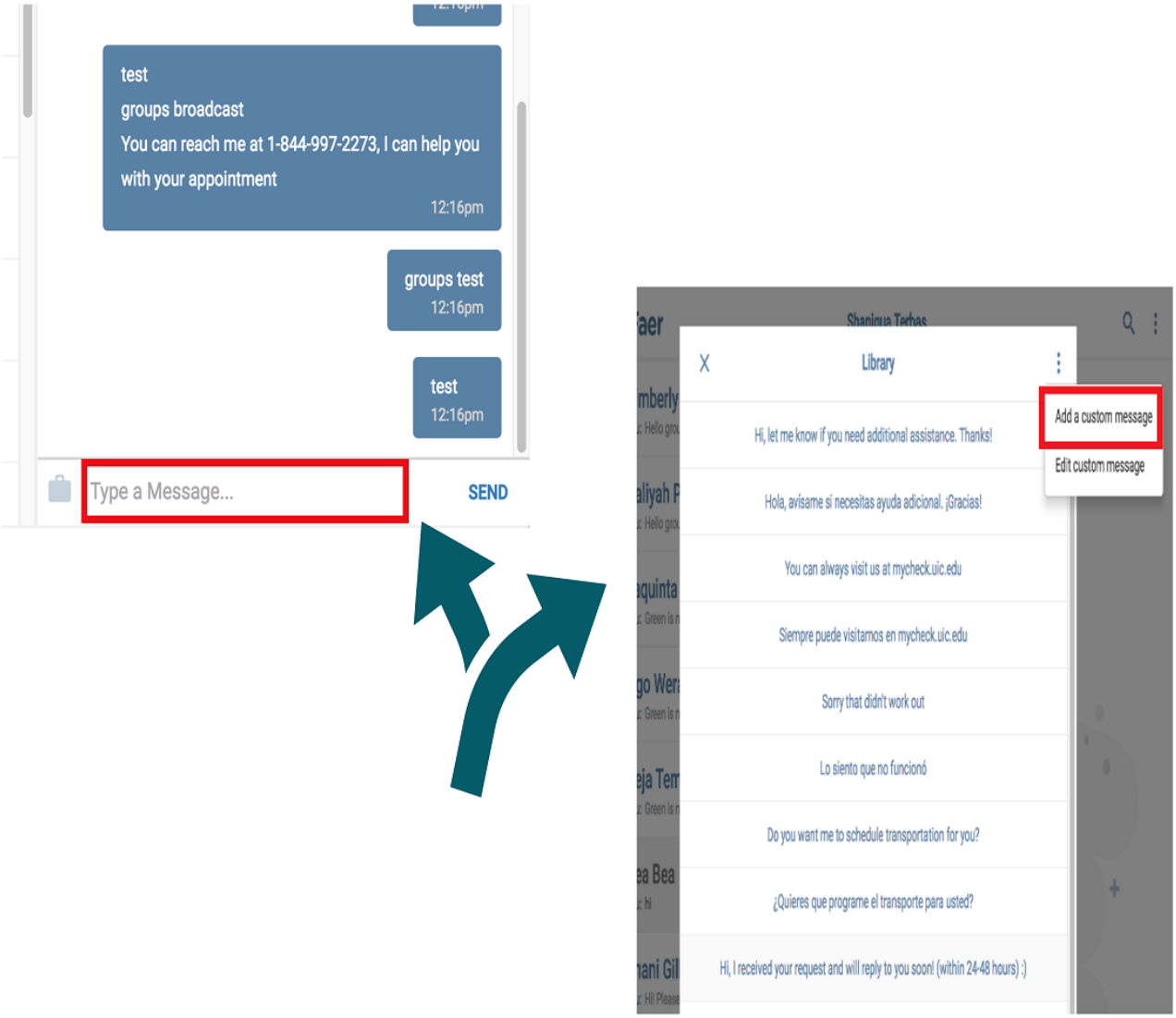
On the patient end, there were no operating system limitations. The patients continued to receive the SMS text messages in their native texting app under 1 unique CHECK SMS text messaging number that they were asked to save and use only for texting.

During the testing of prototype 3, the additional need for managing a higher case load and better scalability was further identified by 1 of the CHWs who said:

> *...with a small number of patients for one or two types of [social] referrals but very time consuming if the number of patients is bigger [than 10] and/or types of referrals expands [beyond transportation and social service delivery]...*

Ultimately, the lack of the texting app tool to fully adapt to the identified operational changes by extending the patient-center model of care coordination into the texting app was identified as a barrier that in turn prevented the app's consistent and uniform implementation by the staff. Full integration of the tool into the care management or electronic health record software was considered to be the biggest limitation across all stakeholder groups.

Figure 5. Customizable and self-learning library of text message templates in English and Spanish.



Ideate, Develop, and Test Prototype 4

Prototype 4 took into consideration the major limitation of the patient-centered approach to care. Prototype 4 (see Figure 2) attempted to address this limitation by expanding several key features, such as creating a Web version of the texting app to ensure adoption of the tool by the staff who interacted with the patients only from the office (eg, mental health staff). To partially address the integration issue of prototype 3, we mapped each patient to their assigned CHWs marking the first interoperating bridge between the care management software and the SMS text messaging app tool. To ensure the staff had the most updated patient list, including the most updated phone numbers, addresses, assigned patient risk scores, and chronic conditions, we automated the process to update on a daily basis. Through the use of a proprietary algorithm that worked in the back-end of the data warehouse (to align with HIPAA requirements), we expanded the patient profiles to include the patients' zip codes, chronic condition, and assigned risk scores to allow patient grouping by service needs for better triaging.

As a result of these changes, prototype 4 featured transportation coordination, social support service delivery, appointment scheduling and reminders, an opt-out and HIPAA disclaimer feature for patients, and a self-managed stop/start enrollment feature for patients. On the field staff side, the case load of patients expanded beyond 200 patients/staff, the prototype provided an automated grouping feature (based on any combination of patients' chronic condition, assigned risk, and zip codes), a technology support group chat feature triaging staff issues in real time (see Figure 3), a self-learning training manual with 24 hour Web access, and an automated scheduling and reminder feature (see Figure 4).

The final testing for prototype 4 began in January 2017 for both Web and Android versions of the texting app with 4 CHWs servicing 800 patients. The testing period continued for 5 months to allow for all the features to be tested and to rule out any additional technical barriers. The feedback provided was that this final texting app product was appealing and user friendly. More specifically, of the 12 staff surveyed, only 66% (8/12) chose to reply and of them only 4 confirmed use of the texting tool. All 4 CHWs agreed (1/4, 25%) or strongly agreed (3/4, 75%) with the following statement: "SMS App is easy to use". A similar number also agreed that the layout design of the texting tool was extremely (3/4, 75%) or moderately (1/4, 25%)

organized. However, they did not share the same strong sentiment on the overall design of the texting app. When asked, only 25% (1/4) agreed it was extremely good and moderately good (2/4, 50%). One CHW listed the design as neither good nor bad. When asked to further elaborate on why they selected that choice, the answers ranged from "It's very user friendly" to "it looks bland" to "There is no way to know if the recipient of the text actually received the text." Regarding the design of the texting app, 3 of the 4 CHWs replied *No* to the question, "Is there anything missing from the SMS App that you were expecting to see (e.g., more text, more images, a FAQ, a question answered, etc.)?" One CHW, who replied in the affirmative, also listed the lack of the feature to allow staff "...to modify patient contact information..." directly in the Web version of the app. Finally, when asked to list their favorite and least favorite things about the texting tool design, 75% (3/4) of the staff listed positive reasons such as "It's one touch access and it loads ... all the information at my fingertips" and less positive ones such as "Not knowing if the recipient actually received the text" (2/4, 50%). The concern about SMS text message receipt was also highlighted as the main technical issue by 75% (3/4) of the staff members with 25% (1/4) listing "cannot easily reset password" as the only technical issue. When asked to provide additional feedback for further improvement, the staff overwhelmingly (3/4, 75%) stated no feedback for further improvement. One CHW reported "...is a good application. I can't think of anything else to improve the application". When asked to list all the barriers that prevented them from using the texting tool, 25% (1/4) checked "Do not feel comfortable with technology", 25% (1/4) checked "I prefer to use other tools better", 25% (1/4) checked technical issues with the texting tool, and 25% (1/4) preferred to use "google voice chat."

Natural Language Understanding and Common Responses

Different approaches for natural language understanding were used depending on the prototype. In phase 2, the first 2 prototypes were built with preset branching logic (drawing from common expected replies) and if the reply from the patient fell outside of the predefined range it was automatically forwarded to the staff on call. The staff would then address the answer according to the protocol in place. This was primarily done to ensure the safety of the patients; hence, no automatic SMS reply was sent back to the patient. On the contrary, the staff member (CHW) on call was instructed to reply back to the patient with a tailored answer appropriate to the situation at hand either immediately (if urgent) or within 24 to 48 hours (if nonurgent). Prototypes 3 and 4 automatized this feature to include an SMS text message with a disclaimer about urgent matters and provided a list of emergency phone numbers tailored to the patients' zip code on file. Furthermore, the last prototype of the tool ensured that the CHW assigned to the patient was notified via text and email. Common responses such as *thanks* or *this wasn't helpful* were designed to be responded to with a preset branching logic to address common and expected replies. Replies falling outside that range were forwarded to the staff to be addressed within 24 to 48 hours, depending on the urgency of the situation.

Discussion

Principal Findings

In this study, we have described the user-informed development and design of a HIPAA secure, two-way interactive SMS text messaging app, serving as the primary mode for information exchange between patients and CHWs. We utilized the design thinking process to take into account multiple perspectives including those of young adults and caregivers to children with chronic health conditions and CHWs in the designing of the tool. The design thinking process centering these users resulted in the development of 4 prototypes leading to a final tool that addressed the needs and barriers during each phase of testing. The final prototype took into account both patient and health care staff preferences to have a tool that is useful for both groups.

The final prototype addressed and included what the patients articulated throughout the testing of other prototypes. The majority preferred texting reminders for tracking goals and appreciated the support in setting them as well. They also wanted the tool to be able to provide appointment scheduling as well as reminders for them. They also emphasized the importance of personalized messaging, human connection, and real time with two-way interaction. Given that the patients were Medicaid recipients, they identified need-based services such as transportation coordination and social support service delivery as helpful to receive through the tool. Finally, HIPAA security was prioritized, given the patient sensitive information exchanged through the tool.

The final prototype also accommodates the changes and preferences of staff feedback given through the testing of other prototypes. For the staff, the tool can be used on the Web and a smartphone/tablet for unlimited access from the office or the field. Assigned patient queues (similar to a contact list in your phone) offer staff the ability to handle a 200+ patient case load. The patient list allows grouping based on factors such as diagnosis and zip code, among others. There is a customizable library of SMS text messaging templates in English and Spanish and an automated patient grouping feature for tailored and efficient mass texting with the ability to receive one-on-one replies. An automated scheduling and reminder feature that is customized and tailored individually for content and frequency ensures that no patient is forgotten. A technology support group chat feature allows for triaging staff issues in real time. This SMS text messaging app has the potential to improve patient engagement with their health and health care.

Limitations and Future Research

There are several study limitations. First, when working to understand the range of stakeholder views to inform the technology development, we were unable to gather direct user feedback from adolescent patients to inform the codesign team on their perspective. Second, we did not analyze data categorizing the participant texting actual content, response sentiment (eg, positive, negative, and neutral), response frequency (ie, how often a user responded), or response time. Future research should examine these variables to better tailor the tool for end-users. Third, although both patients and staff

alike found the tool to be less effective because it was not fully integrated within existing care management technology, the design thinking for tool development did not expand on the established software apps already in use by the program staff. We found that within the context of the larger goals of the CHECK study, emphasis was typically placed on ease of navigation and issues with technical glitches. The integration of this new SMS text messaging tool within existing technology turned out to be extremely challenging because of unprecedented institutional and technological barriers that fall beyond the scope of this study. Despite these barriers, the internal anonymous feedback collected from staff evaluated the tool positively. Fourth, the stakeholder feedback collected from patients' caregivers and CHECK staff occurred independent of each other. Thus, there was no all-inclusive, connected conversation to simultaneously gauge stakeholder interest from each group. This was intended to be sensitive to the patient population's needs [27] where such a crossover may create a more intimidating environment or reduce patients' openness to discuss their experiences with existing services. Fifth, design thinking was the primary framework used to develop the tool, which did not take into account other considerations such as technology acceptance, behavior change, or patient engagement models. Sixth, we were not able to implement a free-to-end-user mechanism for our SMS text messaging tool. Seventh, in adapting the design thinking process to develop the prototypes sequentially instead of in parallel, we limited the product optimization by concentrating on 1 solution. It is possible that sequential prototyping is not as effective as parallel prototyping [63]. Finally, in alignment with the design thinking step 1 of understanding and empathizing with the audience, we put the primary focus on ensuring all stakeholders felt a part of the process and opened up about their experiences without feeling judged. During all phases, we highlighted the importance of anonymity for this purpose and thus did not collect the demographic information of the participants. A strength of this study is the scalability where, in prototype 4, we had 4 CHWs service 800 patients (ie, 200 patients per CHW).

There are a number of areas for future research. First, future research should study if our approach of 1 CHW per approximately 200 patients would work with a larger number of total patients such as 5000 or 10,000 patients. Second, future research also should study if when working with a larger number of patients whether the CHW work load can be increased to more than 200 patients per CHW. Third, in addition to the focus of this study on the process of tool design, it would be useful in future research to determine how this tool or another similar tool would address and potentially positively impact health disparities and social determinants of health. Fourth, with an SMS text messaging tool, there can be many additional ways to tailor content. The final version of the texting tool was developed to meet patient needs on an individual and personalized level. We did not tailor messages based on variables such as diagnosis, response tone, or response frequency as these needs were all addressed on an individual level with personalized texting rather than with large bulk SMS text messages to a larger group of patients. Future research with other SMS text messaging tools that emphasize bulk SMS text messages should consider other areas of tailoring such as tailoring based on disease type, length of sickness, response tone, preferred frequency, and time of day. Finally, with a typical texting tool, it is important to track the *opt-in/opt-out* of patient participants. We did not track this measure for the text messaging tool on the latter prototypes as we relied on the opt-out measure for the CHECK program. Future research should track both measures to accurately measure end-user differences on program preferences versus SMS text messaging tool preferences.

Comparison With Previous Work

In alignment with a previous study on SMS text messaging tool use [27], our stakeholder groups asked for the use of a mobile phone texting feature as a contact method. This may be due to a steady shift in recent years by most health care systems in using one-way texting to confirm or remind patients about their doctor visits. There is also previous research that utilized messaging platform features to improve engagement specifically among Medicaid patients, confirming that texting can be an effective way of reaching low-income populations in particular [64].

In terms of the patient experience, the design thinking process is similar to research highlighting the need to address the barriers with health interventions [36]. Our SMS text messaging tool was adapted to patient-reported needs and included resources to address barriers to engagement and health care access such as transportation coordination and social support service. The final SMS text messaging app product extends the existing texting prototypes [17,27] through the inclusion of dynamic features for a more *real-life* information exchange experience that can promote patient self-efficacy and meet the needs of both patients and health care staff. Similar to previous research where female adolescent patients expressed privacy concerns with texting [65], interviews with stakeholder groups suggested a need for an SMS text messaging tool that remains compliant with HIPAA standards of patient privacy. This concern was addressed in real time during the third phase of the protocol by including privacy protections to ensure all patient communication was HIPAA compliant.

In terms of the staff experience, a major barrier noted by providers in phase 3 was the re-entry of patient name and phone number with each message to ensure that the patient phone number was the most recent on file. Previous research suggests inaccurate phone numbers are a significant concern to some providers and manual re-entry may be 1 way to address this concern [66]. Although manual re-entry was found to be a barrier to use by the professional staff in our project, it could prove to be valuable in addressing the concerns surrounding inaccurate phone numbers by ensuring that the most recent phone number on file is used.

Conclusions

The design thinking process was particularly important for the context of this tool because of the need to accommodate varying users, namely, young adults and caregivers of children with chronic health conditions, along with CHWs. Each group had its own preferences and for the tool to be utilized in an effective way, it needed to incorporate the essential preferences of all.

Our hybrid computer-human two-way interactive SMS text messaging tool may be especially useful to patients who are unable, because of a lack of resources such as time, data, and Wi-Fi, to interact face-to-face with their health care team. We recommend that when developing a technology-based interactive tool, developers use a validated approach such as the design thinking process to ensure that the final product is compatible with the needs of the target population. It is also important that product evaluation needs be ongoing to address the stream of information from multiple systems of stakeholders. Finally, additional focus should be placed on hybrid computer-human interactions that utilize existing low-cost and easy-to-use technology, such as SMS text messaging, to optimize personalized experience and the likelihood of enhanced patient engagement. Future research would be useful to study if the texting tool works with other patient age groups and patient-centered care teams of professional staff.

Acknowledgments

Rachel Jacobs acted as a mentor to the first author and substantially influenced the framing of the manuscript, David Gutelius and Chris Diehl led the efforts for building the AI supported SMS app, Paul Kim and Joanna Tess contributed to data collection and analysis, Melishia Bansa contributed to the collection of the qualitative feedback from the CAB and PAB, and Renee Smith and Alan Schwartz contributed to scientific writing discussions.

This study was funded by Grant Number 1C1CMS331342 from the Department of Health and Human Services, Centers for Medicare & Medicaid Services. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of the US Department of Health and Human Services or any of its agencies. The research presented here was conducted by the awardee. Findings might or might not be consistent with or confirmed by the findings of the independent evaluation contractor.

Authors' Contributions

MMH contributed to the original idea model development, testing and implementation. MMH and SLC cataloged, reviewed and analyzed the metadata of background literature and drafted the manuscript as the first authors. JF and MF contributed to manuscript framing discussions and major editing. ES contributed to editing and reference citation review. BWVV and AEG provided support with concepts surrounding Medicaid and the CHECK overall program.

Conflicts of Interest

None declared.

References

1. Kruse J. Fragmentation in US medical education, research, and practice: the need for system wide defragmentation. Family Medicine 2013;45(1):54-57 [FREE Full text] [Medline: 23334974]
2. Cohen E, Lacombe-Duncan A, Spalding K, MacInnis J, Nicholas D, Narayanan U, et al. Integrated complex care coordination for children with medical complexity: a mixed-methods evaluation of tertiary care-community collaboration. BMC Health Serv Res 2012 Oct 23;12:366 [FREE Full text] [doi: 10.1186/1472-6963-12-366] [Medline: 23088792]
3. Kangovi S, Long JA, Emanuel E. Community health workers combat readmission. Arch Intern Med 2012 Dec 10;172(22):1756-1757. [doi: 10.1001/2013.jamainternmed.82] [Medline: 23128914]
4. The Colorado Trust. Denver, CO; 2011. The Status of Behavioral Healthcare in Colorado: 2011 Update URL: https://www.coloradotrust.org/sites/default/files/ACMHC_2011_Full-Report.pdf [accessed 2019-02-22] [WebCite Cache ID 76MvOG0qq]
5. Berry J, Agrawal R, Cohen E, Kuo D. Columbia University in the City of New York. Overland Park: Children's Hospital Association; 2013. The landscape of medical care for children with medical complexity URL: http://www.columbia.edu/itc/hs/medical/residency/peds/new_compeds_site/pdfs_new/PL3%20new%20readings/Special_Report_The_Landscape_of_Medical_Care_for_Children_with_Medical_Complexity.pdf [accessed 2019-02-22] [WebCite Cache ID 76MvYKZhV]
6. Berry J, Hall M, Neff J, Goodman D, Cohen E, Agrawal R, et al. Children with medical complexity and Medicaid: spending and cost savings. Health Aff (Millwood) 2014 Dec;33(12):2199-2206 [FREE Full text] [doi: 10.1377/hlthaff.2014.0828] [Medline: 25489039]
7. Nguyen T, Henderson D, Stewart O, Hlyva O, Punthakee Z, Gorter JW. You never transition alone! Exploring the experiences of youth with chronic health conditions, parents and healthcare providers on self-management. Child Care Health Dev 2016 Dec;42(4):464-472 [FREE Full text] [doi: 10.1111/cch.12334] [Medline: 27103590]
8. Glassgow A, Martin M, Caskey R, Bansa M, Gerges M, Johnson M, et al. An innovative health-care delivery model for children with medical complexity. J Child Health Care 2017 Sep;21(3):263-272. [doi: 10.1177/1367493517712063] [Medline: 29119824]
9. Gates A, Rudowitz R, Artiga S, Snyder L. The Kaiser Commission on the Medicaid and the Uninsured. Menlo Park, CA: The Henry J Kaiser Family Foundation; 2016. Two Year Trends in Medicaid CHIP Enrollment Data: Findings from the

CMS Performance Indicator Project URL: http://files.kff.org/attachment/Issue-Brief-Two-Year-Trends-in-Medicaid-and-CHIP-Enrollment-Data [accessed 2019-02-19] [WebCite Cache ID 76JK2zcKL]
10. Cohen E, Berry J, Camacho X, Anderson G, Wodchis W, Guttmann A. Patterns and costs of health care use of children with medical complexity. Pediatrics 2012 Dec;130(6):e1463-e1470 [FREE Full text] [doi: 10.1542/peds.2012-0175] [Medline: 23184117]
11. Kuo D, Cohen E, Agrawal R, Berry J, Casey P. A national profile of caregiver challenges among more medically complex children with special health care needs. Arch Pediatr Adolesc Med 2011 Nov;165(11):1020-1026 [FREE Full text] [doi: 10.1001/archpediatrics.2011.172] [Medline: 22065182]
12. Neff J, Sharp V, Muldoon J, Graham J, Myers K. Profile of medical charges for children by health status group and severity level in a Washington State Health Plan. Health Serv Res 2004 Feb;39(1):73-89 [FREE Full text] [doi: 10.1111/j.1475-6773.2004.00216.x] [Medline: 14965078]
13. Hibbard J, Greene J. What the evidence shows about patient activation: better health outcomes and care experiences; fewer data on costs. Health Aff (Millwood) 2013 Feb;32(2):207-214. [doi: 10.1377/hlthaff.2012.1061] [Medline: 23381511]
14. Hibbard J, 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. [doi: 10.1377/hlthaff.2012.1064] [Medline: 23381513]
15. Conway C, Kelechi T. Digital health for medication adherence in adult diabetes or hypertension: an integrative review. JMIR Diabetes 2017 Aug 16;2(2):e20 [FREE Full text] [doi: 10.2196/diabetes.8030] [Medline: 30291093]
16. Martin T. Assessing mHealth: opportunities and barriers to patient engagement. J Health Care Poor Underserved 2012 Aug;23(3):935-941. [doi: 10.1353/hpu.2012.0087] [Medline: 24212144]
17. Shaw R, Bosworth HB, Hess J, Silva S, Lipkus I, Davis L, et al. Development of a theoretically driven mHealth text messaging application for sustaining recent weight loss. JMIR Mhealth Uhealth 2013 May 7;1(1):e5 [FREE Full text] [doi: 10.2196/mhealth.2343] [Medline: 25100678]
18. Ahlers-Schmidt C, Chesser A, Hart T, Paschal A, Nguyen T, Wittler R. Text messaging immunization reminders: feasibility of implementation with low-income parents. Prev Med 2010;50(5-6):306-307. [doi: 10.1016/j.ypmed.2010.02.008] [Medline: 20178813]
19. Evans W, Wallace J, Snider J. Pilot evaluation of the text4baby mobile health program. BMC Public Health 2012 Nov 26;12:1031 [FREE Full text] [doi: 10.1186/1471-2458-12-1031] [Medline: 23181985]
20. Gerber B, Stolley M, Thompson A, Sharp L, Fitzgibbon M. Mobile phone text messaging to promote healthy behaviors and weight loss maintenance: a feasibility study. Health Informatics J 2009 Mar;15(1):17-25 [FREE Full text] [doi: 10.1177/1460458208099865] [Medline: 19218309]
21. Bhat A, Mao J, Unützer J, Reed S, Unger J. Text messaging to support a perinatal collaborative care model for depression: a multi-methods inquiry. Gen Hosp Psychiatry 2018;52:14-20. [doi: 10.1016/j.genhosppsych.2018.01.005] [Medline: 29494854]
22. Prayaga R, Jeong E, Feger E, Noble H, Kmiec M, Prayaga R. Improving refill adherence in medicare patients with tailored and interactive mobile text messaging: pilot study. JMIR mHealth & uHealth 2018;6(1):e30. [doi: 10.2196/mhealth.8930] [Medline: 29382623]
23. Bobrow K, Farmer A, Springer D, Shanyinde M, Yu L, Brennan T, et al. Mobile phone text messages to support treatment adherence in adults with high blood pressure (SMS-Text Adherence Support StAR): a single-blind, randomized trial. Circulation 2016 Feb 9;133(6):592-600 [FREE Full text] [doi: 10.1161/CIRCULATIONAHA.115.017530] [Medline: 26769742]
24. Davidson T, McGillicuddy J, Mueller M, Brunner-Jackson B, Favella A, Anderson A, et al. Evaluation of an mHealth medication regimen self-management program for African American and Hispanic uncontrolled hypertensives. J Pers Med 2015 Nov 17;5(4):389-405 [FREE Full text] [doi: 10.3390/jpm5040389] [Medline: 26593951]
25. Aguilera A, Muñoz RF. Text messaging as an adjunct to CBT in low-income populations: a usability and feasibility pilot study. Prof Psychol Res Pr 2011 Dec 1;42(6):472-478 [FREE Full text] [doi: 10.1037/a0025499] [Medline: 25525292]
26. Chen M, Chang J, Kuo J, Yu J, Huang M, Marks B, et al. A pilot comparitive study of one-way versus two-way text message program to promote physical activity among people with severe mental illness. Ment Health Phys Act 2017;13:143-151 [FREE Full text] [doi: 10.1016/j.mhpa.2017.09.010]
27. Larsen M, Shand F, Morley K, Batterham P, Petrie K, Reda B, et al. A mobile text message intervention to reduce repeat suicidal episodes: design and development of Reconnecting After a Suicide Attempt (RAFT). JMIR Ment Health 2017 Dec 13;4(4):e56 [FREE Full text] [doi: 10.2196/mental.7500] [Medline: 29237584]
28. Perrier T, Dell N, DeRenzi B, Anderson R, Kinuthia J, Unger J, et al. Engaging pregnant women in Kenya with a hybrid computer-human SMS communication system. In: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems. 2015 Presented at: CHI'15; April 18-23, 2015; Seoul, Korea URL: https://www.researchgate.net/publication/282441688_Engaging_Pregnant_Women_in_Kenya_with_a_Hybrid_Computer-Human_SMS_Communication_System [doi: 10.1145/2702123.2702124]
29. Community Need Index. 2018. URL: http://cni.chw-interactive.org/ [accessed 2018-12-03] [WebCite Cache ID 74OTpskS2]

30. Vaughan K, Kok M, Witter S, Dieleman M. Costs and cost-effectiveness of community health workers: evidence from a literature review. Hum Resour Health 2015 Sep 1;13:71 [FREE Full text] [doi: 10.1186/s12960-015-0070-y] [Medline: 26329455]
31. Jack H, Arabadjis S, Sun L, Sullivan E, Phillips R. Impact of community health workers on use of healthcare services in the United States: a systematic review. J Gen Intern Med 2017 Mar;32(3):325-344 [FREE Full text] [doi: 10.1007/s11606-016-3922-9] [Medline: 27921257]
32. Pérez-Escamilla R, Damio G, Chhabra J, Fernandez M, Segura-Pérez S, Vega-López S, et al. Impact of a community health workers-led structured program on blood glucose control among latinos with type 2 diabetes: the DIALBEST trial. Diabetes Care 2015 Feb;38(2):197-205 [FREE Full text] [doi: 10.2337/dc14-0327] [Medline: 25125508]
33. Wennerstrom A, Hargrove L, Minor S, Kirkland A, Shelton S. Integrating community health workers into primary care to support behavioral health service delivery: a pilot study. J Ambul Care Manage 2015;38(3):263-272. [doi: 10.1097/JAC.0000000000000087] [Medline: 26049656]
34. Schroeder S. Shattuck Lecture. We can do better--improving the health of the American people. N Engl J Med 2007 Sep 20;357(12):1221-1228. [doi: 10.1056/NEJMsa073350] [Medline: 17881753]
35. Shah N, Rogers A, Kanter M. NEJM Catalyst.: Massachusetts Medical Society; 2016. Health Care That Targets Unmet Social Needs URL: https://catalyst.nejm.org/health-care-that-targets-unmet-social-needs/ [accessed 2019-02-22] [WebCite Cache ID 76N4bzL7b]
36. Williams D, Costa M, Odunlami A, Mohammed S. Moving upstream: how interventions that address the social determinants of health can improve health and reduce disparities. J Public Health Manag Pract 2008 Nov;14(Suppl):S8-17 [FREE Full text] [doi: 10.1097/01.PHH.0000338382.36695.42] [Medline: 18843244]
37. Gruman J, Holmes-Rovner M, French M, Jeffress D, Sofaer S, Shaller D, et al. Center for Advancing Health. Washington, DC: Center for Advancing Health; 2010. A new definition of patient engagement: what is engagement and why is it important? URL: http://www.cfah.org/pdfs/CFAH_Engagement_Behavior_Framework_current.pdf [accessed 2019-02-22] [WebCite Cache ID 76N5aWZTB]
38. Vahdat S, Hamzehgardeshi L, Hessam S, Hamzehgardeshi Z. Patient involvement in health care decision making: a review. Iran Red Crescent Med J 2014 Jan;16(1):e12454 [FREE Full text] [doi: 10.5812/ircmj.12454] [Medline: 24719703]
39. Greene J, Hibbard J. Why does patient activation matter? An examination of the relationships between patient activation and health-related outcomes. J Gen Intern Med 2012 May;27(5):520-526 [FREE Full text] [doi: 10.1007/s11606-011-1931-2] [Medline: 22127797]
40. Center For Advancing Health. Washington DC: Center For Advancing Health; 2010. A Snapshot of People's Engagement in Their Health Care URL: http://www.cfah.org/file/CFAH_Snapshot_2010_DataTables.pdf
41. Greene J, Hibbard J, Sacks R, Overton V, Parrotta C. 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]
42. Plattner H, Meinel C, Leifer L, editors. Design Thinking: Understand, Improve, Apply. Berlin, Germany: Springer Berlin Heidelberg; 2011:3-219.
43. Bandura A. The explanatory and predictive scope of self-efficacy theory. J Soc Clin Psychol 1986;4(3):359-373. [doi: 10.1521/jscp.1986.4.3.359]
44. Mummah S, Robinson T, King A, Gardner C, Sutton S. IDEAS (Integrate, Design, Assess, and Share): a framework and toolkit of strategies for the development of more effective digital interventions to change health behavior. J Med Internet Res 2016 Dec 16;18(12):e317 [FREE Full text] [doi: 10.2196/jmir.5927] [Medline: 27986647]
45. TechTarget Network. What is use case? URL: https://searchsoftwarequality.techtarget.com/definition/use-case [accessed 2019-02-22] [WebCite Cache ID 76N8NS2IO]
46. Marcu G, Bardram J, Gabrielli S. A Framework for Overcoming Challenges in Designing Persuasive Monitoring and Feedback Systems for Mental Illness. 2011 Jun Presented at: 5th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops; May 23-26, 2011; Dublin, Ireland URL: https://www.researchgate.net/publication/224261456_A_Framework_for_Overcoming_Challenges_in_Designing_Persuasive_Monitoring_and_Feedback_Systems_for_Mental_Illness [doi: 10.4108/icst.pervasivehealth.2011.246097]
47. Otero-Sabogal R, Arretz D, Siebold S, Hallen E, Lee R, Ketchel A, et al. Physician-community health worker partnering to support diabetes self-management in primary care. Qual Prim Care 2010;18(6):363-372 [FREE Full text] [Medline: 21294977]
48. Siedel V, Fixson S. Adopting design thinking in novice multidisciplinary teams: the application and limits of design methods and reflexive practices. J Prod Innov Manage 2013;30(S1):19-33 [FREE Full text] [doi: 10.1111/jpim.12061]
49. Smith A, Page D. Pew Research Center. Washington DC; 2015. US Smartphone Use in 2015 URL: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/ [accessed 2019-02-22] [WebCite Cache ID 76N9ZiGaA]
50. Hoek W, Marko M, Fogel J, Schuurmans J, Gladstone T, Bradford N, et al. Randomized controlled trial of primary care physician motivational interviewing versus brief advice to engage adolescents with an Internet-based depression prevention intervention: 6-month outcomes and predictors of improvement. Transl Res 2011 Dec;158(6):315-325 [FREE Full text] [doi: 10.1016/j.trsl.2011.07.006] [Medline: 22061038]

51. Landback J, Prochaska M, Ellis J, Dmochowska K, Kuwabara S, Gladstone T, et al. From prototype to product: development of a primary care/internet based depression prevention intervention for adolescents (CATCH-IT). Community Ment Health J 2009 Oct;45(5):349-354. [doi: 10.1007/s10597-009-9226-3] [Medline: 19641992]
52. Bernard HR. Research Methods in Anthropology: Qualitative and Quantitative Approaches. Lahman, Maryland: Rowman & Littlefield; 2011:1-570.
53. van der Bijl J, Shortridge-Baggett L. The theory and measurement of the self-efficacy construct. Sch Inq Nurs Pract 2002;15(3):189-207. [Medline: 11871579]
54. Bandura A. Self-efficacy: toward a unifying theory of behavioral change. Psychol Rev 1977 Mar;84(2):191-215. [doi: 10.1016/0146-6402(78)90002-4] [Medline: 847061]
55. Douglas J. Creative Interviewing. Thousand Oaks, CA: Sage Publications; 1985:1-159.
56. Qualtrics. Provo, UT: Qualtrics; 2016. Citing Qualtrics in Academic Research URL: https://www.qualtrics.com/blog/citing-qualtrics/ [accessed 2019-02-22] [WebCite Cache ID 76NBZTGkk]
57. Saldana J. The Coding Manual for Qualitative Researchers. Thousand Oaks, CA: Sage Publications; 2015:1-368.
58. Stata: Software for Statistics and Data Science. College Station, TX: StataCorp LLC; 2017. How do I cite the Stata software, documentation, and FAQs? URL: https://www.stata.com/support/faqs/resources/citing-software-documentation-faqs/ [accessed 2019-02-22] [WebCite Cache ID 76NBt2oMT]
59. Wennerstrom A. Oral Session: Promoting CHWs and their role in addressing and supporting mental health. In: Creating the Healthiest Nation: Ensuring the Right to Health. 2016 Presented at: American Public Health Association 2016 Annual Meeting and Expo; October 29-November 2, 2016; Denver, CO.
60. Gittlen S. NEJM Catalyst.: Massachusetts Medical Society; 2017. Survey snapshot: what patient engagement technology is good for URL: https://catalyst.nejm.org/patient-engagement-technology-good-for/ [accessed 2019-02-22] [WebCite Cache ID 76NDUN3wm]
61. Downing S, Cashman C, McNamee H, Penney D, Russell D, Hellard M. Increasing chlamydia test of re-infection rates using SMS reminders and incentives. Sex Transm Infect 2013 Feb;89(1):16-19. [doi: 10.1136/sextrans-2011-050454] [Medline: 22728911]
62. Koshy E, Car J, Majeed A. Effectiveness of mobile-phone short message service (SMS) reminders for ophthalmology outpatient appointments: observational study. BMC Ophthalmol 2008 May 31;8:9 [FREE Full text] [doi: 10.1186/1471-2415-8-9] [Medline: 18513438]
63. Dow S, Glassco A, Kass J, Schwarz M, Schwartz D, Klemmer S. Parallel prototyping leads to better design results, more divergence, and increased self-efficacy. ACM Trans Comput Hum Interact 2010;17(4):2-24 [FREE Full text] [doi: 10.1145/1879831.1879836]
64. Sigler B. Investigating the perceptions of care coordinators on using behavior theory-based mobile health technology with medicaid populations: a grounded theory study. JMIR Mhealth Uhealth 2017 Mar 21;5(3):e36 [FREE Full text] [doi: 10.2196/mhealth.5892] [Medline: 28325711]
65. Ranney M, Choo E, Cunningham R, Spirito A, Thorsen M, Mello M, et al. Acceptability, language, and structure of text message-based behavioral interventions for high-risk adolescent females: a qualitative study. J Adolesc Health 2014 Jul;55(1):33-40 [FREE Full text] [doi: 10.1016/j.jadohealth.2013.12.017] [Medline: 24559973]
66. Hofstetter A, Vargas CY, Kennedy A, Kitayama K, Stockwell M. Parental and provider preferences and concerns regarding text message reminder/recall for early childhood vaccinations. Prev Med 2013 Aug;57(2):75-80. [doi: 10.1016/j.ypmed.2013.04.007] [Medline: 23624252]

Abbreviations

CAB: Community Advisory Board
CHECK: Coordinated Health Care for Complex Kids
CHW: community health worker
CMC: chronic medical condition
HIPAA: Health Insurance Portability and Accountability Act
mHealth: mobile health
PAB: Parent Advisory Board
SMS: short message service

Edited by G Eysenbach; submitted 04.08.18; peer-reviewed by L Vergani, R Brar Prayaga, G Marton, B Loo Gee; comments to author 12.10.18; revised version received 18.01.19; accepted 30.01.19; published 01.05.19.

Please cite as:
Marko-Holguin M, Cordel SL, Van Voorhees BW, Fogel J, Sykes E, Fitzgibbon M, Glassgow AE
A Two-Way Interactive Text Messaging Application for Low-Income Patients with Chronic Medical Conditions: Design-Thinking Development Approach
JMIR Mhealth Uhealth 2019;7(5):e11833
URL: https://mhealth.jmir.org/2019/5/e11833/
doi:10.2196/11833
PMID:31042152

Original Paper

Assessing the Quality of Mobile Apps Used by Occupational Therapists: Evaluation Using the User Version of the Mobile Application Rating Scale

Kelsea LeBeau[1], MPH; Lauren G Huey[1], MOT, OTR/L; Mark Hart[1], EdD, MALS

College of Public Health and Health Professions, University of Florida, Gainesville, FL, United States

Corresponding Author:
Kelsea LeBeau, MPH
College of Public Health and Health Professions
University of Florida
1225 Center Drive, Health Professions, Nusring & Pharmacy Room 4176
PO Box 100175
Gainesville, FL, 32610
United States
Phone: 1 850 499 1952
Fax: 1 352 273 6048
Email: klebeau@ufl.edu

Abstract

Background: The continuous development of mobile apps has led to many health care professionals using them in clinical settings; however, little research is available to guide occupational therapists (OTs) in choosing quality apps for use in their respective clinical settings.

Objective: The purpose of this study was to use the user version of the Mobile Application Rating Scale (uMARS) to evaluate the quality of the most frequently noted mobile health (mHealth) apps used by OTs and to demonstrate the utility of the uMARS to assess the quality of mHealth apps.

Methods: A previous study surveying OTs' use of apps in therapy compiled a list of apps frequently noted. A total of 25 of these apps were evaluated individually by 2 trained researchers using the uMARS, a simple, multidimensional analysis tool that can be reliably used to evaluate the quality of mHealth apps.

Results: The top 10 apps had a total quality score of 4.3, or higher, out of 5 based on the mean scores of engagement, functionality, and aesthetics. Apps scored highest in functionality and lowest in engagement. Apps noted most frequently were not always high-quality apps; apps noted least frequently were not always low-quality apps.

Conclusions: Determining the effectiveness of using apps in clinical settings must be built upon a foundation of the implementation of high-quality apps. Mobile apps should not be incorporated into clinical settings solely based on frequency of use. The uMARS should be considered as a useful tool for OTs, and other professionals, to determine app quality.

(JMIR Mhealth Uhealth 2019;7(5):e13019) doi:10.2196/13019

KEYWORDS

occupational therapy; occupational therapists; mobile apps; mobile applications; technology; computers, handheld; mHealth; mobile health

Introduction

Background

The field of occupational therapy is continuing to increase throughout the United States. According to the United States Department of Labor Bureau of Labor Statistics [1], there are approximately 130,000 occupational therapists (OTs) working in offices, hospitals, schools, nursing homes, and home health services. The projected percent change in employment from 2016 to 2026 for occupational therapy is expected to be 21%, which is much higher than the average growth rate for all occupations of 7% [1].

Occupational therapy is a client-centered health profession, which aims to facilitate rehabilitation, health, and overall

well-being through occupation. OTs are allied health professionals who provide the care and support needed to enable injured, ill, or disabled patients to participate in activities of everyday life. They assist clients across all age groups to promote, develop, recover, improve, and maintain abilities allowing them to engage in the occupations needed for daily living and working. Moreover, OTs work with clients to develop ways in which the occupation or the environment can be modified to better support occupational engagement [2,3].

The world of technology continues to grow exponentially every year, and the use of mobile apps is becoming increasingly commonplace. According to the Pew Research Center, the number of US adults who own a smartphone has increased over the past 5 years from 51% in 2013 to 77% in 2018 [4]. With the continuous growth in smartphone usage, an entire new industry for third-party apps has formed. An "app," short for application, is defined as a self-contained program for smartphones designed to fulfill a particular purpose [5]. As of March 2017, there were 2.2 million and 2.8 million apps available in the Apple App Store and Google Play Store, respectively [6]. Of these 5 million apps, there are more than 318,000 health-related mobile apps with an estimated 200 health apps being added each day [7]. Given the versatility of apps, possibilities for mobile app integration into the field of occupational therapy continues to expand.

Mobile health (mHealth) is the application of technologies to improve health outcomes [8]. mHealth apps have the potential to be extremely beneficial in multiple health fields for several reasons: (1) They are cost effective, (2) They are accessible at any time, (3) They are convenient to the client, (4) They do not require assistance from health care professionals to operate successfully, and (5) They can allow the user to personalize their experience. In addition, mHealth apps have the potential to increase access to evidence-based health information.

As a result of the rise in everyday mobile technology use, the health care system has seen incredible growth in the use, and integration, of mobile apps for the promotion of health and well-being [9,10]. Although there is room for improvement, occupational therapy is among one of the fields that has begun to slowly integrate technology into its everyday practice. Apps are being used as intervention tools, education aides, and for support purposes [11]. A blog poll taken by the American Occupational Therapy Association (AOTA) indicated more than half of practitioners use apps at least occasionally in the clinic [12]. Many websites, including the AOTA website, provide lists of popular apps being used in occupational therapy.

OTs use apps as therapeutic tools inside and outside the clinical setting. Apps can be used to address problems in patients with developmental disorders, traumatic brain injury, stroke, and spinal cord injury. Some therapeutic apps address improving handwriting, fine motor dexterity, motor planning, sequencing, memory recall, social interaction, visual perception skills, and more [11]. Due to the constant development of apps, OTs need to be aware of the quality, efficacy, reliability, and security of apps they are using to ensure best practices and results [13]. It is especially important to consider the reliability and consistency of apps being used, both for clinical and evaluation purposes, as well as the safeguarding of data.

Scales classifying, and rating, the quality of apps are crucial to pinpoint top apps that should be promoted. The 5-star rating scale seen in many app stores has not been empirically proven to enable a potential user to decipher the quality of 1 app versus another [10]. The majority of rating scales aim to understand the user's perception of the app leading to subjective reviews and selection of apps based on popularity as opposed to quality [10]. Scales need to focus more on classifying the quality of the app and less on rating the developer's description. Therefore, researchers must take the time to download the app, familiarize themselves with its features, and rate the app accordingly using empirically supported evaluation tools.

Several studies have been conducted regarding mobile device use and decision making, mobile apps use, decision making, and opinions, and technology acceptance and utilization among OTs. Kelly Erickson [14] performed a systematic review of the literature specifically looking at the use of mobile app–based devices in occupational therapy practice. Erickson found limited research evidence related to the use of apps in practice. A total of 3 areas of focus for practitioners were established from the literature review: (1) The mobile app–based devices should be client centered, (2) The role of the OT is to show an individual the possibilities of mobile app–based devices, and (3) The OT should consider features of the mobile app–based device and the chosen apps. Erickson touches on the importance of choosing quality apps; however, the study does not specifically examine quality criteria for apps being chosen by OTs for their clients.

A study conducted by Ravenek and Alvarez [15] developed detailed considerations that can be used to inform OTs' decisions related to selecting appropriate apps for clinical practice. The considerations proposed allow OTs to weigh therapist, client, and app characteristics so that a specific app can be chosen for a specific client. Although these considerations are valuable for all OTs and important to consider for each client, the study does not discuss how OTs can objectively evaluate the quality of the apps they are individually choosing for their clients.

Objective

There is a paucity of knowledge surrounding OTs and their use of quality apps in the clinical setting. Although therapists do use their clinical judgement when using apps in therapy settings, such as target skills and preferred features, the objective quality of apps is still unknown [11]. Therapists are likely to achieve better results with clients if they choose to use quality apps; therefore, it is crucial for therapists to be able to distinguish low-quality apps from high-quality apps. To do so, OTs must be aware of resources available to assist them in evaluating apps they frequently use in their practices. The purpose of this study was to use the user version of the Mobile Application Rating Scale (uMARS) to investigate the quality of apps most frequently used by OTs and demonstrate the utility of the uMARS to assess the quality of mHealth apps.

Methods

Theoretical Perspective

Before mobile apps are adopted into clinical settings, they first should be evaluated to assess their overall quality. Doing so could help ensure high-quality apps are being used by practitioners and could eliminate some of the cost and time demands associated with the trial and error process of finding apps to use in therapeutic settings. However, there is limited information to guide evaluation of the quality of mHealth apps, and the evaluation tools that are available may not be well known or widely used by occupational therapy practitioners. To facilitate the use of mHealth app quality evaluation tools and the adoption of high-quality apps in therapeutic settings, a diffusion theory approach might be useful.

The diffusion of innovations theory is concerned with how ideas and practices are adopted over time through both formal and informal communication channels and processes. More specifically, diffusion theory can be used to explain how technology spreads as well as the rate at which technology is adopted by its intended audiences in higher education and clinical environments [16,17]. The use of mobile apps is still relatively new within the field of occupational therapy; therefore, not all therapists are willing to adopt and embrace this technology [15]. It is also possible there is a fair amount of uncertainty (a barrier to adopting an innovation) surrounding mHealth apps given the large number of mHealth apps currently available and the limited number of reliable app-quality assessment tools. According to diffusion of innovations theory, the rate of adoption of innovations is largely influenced by 5 innovation characteristics: (1) Relative advantage, (2) Compatibility, (3) Complexity, (4) Trialability, and (5) Observability [16,17]. By determining the quality and effectiveness of mobile apps with a reliable assessment tool, each of these characteristics can be targeted to resolve uncertainties and help facilitate the adoption of quality mobile app use in clinical settings through empirical evidence and informed decision making by OTs.

Diffusion theory also emphasizes the importance of a 2-stage communication approach. That is, information about an innovation "communicated to influential others, friends, relatives, opinion leaders, in the recognition that adoption can be influenced indirectly in this manner" [16,17]. Opinion leaders, also known as early adopters, are well respected and admired by others in their social network and generally possess a large amount of social influence. They are particularly important to the diffusion process because their opinions and adoption of certain behaviors influences the opinions and adoption of certain behaviors of others. Thus, opinion leaders play a crucial role in promoting adoption. It is important to consider the influence of such opinion leaders in the occupational therapy clinical setting and the impact they can have on the adoption of using a tool such as the uMARS to identify high-quality mobile apps. If the usefulness of the uMARS in identifying high-quality apps can be demonstrated to occupational therapy opinion leaders, it is likely they might adopt this innovation. In line with diffusion theory, their adoption of the uMARS might influence the adoption of the uMARS by other OTs [16,17].

Before development of the Mobile Application Rating Scale (MARS), few health-related app-quality assessment tools existed beyond the 5-star rating scale seen in application stores. Although the US Food and Drug Administration (FDA) does provide guidance regarding the development and use of mobile medical apps, as well as considerations for practitioners and clients, the FDA does not specifically provide a user-friendly assessment rating tool. Other organizations, such as the American Medical Association and the Healthcare Information and Management Systems Society, have developed mHealth app guidelines. There is no consensus, however, about which apps are the best to use or the highest quality. Some countries have developed systems or processes for health apps to be assesses for safety and/or quality. For example, Spain has the AppSoludable Quality Seal and the United Kingdom has their NHS Digital Apps Library.

Mobile Application Rating Scale and User Version of Mobile Application Rating Scale

The initial goal of the MARS was to create a tool that trained researchers could utilize to determine whether mHealth apps satisfied certain quality criteria instead of relying on the subjective 5-star rating system. Thus, the MARS was created as one of the first reliable and objective instruments for trialing, classifying, and rating the quality of mHealth apps [10]. The MARS provides a multidimensional measure of 4 objective quality app indicators: engagement, functionality, aesthetics, and information quality. It also includes a subject quality indicator. In addition to being easy-to-use, the MARS is widely applicable to various health domains and can be modified to measure the quality of apps with no relation to health. The MARS has demonstrated excellent internal consistency, interrater reliability, and validity [10].

Adapted from the 23-item MARS rating tool, the uMARS was developed as a simpler, more user-friendly alternative to the MARS tool. The MARS requires training and expertise in mHealth and the relevant health field to be administered [10]. The uMARS eliminates the need for trained experts and provides a reliable tool to assist app developers and researchers with assessing the quality of mHealth apps [10,13]. The scale consists of a 20-item measure including 4 objective quality subscales: engagement, functionality, aesthetics, and information quality. In addition, 1 subjective quality subscale and 1 6-item perceived impact subscale is included. The uMARS has good reliability, proven through test-retest studies, and excellent internal consistency (full scale Cronbach alpha=.90), with high individual alphas for all subscales [13]. The subjective subscales of the uMARS also have very high internal consistencies, with an engagement alpha of .80, a functionality alpha of .70, an aesthetics alpha of .71, and an information alpha of .78 [13]. The reliability for each subscale was highest for engagement, functionality, and aesthetics [13]. This indicates the uMARS provides an accurate measure of app quality for target users.

Data Collection

A total of 30 mobile apps were selected for evaluation. Mobile apps were initially chosen based on their frequency noted in a peer-reviewed research study performed by Seifert et al (2017) titled, "Apps in therapy: OTs' use and opinions" [11]. The cut-off for frequency noted was 4; therefore, OTs surveyed in the mentioned study [11] must have noted use of the mobile app 4 or more times for it to be included for evaluation. Additional inclusion criteria for app selection included the app was in English, available through the US Apple App store, compatible with iPad, and US $9.99 or under. A total of 5 apps were excluded because they did not meet these additional inclusion criteria. Thus, a total of 25 apps were selected for review. The 25 apps were reviewed by 2 trained uMARS evaluators. All apps were reviewed and evaluated on iPads because these are often the mobile devices used in clinical settings.

The principal researcher and another recruited student from the University of Florida evaluated each app (N=25) using the uMARS tool. The student evaluator was selected through convenience sampling. Both researchers of this study had previously attended a uMARS training session facilitated by researchers from a different study at the University of Florida. During the session, the trainees watched 3 video tutorials detailing the procedure for evaluating apps using the uMARS. During the training session, the researchers reviewed the uMARS rating tool and evaluated 2 trial apps to demonstrate appropriate mastery of the uMARS. Each trial app was examined for 10 min and then independently rated by the trainees. The training session lasted 60 min.

Upon completion of the training session, the principal researcher distributed the list of 25 mobile apps to the other evaluator. All 25 apps were then individually assessed by both evaluators according to 3 validated subscales on the uMARS: (A) Engagement, (B) Functionality, and (C) Aesthetics. These 3 subscales were chosen because of their internal consistencies and their test-retest reliabilities [13]. There was also a system in place in case discrepancies occurred among reviewers. To enhance reliability of evaluators' scores, both reviewers evaluated the apps separately then came together to discuss scores. Any discrepancies were then discussed until an agreement was made between the reviewers. In a study with more than 100 apps, inter-rater reliability is, and should be, measured.

Data Analysis

The uMARS rating tool was used to evaluate the quality of mobile apps most frequently used by OTs. Apps were rated using iPads, as they are the most common mobile device used in clinical settings. To collect and analyze descriptive and technical information about each app, 3 of the uMARS subscales were used: (A) Engagement, (B) Functionality, and (C) Aesthetics. For uMARS sections A, B, and C, items are rated on a 5-point scale (1-inadequate, 2-poor, 3-acceptable, 4-good, and 5-excellent). This study was performed to understand how user-friendly the uMARS could be for OTs (and the broader health professionals). This was based on the validated measurements of subscales A, B, and C, as explained in the previous section [13]. The researchers chose not to include the subjective measures of the uMARS. However, these are options for users to rate if they so desire.

The individual scores for each section were determined by calculating the mean of the ratings for each question in that designated section. This calculation provided the top ranked apps in each section. The total quality score of the app was determined by averaging the mean scores of the engagement, functionality, and aesthetics sections. This calculation provided the highest scoring apps overall.

Results

App Inclusion

A total of 30 mobile apps were considered at the start of this study. On the basis of inclusion criteria established by the researchers, 5 mobile apps were not included. Thus, a total of 25 mobile apps were evaluated using the uMARS for this study. Scores for engagement, functionality, and aesthetics were calculated using the 5-point scale. The scores described in the upcoming sections are somewhat high. As a reminder, these apps were already prescreened as apps used by the OT population [11]. It is likely these apps were of higher quality to begin with, which would help explain the high scores.

Engagement

Engagement criteria were evaluated based on entertainment, interest, customization, interactivity, and target group appeal. Entertainment means how fun and entertaining the app is to use and how the components making the app fun compare with similar apps. Interest means how interesting the app is to use and how the information is presented compared with similar apps. Customization means whether the app allows the user to customize settings and preferences such as sound, content, and notifications. Interactivity means whether the app allows user input, provides feedback, and contains prompts such as reminders, sharing options, and notifications. Target group appeal means whether the content (visuals, language, and design) is appropriate for the target audience.

Averages for engagement were calculated for all apps included in the study. The 10 apps with the highest average mean scores in terms of engagement can be found in Table 1. Fit Brains scored a perfect 5 for engagement. Fit Brains is highly entertaining and interesting to use, would stimulate repeated use, allows for the user to tailor all preferences and settings, has a high level of responsiveness through interactive features and feedback, and is designed specifically for its target audience. Apps scoring close to 5, Lumosity and Bugs & Buttons, demonstrated many of the same qualities previously mentioned.

Functionality

Functionality criteria were evaluated based on app performance, ease of use, navigation, and gestural design. Performance means how accurately and quickly the app features (functions) and components (buttons, menu) work. Ease of use means how easy it is to learn how to use the app and how clear the menu labels, icons, and instructions are to the user. Navigation means how logical the flow and movement between screens is for users and

whether the app has all the necessary links to navigate between screens. Gestural design means the taps, swipes, pinches, and scrolls make sense to the user and are consistent across all components and screens. The 10 apps with the highest average mean scores in terms of functionality can be found in Table 2.

Table 1. Top 10 averaged scores of occupational therapy apps evaluated in section A: engagement of the user version of the Mobile Application Rating Scale evaluation.

Occupational therapy app title	Average score
Fit Brains	5
Lumosity	4.7
Bugs & Buttons	4.4
Bugs & Bubbles	4.3
Writing Wizard	4.2
Ready to Print	4.1
Peekaboo Barn	4.1
Choice Works	4.1
Visual Attention	3.9
Handwriting Without Tears Wet, Dry, Try	3.8

Table 2. Top 10 averaged scores of occupational therapy apps evaluated in section B: functionality on the user version of the Mobile Application Rating Scale evaluation.

Occupational therapy app title	Average score
Bugs & Buttons	5
Ready to Print	5
Writing Wizard	5
Bugs & Bubbles	5
Handwriting Without Tears Wet, Dry, Try	5
Letter School	4.875
Letter Reflex	4.75
Cursive Writing Handwriting Without Tears HD Style	4.75
Fit Brains	4.75
Toca Kitchen	4.75

The mean scores for functionality were calculated to determine the top 10 apps in this category (Table 2). A total of 5 of the apps scored a perfect 5 (Bugs & Buttons, Ready to Print, Writing Wizard, Bugs & Bubbles, and Handwriting Without Tears [HWT] Wet, Dry, Try), which indicates a highly functional app according to uMARS criteria. The functionality category had the highest overall scores on the uMARS. The average of all 25 apps was 4.535, and none of the apps evaluated scored below a 3.625. This indicates a majority of the apps have fast performance features; have little or no technical bugs; are simple, intuitive, and easy to use immediately without instructions; have logical, easy, and clear screen flow; and have consistent gestural designs.

Aesthetics

Aesthetics of apps were evaluated based on layout, graphics, and visual appeal. Layout means how appropriate the arrangement and size of buttons, icons, menus, and content on the screen are for the user. Graphics means how high the quality and resolution of the graphics for buttons, icons, menus, and content appear in the app. Visual appeal is described as how good the app looks, as well as the overall stylistic consistency of the app. The 10 apps with the highest average mean scores in terms of aesthetics can be found in Table 3.

Over half (64%) of the 25 apps scored over a 4 on aesthetics. As seen in Table 3, 3 apps scored a perfect 5 (Bugs & Buttons, Bugs & Bubbles, and Toca Kitchen). These 3 apps had appropriate layouts with professional, simple, clear, orderly, and logically organized designs; very high-quality graphics, visual designs, and resolutions; and very attractive, memorable, and outstanding visual appeal. The lowest scoring app, Visual Timer, scored a 3 because of satisfactory layout; few problems selecting, locating, seeing, and reading items; moderate-quality graphic and visual design; and average visual appeal, which is neither pleasant nor unpleasant.

Table 3. Top 10 averaged scores of occupational therapy apps evaluated in section C: aesthetics on the user version of the Mobile Application Rating Scale evaluation.

Occupational therapy app title	Average score
Bugs & Buttons	5
Bugs & Bubbles	5
Toca Kitchen	5
Lumosity	4.83
Peekaboo Barn	4.83
Fit Brains	4.66
Letter School	4.5
Cursive Writing Handwriting Without Tears HD Style	4.46
Cursive Touch & Right	4.3
Write My Name	4.3

Table 4. Top 10 occupational therapy apps based on the overall average score from the user version of the Mobile Application Rating Scale evaluation.

Occupational therapy app title	Average score
Fit Brains	4.803
Bugs & Buttons	4.8
Bugs & Bubbles	4.766
Lumosity	4.718
Toca Kitchen	4.483
Writing Wizard	4.45
Peekaboo Barn	4.435
Ready to Print	4.366
Letter School	4.325
Handwriting Without Tears Wet, Dry, Try	4.31

Top Occupational Therapy Apps

The top 10 occupational therapy apps were chosen according to their average scores derived from the uMARS subscales of engagement, functionality, and aesthetics, each with its own independent validity (Table 4). A 5-point rating scale (1-inadequate, 2-poor, 3-acceptable, 4-good, 5-excellent) was utilized for each subscale, and then all 3 subscales were averaged to determine the final score for each app. Many of the highest scoring apps overall appeared on the top 10 lists of apps for each individual uMARS category.

The top apps all scored above a 4.3, with only a small amount of variance between each one. These apps are geared toward improving fine motor skills, spatial reasoning skills, or cognitive functioning skills. Overall, the apps scoring highest in functionality often scored lowest in engagement. In addition, 4 out of the 10 apps (Writing Wizard, Ready to Print, Letter School, and HWT Wet, Dry, Try) are all focused solely on improving handwriting.

Fit Brains was the highest scoring app overall with a uMARS score of 4.803 (Table 4). Fit Brains had high scores in all 3 categories, indicating a high-quality app according to the uMARS rating tool. The lowest scoring app overall had a uMARS score of 3.33. Out of the 25 apps scored for this study, all of their overall scores ranged from 3.33 to 4.03, meaning most of these apps have a relatively average or above average score.

Discussion

Principal Findings

The uMARS is a simple tool that can easily be used by end users to evaluate the quality of mHealth apps, including end users in clinical settings such as OTs. The uMARS utilizes multidimensional analyses to measure certain qualities of mobile apps, including engagement, functionality, and aesthetics. According to the study performed by Stoyanov et al [9], the uMARS has good internal consistency (alpha=.90) and high inter-rater reliability, thus indicating the uMARS is a reliable tool for quality ratings of apps used by OTs.

The purpose of this study was to evaluate the quality of the most frequently noted mobile apps used by OTs based on a previous study that surveyed OTs most frequently used mobile apps in therapy [11], as well as demonstrate the utility of the uMARS to assess the quality of mHealth apps. The results of this study indicate mobile apps should not be incorporated into clinical

settings solely based on frequency of use by OTs. Many of the apps analyzed in this study were not necessarily high-quality apps according to the uMARS. In the same way, many of the apps noted least frequently were not necessarily low-quality apps according to the uMARS analysis tool. The results also show how the uMARS can be used to score the quality of mHealth apps in an objective manner.

Multimedia Appendix 1 shows the comparison between apps most frequently used by OTs [11] and the uMARS scores of 25 of those most frequently used apps. The apps (middle column) in Multimedia Appendix 1 are presented in order of frequency (left column) with their respective uMARS score and rank (right column). On the basis of analyses performed in this study, there is no clear relationship between apps being noted more frequently and the quality of those apps being higher. Letter School, an app noted by OTs 69 times in the study by Seifort et al, scored an overall 4.325 and ranked 9th out of 25 apps. The overall highest scoring app on the uMARS, Fit Brains was only noted by OTs 5 times in the study by Seifert et al [11]. The lowest scoring app on the uMARS, however, was noted 6 times. Therefore, it seems there is an apparent disconnect between high-quality apps and their usage by OTs. It is imperative OTs utilize more appropriate quality measurements to determine which apps are best to use in a clinical setting. The uMARS offers a quick and easy way for OTs to measure such quality and select high-quality apps because just as simplicity is important so is a tool that requires little time to use.

Limitations

This study is not without limitations. As is the nature of mobile apps and mobile technology, apps and technologies are regularly undergoing changes and updates. Apps in the Apple App store are no exception. Since conducting the uMARS evaluation, it is possible many of the apps reviewed have been updated to newer versions. The updates on the most up-to-date versions could alter the results of this analysis. New features could have been added, aesthetic elements could have been changed, and glitches could have been fixed. This study also did not compare the market rating (5-star app store rating) with the quality scores from the uMARS. This information could provide important information about the accuracy and trustworthiness of app store ratings. Future research could be conducted comparing the market rating with an app's uMARS score. Moreover, new mobile apps could have been developed since the compilation of the apps reportedly used by OTs [11]. Updates and changes to mobile apps and technologies must be kept in mind when performing studies such as this one. If OTs are choosing to utilize mobile app technology in their clinical settings, however, it is also their responsibility to attempt to keep up with the ever-changing mobile app industry. Future research should build off of this study by using the uMARS tool to analyze the quality of new and improved mobile apps. This study only used 3 subscales of the uMARS tool to evaluate mobile apps. To add to this area of research, the uMARS tool in its entirety should be used to evaluate mHealth apps. There are also other considerations that go into incorporating mHealth apps into a clinical setting. This study looked at 1 aspect of these considerations (the quality of mHealth apps being used); however, it is important to note other aspects of mHealth apps should be considered. One aspect of importance is the security of data collected within mHealth apps. This is a critical feature not specifically measured through the uMARS but that must be considered when using mHealth apps in clinical settings.

The list of apps used in this study came from a previous study that surveyed 20 OTs in Ohio [11]. The majority (40.5%) of these OTs worked in pediatric settings. This resulted in a majority of the reported apps being used for a patient population aged younger than 12 years. Therefore, the results of this study may not be generalizable to mobile apps used by OT populations outside of pediatric settings. To address this limitation, future research should be expanded to include OTs and mobile apps from a variety of settings. The AOTA website offers a comprehensive list of apps for occupational therapy practitioners [12]. Although not quality reviewed, this might be a good place to start.

Conclusions

As the use, and introduction, of mobile apps continues to grow, it will become increasingly important for therapists to adopt high-quality apps. Consistent with diffusion theory, by determining the quality and effectiveness of mobile apps with a reliable assessment tool such as the uMARS, the 5 attributes of innovation (relative advantage, compatibility, complexity, trialability, and observability) can be targeted to help facilitate widespread adoption of quality mobile app use in occupational therapy clinical settings [16,17]. To promote the use of the uMARS tool in clinical settings, the influences of opinion leaders (eg, early adopters) and their respective social networks should be considered [16,17]. By encouraging OT opinion leaders to adopt the use of the uMARS in their clinical setting, it is likely these influencers will intentionally or inadvertently influence other OTs to also adopt the use of the uMARS [16,17].

Future research can focus on the top-rated apps found in this study to determine effectiveness in therapy. It is important to understand the implications a high-quality app can have as a complement to the occupational therapy services provided to a patient. Investigating whether or not utilizing mobile apps in clinical settings is helpful to the overall rehabilitation of a patient is crucial. These apps could also be compared in different settings to examine if some mobile apps are best suited for 1 setting instead of another. In addition, it is important to determine if using higher rated apps results in increased occupational therapy gains. As the overall goal is to benefit the OTs and the patients receiving therapy, future research is needed to understand if the quality of the app is directly related to how well a patient does in therapy. The evidence supporting the effectiveness of using higher quality apps has been addressed minimally in research. Understanding the effectiveness of high-quality apps compared with low-quality apps could inform practitioner's decisions about using mHealth apps in therapy. Furthermore, qualities of apps that ranked higher could be used to improve existing apps or help with the development of new ones. As apps in this study scored lowest in engagement, focusing on improving engagement among existing apps could be a good place to start. Future apps could learn from this research by ensuring user engagement is prioritized during the development phase. In addition, this study can be used as a

guide for using uMARS for future evaluations of apps utilized in occupational therapy clinical settings. In this way, OTs will have access to a validated tool to help them choose high-quality apps to use for their specific patient population.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Comparison of the top 25 most frequently noted apps by occupational therapists with their total user version of the Mobile Application Rating Scale (uMARS) score and respective rank based on uMARS score.

[PDF File (Adobe PDF File), 42KB - mhealth_v7i5e13019_app1.pdf]

References

1. Bureau of Labor Statistics. Occupational Outlook Handbook, Occupational Therapists. US Department of Labor URL: https://www.bls.gov/ooh/ [accessed 2018-11-27] [WebCite Cache ID 74EMiqm3S]
2. World Federation of Occupational Therapists (WFOT). About Occupational Therapy URL: https://www.wfot.org/about-occupational-therapy [accessed 2019-04-24] [WebCite Cache ID 77sUr100l]
3. Schaper LK, Pervan GP. ICT and OTs: a model of information and communication technology acceptance and utilisation by occupational therapists. Int J Med Inform 2007 Jun;76(Suppl 1):S212-S221. [doi: 10.1016/j.ijmedinf.2006.05.028] [Medline: 16828335]
4. Pew Research Center. 2018 Feb 05. Mobile Fact Sheet URL: https://www.pewinternet.org/fact-sheet/mobile/ [accessed 2019-04-24] [WebCite Cache ID 77sVIL7mA]
5. Higgins JP. Smartphone Applications for Patients' Health and Fitness. Am J Med 2016 Jan;129(1):11-19. [doi: 10.1016/j.amjmed.2015.05.038] [Medline: 26091764]
6. Statista. 2017. Mobile App Usage-Statistics & Facts URL: https://www.statista.com/topics/1002/mobile-app-usage/ [accessed 2018-11-27] [WebCite Cache ID 74ENxiL8s]
7. Aitken M, Clancy B, Nass D. IQVIA Institute Report. 2017 Nov 07. The growing value of digital health: evidence and impact on human health and the healthcare system URL: https://www.iqvia.com/gated-form-page/institute-gated-form-page?redirectUrl=%2f-%2fmedia%2fiqvia%2fpdfs%2finstitute-reports%2fthe-growing-value-of-digital-health.pdf%3f_%3d1548711898666&Name=The+Growing+Value+of+Digital+Health [accessed 2019-04-24] [WebCite Cache ID 77sW4aVyA]
8. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
9. Batista MA, Gaglani SM. The future of smartphones in health care. Virtual Mentor 2013 Nov 01;15(11):947-950. [doi: 10.1001/virtualmentor.2013.15.11.stas1-1311] [Medline: 24257085]
10. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. JMIR Mhealth Uhealth 2015 Mar 11;3(1):e27 [FREE Full text] [doi: 10.2196/mhealth.3422] [Medline: 25760773]
11. Seifert AM, Stotz N, Metz AE. Apps in therapy: occupational therapists' use and opinions. Disabil Rehabil Assist Technol 2017 Dec;12(8):772-779. [doi: 10.1080/17483107.2016.1262912] [Medline: 27982716]
12. Yamkovenko S. American Occupational Therapy Association. Apps for Occupational Therapy: Find Apps for Your Practice area URL: https://www.aota.org/Practitioners/Resources/Apps.aspx [accessed 2018-11-27] [WebCite Cache ID 74EQG05gc]
13. Stoyanov SR, Hides L, Kavanagh DJ, Wilson H. Development and validation of the user version of the Mobile Application Rating Scale (uMARS). JMIR Mhealth Uhealth 2016 Jun 10;4(2):e72 [FREE Full text] [doi: 10.2196/mhealth.5849] [Medline: 27287964]
14. Erickson K. Evidence considerations for mobile devices in the occupational therapy process. Open J Occup Ther 2015;3(2) [FREE Full text] [doi: 10.15453/2168-6408.1132]
15. Ravenek M, Alvarez L. Use of mobile ‘apps’ in occupational therapy: therapist, client and app considerations to guide decision-making. World Federation of Occupational Therapists Bulletin 2016 Mar 31;75(1):43-49. [doi: 10.1080/14473828.2016.1162430]
16. Rogers EM. Diffusion of Innovations, Fourth Edition. New York, NY: The Free Press; 1995.
17. Simons-Morton B, McLeroy KR, Wendel ML. Behavior Theory in Health Promotion Practice and Research. Burlington, MA: Jones & Bartlett Learning; 2012.

Abbreviations

AOTA: American Occupational Therapy Association
HWT: Handwriting Without Tears
MARS: Mobile Application Rating Scale
mHealth: mobile health
OT: occupational therapist
uMARS: user version of the Mobile Application Rating Scale

Edited by G Eysenbach; submitted 03.12.18; peer-reviewed by K Erickson, F López Seguí; comments to author 10.01.19; revised version received 05.03.19; accepted 24.03.19; published 01.05.19.

Please cite as:
LeBeau K, Huey LG, Hart M
Assessing the Quality of Mobile Apps Used by Occupational Therapists: Evaluation Using the User Version of the Mobile Application Rating Scale
JMIR Mhealth Uhealth 2019;7(5):e13019
URL: https://mhealth.jmir.org/2019/5/e13019/
doi:10.2196/13019
PMID:31066712

Original Paper

Smartphone Apps to Support Self-Management of Hypertension: Review and Content Analysis

Tourkiah Alessa[1,2], MSc; Mark S Hawley[1], PhD; Emma S Hock[3], PhD; Luc de Witte[1], PhD

[1]Centre for Assistive Technology and Connected Healthcare, School of Health and Related Research, University of Sheffield, Sheffield, United Kingdom
[2]Biomedical Technology Department, College of Applied Medical Sciences, King Saud University, Riyadh, Saudi Arabia
[3]Health Economics and Decision Science, School of Health and Related Research, University of Sheffield, Sheffield, United Kingdom

Corresponding Author:
Luc de Witte, PhD
Centre for Assistive Technology and Connected Healthcare
School of Health and Related Research
University of Sheffield
30 Regent Street
Sheffield,
United Kingdom
Phone: 44 1142224399
Email: l.p.dewitte@sheffield.ac.uk

Abstract

Background: Hypertension is a widespread chronic disease, and its effective treatment requires self-management by patients. Health-related apps provide an effective way of supporting hypertension self-management. However, the increasing range and variety of hypertension apps available on the market, owing to the global growth in apps, creates the need for patients and health care professionals to be informed about the effectiveness of these apps and the levels of privacy and security that they provide.

Objective: This study aimed to describe and assess all available apps supporting hypertension self-management in the most popular app stores and investigate their functionalities.

Methods: In January 2018, the UK Apple and Google Play stores were scanned for all free and paid apps supporting hypertension self-management. Apps were included if they were in English, had functionality supporting hypertension self-management, and targeted adult users with hypertension. The included apps were downloaded and their functionalities were investigated. Behavior change techniques (BCTs) linked with the theoretical domain framework (TDF) underpinning potentially effective apps were independently coded by two reviewers. The data privacy and security of the apps were also independently assessed.

Results: A total of 186 hypertension apps that met the inclusion criteria were included in this review. The majority of these apps had only one functionality (n=108), while the remainder offered different combinations of functionalities. A small number of apps had comprehensive functionalities (n=30) that are likely to be more effective in supporting hypertension self-management. Most apps lacked a clear theoretical basis, and 24 BCTs identified in these 30 apps were mapped to 10 TDF mechanisms of actions. On an average, 18.4 BCTs were mapped to 6 TDF mechanisms of actions that may support hypertension self-management behaviors. There was a concerning absence of evidence related to the effectiveness and usability of all 186 apps, and involvement of health care professionals in the app development process was minimal. Most apps did not meet the current standards of data security and privacy.

Conclusions: Despite the widespread accessibility and availability of smartphone apps with a range of combinations of functionalities that can support the self-management of hypertension, only a small number of apps are likely to be effective. Many apps lack security measures as well as a clear theoretical basis and do not provide any evidence concerning their effectiveness and usability. This raises a serious issue, as health professionals and those with hypertension have insufficient information to make decisions on which apps are safe and effective.

(JMIR Mhealth Uhealth 2019;7(5):e13645) doi:10.2196/13645

KEYWORDS

smartphone apps; mobile apps; self-management; hypertension; blood pressure; mobile applications

Introduction

Internationally, hypertension is one of the most common chronic diseases in adults and is considered one of the main risk factors for numerous diseases such as stroke, heart disease, and renal failure [1]. It is estimated that around one billion individuals live with hypertension worldwide, and the majority of people are not proficient at controlling their blood pressure (BP) through medication or lifestyle choices, despite the fact that lowering BP decreases the risk of renal and cardiovascular disease [2]. Self-management is considered to be among the most effective methods of coping with hypertension, helping individuals with hypertension be more responsible for their own health [3].

The recent emergence of information and communication technologies such as mobile health supports the self-management of chronic conditions [4-7]. The increase in smartphone devices over the past decades has been rapid: By 2022, it is predicted that there will be 6.8 billion smartphone users [8]. This rapid increase of smartphone users corresponds with an increase in health apps offering health services and information [9,10].

Many apps have become available for patients with hypertension, and their number is increasing rapidly [11,12]. The majority of these smartphone apps are aimed at helping people manage and control their hypertension [11,12], but it is currently unclear to what extent the evidence supports their effectiveness. A recent systematic review of apps aimed at supporting the self-management of hypertension found few studies reporting the effectiveness of apps [13]. The majority of the apps in this review were study-specific and are therefore not available commercially in the app stores. The review concluded that apps containing more comprehensive functionalities, defined as three or more functionalities, are more likely to be effective in lowering BP [13] than apps with only one or two functionalities.

Even though many of the apps lacked evidence of theoretical underpinning, an examination of their functionalities revealed recognizable elements of behavioral change strategies [13]. Studies have shown that self-management programs are more likely to be effective if they are supported by theory-based interventions [14-16]. Theory allows identification of target behavior and strategies of behavioral changes needed to achieve desirable health outcomes. However, research has revealed that many commercial health apps lack theoretical underpinnings and theoretically consistent use of behavior change techniques (BCTs) [17-19]. In addition, the majority of health apps lack privacy and security measures that adequately ensure protection of users' data, posing risks to user confidentiality [19,20]. This is problematic, as it compromises both the personal data of the user as well as their trust in the app.

These shortcomings might lead to significant concerns about apps having little to no benefit, or even presenting a risk to users [17], highlighting the necessity of providing adequate information about the effectiveness of these apps and the robustness of their privacy and security features for patients and health care professionals. As such, these findings increase the importance of characterizing and investigating potential theoretical mechanisms of action in existing commercial apps with comprehensive functionalities as well as assessing the privacy and security of such apps. A method of investigating potential theoretical mechanisms of action by grouping BCTs with theoretical domain framework (TDF) mechanisms of actions, using the TDF and BCT Taxonomy (v1) (BCTTv1), has been extensively employed to characterize BCTs in health interventions [17,21-23], especially those relating to chronic diseases.

A review by Kumar et al [12] searched for the most downloaded and popular apps in May 2014 and found there are many apps that support the self-management of hypertension by offering self-monitoring activities, feedback, reminders, and tailored information. However, the search was restricted to the 50 most popular apps for every search term (high blood pressure and hypertension) on the two smartphone platforms. As a result, only 200 apps were screened, excluding many apps that might be suitable to support people with hypertension in their self-management. Furthermore, this review excluded smartphone app–based BP-measuring devices, arguing that they lacked accuracy, despite evidence that some of these specific devices used for measuring BP have been found to be accurate [24,25].

This study has reviewed all the available apps, updated our knowledge of new apps related to hypertension, and described their main functionalities as well as functionality combinations. Even though apps have numerous potential benefits and are used by an increasing number of patients, to the best of our knowledge, no previous review has analyzed all available apps; considered functionality combinations; included apps associated with accessory devices; considered the link between BCTs and the TDF mechanisms of action, which underpin the potentially effective apps; and considered privacy and security assessment of the potentially effective apps. The aim of this study was to fill this knowledge gap by addressing these points.

Methods

Study Design

This study is a content analysis and review of apps supporting hypertension self-management available in the most popular app stores. The *Quality and Risk of Bias Checklist for Studies That Review Smartphone Applications* [26] was utilized to ensure the adequate description of the app review's methods.

App Identification

Overview

In January 2018, an electronic search of apps was undertaken on the app stores of the two major types of smartphones in the United Kingdom—the iPhone (Apple Store) and Android (Google Play). These two platforms were searched because they were the world's most used operating systems in 2017 [27]. The terms "hypertension" and "high blood pressure" were separately searched for in both stores. There were no restrictions concerning subcategories like "health and wellness."

Inclusion and Exclusion Criteria

An app was included based on the following criteria: (1) The description was written in English, and "hypertension" or "high blood pressure" was included in the keywords or the accompanying description. (2) The collected data provided feedback, connected with health care professionals, or informed the patient about hypertension and self-management tasks related to hypertension; such tasks involve the self-monitoring of BP and other health data including healthy diet, exercising, taking medications, maintaining an appropriate weight, and managing stress. (3) The app was aimed at adults, in general, rather than health care providers (HCPs) specifically. Both paid and free apps were considered in the study.

An app was excluded based on one of the following criteria: (1) if it was not targeted at hypertension or if it focused solely on hypertension during pregnancy or primary prevention of hypertension; (2) if it was described in the app store catalogue as a "prank app" because it was not designed for medical purpose, but for entertainment; (3) if it was not designed for general use, for example, if it only provided services offered by particular hospitals or was designed only to be used as part of a specific study; and (4) if it did not run properly or required identification access after downloading the app, such as personal identification or primary care/hospital number.

The researcher selected the basic, completely functional version of an app if it had more than one version, such as high definition, lite, or pro. Apps appearing in both stores were rated independently to account for differences in features supported by various mobile operating systems. If an app appeared in response to searches for both "hypertension" and "high blood pressure" by a platform, it was included once, not twice.

Screening and Selection of Apps

All apps that were identified through the search and met the inclusion criteria based on their title and description were downloaded by the researcher (TA) onto an iPhone 6 (running iPhone operating system, version 11.2.2; iOS, Apple Inc, Cupertino, CA) and Android Samsung Galaxy S7 (running 8.1 software; Seoul, South Korea). The apps were then screened for all exclusion and inclusion criteria. If they met the criteria, the apps were run for 2 days, so that the researcher could investigate all reminders or notifications that appeared. Data on all the included apps were charted.

Data Abstraction

Overview

Abstracted data for all identified apps involved the name of the app, developer, version date, price and functions, available languages, and number of downloads. The involvement of health care professionals (eg, medical/health care professionals and behavior change specialists) was determined based on whether health professionals were involved in the development of the app as well as user involvement, which was included in the description on the app store. Following data abstraction, potentially effective apps (apps that were found to have comprehensive functionalities) were selected and considered for further analysis.

Functionalities

App functionalities were categorized based on the functionalities of hypertension self-management that have been determined in several previous studies about hypertension apps [11-13] and examined for effectiveness in scientific trials [28,29]. The functionalities that were considered in this study are self-monitoring, goal setting, reminders, educational information, feedback, stress management, communication with HCPs and others, and export of users' data to others via email.

Apps Considered for Further Analysis

According to Alessa et al [13], apps with comprehensive functionalities are more likely to be effective. Such apps were identified on the basis of the presence of three or more functionalities, including (but not limited to) self-monitoring, reminders, and educational information or automatic feedback. Therefore, of the apps originally identified, apps that were found to have comprehensive functionalities were considered for further analysis. These apps were then analyzed to assess their privacy, security, and theoretical underpinning. This is because theoretical underpinning and privacy as well as security measures are essential criteria for apps to be used in health care [19].

Privacy and Security

Privacy and security were assessed based on the Online Trust Alliance [30] and the recommendations of the Information Commissioner's Office [31]. These recommendations consist of seven questions examining the accessibility and availability of the privacy policy, the practices of data sharing and collecting, and data security as interpreted in the privacy and security statement (Table 3). These assessment questions and recommendations have also been previously used to assess privacy and security of existing health apps [19]. The assessment was conducted by two independent reviewers (TA and EH). Any discrepancies were resolved by discussion with the other researchers (LdW and MSH).

Theoretical Underpinning

To identify the mechanisms of action underpinning the existing apps with comprehensive functionalities, the BCT v1 Taxonomy (BCTTv1) was used to code the content of the app and extract the number of BCTs in each app and the frequency of use of each BCT in the apps. Each BCT was coded with "0" as Absent or "1" as Present [18]. The coding was undertaken by the two reviewers. Any disagreements were resolved by discussion with the other researchers. Interrater reliability for the presence or absence of the BCTs was assessed by calculating Cohen kappa for each item.

The present BCTs were then mapped to mechanisms of action of the TDF, based on several previously published expert consensuses linking BCTs to TDFs domains for health interventions, and the agreed judgement (consensus) of the study's researchers [23,32-34]. The linking of BCTs to TDF was conducted independently by the two reviewers, and any discrepancies were resolved by discussion with other researchers of the study team. The final results were then agreed upon by the research team.

Additional Aspects

Two additional characteristics/aspects for the selected apps were also described—US Food and Drug Administration (FDA) or European Union Conformite Europeene (CE) approval and their individual user rating.

Statistical Analysis

The number and frequency of BCTs and TDF used in the reviewed apps were summarized as the SD, mean, and median using Microsoft Excel (Microsoft Corp, Redmond, WA). Proportions were also used to summarize the variables, including app functionalities, user ratings, and data privacy and security.

Results

Summary of Search Results

The study steps are summarized in Figure 1. A search of the two app stores yielded a total of 775 apps (495 in Android Google Play Store and 280 in iPhone Apple Store). The titles and descriptions of these apps were screened for eligibility. A total of 564 apps were excluded because they did not meet the specific inclusion criteria. The 211 remaining apps (116 in Google Play and 95 in Apple Store) were considered for further analysis (installed). Subsequently, a total of 25 apps (11 in Google Play and 14 in App Store) were excluded because of registration problems (eg, requiring specific identification access such as hospital or primary care identification) or installation failure. The remaining 186 apps (106 in Google Play and 80 in App Store) were included in the review.

The cost of the apps varied. Over a quarter of the apps (27.9%) cost between £0.59 to £17. Most apps (134/186, 72.1%) were free to download. Of the apps that were free to download, 19 either were trials of the complete app or required subscription fees. All apps (n=186, 100%) were in English, although some also supported other languages such as Chinese, German, and Russian.

General App Characteristics

Of the 186 apps that met the selection criteria, more than half (106/186, 57%) were available through the Android operating system. The remaining apps were available through the iPhone operating system (80/186, 43%). Only 11 apps were found to be available on both platforms (Multimedia Appendix 1).

The version date of the reviewed apps ranged from February 8, 2012, to February 13, 2018. According to the number of downloads, more than half of the included Android apps (60/106, 57%) had over 500 downloads. Information on the number of downloads was not available for Apple apps.

Apps' Purpose and Functionalities

All apps could be classified according to their functionalities, including stress management, communication with HCPs and others, self-monitoring abilities, reminders, automatic feedback, educational information, and goal setting. Each app had at least one of these functionalities (Multimedia Appendix 1).

Table 1 summarizes the frequency of functionalities across the included apps. The most common self-management functionality was educational information about hypertension (110/186, 59.1%). Educational content varied across apps. Most included basic educational information on high BP or information on diet and food (eg, dietary approaches to stop hypertension). Some apps contained general information on hypertension or alternative treatments. Although the majority of educational material was text-based, several apps contained video and images to depict their content.

The second most common functionality was self-monitoring (99/186, 53.2%), which allows users to monitor their BP and other data over a period of time presented in different forms, including graphs or tables, and to see an overview. The majority of these apps (n=94) aided BP tracking, while some of them also supported the self-monitoring of other data concerning medication, nutrition, physical activity, weight, and emotions. A few apps (n=5) only focused on tracking medication compliance, potassium intake, or sodium intake. Seven of these apps received BP readings automatically from the BP measurement device but do not provide a manual entry function. Of the remaining 84 apps, 73 necessitated manual entry of BP data, and 11 allowed both manual and automatic data transfer. Notably, a few apps (3/186, 1.6%) claimed that they turn the smartphone into a device capable of recording BP data. This was presumably achieved by using a "cuffless technique" in which the user puts a finger over the camera of their smartphone. Despite most of these apps claiming to measure BP, they did not report any evidence of their reliability and validity.

The third most common functionality was the provision of automatic feedback (52/186, 28%). This feedback was provided to users in different ways, either through self-care messages and notifications or by representing data in distinct color codes to inform the user of whether measurements have diverged from the average level.

One-fifth of the apps (39/186, 21%) had a functionality to remind users about BP measurements, their hospital appointments, their medication time(s), and personal goals. Certain apps (10/186, 5.4%) included BP goal setting, and a few also enabled the user to set other goals such as blood glucose levels, weight, and physical activity. A few apps (5/186, 2.7%) provided a tool for communication with others, including HCPs or friends, through text messaging, chats, or virtual meetings with coaches. Five apps (2.7%) also supported stress management by providing relaxation tips or other therapies.

Around one-fourth of the apps (51/186, 27.4%) allowed users to export their entered data over time directly to others, including physicians, via email and other apps such as "WhatsApp," thus facilitating patient-physician communication.

As shown in Table 2, the majority of the apps (n=108) included only one functionality such as educational information (n=82), self-monitoring (n=25), or stress management (n=1). Almost one-fourth (45/186, 24.1%) of the apps combined two functionalities, while a small number of apps (33/186, 17.8%) included comprehensive functionalities (ie, three or more functionalities). None of the 33 apps included all 8 abstracted functionalities. Thirty of these apps included, among other functionalities, self-monitoring and reminders, with educational information (5/186, 2.7%), automatic feedback (16/186, 8.6%), or both (9/186, 4.8%).

Figure 1. Flow diagram of the app-search process. ID: identification; HCP: health care provider.

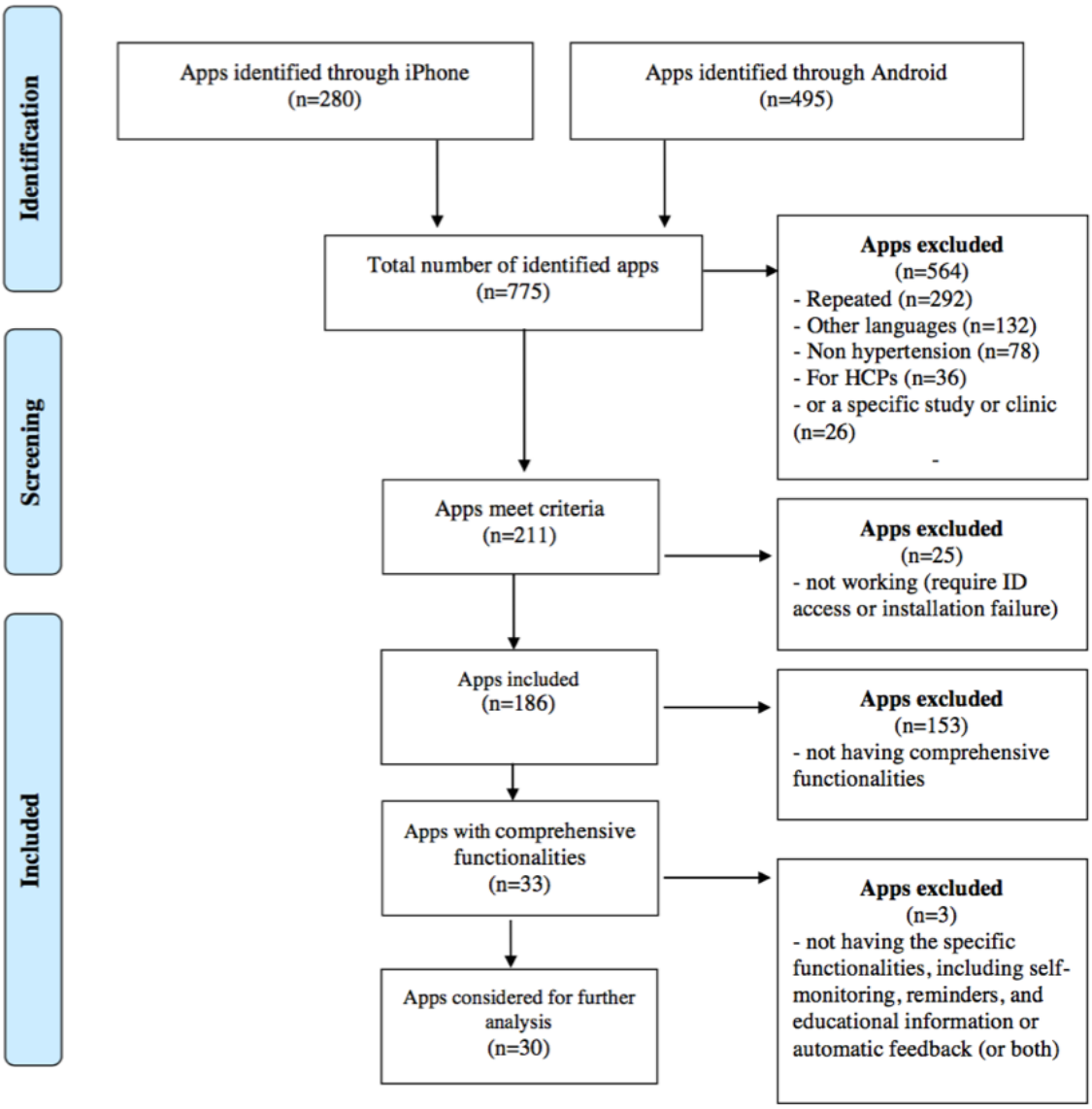


Table 1. Frequency of app functionalities.

Functionality	iPhone (Apple; N=80), n (%)	Android (Google Play; N=106), n (%)	Total (N=186), n (%)
Educational information	34 (43.8)	76 (70.8)	110 (59.1)
Self-monitoring	56 (70)	43 (40.6)	99 (53.2)
Feedback	36 (43.8)	16 (15.1)	52 (28)
Export	29 (36.3)	22 (20.8)	51 (27.4)
Reminder	23 (28.8)	16 (15)	39 (21)
Goal setting	8 (10)	2 (1.9)	10 (5.4)
Stress management	3 (3.8)	2 (1.9)	5 (2.7)
Communication with others	4 (5)	1 (0.9)	5 (2.7)

Table 2. Common combinations of app functionalities.

Functionality combinations	Number of functionalities	iPhone (N=80), n (%)	Android (N=106), n (%)	Total number of apps used in combination, n (%)
Educational informational	1	21 (26.3)	61 (57.5)	82 (44.1)
Self-monitoring	1	10 (12.5)	15 (14.1)	25 (14)
Stress management	1	1 (1.3)	0 (0)	1 (0.5)
Self-monitoring + Feedback	2	18 (22.5)	6 (5.7)	24 (12.9)
Self-monitoring + Educational information	2	3 (3.8)	6 (5.7)	9 (4.8)
Self-monitoring + Reminder	2	5 (6.25)	3 (2.8)	8 (3.8)
Educational information + Communication with others	2	2 (2.5)	0 (0)	2 (1.1)
Educational information + Stress management	2	0 (0)	1 (0.9)	1 (0.5)
Educational information + Reminders	2	0 (0)	1 (0.9)	1 (0.5)
Self-monitoring + Reminder + Feedback	3	5 (6.3)	3 (2.8)	8 (4.3)
Self-monitoring + Reminder + Educational information	3	2 (2.5)	3 (2.8)	5 (2.7)
Self-monitoring + Feedback + Communication with others	3	1 (1.3)	1 (0.9)	2 (1.1)
Self-monitoring + Reminder + Feedback + Goal setting	4	5 (6.3)	2 (1.9)	7 (3.8)
Self-monitoring + Reminder + Feedback + Educational information	4	3 (3.8)	3 (2.8)	6 (3.2)
Self-monitoring + Feedback + Educational information + Goal setting	4	1 (1.3)	0 (0)	1 (0.5)
Self-monitoring + Reminder + Feedback + Goal setting + Communication with others	5	1 (1.3)	0 (0)	1 (0.5)
Self-monitoring + Reminder + Feedback + Educational information + Stress management	5	1 (1.3)	1 (0.9)	2 (1.1)
Self-monitoring + Reminder + Feedback + Educational information + Stress management + Goal setting	6	1 (1.3)	0 (0)	1 (0.5)

The most frequently used combination of functionalities was self-monitoring with automatic feedback (24/186, 12.9%). The second most common combination was self-monitoring and educational information (9/186, 4.8%).

Involvement of Health Care Professionals and Users in App Development and Scientific Evaluation

Six apps (3.2%) claimed to have had contributions from an HCP or medical organizations during their development; the other apps did not. No apps reported end-user involvement (eg, hypertensive patients) in their development. None of the apps appeared to have been scientifically evaluated. The description provided indicates that there is an absence of evidence concerning the effectiveness or usability of apps designed to help manage hypertension.

Data Security and Privacy

Accessibility and Availability of Privacy Policy

Of the 30 apps in the study that had comprehensive functionalities, the availability of a privacy policy in English was found in 20 apps (66.6%; Table 3). Of the 10 apps without an English-language privacy policy, only one provided a link to such a policy, but the link was not functional. Further, 4 of these apps provided a privacy policy in non-English languages.

The short-form notice indicating key data practices was not applicable to the 20 apps that provided a privacy policy, since the policies were already concise. Apps rarely offered multilingual policies, with only one app offering a policy in two other languages.

Table 3. Data privacy and security assessment of apps (data gathering, sharing, and security) on the basis of the description in the privacy policy.

Privacy and security question	iPhone (N=12), n (%)	Android (N=8), n (%)	Total (N=20)[a], n (%)
Is the privacy policy available without the need to download the app?			
No	0 (0)	0 (0)	0 (0)
Yes	12 (100)	8 (100)	20 (100)
Is the privacy policy available within the app?			
No	5 (42)	2 (25)	7 (35)
Yes	7 (58)	6 (75)	13 (65)
Is there a short form notice (in plain English) highlighting key data practices?			
No	0 (0)	0 (0)	0 (0)
Yes	0 (0)	0 (0)	0 (0)
Not applicable	12 (100)	8 (100)	20 (100)
Is the privacy policy available in any other languages?			
No	11 (92)	8 (100)	19 (95)
Yes	1 (8)	0 (0)	1 (5)
Does the app collect personally identifiable information?			
No	1 (8)	0 (0)	1 (5)
Yes	10 (83)	6 (75)	16 (80)
Not specified	1 (8)	2 (25)	3 (15)
Does the app share users' data with a 3rd party?			
No	0 (18)	0 (11)	0 (15)
Yes	8 (67)	6 (75)	14 (70)
Not specified	4 (33)	2 (25)	6 (30)
Does the app say how the users' data security is ensured? For example, encryption, authentication, and firewall			
No	6 (50)	5 (62)	11 (55)
Yes	6 (50)	3 (38)	9 (45)

[a]Only 20 apps had a privacy policy; 10 apps did not have a privacy policy available.

Data Gathering and Sharing

Sixteen of the 20 apps with a privacy policy in English (80%) disclosed the collection of personally identifiable information such as age. In three other apps, the data-gathering practices were not discussed. One app did not report personal data gathering.

The developers of 14 apps revealed that they shared the data they gathered with third parties and discussed sharing practices. In three apps, data-sharing practices were not discussed. In three other apps, the policies stated that data would not be shared, except in exceptional cases that were general and vague. Despite reporting that they did not share data, except in exceptional circumstances, we believe that they share data without specifically discussing their data-sharing practices.

Data Security

Almost half (9/20) of the apps reported how consumer data were secured. In these cases, the privacy policies explained that data safety and security are essential to their practices and that users' data have been encrypted, anonymized, or accessed only by authorized persons.

Behavior Change Techniques and Theoretical Domain Framework

Presence of Behavior Change Techniques

We identified 24 BCTs in the 30 reviewed apps featuring comprehensive functionalities (Table 4). The Cohen kappa for agreement between the two reviewers for coding BCTs was 0.85.

The total number of BCTs in each app ranged between 6 and 17 BCTs, with a mean of 18.4 (SD 2.6) and a median of 9. The most frequently used BCTs were "Self-monitoring of behavior," "Prompts/cues," and "Action planning." These were present in all 30 reviewed apps. The next most frequently used BCTs were "Feedback on behavior" and "Monitoring of behavior by others without feedback," which were present in 25 and 24 apps, respectively. Two of these 24 BCTs ("Social comparison" and "Demonstration of the behavior") were present only once. Table 4 presents the frequency of BCTs used in these 30 apps.

Table 4. Behavior change techniques (N=24) used in the reviewed apps (N=30).

Behavior change technique	Most common function of the app	Number of apps
Self-monitoring of behavior	Allows users to frequently record and self-monitor the performed behavior of their health in a diary, including measuring BP[a], weight, emotion, and record if they took medication or other behaviors	30
Prompts/cues	An alarm is activated when it is time to perform a task with the purpose of cueing and promoting the behavior	30
Action planning	Setting a reminder to perform a task(s) (taking medication, measuring BP, exercise, etc) at a specific time	30
Feedback on behavior	Provides feedback on the entered data by representing data in different color codes or through self-care messages	25
Monitoring of behavior by others without feedback	Allows others to consensually observe the performed behavior to support the management of hypertension	24
Self-monitoring of outcomes of behavior	Allows users to monitor BP readings (eg, average) and view trends over time	22
Feedback on outcomes of behavior	Provides feedback on behavior outcomes over time (eg, average) to allow users to view their health status	17
Pharmacological support	Setting a specific reminder to facilitate medication adherence at a specific time	15
Information about health consequences	Offers educational information about the health benefits and consequences of controlling and managing hypertension	14
Instruction on how to perform a behavior	Provides overall orientation on hypertension management (including how to self-monitor BP) as well as other behaviors	10
Focus on past success	Offers the number of cases where the BP level has been successfully controlled or the successful achievement specific goals	10
Credible source	Contains information from credible sources	9
Goal setting (outcome)	Allows users to set health goals for controlling BP	8
Problem solving	Enables users to analyze BP with other behaviors trends and develop the knowledge required to understand how to achieve optimal BP level by identifying problems that hinder BP control	8
Goal setting (behavior)	Allows users to set goals for the behavior to be attained, including times to measure BP, weight, exercise, and food goals	7
Review outcome goal(s)	Allows users to examine how well the BP level was controlled according to the agreed goal and consider modifying outcome goals accordingly	5
Review behavior goal(s)	Allows users to modify their goals according to their achievements	5
Social support (unspecified)	Offer a space to chat with others (eg, friends or families)	3
Social incentives	Add points or provides badges for users when achieving their target goals and tasks	3
Salience of consequences	Shows pictures of health consequences such as dizziness to shed light the dangers of uncontrolled BP	2
Monitoring of emotional consequences	Allows users to record their feeling after performing tasks including measuring BP or exercises	2
Reduce negative emotions	Provides advice on the ways to minimize negative emotions	2
Demonstration of the behavior	Offers an observable sample of the performance of the behavior with the help of pictures for the person to aspire to or imitate	1
Social comparison	Allows comparison of the user's own performance with others by sharing his/her performance to draw attention to the performance of others	1

[a]BP: blood pressure.

Mechanisms of Action of the Theoretical Domain Framework

BCTs present in the 30 reviewed apps could be linked to 10 TDF mechanisms of action. The number of TDF mechanisms of action underlying each app varied, ranging from 5 to 9, with a mean (SD) of 6 (1) and a median of 6.

The most common TDF mechanisms of action were "Behavior regulation" (30/30, 100%), "Knowledge" (30/30, 100%), "Goals' (30/30, 100%), "Memory attention and decision process" 30/30, 100%), and "Beliefs about consequences" (30/30, 100%), which were present in all studied apps. The "Behavior regulation" mechanism of action was typically targeted by the "Self-monitoring of behavior" and

"Self-monitoring of outcome(s) of behavior"BCTs, while "Knowledge" was mostly targeted by "Feedback on outcome(s) of behavior," "Instruction on how to perform the behavior," "Information about health consequences," and "Feedback on behavior." The "Goals" and "Memory attention and decision process" were mostly targeted using "Action planning" and "Prompts/cues" BCTs, respectively. The next most common mechanisms of action identified were "Beliefs about capabilities" (16/30), which was mostly targeted using BCTs "Problem solving," "Focus on past success," and "Social incentive." Almost one-third of the apps (9/30) had "Skills" as a mechanism of action, which was mostly targeted using "Problem solving" and "Demonstration of the behavior." "Social influences" (4/30 13%) was an infrequently used mechanism of action. The least common mechanisms were "Reinforcement" (3/30, 10%) and "Emotion" (3/30, 10%), which were present in only three apps. The mechanisms of action "Intention," "Optimism," "Professional role and identity," and "Environmental context and resources" were not presented in any app.

Additional Aspects

None of the 30 apps were FDA or CA approved. Eighteen apps (60%) were found to have information available concerning their user rating. Of these 18 apps, 13 (72.2%) scored 4 or more stars (of 5). Only 5 (27.8%) app ratings were below 4 (Multimedia Appendix 1).

Discussion

Principal Findings

This study aimed to review all apps developed to support the self-management of hypertension, which are available on the two most popular app stores Google Play (Android) and Apple Store (iPhone).

The review showed that a significant number of apps (n=186) are available to support the self-management of hypertension. These apps had similar functionalities, although they differed in terms of the combination of functionalities provided. The majority of these apps had only one function (n=108, 58.1%), while the remaining offered different combinations of functionalities. This review indicated that there were few apps with comprehensive functionalities. Apps with comprehensive functionalities are potentially more effective [13].

There are also serious issues regarding the privacy and security of the apps and inconsistencies in apps' theoretical underpinning, where in many cases, apps were developed without an explicitly clear theoretical basis. The evaluation of the selected apps' data security and privacy revealed that the privacy policy was not available for 35% of the apps assessed in detail. Most apps gathered identifiable personal information and engaged in sharing user data with third parties and almost half of the selected apps (45%) did so without clearly disclosing how data security was ensured. The evaluation of the theoretical underpinning of apps revealed that a total of 24 BCTs, ranging from 6 to 17 (with median of 9), identified in the 30 reviewed apps mapped to 10 TDF mechanism of actions, ranging from 5 to 9 (with a median of 6), may have supported hypertension self-management behaviors. These findings are similar to reviews of apps related to other chronic diseases that have reported that few apps contain both comprehensive functionalities [10,35] and inconsistent BCTs [17,21]. Despite much research that code BCTs underpinning health apps [17,19,21,22], there is little research reviewing how existing apps map BCTs to TDF domains. The linking of BCTs to TDF conducted in the present study may help developers and researchers in selecting appropriate BCTs when developing apps. It may also help researchers understand which BCTs are effective and how they exert their effects [36-38].

None of the reviewed apps made claims based on behavioral theories or strategies relating to various self-management interventions, although self-management programs are likely to be effective if they are supported by theory-based interventions [14-16]. This may be because the expertise of health professionals was not factored into the development of the majority of these apps [35,39], despite the stressed importance of involving multidisciplinary perspectives and skills in developing a product within a user-center design framework [19]. However, for the 30 reviewed apps with comprehensive functionalities, the examination of the BCT and TDF domains underpinning them shows that a number of BCTs and TDF mechanisms of action were present. There is still no conclusive evidence for which combinations of BCTs or TDFs are the key moderators for effective chronic disease self-management, especially hypertension [21,40,41]. This is an area that requires further research. However, all present TDFs in these 30 apps have the potential to stimulate hypertension self-management activities through different mechanisms of action, particularly those of "Behavior regulation," "Knowledge," "Memory, attention and decision processes," and "Goals." These mechanisms of action are supported by studies identifying the key TDF domains that need to be targeted to influence patient behaviors and support self-management in chronic diseases [42,43]. Although other studies have also found that "Skills," "Emotions," "Reinforcement," and "Belief of capabilities" are essential to increase people's motivation in managing their health, many of the reviewed apps lack these characteristics [44].

The evaluation of the privacy policy showed that the security and privacy of consumers could be substantially improved. Our findings are in accord with those of earlier studies that have evaluated data security and safety of existing apps [19,20]. Huckvale et al revealed that one-fifth of apps in the National Health Service Apps Library lacked privacy policies, and the majority of the apps violated user data privacy and security [20]. Practices of data gathering and analysis by app developers can be advantageous to users, providing greater levels of personalization and data-informed improvements to the app. However, such practices of data gathering should be disclosed clearly, so that a potential user is aware of the possible risks to their data security [45]. To ensure users are able to make fully informed decisions, they must be equipped with the skills and information necessary to scrutinize these privacy and security policies. Because of the large scale of the app market, the regulation and preservation of data protection is difficult. As a

result, the management of data privacy and security is entrusted to the developers of apps [46].

This review identified a small number of apps that are able to use smartphones as a medical device (cuff-less device) to measure BP. However, none of these apps were approved as a validated medical device. Indeed, cuff-less devices for measuring BP based on smartphone apps have recently been shown to be highly inaccurate and unfeasible [47] and may negatively affect patients' health and safety. This is of particular concern, since a recent study by Kumar et al (2016) found that even though only a small number of apps have this feature, users have a strong inclination to download and favorably rate these types of apps [12]. This highlights the need for extensive clinical validation studies in different patient populations before such technology is used in commercial and clinical capacities. As such, physicians should currently be aware of the use of such apps by their patients and should promote only the use of validated devices for BP measurement.

Apps with more comprehensive functionalities have the potential to be more difficult for patients to use. This study found that there was an absence of evidence concerning the usability of the apps in the apps' descriptions. Although this study did not directly evaluate the apps' usability, user ratings were used as a proxy of the apps' usability. The usefulness of user ratings as a measure of apps' technical usability is questionable. In a review of mobile apps for the self-management of diabetes, Hood et al [48] found that the user rating was poorly correlated with the results of the study's usability evaluation. However, in a general sample of health apps, Mendiola et al found that user ratings could be related to an app's technical usability regarding aspects such as layout, interactive features, and general ease of use [49]. The user ratings for the apps considered in this study were high, with 73% obtaining 4 or more stars. This is in line with previous studies [50,51] where participants reported that they were satisfied with apps that include comprehensive functionalities, finding them easy to use.

The majority of apps identified in the recent systematic review of Alessa et al [13] were study specific, that is, developed for the aims of the study alone [13]. However, the apps considered in this review were commercially available apps in app stores. The descriptions of these apps lacked evidence about their effectiveness and did not even mention or consider the importance of such evidence. None of them were approved by the FDA or CE as a medical device. This is in line with previous reviews, which reported that the rapid growth of the commercial market for such apps has created an overabundance of apps that lack readily available evidence of their effectiveness [25] and lack FDA or CE approval [52,53]. Applying the findings of this recent systematic review within this review of commercial apps indicated that few apps (30/186) seem to have the potential to be effective. Apps that have this potential would need to be scientifically evaluated to ensure that this potential for effectiveness and usability is realized in practice. This indicates a critical gap between the research domain and the work of commercial app developers, emphasizing the importance of cooperation between them.

Limitations and Strengths of the Study

This review has a number of limitations. First, the review only included apps that were developed to be used by English-speaking users, excluding apps in other languages such as Chinese. Second, since these apps are tailored for the self-management of hypertension, they support a wide range of different behaviors such as medication adherence, weight, diet, and physical activity in addition to the self-monitoring of BP, which makes it challenging to code them according to a single specific behavior and exclude other behaviors. This may be attributed to the complexity of the self-management process, which encompasses an array of behaviors and activities to effectively control BP. Third, this study excluded apps that require identification access. Moreover, the content of educational information of included apps was not checked to ensure they were up to date and met medical standards and health literacy guidelines. Finally, data privacy and security were assessed in relation to policy statements rather than practices. There is evidence of inconsistencies in some cases between the real practices of app developers and policy statements [20].

Despite these limitations, the study has several strengths. As this study reviewed all apps supporting the self-management of hypertension, rather than limiting itself to only the most popular apps [12], the results of this review offer a general picture of the present status of smartphone app stores in the field of hypertension. This comprehensive review will guide further research and development of these tools in different ways, for example, by encouraging developers and researchers to assess commercially available apps' effectiveness and usability among potential users and by urging app developers to be more transparent about privacy and security. The study reviewed apps on the two most common smartphone platforms; it thus considered a large user base. Furthermore, this study is the first systematic review to explore the theoretical underpinning of apps by seeking to map BCTs to TDF domains in apps containing comprehensive functionalities. The insights could be useful for content developers designing apps in the area of hypertension or other chronic diseases that aim to engage both users and health care personnel who are likely to encourage patients to utilize these technologies.

Recommendations

Based on the result of this review, some recommendations can be made. Despite the widespread availability of apps, potential users and health care providers should be made aware of the shortcomings in the security of private data as well as in the potential effectiveness of the apps in supporting hypertension self-management. Future efforts (and collaborations) should also be made by both researchers and commercial developers to encourage the development of apps that demonstrate scientific evidence of their effectiveness and usability to the public. The importance of involving end users in app development should be noted, as it helps improve user satisfaction and acceptance. This study's findings encourage further research to evaluate app effectiveness and technical usability. It is important to assess the effectiveness of commercially available apps in order to

determine the positive and potential negative effects of using the app.

Conclusions

The review identified the widespread accessibility and availability of smartphone apps with a range of combinations of functionalities that can support the self-management of hypertension. However, relatively few of these apps contained comprehensive functionalities, which are more likely to be effective in lowering blood pressure; many lacked security measures; and most lacked a clear theoretical basis. Furthermore, there is a concerning absence of evidence with regard to their effectiveness and usability and involvement of health care professionals in the development process. This raises a serious practical issue for health care professionals and patients in determining which app to choose or use, as there are no specific criteria for them to make an informed selection. These findings demonstrate that the technical usability and effectiveness of apps in supporting the self-management of hypertension urgently need to be evaluated and that clear criteria need to be established to guide the selection of the most suitable app.

Acknowledgments

We thank all authors for their significant contribution. TA is a PhD student at Sheffield University and is funded by King Saud University, Saudi Cultural Bureau..

Conflicts of Interest

None declared.

Multimedia Appendix 1

Data of the included apps.

[PDF File (Adobe PDF File), 269KB - mhealth_v7i5e13645_app1.pdf]

References

1. Olsen MH, Angell SY, Asma S, Boutouyrie P, Burger D, Chirinos JA, et al. A call to action and a lifecourse strategy to address the global burden of raised blood pressure on current and future generations: the Lancet Commission on hypertension. Lancet 2016 Dec 26;388(10060):2665-2712. [doi: 10.1016/S0140-6736(16)31134-5] [Medline: 27671667]
2. Chow CK, Teo KK, Rangarajan S, Islam S, Gupta R, Avezum A, PURE (Prospective Urban Rural Epidemiology) Study investigators. Prevalence, awareness, treatment, and control of hypertension in rural and urban communities in high-, middle-, and low-income countries. JAMA 2013 Sep 04;310(9):959-968. [doi: 10.1001/jama.2013.184182] [Medline: 24002282]
3. Shahaj O, Denneny D, Schwappach A, Pearce G, Epiphaniou E, Parke HL, et al. Supporting self-management for people with hypertension: a meta-review of quantitative and qualitative systematic reviews. J Hypertens 2019 Feb;37(2):264-279. [doi: 10.1097/HJH.0000000000001867] [Medline: 30020240]
4. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The Impact of mHealth Interventions: Systematic Review of Systematic Reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
5. Bonoto BC, de Araújo VE, Godói IP, de Lemos LLP, Godman B, Bennie M, et al. Efficacy of Mobile Apps to Support the Care of Patients With Diabetes Mellitus: A Systematic Review and Meta-Analysis of Randomized Controlled Trials. JMIR Mhealth Uhealth 2017 Mar 01;5(3):e4 [FREE Full text] [doi: 10.2196/mhealth.6309] [Medline: 28249834]
6. Whitehead L, Seaton P. The Effectiveness of Self-Management Mobile Phone and Tablet Apps in Long-term Condition Management: A Systematic Review. J Med Internet Res 2016 May 16;18(5):e97 [FREE Full text] [doi: 10.2196/jmir.4883] [Medline: 27185295]
7. 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]
8. Barnes SJ, Pressey AD, Scornavacca E. Mobile ubiquity: Understanding the relationship between cognitive absorption, smartphone addiction and social network services. Computers in Human Behavior 2019 Jan;90:246-258. [doi: 10.1016/j.chb.2018.09.013]
9. Zhang L, Hua K, Wang H, Qian G, Zhang L. Sentiment Analysis on Reviews of Mobile Users. Procedia Computer Science 2014;34:458-465. [doi: 10.1016/j.procs.2014.07.013]
10. Rivera J, McPherson A, Hamilton J, Birken C, Coons M, Iyer S, et al. Mobile Apps for Weight Management: A Scoping Review. JMIR Mhealth Uhealth 2016 Jul 26;4(3):e87 [FREE Full text] [doi: 10.2196/mhealth.5115] [Medline: 27460502]
11. Goldberg EM, Levy PD. New Approaches to Evaluating and Monitoring Blood Pressure. Curr Hypertens Rep 2016 Dec;18(6):49. [doi: 10.1007/s11906-016-0650-9] [Medline: 27137524]
12. Kumar N, Khunger M, Gupta A, Garg N. A content analysis of smartphone-based applications for hypertension management. J Am Soc Hypertens 2015 Feb;9(2):130-136. [doi: 10.1016/j.jash.2014.12.001] [Medline: 25660364]

13. Alessa T, Abdi S, Hawley MS, de Witte L. Mobile Apps to Support the Self-Management of Hypertension: Systematic Review of Effectiveness, Usability, and User Satisfaction. JMIR Mhealth Uhealth 2018 Jul 23;6(7):e10723 [FREE Full text] [doi: 10.2196/10723] [Medline: 30037787]
14. Al-Washali AA, Kadri H, Ismail S, Rahman H, Elezzy Y. Systematic review on effectiveness of theory-based intervention on self-care behaviors among patients with type 2 diabetes. Clin Trials Degener Dis 2018;3(3):88. [doi: 10.4103/2542-3975.242955]
15. Band R, Bradbury K, Morton K, May C, Michie S, Mair FS, et al. Intervention planning for a digital intervention for self-management of hypertension: a theory-, evidence- and person-based approach. Implement Sci 2017 Dec 23;12(1):25 [FREE Full text] [doi: 10.1186/s13012-017-0553-4] [Medline: 28231840]
16. Lepard MG, Joseph AL, Agne AA, Cherrington AL. Diabetes self-management interventions for adults with type 2 diabetes living in rural areas: a systematic literature review. Curr Diab Rep 2015 Jun;15(6):608 [FREE Full text] [doi: 10.1007/s11892-015-0608-3] [Medline: 25948497]
17. Edwards EA, Lumsden J, Rivas C, Steed L, Edwards LA, Thiyagarajan A, et al. Gamification for health promotion: systematic review of behaviour change techniques in smartphone apps. BMJ Open 2016 Dec 04;6(10):e012447 [FREE Full text] [doi: 10.1136/bmjopen-2016-012447] [Medline: 27707829]
18. 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]
19. Bondaronek P, Alkhaldi G, Slee A, Hamilton FL, Murray E. Quality of Publicly Available Physical Activity Apps: Review and Content Analysis. JMIR Mhealth Uhealth 2018 Mar 21;6(3):e53 [FREE Full text] [doi: 10.2196/mhealth.9069] [Medline: 29563080]
20. Huckvale K, Prieto JT, Tilney M, Benghozi P, Car J. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. BMC Med 2015 Sep 25;13(1):214 [FREE Full text] [doi: 10.1186/s12916-015-0444-y] [Medline: 26404673]
21. Morrissey EC, Corbett TK, Walsh JC, Molloy GJ. Behavior Change Techniques in Apps for Medication Adherence: A Content Analysis. Am J Prev Med 2016 May;50(5):e143-e146. [doi: 10.1016/j.amepre.2015.09.034] [Medline: 26597504]
22. Heron N, Kee F, Donnelly M, Cardwell C, Tully MA, Cupples ME. Behaviour change techniques in home-based cardiac rehabilitation: a systematic review. Br J Gen Pract 2016 Oct;66(651):e747-e757 [FREE Full text] [doi: 10.3399/bjgp16X686617] [Medline: 27481858]
23. Steinmo S, Fuller C, Stone SP, Michie S. Characterising an implementation intervention in terms of behaviour change techniques and theory: the 'Sepsis Six' clinical care bundle. Implement Sci 2015 Aug 08;10:111 [FREE Full text] [doi: 10.1186/s13012-015-0300-7] [Medline: 26253306]
24. Mazoteras Pardo V, Losa Iglesias ME, López Chicharro J, Becerro de Bengoa Vallejo R. The QardioArm App in the Assessment of Blood Pressure and Heart Rate: Reliability and Validity Study. JMIR Mhealth Uhealth 2017 Dec 15;5(12):e198 [FREE Full text] [doi: 10.2196/mhealth.8458] [Medline: 29246880]
25. Parati G, Torlasco C, Omboni S, Pellegrini D. Smartphone Applications for Hypertension Management: a Potential Game-Changer That Needs More Control. Curr Hypertens Rep 2017 Jun;19(6):48. [doi: 10.1007/s11906-017-0743-0] [Medline: 28501982]
26. BinDhim NF, Hawkey A, Trevena L. A systematic review of quality assessment methods for smartphone health apps. Telemed J E Health 2015 Feb;21(2):97-104. [doi: 10.1089/tmj.2014.0088] [Medline: 25469795]
27. Karthick S, Binu S. Android security issues and solutions. 2017 Presented at: Proceedings of the IEEE Int Conf Innov Mech Ind Appl; 2017; Bangalore, India p. 686. [doi: 10.1109/ICIMIA.2017.7975551]
28. Or C, Tao D. A 3-Month Randomized Controlled Pilot Trial of a Patient-Centered, Computer-Based Self-Monitoring System for the Care of Type 2 Diabetes Mellitus and Hypertension. J Med Syst 2016 Apr;40(4):81. [doi: 10.1007/s10916-016-0437-1] [Medline: 26802011]
29. Logan AG, Irvine MJ, McIsaac WJ, Tisler A, Rossos PG, Easty A, et al. Effect of home blood pressure telemonitoring with self-care support on uncontrolled systolic hypertension in diabetics. Hypertension 2012 Jul;60(1):51-57 [FREE Full text] [doi: 10.1161/HYPERTENSIONAHA.111.188409] [Medline: 22615116]
30. ICO. Privacy in mobile apps: Guidance for app developers URL: https://ico.org.uk/media/for-organisations/documents/1596/privacy-in-mobile-apps-dp-guidance.pdf [accessed 2019-05-10] [WebCite Cache ID 75zEqTQ1h]
31. TA: Online Trust Alliance. 2014. Mobile App Privacy & Security URL: https://otalliance.org/best-practices/mobile-app-privacy-security [accessed 2019-02-06] [WebCite Cache ID 75zF82lz3]
32. Cane J, Richardson M, Johnston M, Ladha R, Michie S. From lists of behaviour change techniques (BCTs) to structured hierarchies: comparison of two methods of developing a hierarchy of BCTs. Br J Health Psychol 2015 Feb;20(1):130-150. [doi: 10.1111/bjhp.12102] [Medline: 24815766]
33. Michie S, Atkins L, West R. The behaviour change wheel book - a guide to designing interventions. London: Silverback; 2014:9781912141081.

34. Michie S, Johnston M, Francis J, Hardeman W, Eccles M. From theory to intervention: mapping theoretically derived behavioural determinants to behaviour change techniques. Appl Psychol 2008;57(4):660. [doi: 10.1111/j.1464-0597.2008.00341.x]
35. Lalloo C, Shah U, Birnie KA, Davies-Chalmers C, Rivera J, Stinson J, et al. Commercially Available Smartphone Apps to Support Postoperative Pain Self-Management: Scoping Review. JMIR Mhealth Uhealth 2017 Oct 23;5(10):e162 [FREE Full text] [doi: 10.2196/mhealth.8230] [Medline: 29061558]
36. Atkins L, Francis J, Islam R, O'Connor D, Patey A, Ivers N, et al. A guide to using the Theoretical Domains Framework of behaviour change to investigate implementation problems. Implement Sci 2017 Dec 21;12(1):77 [FREE Full text] [doi: 10.1186/s13012-017-0605-9] [Medline: 28637486]
37. Carey RN, Connell LE, Johnston M, Rothman AJ, de Bruin M, Kelly MP, et al. Behavior Change Techniques and Their Mechanisms of Action: A Synthesis of Links Described in Published Intervention Literature. Ann Behav Med 2018 Oct 10:A. [doi: 10.1093/abm/kay078] [Medline: 30304386]
38. Michie S, Carey RN, Johnston M, Rothman AJ, de Bruin M, Kelly MP, et al. From Theory-Inspired to Theory-Based Interventions: A Protocol for Developing and Testing a Methodology for Linking Behaviour Change Techniques to Theoretical Mechanisms of Action. Ann Behav Med 2018 May 18;52(6):501-512 [FREE Full text] [doi: 10.1007/s12160-016-9816-6] [Medline: 27401001]
39. Rosser BA, Eccleston C. Smartphone applications for pain management. J Telemed Telecare 2011;17(6):308-312. [doi: 10.1258/jtt.2011.101102] [Medline: 21844177]
40. Michie S, West R, Sheals K, Godinho CA. Evaluating the effectiveness of behavior change techniques in health-related behavior: a scoping review of methods used. Transl Behav Med 2018 Dec 01;8(2):212-224 [FREE Full text] [doi: 10.1093/tbm/ibx019] [Medline: 29381786]
41. Kebede MM, Liedtke TP, Möllers T, Pischke CR. Characterizing Active Ingredients of eHealth Interventions Targeting Persons With Poorly Controlled Type 2 Diabetes Mellitus Using the Behavior Change Techniques Taxonomy: Scoping Review. J Med Internet Res 2017 Dec 12;19(10):e348 [FREE Full text] [doi: 10.2196/jmir.7135] [Medline: 29025693]
42. McCullough AR, Ryan C, O'Neill B, Bradley JM, Elborn JS, Hughes CM. Defining the content and delivery of an intervention to Change AdhereNce to treatment in BonchiEctasis (CAN-BE): a qualitative approach incorporating the Theoretical Domains Framework, behavioural change techniques and stakeholder expert panels. BMC Health Serv Res 2015 Aug 22;15:342 [FREE Full text] [doi: 10.1186/s12913-015-1004-z] [Medline: 26297038]
43. Connell LA, McMahon NE, Tyson SF, Watkins CL, Eng JJ. Mechanisms of action of an implementation intervention in stroke rehabilitation: a qualitative interview study. BMC Health Serv Res 2016 Dec 30;16(1):534 [FREE Full text] [doi: 10.1186/s12913-016-1793-8] [Medline: 27716247]
44. Thomas S, Mackintosh S. Use of the theoretical domains framework to develop an intervention to improve physical therapist management of the risk of falls after discharge. Phys Ther 2014 Nov;94(11):1660-1675. [doi: 10.2522/ptj.20130412] [Medline: 25035266]
45. 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]
46. Wicks P, Chiauzzi E. 'Trust but verify'--five approaches to ensure safe medical apps. BMC Med 2015 Sep 25;13:205 [FREE Full text] [doi: 10.1186/s12916-015-0451-z] [Medline: 26404791]
47. Plante TB, Urrea B, MacFarlane ZT, Blumenthal RS, Miller ER, Appel LJ, et al. Validation of the Instant Blood Pressure Smartphone App. JAMA Intern Med 2016 Dec 01;176(5):700-702 [FREE Full text] [doi: 10.1001/jamainternmed.2016.0157] [Medline: 26938174]
48. Hood M, Wilson R, Corsica J, Bradley L, Chirinos D, Vivo A. What do we know about mobile applications for diabetes self-management? A review of reviews. J Behav Med 2016 Dec;39(6):981-994. [doi: 10.1007/s10865-016-9765-3] [Medline: 27412774]
49. Mendiola MF, Kalnicki M, Lindenauer S. Valuable features in mobile health apps for patients and consumers: content analysis of apps and user ratings. JMIR Mhealth Uhealth 2015 May 13;3(2):e40 [FREE Full text] [doi: 10.2196/mhealth.4283] [Medline: 25972309]
50. Hallberg I, Ranerup A, Kjellgren K. Supporting the self-management of hypertension: Patients' experiences of using a mobile phone-based system. J Hum Hypertens 2016 Feb;30(2):141-146 [FREE Full text] [doi: 10.1038/jhh.2015.37] [Medline: 25903164]
51. McGillicuddy JW, Weiland AK, Frenzel RM, Mueller M, Brunner-Jackson BM, Taber DJ, et al. Patient attitudes toward mobile phone-based health monitoring: questionnaire study among kidney transplant recipients. J Med Internet Res 2013 Jan 08;15(1):e6 [FREE Full text] [doi: 10.2196/jmir.2284] [Medline: 23305649]
52. Wisniewski H, Liu G, Henson P, Vaidyam A, Hajratalli NK, Onnela J, et al. Understanding the quality, effectiveness and attributes of top-rated smartphone health apps. Evid Based Ment Health 2019 Feb;22(1):4-9. [doi: 10.1136/ebmental-2018-300069] [Medline: 30635262]
53. Eng DS, Lee JM. The promise and peril of mobile health applications for diabetes and endocrinology. Pediatr Diabetes 2013 Jun;14(4):231-238 [FREE Full text] [doi: 10.1111/pedi.12034] [Medline: 23627878]

Abbreviations

BCT: behavior change technique
BCTTv1: behavior change technique taxonomy v1
BP: blood pressure
CE: European Union Conformite Europeene
FDA: Food and Drug Administration
HCP: health care provider
TDF: theoretical domain framework

Edited by G Eysenbach; submitted 06.02.19; peer-reviewed by S Donevant, B Green, L Buis, A Maeder, S Omboni, E Gong; comments to author 02.03.19; revised version received 17.04.19; accepted 18.04.19; published 28.05.19.

Please cite as:
Alessa T, Hawley MS, Hock ES, de Witte L
Smartphone Apps to Support Self-Management of Hypertension: Review and Content Analysis
JMIR Mhealth Uhealth 2019;7(5):e13645
URL: http://mhealth.jmir.org/2019/5/e13645/
doi:10.2196/13645
PMID:31140434

Original Paper

User Preferences and Persona Design for an mHealth Intervention to Support Adherence to Cardiovascular Disease Medication in Singapore: A Multi-Method Study

Victoria Haldane[1], MPH; Joel Jun Kai Koh[2], MSc; Aastha Srivastava[2], MPH; Krichelle Wei Qi Teo[3], BSc; Yao Guo Tan[3], BSc; Rui Xiang Cheng[3], BSc; Yi Cheng Yap[3], BSc; Pei-Shi Ong[3], BSc, PhD; Rob M Van Dam[2], MSc, PhD; Jie Min Foo[2], BSc; Falk Müller-Riemenschneider[2], MD, PhD; Gerald Choon-Huat Koh[2], MD, PhD; Pin Sym Foong[2], MSc, PhD; Pablo Perel[4,5], MD, PhD; Helena Legido-Quigley[2,4], PhD

[1]Institute of Health Policy, Management and Evaluation, University of Toronto, Toronto, ON, Canada

[2]Saw Swee Hock School of Public Health, National University of Singapore, Singapore, Singapore

[3]Department of Pharmacy, National University of Singapore, Singapore, Singapore

[4]London School of Hygiene and Tropical Medicine, London, United Kingdom

[5]World Heart Federation, Geneva, Switzerland

Corresponding Author:
Victoria Haldane, MPH
Institute of Health Policy, Management and Evaluation
University of Toronto
155 College St, 4th Floor
Toronto, ON,
Canada
Phone: 1 416 978 4326
Email: v.haldane@mail.utoronto.ca

Abstract

Background: The use of mobile health (mHealth) has gained popularity globally, including for its use in a variety of health interventions, particularly through short message service (SMS) text messaging. However, there are challenges to the use of mHealth, particularly among older users who have a large heterogeneity in usability and accessibility barriers when using technology.

Objective: In order to better understand and conceptualize the diversity of users and give insight into their particular needs, we turned to persona creation. Personas are user archetypes created through data generated from multi-method inquiry with actual target users. Personas are an appropriate yet largely underutilized component of current mHealth research.

Methods: Leveraging data from a multi-method study conducted in Singapore with an ethnically diverse population including Chinese, Malay, and Indian participants, we used a proforma to analyze data from the qualitative component (ie, 20 in-depth interviews) and quantitative component (ie, 100 interviewer-guided surveys). We then identified key characteristics, including technology use and preferences as well as adherence factors, to synthesize five personas reflective of persons over the age of 40 years in Singapore with atherosclerotic cardiovascular disease (ASCVD) or ASCVD risk factors, such as hypertension.

Results: We present five personas typologized as (1) The Quiet Analog, (2) The Busy Grandparent, (3) The Socializer, (4) The Newly Diagnosed, and (5) The Hard-to-Reach. We report on four key characteristics: health care access, medication adherence, mobile phone technology usage (ie, ownership, access, and utilization), and interest in mHealth. Finally, we provide insights into how these personas may be used in the design and implementation of an mHealth intervention. Our work demonstrates how multi-method data can create biopsychosocial personas that can be used to explore and address the diversity in behaviors, preferences, and needs in user groups.

Conclusions: With wider adoption of mHealth, it is important that we consider user-centered design techniques and design thinking in order to create meaningful, patient-centered interventions for adherence to medications. Future research in this area should include greater exploration of how these five personas can be used to better understand how and when is best to deliver mHealth interventions in Singapore and beyond.

(JMIR Mhealth Uhealth 2019;7(5):e10465) doi:10.2196/10465

KEYWORDS

personas; biopsychosocial personas; qualitative; ASCVD; adherence; patient perspectives

Introduction

The use of mobile phones for health-related purposes (ie, mobile health [mHealth]) has gained popularity globally, and mobile phones are becoming widely used in a variety of health interventions [1]. In particular, short message service (SMS)-based interventions have shown promise due to affordability and wide outreach and have, therefore, been applied in health promotion and medication adherence [2-6]. However, there are challenges to the use of mHealth, particularly among older adult users who may face a wide variety of perceptual and impairment barriers to access, as well as technology usability challenges [7,8]. Efforts to address these challenges are complicated by differing degrees of effect on each user, as older adults are highly heterogenous in their experience and ability to use technology [9]. In light of this, there is a need for user-centered design in mHealth interventions in order to bridge the digital divide and enable appropriate and effective interventions that reach users who need them most [10].

Personas are a key feature of user-centered design and are used widely in Web design fields, including product evaluation, design, usability testing, and marketing and product development, to enable technology developers to have a deeper understanding of sometimes large and diverse target audiences. These user archetypes are developed through quantitative and qualitative user experience research and are particularly useful communication tools for developers to grasp the needs of marginalized or hard-to-reach target users [11-13]. As such, personas, especially ones that leverage biopsychosocial data from multi- or mixed-method inquiry, are an appropriate yet largely underutilized component of current mHealth development [13]. For health services, the value of developing personas lie in their potential to allow for better design of future user-centric mHealth interventions.

This study is the development phase of a proposed mHealth intervention, the txt2heart trial, to support patient adherence to medications for atherosclerotic cardiovascular disease (ASCVD). This is an area of much-needed support as, globally, many individuals with risk factors for ASCVD remain undiagnosed, untreated, and uncontrolled [14-16]. Although medication has shown to be effective in preventing cardiovascular events associated with ASCVD, patients may exhibit suboptimal adherence due to the long-term management required, disease misperceptions, poor health literacy, poor patient-physician relationship, weak social support, and comorbidities [17-21]. The txt2heart trial is an international collaboration evaluating the efficacy and safety of SMS text messaging on clinical outcomes and adherence in different countries, including Colombia, Ghana, India, and Singapore. This paper focuses on work conducted in Singapore where, based on the high mobile penetration locally and available data, we may be able to employ personas to gain a greater understanding of the preferences and behaviors of the target population.

Our multi-method behavioral study sought to explore a variety of participant experiences with ASCVD and associated risk factors such as hypertension. Our paper aims to put forth personas that illustrate the variety of both adherence and technology behaviors in Singapore and to propose ways forward for the use of personas in the design and implementation of mHealth interventions.

Methods

Sampling and Recruitment

The study took place in Singapore and was comprised of an exploratory quantitative survey and qualitative study using in-depth interviews. For both the qualitative and quantitative component, we used purposive sampling from an existing community health study patient list. This list comprised participants from two neighborhoods in Singapore who participated in a community health study aiming to identify health needs of residents. Of those who consented to recontact, we recruited via telephone those aged 40 years and above with ASCVD or associated risk factors; Textbox 1 lists the complete inclusion and exclusion criteria. Once recruited, surveys and interviews were conducted in person at places of convenience to each participant. There was no overlap between recruitment for the qualitative and quantitative studies.

Textbox 1. Study inclusion and exclusion criteria.

Inclusion criteria:

1. Patients with a history of atherosclerotic cardiovascular disease (ASCVD): coronary artery disease, ischemic stroke, peripheral artery disease, and atherosclerotic aortic disease OR
2. Patients with at least one risk factor such as hypertension or hyperlipidemia, where antiplatelet, antihypertensives, and/or statins are recommended.

Exclusion criteria:

1. Participants unable to participate in a verbal interview.
2. Participants who did not speak Mandarin, English, or Malay languages.

Data Collection

Research team members administered the survey instrument and conducted semistructured, in-depth interviews in the participant's preferred language by staff fluent in that language. Both the survey instrument and interview guide were developed as part of the larger txt2heart collaboration, which we then adapted to the Singapore context (see Multimedia Appendix 1). Both covered topics such as patients' sociodemographic characteristics, health care access, medication adherence, mobile phone technology usage (ie, ownership, access, and utilization), and interest in mHealth. We have included a consolidated criteria for reporting qualitative research (COREQ) checklist in Multimedia Appendix 2.

The quantitative exploratory component involved 100 participants who took part in either a face-to-face or phone survey as per their convenience and availability. The qualitative component involved face-to-face, semistructured, in-depth interviews with 20 participants. Out of the 20 participants interviewed, 19 agreed to audio recordings and 1 participant declined. For the latter, detailed field notes and an extensive memo were taken for inclusion in data analysis.

Ethical Considerations

Informed consent for participation and recording was obtained before both the quantitative surveys and qualitative interviews using a participant consent form and information sheet. Participants could refuse to answer any of the questions and/or discontinue participation in the research at any time. Ethical approval was obtained from the Institutional Review Board at the National University of Singapore.

Analysis

For the quantitative component, all data were analyzed using SPSS Statistics for Windows, version 24.0 (IBM Corp). Frequencies (n) and percentages (%) were used to summarize sociodemographic characteristics, the pattern of mobile phone technology usage, and interest in mHealth.

For the qualitative component, interviews were recorded and transcribed in full. Two research team members coded interviews using NVivo 11 software (QSR International), applying inductive approaches, thematic analysis, and techniques from the constant comparative method, where line-by-line analysis of early interviews is used on subsequent interviews to test preliminary assumptions [22,23]. Reviewers agreed on identified codes and themes, namely barriers and facilitators to adherence, mobile phone usage, attitudes and preferences, health information-seeking behaviors, and willingness to learn. Through these discussions, reviewers agreed that thematic saturation had been reached.

Persona Development

Using emergent themes from the analysis of the qualitative study, we developed a proforma to guide persona development (see Multimedia Appendix 3). The proforma captures participant demographics, socioeconomic factors, social supports, treatment adherence, health literacy and information seeking, and mobile phone usage; using standardized definitions (see Multimedia Appendix 3), each qualitative interview was analyzed against the proforma to bring forth the salient attributes of each persona.

This proforma enabled us to cluster participants into groups, taking into consideration low-to-high social support, treatment adherence habits, and mobile phone usage, as well as identification of common factors contributing to these. Using both qualitative and quantitative data analysis, we constructed personas that reflect behaviors and characteristics of the user population in Singapore with relevant codes identified by participant ID number to maintain confidentiality and exemplary quotes from the 20 qualitative interviews. While some studies opt to give personas names, faces, and individual characteristics, we have kept our personas as icons and titles in order to maintain the focus on broader trends in user groups rather than construction of individuals.

Results

The Population Context

Participant characteristics from the qualitative study are outlined in Table 1. The quantitative component involved a total of 100 interviewer-administered surveys, inclusive of one incomplete survey that had missing data on mobile phone technology usage. Participant characteristics from the quantitative study are outlined in Table 2.

The Singapore Context

As persona creation must account for larger socioeconomic processes, it is important to take note of key health systems factors in Singapore. Health care in Singapore is generally accessible and affordable, with patients availing themselves of largely subsidized care offered in polyclinics (ie, government-subsidized general practice clinics), by private general practitioners, and even in tertiary care facilities for primary care and management of chronic conditions. Within primary care settings, it is a common practice for doctors to prescribe multiple months' worth of medication to regular patients; this availability of medication is reported by our participants to facilitate their adherence. Further, the Singapore government subsidizes and offers various accessible schemes for those with chronic conditions; by and large, our participants reported using these schemes to fund their medications and most reported being able to afford their medications.

However, despite these systems-level facilitators to adherence, 65 out of 100 (65.0%) survey respondents reported suboptimal adherence to ASCVD medication as defined by the Medication Adherence Report Scale (MARS-5). Further, our qualitative findings highlighted different patient experiences, including self-titration of medications in response to symptoms and forgetfulness, as playing a role in suboptimal adherence.

The Technological Context: Mobile Phone Behaviors and Interest

We analyzed our survey data to report mobile phone ownership, access, use, and preferences in our target population, as well as to inform personas that reflect the technological context of a wider population in Singapore.

Table 1. Sociodemographic characteristics of participants of the qualitative study (N=20).

Sociodemographic characteristics	Value
Age (years), mean (SD)	72.5 (7.5)
Age (years), n (%)	
61-70	7 (35)
71-80	8 (40)
81-90	4 (20)
Unreported	1 (5)
Gender, n (%)	
Male	12 (60)
Female	8 (40)
Ethnicity, n (%)	
Chinese	15 (75)
Indian	4 (20)
Malay	1 (5)

Table 2. Sociodemographic characteristics of participants of the quantitative study (N=100).

Sociodemographic characteristics	Value
Age (years), mean (SD)	65.3 (9.6)
Age (years), n (%)	
<65	45 (45.0)
≥65	55 (55.0)
Gender, n (%)	
Female	30 (30.0)
Male	70 (70.0)
Ethnicity, n (%)	
Chinese	66 (66.0)
Malay	19 (19.0)
Indian	14 (14.0)
Other	1 (1.0)

Among 100 surveyed participants, 99 (99.0%) fully completed the survey; 90 out of 99 (91%) participants owned a mobile phone, either an advance-feature or basic-feature mobile phone. Of those mobile phone owners, 70 out of 90 (78%) reported accessing their mobile phones in general at least once a day; the same proportion of patients—70 out of 90 (78%)—reported using SMS text messaging at least once a day. In general, participants predominantly used their phones for phone calls (87/90, 97%), SMS text messaging (60/90, 67%), and other text messaging services, such as WhatsApp (54/90, 60%).

Among mobile phone owners, 48 out of 90 (53%) indicated interest in receiving medication information via mobile phone. Among those interested in receiving medication information via mobile phone, SMS text messaging was the preferred mode of delivery (40/48, 83%).

Personas

Overview

We report our findings as classified into five personas (see Table 3) that have differing awareness of lifestyle factors, such as what constitutes a healthy diet and regular exercise; differing social milieu and social connectedness; a variety of factors contributing to suboptimal treatment adherence; and mobile phone usage, mobile phone usability issues, and willingness to learn about mobile phone features. These factors offer insight into user behaviors in Singapore that may impact user uptake and may be used to inform the design and implementation of mHealth interventions in this population. We have typologized our personas as The Quiet Analog, The Busy Grandparent, The Socializer, The Newly Diagnosed, and The Hard-to-Reach. A visual compilation of persona types can be found in Multimedia Appendix 4.

Table 3. Personas overview.

Persona features	Personas				
	The Quiet Analog	The Busy Grandparent	The Socializer	The Newly Diagnosed	The Hard-to-Reach
Lifestyle factors					
Knowledge level of lifestyle factors	Intermediate	Intermediate	High	Low	Intermediate
Level of social connectedness	Intermediate	High	High	Unclear	Low
Suboptimal adherence factors					
Level of adherence	Suboptimal	Suboptimal	Suboptimal	N/A[a]	Suboptimal
Suboptimal adherence behaviors	Side effects and polypharmacy	Forgetfulness	Self-titration	Learning to take medications after an acute event	Afraid of acute event
Example quotes	"I am not taking it...as long as my readings are fine...you know why? Sometimes it gives you aches in your bones. It starts to hurt." [ID011]	"I tend to forget my medications in the morning on weekends. I take care of my grandchildren; I go to where my daughter stays and so I forget." [ID004]	"I'm very health conscious. I know how to monitor and adjust...not like those uneducated people who don't know." [IDI003]	"I was taking about 11 different medications...At the beginning, there was 14, when I was in the hospital." [ID007]	"Sometimes I don't take one. I'm so frightened, you know. I can take two or what, I don't know but I then forget thinking I got take or what." [ID014]
Mobile phone use and needs					
Level of mobile phone usage	Low	Intermediate	High	High	Low
Mobile phone needs	Phone calls only	Calls and SMS[b]	Calls, SMS, WhatsApp, and video	Calls, SMS, WhatsApp, and video	Only for calls when out
Level of interest in using a mobile phone	Low	Intermediate	High	High	Low
Level of interest in mHealth	Low	Low	Intermediate	High	Intermediate
Desires	Wants to maintain own routine	Wants information from a trusted source	Wants tailored content	Wants to learn how to manage	Wants information for peace of mind
Example quotes	"There is a lot of stuff in the mobile phone but it's very difficult for us to learn how to use them. It's not easy." [ID001]	"Maybe when they [government] send, I am more comfortable." [ID016]	"It would help if it's in Chinese, Malay, Tamil. Those of us who are older don't understand English and thus the content in the SMS." [ID009]	"It [SMS reminder] will be useful for patients who just came out of the ward." [ID007]	"I find myself having the habit of worrying while taking care of myself. Thus, I don't listen much to the words of others." [ID005]

[a]N/A: not applicable.

[b]SMS: short message service.

Persona A: The Quiet Analog

The first persona explores a patient who largely shies away from uptake of technology and has low interest in an mHealth intervention (see Table 4). Based on the 20 qualitative interviews, 6 (30%) of our participants exhibited low mobile use and a preference for face-to-face contact with health care providers or traditional media for consumption of health information. Participants also reported physical barriers to using their phones such as poor eyesight, which may make it difficult to read SMS text messages. While these patients exhibited suboptimal adherence for a variety of reasons, they did not see mHealth as a solution to these issues; indeed, one participant indicated that medication consumption was a choice dictated entirely by the importance one places on that medication. As these patients exhibit lower mobile phone usage and are not actively seeking health information, they may require a longer time to familiarize themselves with an mHealth intervention and may benefit from being introduced to mHealth by a health care provider.

Table 4. Persona A: The Quiet Analog.

Persona features	Details
Lifestyle	
Features	May have some mobility issues but attempts to have a good diet and get exercise Independent in managing their conditions
Example quote	"I always go [to the clinic] by myself. No need to make trouble for others." [ID001]
Social supports	
Family support	High
Friend support	Unclear
Adherence factors	
Barriers	Forgetfulness Side effects Polypharmacy Self-titration
Facilitators	Medicine supply Cost
Example quote	"But if I go out, I won't bring and eat it. It's very troublesome to take it when I'm on the bus or MRT[a]. I'll just miss taking it." [ID018]
Health information seeking	
Level	Intermediate
Actions	Gets information during consultations with general practitioners and from pharmacists Reads about health topics in the newspaper Does not seek out health information
Example quote	"They can be seen on the newspaper, but you cannot believe in all of them. I think the most important thing is that they are not trustworthy. You need to think about whether what they say is correct and true." [ID001]
Mobile usage	
Level	Low
Example quote	"I have a mobile phone but don't really use it. My children gave it to me. I didn't put the phone card into it." [ID001]
Usability concerns	
Concerns	Low interest in using phone Physical barriers
Example quote	"Our eyesight is also failing. Sometimes we see an '8' as an 'S.' It also makes it difficult to read our SMS[b]." [ID014]
Attitudes toward mHealth	
Attitudes	No interest in mHealth Happy with current phone usage
Example quote	"No. As if they want to take medication, they will. Some people, they're not taking it on purpose. I don't take some of my medication too. We take only the important medicine and don't take those that we feel aren't important." [ID018]

[a]MRT: mass rapid transit.

[b]SMS: short message service.

Persona B: The Busy Grandparent

The second persona explores a patient who uses their mobile phone regularly and expresses interest in an mHealth intervention but wants information from a trusted source (see Table 5). A total of 4 of our participants out of 20 (20%) exhibited or largely exhibited intermediate mobile phone use and a diverse range of health information sources, including online sources. While this persona may use their mobile phone, they may not be interested in using it for mHealth, as their phones may be largely for direct communication with their children. Other usability barriers included receiving spam, which may cause them to inadvertently filter out mHealth SMS text messages. This also speaks to their need to trust the information

source, as they will only pay attention to known messengers. These patients reported suboptimal adherence, often due to busy schedules from taking care of grandchildren. As these patients are already on their mobile phones, it is important to leverage the ways they currently interact with SMS text messaging. Participants reported liking the SMS text message reminder services from the polyclinic, so while they may report that they do not think an mHealth intervention is useful, in practice they may find similar utility in such a service as they do with their current SMS text message reminders.

Persona C: The Socializer

The third persona explores a patient who uses their mobile phone regularly and they may, themselves, disseminate health information to their diverse social circles (see Table 6). A total of 5 of our participants out of 20 (25%) exhibited intermediate or high phone use and described their large family and social networks. This persona is physically and socially active and strives to lead a healthy life. This persona uses their phone to stay connected, browse the Internet, and regularly interacts with a host of traditional and digital information sources; thus, they may prefer content that is tailored to their needs. One usability concern is the need for services in languages other than English.

Table 5. Persona B: Busy Grandparent.

Persona features	Details
Lifestyle features	Active and advocates for good diet and exercise
	Mostly independent in managing their conditions
Social supports	
Family support	High
Friend support	Intermediate
Adherence factors	
Barriers	Busy with other activities
	Forgetfulness
	Fear of side effects from long-term use
Facilitators	Medicine supply
	Cost
	Habit and self-organization
	Proximity to polyclinic
Example quote	"I tend to forget my medications in the morning on weekends. I take care of my grandchildren; I go to where my daughter stays and so I forget." [ID004]
Health information seeking	
Level	Intermediate
Actions	Gets information from various health providers as well as family and friends
	Reads up online about conditions and confident in coping with chronic conditions
Example quote	"I've been sick for so long, I'm used to taking medications once in the morning, once in the evening, etc. I place the medication on the table every night so that I see it in the morning." [ID002]
Mobile usage	
Level	Intermediate
Actions	Uses phone to stay in touch with family and friends, primarily through phone calls and SMS[a]
Usability concerns	
Concerns	Receives a lot of spam messages
	Low interest in using phone for mHealth
	Physical barriers
Example quote	"Firstly, the numbers are very small, my fingers are very thick, so two numbers would be dialed at the same time. Second, I didn't buy my own mobile phone, my children bought it for me." [ID002]
Attitudes toward mHealth	
Attitudes	Interested in mHealth SMS text messages from polyclinic but does not think the intervention is useful
Example quote	"I think whatever method you want to use, it's nothing compared to convincing that patient that it is in their own interest and it is important for them to take the medicine regularly according to the schedule. Right?" [ID019]

[a]SMS: short message service.

Table 6. Persona C: The Socializer.

Persona features	Details
Lifestyle	
Features	Active lifestyle, both physically and socially Independent in managing their conditions
Example quote	"People don't need to tell me. I know myself." [ID006]
Social supports	
Family support	High
Friend support	High
Adherence factors	
Barriers	Busy with other activities Side effects
Facilitators	Habit and self-organization Medicine supply Cost Proximity to polyclinic Instructions written in preferred language
Example quote	"The pharmacist told me how to take them so I can manage them myself. Since the medication is meant for yourself, then you need to manage it yourself as well." [ID009]
Health information seeking	
Level	High
Actions	Seeks information from multiple sources (eg, health care providers, media, the Internet, friends, and family) and often disseminates that information to their family and various social circles
Example quote	"Cause why you know when they see my SMS[a], I said, 'Oh, she said, don't take this.' Then after they see me, they say, 'Good you remind me of visit, my cough lessen.'" [ID012]
Mobile usage	
Level	High
Actions	Uses phone to stay in touch with family and friends through calls and WhatsApp Uses various apps
Usability concerns	
Concerns	Language concerns
Example quote	"It would help if it's in Chinese, Malay, Tamil. Those of us who are older don't understand English and thus the content in the SMS." [ID009]
Attitudes toward mHealth	
Attitudes	Interest in mHealth
Example quote	"This is the, what is it, the IT world. Everything must learn." [ID012]

[a]SMS: short message service.

These patients have a high interest in mHealth, some already even use health apps, and would require an intervention that does not simply tell them what to do, as they already feel a high degree of agency in self-management.

Persona D: The Newly Diagnosed

The fourth persona explores a patient who has been newly diagnosed with a chronic condition (see Table 7). While only 1 participant out of 20 (5%) described the challenges of the learning curve in medication use, it is an important stage to consider for an mHealth intervention for treatment adherence. The majority of our participants attributed their ability to adhere to habit formation and a self-determined system for organizing and coordinating their medication. The time after diagnosis is a crucial time for the development of these habits and one that may require more intense support. This persona is coming to terms with a new diagnosis and, after being discharged, is managing multiple medications on their own for the first time. This persona may need reassurance and support to enable habit formation and to understand any side effects; thus, an mHealth intervention may meet their needs and allow continuity of care from inpatient to outpatient settings.

Table 7. Persona D: The Newly Diagnosed.

Persona features	Details
Lifestyle features	May have low knowledge of lifestyle factors or an inactive lifestyle
	Unsure of their condition and may be dependent on others to help manage
Social supports	
Family support	Unclear
Friend support	Unclear
Adherence factors	
Barriers	Transition from inpatient to outpatient care
	Polypharmacy
	Creating new routines
Facilitators	Fear of acute event or recurrence of acute event
	Contact with health care professionals
	Medicine supply
	Cost
Example quote	"I was taking about 11 different medications...At the beginning there was 14, when I was in the hospital." [ID007]
Health information seeking	
Level	High
Actions	Learning how to manage
	Receiving information from multiple sources
Mobile usage	
Level	High
Actions	Uses phone regularly
Usability concerns	No specific concerns
Attitudes toward mHealth	
Attitudes	High interest in mHealth to help learn how to manage
Example quote	"It [SMS[a] text message reminder] will be useful for patients who just came out of the ward." [ID007]

[a]SMS: short message service.

Persona E: The Hard-to-Reach

The final persona explores a patient who is disconnected, both technologically as well as socially (see Table 8). This persona may have limited support to lead an active lifestyle and make healthy diet choices; while they are independent in managing their condition, it may not be by choice, due to low family and social supports. A total of 4 of our 20 participants (20%) exhibited characteristics that typify this persona. This persona may be motivated by fear; this fear may indeed facilitate adherence as a way to prevent acute events, but it may also cause them to worry and, thus, skip doses if they forget if they have taken their medication. While they may seek information from health care providers, these interactions are often sporadic, and they may lack other social touch points for health information. This persona may have low literacy, both technological as well as reading literacy, and may have challenges learning to use their mobile phones, as they may not have the resources to teach themselves how to use SMS text messaging or other apps. Some participants in this category reported using and liking their current appointment SMS text message reminder services; it may be beneficial to use a similar format and content, as well as to explain an mHealth intervention for adherence or information in the health care setting from a trusted source.

Table 8. Persona E: The Hard-to-Reach.

Persona features	Details
Lifestyle	
Features	Limited support to lead active lifestyle
	Independent in managing their conditions but not always by choice
Example quote	"But I have no choice. Who should I call? My niece or nephews? They need to work. It's better that I go on my own." [ID005]
Social supports	
Family support	Low
Friend support	Low
Adherence factors	
Barriers	Side effects
	Polypharmacy
Facilitators	Habit and self-organization
	Medicine supply
	Cost
	Motivated by fear
Example quotes	"Sometimes I don't take one. I'm so frightened, you know. I can take two or what, I don't know but I then forget thinking I got take or what." [ID014]
	"Because I'm scared to die, it motivates me to eat my medication, without even any reminder...I used to take care of old people, my grandmother, and I saw them suffering." [ID008]
Health information seeking	
Level	Intermediate
Actions	Seeks information from health care providers
	Has low connectedness to other sources beyond the media
	May have lost or not have access to important sources of health information and reminders
Example quote	"Every time she gives me call and I remember [to take]; she suddenly pass away, one month now. That's why I think, nobody helps me." [ID013]
Mobile usage	
Level	Low
Actions	Calls only
	May read SMS[a] text message but cannot reply
Usability concerns	
Concerns	Low literacy: cannot read the characters and limited English
	Challenges to learning
Example quote	"I've heard that we need to chase up with technology. But do you think everyone went to school? If you teach me what this is, I'll forget what that was." [ID005]
Attitudes toward mHealth	
Attitudes	Interest in mHealth, but only for SMS appointment reminders
Example quote	"Every time I cannot see the [appointment] card...so the [text] message comes, it's easy for me." [ID014]

[a]SMS: short message service.

Persona Application

These five personas present insight into our target user group for a treatment adherence intervention in Singapore. Based on these personas, we can see variability in multiple dimensions, including lifestyle and social factors, adherence factors, and mobile technology use and needs. These will be useful across the project process and will help to better refine the intervention implementation and dissemination strategy [11,24].

For example, within our population, The Socializer persona may provide an accessible group to initially focus on and engage with, as members of this group seek and disseminate health information via mobile phone. As this group has high social connectedness, they may also act as *influencers* and provide a

platform for scale-up and buy-in from their social networks over time. Thus, initial message content design may be tailored to and tested with this group and may account for their belief in their abilities to self-manage and their perceived high health literacy.

Conversely, both The Busy Grandparent and The Hard-to-Reach personas more keenly seek information from familiar trusted sources as compared to The Socializer; thus, when considering how to onboard users, these groups may prefer to be introduced to the intervention through a trusted health care provider. This information helps clarify stakeholders to be involved in the dissemination and diffusion of the intervention.

From an implementation perspective, our personas show that users are not only using SMS text messaging, but also other messaging platforms such as WhatsApp. The feasibility of using such apps for an intervention may be worth exploring in the Singapore context. We also see how The Quiet Analog persona prefers to use their phone for calls only, thus the intervention may require a multiplatform approach in order to reach the widest cross-section of users.

Finally, The Newly Diagnosed persona points toward the utility of the intervention in a previously unconsidered group; this provides the opportunity to explore a modified or targeted intervention to onboard patients at the point of diagnosis or shortly thereafter. While the different personas provide insight into the different ways that users engage with mobile phone use, they should be used as a guide rather than a direct prescription of how an intervention should change, given that the context in which these personas operate do influence an intervention's efficacy.

Discussion

Principal Findings

Our work explores the creation of biopsychosocial personas and how these may highlight different patterns of use and inform the design of mHealth interventions for adherence, as they assist in communicating the diversity in behaviors, preferences, and needs that can be represented through their use. Further, our study explores these personas in the unique context of Singapore.

Singapore has a rapidly aging population and increasing prevalence of chronic disease, which require innovative approaches to disease management [25]. With a high rate of mobile phone usage and digital penetration, as demonstrated by the 91% mobile phone ownership in our sample, Singapore is an ideal candidate for mHealth interventions. Therefore, our research provides a handle for technology designers to tackle behavioral variability in adherence among older adults.

Adherence behaviors are inherently complex and dependent on multiple factors, including the health system, socioeconomic factors, the medication itself, patient factors, as well as factors related to the condition being treated [26]. These factors were evident in our study sample, particularly with regard to health system factors, including ease of access and affordability of medications. Of note is the challenge in promoting adherence for asymptomatic chronic conditions such as hypertension and hyperlipidemia; indeed, our participants demonstrated self-titration and suboptimal adherence to these medications. Other studies have found similar findings, pointing to the challenge of promoting adherence to these medications and the need for innovative patient supports for adherence [27].

While mHealth shows promise in providing a far-reaching and cost-effective approach to facilitate adherence [28-30], caution is warranted in implementing these interventions indiscriminately to target populations. For example, in our study, participants reported a wide range of mobile phone usage behaviors, and even those who had high mobile use did not necessarily want to receive an mHealth intervention for their ASCVD medications. Patient-centered care seeks to provide appropriate services that address the unique needs and preferences of individuals; as such, mHealth approaches should adopt the user-centered design tenet of "know thy user" and use best practices in design to create appropriate and effective interventions in line with user needs and preferences [31]. For example, the majority of participants, when asked in the interview, reported not wanting an mHealth intervention for adherence; however, many reported liking an SMS text message appointment reminder. Thus, there is the potential to gain buy-in from these users as they are already using a similar service.

There is no one solution for digital approaches and users must be understood and contextualized. This is important when attempting to bridge the digital divide and provide mHealth interventions to aging populations who are not "digital natives" [32-34]. For example, our personas highlight the multiple usability issues reported by this population, including wanting content in languages other than English and physical barriers to use such as poor eyesight. We also highlight users who use their phones in specific ways, such as only for phone calls to family or for a wide range of activities such as games and videos, but who feel empowered to manage their own medications. To this end, health researchers must strive to consider a diversity of older adult user behaviors, including user motivations, attitudes, and preferences, in order to design effective and appropriate interventions for adherence to ASCVD medications.

Personas allow a middle path to highlighting the diversity of the user base in a manageable and actionable way. While it may not be feasible to design and implement mHealth interventions for the preferences or adherence patterns of each individual, as evidence has shown, it is also unwise to design for a user group that has been reduced to a homogenous typology, such as *adherent/nonadherent* or *elderly* [35]. Personas can then be used to begin discussions on appropriateness during the intervention design phase, enabling the creation of more robust user journeys by prompting discourse on how a given persona may react to a certain task, prompt, or call to action, and by providing reminders of who is being designed for throughout the research process [24,36]. As noted in our personas, a few groups expressed low interest in using their phone to assist in their medication taking. We include these personas as it is important to consider why certain groups may not want to take up an mHealth intervention and to move toward either strategies to gain their buy-in or different interventions to support them.

As the use of mHealth becomes increasingly popular, health researchers will need to adopt *design thinking* approaches and research methods, such as persona creation, in order to better provide contextually appropriate health interventions [13,24]. For example, a study by LeRouge et al explored the creation and use of personas in the Chinese Aged Diabetic Assistant app and showed how devoting time to the creation of personas in the early stages of an mHealth or eHealth (electronic health) intervention can help to shape design requirements, influence functional requirements, and guide the choice of appropriate channels for implementation and diffusion resulting in a contextually appropriate and acceptable intervention [24]. Thus, personas are an important bridging step to facilitate design discussions and allow important data on user experiences and perception to inform and shape the design process of mHealth interventions. Indeed, user-centered design aligns with the ethos of people-centered care; therefore, methods that better enable program designers and implementers to frame programs with diverse user groups in mind are an important piece to creating meaningful and effective mHealth services to promote adherence [37].

Strengths and Limitations

A strength of our study is the multidisciplinary project team and multi-method study design, which provides greater exploration of nuanced personal experiences and perspectives impacting adherence and preferences. Further, our study includes participants from multiple ethnic backgrounds, giving perspectives in both English, Mandarin, and Chinese dialects. Finally, to our knowledge, this is the first study of its kind in Singapore to construct personas to inform the design of an mHealth intervention for adherence among ASCVD patients. Further, the results of this study may be of use to inform designers and implementers of mHealth interventions in other Asian urban centers with high mobile phone use.

A limitation is that our study excluded individuals with disabilities that prevented them from participating in verbal interviews. Thus, we are missing the views of persons who are hard of hearing or with other disabilities who may benefit from an mHealth intervention for adherence. More research is needed in designing mHealth interventions and creating personas that reflect the unique needs of individuals living with disabilities. Also, our research focused primarily on mobile phone use for the purpose of an SMS text messaging intervention; with increasing mobile phone penetration, it would be worthwhile to explore in greater depth the use and feasibility of other messaging services (eg, WhatsApp and Line) for mHealth interventions. Further, as The Newly Diagnosed persona exemplifies, there is a need to explore interventions for those who are newly diagnosed to support them in achieving optimal adherence.

Implications for Future Research and Development

Future research in this area should include greater exploration of user journeys among diverse elderly groups to better understand how and when is best to deliver mHealth interventions. These studies should leverage additional user experience research techniques, including user acceptance and usability testing, as well as traditional public health study designs, such as longitudinal studies with routine follow-up, to determine the impact of messaging to different groups. Unlike persona use in other fields such as marketing [38], there is currently a dearth of robust health research on outcomes and evaluations of interventions that use personas. There is also a need for more multi- and mixed-method, user-centered research that synthesizes evidence on user behaviors and preferences for mHealth interventions for adherence among older persons. Finally, there is a need for greater collaboration between health researchers and those working in digital and creative fields to develop robust and feasible interventions that leverage innovative techniques and best practices.

Conclusions

As public health and medicine move toward wider adoption of mHealth, it is increasingly important that we pay due consideration to user-centered design techniques, and *design thinking* in order to create meaningful, patient-centered interventions for adherence to medications, as well as to evaluative if these methods result in more effective interventions. Our study leverages multi-method data to explore context, user behaviors, and preferences, and to put forth personas that typify a range of user group experiences to inform the design and creation of mHealth interventions to facilitate adherence to ASCVD medication.

Acknowledgments

This research was supported by the Humanities and Social Sciences Seed Fund (R-608-000-137-646) and the National University Health System (NUHS) Singapore Population Health Improvement Centre (SPHERiC).

Conflicts of Interest

None declared.

Multimedia Appendix 1

Interview guide.

[DOCX File, 15KB - mhealth_v7i5e10465_app1.docx]

Multimedia Appendix 2

Consolidated criteria for reporting qualitative research (COREQ) checklist.

[DOCX File, 18KB - mhealth_v7i5e10465_app2.docx]

Multimedia Appendix 3

Proforma legend and example.

[DOCX File, 503KB - mhealth_v7i5e10465_app3.docx]

Multimedia Appendix 4

Overview of personas.

[PDF File (Adobe PDF File), 4MB - mhealth_v7i5e10465_app4.pdf]

References

1. mHealth: New Horizons for Health Through Mobile Technologies. Based on the Findings of the Second Global Survey on eHealth. Geneva, Switzerland: World Health Organization; 2011. URL: https://www.who.int/goe/publications/goe_mhealth_web.pdf [accessed 2019-05-14] [WebCite Cache ID 78MuuXkNG]
2. Ali EE, Chew L, Yap KYL. Evolution and current status of mhealth research: A systematic review. BMJ Innov 2016 Jan 05;2(1):33-40. [doi: 10.1136/bmjinnov-2015-000096]
3. Asiimwe C, Gelvin D, Lee R, Ben Amor Y, Quinto E, Katureebe C, et al. Use of an innovative, affordable, and open-source short message service-based tool to monitor malaria in remote areas of Uganda. Am J Trop Med Hyg 2011 Jul;85(1):26-33 [FREE Full text] [doi: 10.4269/ajtmh.2011.10-0528] [Medline: 21734120]
4. Fassnacht D, Ali K, Silva C, Gonçalves S, Machado P. Use of text messaging services to promote health behaviors in children. J Nutr Educ Behav 2015;47(1):75-80. [doi: 10.1016/j.jneb.2014.08.006] [Medline: 25282200]
5. Fiordelli M, Diviani N, Schulz P. Mapping mHealth research: A decade of evolution. J Med Internet Res 2013 May 21;15(5):e95 [FREE Full text] [doi: 10.2196/jmir.2430] [Medline: 23697600]
6. Leon N, Surender R, Bobrow K, Muller J, Farmer A. Improving treatment adherence for blood pressure lowering via mobile phone SMS-messages in South Africa: A qualitative evaluation of the SMS-text Adherence SuppoRt (StAR) trial. BMC Fam Pract 2015 Jul 03;16:80 [FREE Full text] [doi: 10.1186/s12875-015-0289-7] [Medline: 26137844]
7. Burdick DC, Kwon S. Gerotechnology: Research and Practice in Technology and Aging. New York, NY: Springer Publishing Company; 2004.
8. Fisk AD, Rogers WA, Charness N, Czaja SJ, Sharit J. Designing for Older Adults: Principles and Creative Human Factors Approaches. 2nd edition. Boca Raton, FL: CRC Press; 2009.
9. Hanson VL. Age and Web access: The next generation. In: Proceedings of the International Cross-Disciplinary Conference on Web Accessibility (W4A). 2009 Presented at: International Cross-Disciplinary Conference on Web Accessibility (W4A); April 20-21, 2009; Madrid, Spain p. 7-15. [doi: 10.1145/1535654.1535658]
10. OECD. OECD Digital Economy Papers No. 49. Paris, France: OECD Publishing; 2001 Jan 01. Understanding the Digital Divide URL: https://www.oecd-ilibrary.org/understanding-the-digital-divide_5kzd927kqhd6.pdf?itemId=%2Fcontent%2Fpaper%2F236405667766&mimeType=pdf [accessed 2019-05-14] [WebCite Cache ID 78Mvo3fe1]
11. Pruitt J, Adlin T. The Persona Lifecycle: Keeping People in Mind Throughout Product Design. San Francisco, CA: Morgan Kaufmann Publishers; 2006.
12. Cooper A. The Inmates Are Running the Asylum. Indianapolis, IN: Sams Publishing; 1999.
13. Holden R, Kulanthaivel A, Purkayastha S, Goggins K, Kripalani S. Know thy eHealth user: Development of biopsychosocial personas from a study of older adults with heart failure. Int J Med Inform 2017 Dec;108:158-167 [FREE Full text] [doi: 10.1016/j.ijmedinf.2017.10.006] [Medline: 29132622]
14. Perel P, Avezum A, Huffman M, Pais P, Rodgers A, Vedanthan R, et al. Reducing premature cardiovascular morbidity and mortality in people with atherosclerotic vascular disease: The World Heart Federation Roadmap for secondary prevention of cardiovascular disease. Glob Heart 2015 Jun;10(2):99-110. [doi: 10.1016/j.gheart.2015.04.003] [Medline: 26213297]
15. Perel P, Casas J, Ortiz Z, Miranda J. Noncommunicable diseases and injuries in Latin America and the Caribbean: Time for action. PLoS Med 2006 Sep;3(9):e344 [FREE Full text] [doi: 10.1371/journal.pmed.0030344] [Medline: 16953660]
16. A Global Brief on Hypertension: Silent Killer, Global Public Health Crisis. Geneva, Switzerland: World Health Organization; 2013 Apr. URL: https://www.who.int/iris/bitstream/10665/79059/1/WHO_DCO_WHD_2013.2_eng.pdf?ua=1 [accessed 2019-05-14] [WebCite Cache ID 78MwySiWc]
17. Prevention of Cardiovascular Disease: Pocket Guidelines for Assessment and Management of Cardiovascular Risk. Geneva, Switzerland: World Health Organization; 2007. URL: https://www.who.int/cardiovascular_diseases/guidelines/PocketGL.ENGLISH.AFR-D-E.rev1.pdf [accessed 2019-05-08] [WebCite Cache ID 78DxlwoNA]

18. Bansilal S, Castellano J, Fuster V. Global burden of CVD: Focus on secondary prevention of cardiovascular disease. Int J Cardiol 2015 Dec;201 Suppl 1:S1-S7 [FREE Full text] [doi: 10.1016/S0167-5273(15)31026-3] [Medline: 26747389]
19. Legido-Quigley H, Camacho Lopez PA, Balabanova D, Perel P, Lopez-Jaramillo P, Nieuwlaat R, et al. Patients' knowledge, attitudes, behaviour and health care experiences on the prevention, detection, management and control of hypertension in Colombia: A qualitative study. PLoS One 2015;10(4):e0122112 [FREE Full text] [doi: 10.1371/journal.pone.0122112] [Medline: 25909595]
20. Rashid M, Edwards D, Walter F, Mant J. Medication taking in coronary artery disease: A systematic review and qualitative synthesis. Ann Fam Med 2014;12(3):224-232 [FREE Full text] [doi: 10.1370/afm.1620] [Medline: 24821893]
21. Zhang Y, Bai L, Shi M, Lu H, Wu Y, Tu J, et al. Features and risk factors of carotid atherosclerosis in a population with high stroke incidence in China. Oncotarget 2017 Aug 22;8(34):57477-57488. [doi: 10.18632/oncotarget.15415] [Medline: 28915687]
22. Charmaz K. Constructing Grounded Theory: A Practical Guide through Qualitative Analysis. London, UK: SAGE Publications; 2006.
23. Strauss AL. Qualitative Analysis for Social Scientists. Cambridge, UK: Cambridge University Press; 1987.
24. LeRouge C, Ma J, Sneha S, Tolle K. User profiles and personas in the design and development of consumer health technologies. Int J Med Inform 2013 Nov;82(11):e251-e268. [doi: 10.1016/j.ijmedinf.2011.03.006] [Medline: 21481635]
25. Ho K, Suok KC, Chien EL, Hong CL, Kee TG. State of Health: Report of the Director of Medical Services 2003-2012. Singapore: Singapore Ministry of Health; 2013.
26. Sabaté E. Adherence to Long-Term Therapies: Evidence for Action. Geneva, Switzerland: World Health Organization; 2003. URL: https://www.who.int/chp/knowledge/publications/adherence_full_report.pdf?ua=1 [accessed 2019-05-14] [WebCite Cache ID 78MyH2w2N]
27. Adler A, Martin N, Mariani J, Tajer CD, Owolabi OO, Free C, et al. Mobile phone text messaging to improve medication adherence in secondary prevention of cardiovascular disease. Cochrane Database Syst Rev 2017 Apr 29;4:CD011851 [FREE Full text] [doi: 10.1002/14651858.CD011851.pub2] [Medline: 28455948]
28. Catalani C, Philbrick W, Fraser H, Mechael P, Israelski D. mHealth for HIV treatment & prevention: A systematic review of the literature. Open AIDS J 2013;7:17-41 [FREE Full text] [doi: 10.2174/1874613620130812003] [Medline: 24133558]
29. 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]
30. Marcolino MS, Oliveira JAQ, D'Agostino M, Ribeiro AL, Alkmim MBM, Novillo-Ortiz D. The impact of mHealth interventions: Systematic review of systematic reviews. JMIR Mhealth Uhealth 2018 Jan 17;6(1):e23 [FREE Full text] [doi: 10.2196/mhealth.8873] [Medline: 29343463]
31. Miaskiewicz T, Kozar KA. Personas and user-centered design: How can personas benefit product design processes? Des Stud 2011 Sep;32(5):417-430. [doi: 10.1016/j.destud.2011.03.003]
32. Heinz M, Martin P, Margrett J, Yearns M, Franke W, Yang HI, et al. Perceptions of technology among older adults. J Gerontol Nurs 2013 Jan;39(1):42-51. [doi: 10.3928/00989134-20121204-04] [Medline: 23244061]
33. Levine D, Lipsitz SR, Linder J. Trends in seniors' use of digital health technology in the United States, 2011-2014. JAMA 2016 Aug 02;316(5):538-540. [doi: 10.1001/jama.2016.9124] [Medline: 27483069]
34. Prensky M. H sapiens digital: From digital immigrants and digital natives to digital wisdom. Innovate (North Miami Beach) 2009;5(3):1-9 [FREE Full text]
35. Dinsmore J, McDonnell R. mHealth translation framework: Behavioral design and development to improve self-management in chronic health conditions. In: Adibi S, editor. mHealth Multidisciplinary Verticals. Boca Raton, FL: CRC Press; 2015:181-198.
36. Williams I, Brereton M, Donovan J, McDonald K, Millard T, Tam A, et al. A collaborative rapid persona-building workshop: Creating design personas with health researchers. Int J Sociotechnol Knowl Dev 2014 Apr;6(2):17-35. [doi: 10.4018/ijskd.2014040102]
37. People-Centered Health Care: A Policy Framework. Manila, Philippines: World Health Organization, Western Pacific Region; 2007. URL: http://www.wpro.who.int/health_services/people_at_the_centre_of_care/documents/ENG-PCIPolicyFramework.pdf?ua=1 [accessed 2019-05-14] [WebCite Cache ID 78MzXBt6X]
38. Martell K. Cintell. 2016 Mar 10. 38 handy stats to prove the value of personas URL: https://www.cintell.net/38-handy-stats-to-prove-the-value-of-personas/ [accessed 2019-05-14] [WebCite Cache ID 78Mzs7mox]

Abbreviations

ASCVD: atherosclerotic cardiovascular disease
COREQ: consolidated criteria for reporting qualitative research
eHealth: electronic health
MARS-5: Medication Adherence Report Scale
mHealth: mobile health

MRT: mass rapid transit
NUHS: National University Health System
SMS: short message service
SPHERiC: Singapore Population Health Improvement Centre

Edited by G Eysenbach; submitted 21.03.18; peer-reviewed by T Powell-Wiley, M Harris; comments to author 06.09.18; revised version received 07.12.18; accepted 10.12.18; published 28.05.19.

Please cite as:
Haldane V, Koh JJK, Srivastava A, Teo KWQ, Tan YG, Cheng RX, Yap YC, Ong PS, Van Dam RM, Foo JM, Müller-Riemenschneider F, Koh GCH, Foong PS, Perel P, Legido-Quigley H
User Preferences and Persona Design for an mHealth Intervention to Support Adherence to Cardiovascular Disease Medication in Singapore: A Multi-Method Study
JMIR Mhealth Uhealth 2019;7(5):e10465
URL: http://mhealth.jmir.org/2019/5/e10465/
doi:10.2196/10465
PMID:31140445

Original Paper

Who is Tracking Health on Mobile Devices: Behavioral Logfile Analysis in Hong Kong

Lu Guan[1], MA; Tai-Quan Peng[2*], PhD; Jonathan JH Zhu[1,3*], PhD

[1]Department of Media and Communication, City University of Hong Kong, Hong Kong, China (Hong Kong)

[2]Department of Communication, Michigan State University, East Lansing, MI, United States

[3]School of Data Science, City University of Hong Kong, Hong Kong, China (Hong Kong)

[*]these authors contributed equally

Corresponding Author:
Tai-Quan Peng, PhD
Department of Communication
Michigan State University
404 Wilson Road
East Lansing, MI, 48824
United States
Phone: 1 5173550221
Email: winsonpeng@gmail.com

Abstract

Background: Health apps on mobile devices provide an unprecedented opportunity for ordinary people to develop social connections revolving around health issues. With increasing penetration of mobile devices and well-recorded behavioral data on such devices, it is desirable to employ digital traces on mobile devices rather than self-reported measures to capture the behavioral patterns underlying the use of mobile health (mHealth) apps in a more direct and valid way.

Objective: The objectives of this study were to (1) assess the demographic predictors of the adoption of mHealth apps; (2) investigate the temporal pattern underlying the use of mHealth apps; and (3) explore the impacts of demographic variables, temporal features, and app genres on the use of mHealth apps.

Methods: Logfile data of mobile devices were collected from a representative panel of about 2500 users in Hong Kong. Users' mHealth app activities were analyzed. We first conducted a binary logistic regression analysis to uncover demographic predictors of users' adoption status. Then we utilized a multilevel negative binomial regression to examine the impacts of demographic characteristics, temporal features, and app genres on mHealth app use.

Results: It was found that 27.5% of mobile device users in Hong Kong adopt at least one genre of mHealth app. Adopters of mHealth apps tend to be female and better educated. However, demographic characteristics did not showcase the predictive powers on the use of mHealth apps, except for the gender effect (B_{female} vs B_{male}=–0.18; P=.006). The use of mHealth apps demonstrates a significant temporal pattern, which is found to be moderately active during daytime and intensifying at weekends and at night. Such temporal patterns in mHealth apps use are moderated by individuals' demographic characteristics. Finally, demographic characteristics were also found to condition the use of different genres of mHealth apps.

Conclusions: Our findings suggest the importance of dynamic perspective in understanding users' mHealth app activities. mHealth app developers should consider more the demographic differences in temporal patterns of mHealth apps in the development of mHealth apps. Furthermore, our research also contributes to the promotion of mHealth apps by emphasizing the differences of usage needs for various groups of users.

(JMIR Mhealth Uhealth 2019;7(5):e13679) doi:10.2196/13679

KEYWORDS

mobile apps; mHealth; circadian rhythm

Introduction

Background

Mobile technologies, including wireless devices and sensors that are intended to be worn, carried, or accessed by users during normal daily activities [1], have brought significant transformations to many practices in public health. In particular, because of *always-on, always-on you* [2] quality of mobile technologies, health apps on mobile devices (hereinafter labeled "mHealth apps") have provide unprecedented opportunities for ordinary people to develop social connections revolving around health issues, seek health-related information, track daily workouts, manage chronic health conditions, and make medical appointments [3,4]. mHealth apps are found to play important roles in changing users' health behaviors such as smoking cessation, weight management, and physical activity [5-8].

Understanding the adoption and use of mHealth apps has become an increasingly important research question in public health research. According to a population-based survey in Germany in 2015, 21% of smartphone users adopt mHealth apps [9]. Demographic factors are found to predict the adoption and use of mHealth apps. Users of mHealth apps tend to be younger, better educated, and in more healthy conditions [10]. However, previous research on mHealth apps was dominated by self-reported measures in surveys or interviews, in which participants were solicited to report their use of mHealth apps and other health-related behaviors. It has been widely acknowledged that the validity of self-reported measurements is threatened by many confounding factors such as memory errors, intentional distortion, cognitive factors, and social context factors [11]. Individuals may underestimate their health-risk behaviors to avoid the pressure of breaking social norms, whereas they may overestimate their health-conducive behaviors to meet social desirability.

With well-recorded behavioral data on mobile devices, it is desirable to employ digital traces on mobile devices rather than self-reported measures to capture behavioral patterns underlying the use of mHealth apps in a more direct and valid way. More importantly, the time-stamped digital traces can empower researchers to understand the temporal patterns, such as circadian rhythms and weekday-weekend difference, in the use of mHealth apps. The circadian rhythm in human behavior plays a crucial role in guiding people's daily lives [12]. Empirical studies have found that maintaining the biological circadian rhythm is a necessity for human health [13]. Users of mHealth apps, who have relatively high health consciousness and health literacy [14], are expected to follow certain circadian rhythm in their use of mHealth apps. However, the fine-grained temporal information is extremely difficult, if not impossible, to be garnered in traditional self-reported survey research. The time-stamped behavioral data on mobile devices provide an unprecedented opportunity to understand how and when individuals will use mHealth apps.

Hong Kong provides a very good testbed to understand the adoption and use of mHealth apps. Hong Kong is one of the most populous cities in the world but also among the top 10 regions with the greatest life expectancy at birth [15]. Mobile cellular and the internet have been widely adopted in Hong Kong, whose penetration rates were 250% for mobile cellular by 2017 [16] and 87% for the internet by 2018 [17]. Among all internet users in Hong Kong, 98% of them go on the internet via mobile devices.

Objectives

By analyzing the behavioral log data of mobile devices collected from a representative panel of ordinary users in Hong Kong, we aim to (1) assess the demographic predictors of the adoption of mHealth apps, (2) investigate the temporal pattern underlying the use of mHealth apps, and (3) disentangle the intertwined effects of demographic factors, temporal features, and app genres on the use of mHealth apps. The findings of this study advance our knowledge of demographic profiles of the adopters of mHealth apps and enhance our understanding of behavioral patterns underlying the use of mHealth apps, which possesses significant implications for promoting the adoption and use of mHealth apps in different dimensions of public health.

Methods

Data Collection

The data of the study were collected from a mobile audience study in Hong Kong. Users in the panel were recruited by a marketing research firm for local media organizations. To assure the representativeness of the panel sample, the panel was weighted against Hong Kong population census estimates in terms of the cross-distribution of age and gender. An on-device meter was installed on mobile devices (using Android and iPhone) of the users in the sample, to passively track all the activities on the devices without interrupting the users, who were informed about and agreed to the tracking and subsequent analyses of mobile app uses under the anonymous condition. The starting and ending time of each app use was recorded by the installed meter. The observation period for this study was from July to November 2017. All personally identifiable information about the users was removed before we were given access to the data. To further protect user privacy, the specific name of all mobile apps was removed, although the generic category (eg, Communication, News, Health, Games) is available, which provides the basis for this study to measure mHealth app use.

The apps adopted by all users were categorized into dozens of genres based on Google mobile app categories. Only the adoption and use of health apps are examined in this study. Of all the apps adopted by the users, 155 are health apps, which are categorized into 5 genres based on their functional features: generic activity tracking apps (51/155, 32.9%), health records log apps (36/155, 23.2%), sleep management and relaxation apps (29/155, 18.7%), training and coaching apps (23/155, 14.8%), and weight and diet management apps (16/155, 10.3%).

Data Analysis

To understand who is more likely to adopt mHealth apps, a binary logistic regression is run to predict the adoption status of mHealth apps among mobile device users, with adopters of mHealth apps coded as 1 and nonadopters coded as 0. This analysis includes the whole panel of all users (N=2,591).

Demographic variables, including gender, age, education degree, occupation status, parenting status, and marital status, are included as explanatory variables in the model.

Furthermore, following Dutton et al's conceptualization of patterns of use for technologies [18], the use of mHealth apps is operationalized as the duration of each app use, which is calculated as the time lag between the start time and end time of each app use. To understand the temporal patterns underlying the use of mHealth apps, 2 temporal features are constructed based on the temporal information of each app use. First, the hour of a day when an mHealth app is triggered is recoded into 5 time windows on a 24-hour cycle: morning (6-10), noon (11-13), afternoon (14-17), evening (18-21), and night (22-5). Second, the weekday for the occurrence of an mHealth app use is also recoded into a binary variable with weekday (ie, Monday to Friday) being coded as 0 and weekends (ie, Saturday and Sunday) being coded as 1.

Then, a multilevel negative binomial regression is run to account for within-individual and between-individual level differences on the use of mHealth apps. This analysis only includes participants who adopted the mHealth apps (N=713). Negative binomial regression is employed here to handle the dependent variable with discrete probability distributions, such as count data and time data [19]. As the use records under the same user could be influenced by the unique characteristics of that particular user, a multilevel design is employed to unravel the within-individual and between-individual variations in the use of mHealth apps. Specifically, users' demographic characteristics are included as independent variables at the between-individual level, whereas the 2 constructed temporal features (ie, daily time windows and weekday-weekend) and the genres of mHealth apps are included as independent variables at the within-individual level.

Results

Descriptive Statistics

Table 1 shows the descriptive statistics of demographic variables for all users in the panel. The average age of all users in the panel is 34 years, and 47.20% (1223/2591) are women. Among 2591 users in the panel, 40.80% (1057/2591) have high education level (ie, college degree or above), 20.57% (533/2591) medium education level (ie, associate degree), and 38.52% (998/2591) low education level (ie, high school education or lower). Users in the panel work as managers, administrators, or professionals (746/2591, 28.79%); clerks (640/2591, 24.70%); workers (482/2591, 18.60%); and students (396/2591, 15.28%), whereas 12.50% (583/2591) are unemployed (including housewives and retired). A total of 41.10% (1065/2591) are married, and 33.19% (860/2591) have at least 1 kid at home.

It was found that 27.52% (713/2591) of the users in the study adopt at least one genre of mHealth apps. The majority of the users (578/713, 81.1%) adopt only 1 genre of mHealth apps, 15.0% (107/713) adopt 2 genres, and 3.8% (27/713) adopt at least 3 genres.

The most popular genre of mHealth apps is the generic activity tracking, followed by health records log apps, weight and diet management apps, training and coaching apps, and sleep management and relaxation apps. On average, users spend 85.8 seconds in 1-time health app usage (SD=238.2 seconds). Users spend the longest time on sleep management and relaxation apps (mean=160.8 seconds, SD=435 seconds), followed by health records log apps (mean=94.8 seconds, SD=261 seconds), training and coaching apps (mean=91.8 seconds, SD=208 seconds), generic activity tracking (mean=79.8 seconds, SD=220 seconds), and training and coaching apps (mean=70.8 seconds, SD=169.8 seconds).

Who Adopts Mobile Health Apps?

The binary logistic regression analysis results are summarized in Table 2. Female users are more likely to adopt mHealth apps than males (odds ratio, OR=1.44; $P<.001$). Better-educated users are more likely to adopt mHealth apps than less-educated users. Clerks are more likely to adopt mHealth apps, in comparison with workers (OR=1.68; $P<.001$). However, age, marital status, and parenting status of mobile device users do not significantly influence the adoption of mHealth apps.

Table 1. Descriptive statistics of demographic variables (N=2591).

Demographic variables	Frequency, n (%)
Gender	
Male	1368 (52.80)
Female	1223 (47.20)
Age group (years)	
18-30	1164 (44.92)
31-45	993 (38.32)
46-64	434 (16.75)
Education	
Low education level	998 (38.52)
Medium education level	534 (20.61)
High education level	1059 (10.87)
Occupation	
Managers, administrators, and professionals	746 (28.79)
Clerks	641 (24.74)
Workers	482 (18.60)
Students	397 (15.32)
Unemployed	325 (12.54)
Parenting status	
With kids at home	1730 (66.77)
Without kids	861 (33.23)
Marital status	
Married	1525 (58.86)
Single, divorced or widowed	1066 (41.14)

Table 2. Binary logistic regression coefficients predicting the adoption of mobile health apps[a] (N=2591).

Independent variables	Odds ratio (95% CI)	*P* value
Female (vs male)	1.44 (1.20-1.73)	<.001
Age (years)	0.99 (0.98-1.00)	.09
Education (vs low education level)		
Medium education level	1.34 (1.04-1.71)	.02
High education level	1.45 (1.16-1.81)	<.001
Occupation (vs workers)		
Managers, administrators, and professionals	1.27 (0.94-1.72)	.12
Clerks	1.69 (1.27-2.24)	.02
Students	1.17 (0.83-1.65)	.37
Unemployed	1.21 (0.86-1.69)	.28
Parenting status (with kids at home=1, without kids=0)	0.95 (0.77-1.18)	.66
Marital status (married=1, single/divorced/widowed=0)	1.24 (0.98-1.57)	.07

[a]Cox and Snell R^2=0.022.

What Accounts for the Duration of Mobile Health App Use?

A null model with the intercept only is first estimated to get the intraclass correlation coefficient (ICC). The ICC is 0.36, indicating that 36% of the variance in the use of mHealth apps could be attributed to between-individual differences (eg, the demographic characteristics of individuals). The high ICC detected here also implies that a multilevel design can provide adequate statistical power to account for the use of mHealth apps. A series of 3 multilevel negative binomial regressions are run to uncover the between-individual and within-individual antecedents underlying the use of mHealth apps. The analytical results of main effects are summarized in the Multimedia Appendix 1.

The conditional R^2 is estimated for the 3 multilevel models reported in Multimedia Appendix 1. The conditional R^2 describes the proportion of variance explained by both the between-level and within-individual factors included in the 3 models. Model 1, which includes the between-individual variables as independent variables, explains 35.6% of the variance in the use of mHealth apps. The addition of within-individual factors in model 2 increases the explanatory power by 0.8% (P<.001; likelihood-ratio, LR $\chi^2_{9=}$ 145). Model 3, which incorporates the interactions between within-level factors (ie, daily time window, weekday-weekend, and health app genre) and between-individual factors, adds 1.1% of the variance (P<.001; LR $\chi^2_{90=}$ 609).

Gender is the only significant demographic predictor of use of mHealth apps. Female users (estimated duration=80.2 seconds, 95% CI [72.9-87.6]) tend to spend less time on mHealth apps than male users (estimated duration=96.1 seconds, 95% CI [86.3-105.9]). It is surprising to find that all other between-individual variables, including age, education degree, occupation status, parenting status, and marital status, do not exert statistically significant main effects on the use of mHealth apps.

Temporal patterns have significant effects on the use of mHealth apps. Users are found to spend more time on mHealth apps during weekends (estimated duration=88.3 seconds) than during weekdays (estimated duration=85.2 seconds). Meanwhile, the use of mHealth apps tends to be more enduring at night (estimated duration=93.2 seconds), whereas less enduring in the morning (estimated duration=82.9 seconds).

There are significant differences among the use of different genres of mHealth apps. Training and coaching apps is the most enduring (estimated duration=126 seconds), followed by sleep management and relaxation apps (estimated duration=110 seconds), health records log apps (estimated duration=95.2 seconds), weight and diet management apps (estimated duration=90.2 seconds), and generic activity tracking apps (estimated duration=79.3 seconds).

The temporal patterns in the use of mHealth apps are found to be moderated by between-individual variables, as suggested by the significant cross-level interaction effects between temporal features, app genres, and demographic characteristics.

First, the gender difference on the use of mHealth apps is significantly greater in the morning than that in other time windows of the day, as shown in Figure 1. Second, users in different age groups demonstrate different temporal patterns in the use of mHealth apps. Users of different age groups spend significantly similar amount of time on mHealth apps at noon and night, whereas older users spend less time on mHealth apps in the morning and evening compared with the youngsters.

Third, users' education level is found to condition the temporal patterns in the use of mHealth apps. Users with different education levels tend to spend similar amount of time on mHealth apps in the morning and at night, whereas better-educated users spend significantly less time on mHealth apps in the daytime and during the weekends.

Besides, occupation is found to moderate the temporal patterns in the use of mHealth apps. Users of different occupations spend similar amount of time on mHealth apps at noon and night, whereas in the afternoon and evening time, housewives and retired users spend notably more time, and students spend significantly less time. Besides, student and unemployed users are found to have more stable use time schedules on mHealth apps no matter it is weekday or weekend, whereas workers, managers, administrators, and professionals spend significantly less time on mHealth apps during the weekdays than weekends.

Finally, our work also observed that demographics could differ the use of different genres of mHealth apps. First, the gender difference on the use of sleep management and relaxation apps is significantly greater than that of generic activity tracking apps. Second, education level is found to moderate the use of different genres of mHealth apps. Users with different education levels tend to spend similar amount of time on the use of generic activity tracking apps, whereas users with middle-level education spend significantly less time on training and coaching apps.

As shown in Figure 2, users' occupation is found to condition the use of different genres of mHealth apps. Managers, administrators, and professionals spend longer time on training and coaching apps and less time on sleep management and relaxation apps. Clerks are found to spend the longest time on sleep management and relaxation apps among all the occupations.

Parenting and marital status are also found to be moderate the use of different genres of mHealth apps. The differences between users with different parenting status are larger in the use of training and coaching apps compared with the use of generic activity tracking apps, whereas the differences between users with different marital status are larger in the use of health records log apps and sleep management apps.

Figure 1. Interaction effect of gender and time window of the day on the length of use of mobile health apps.

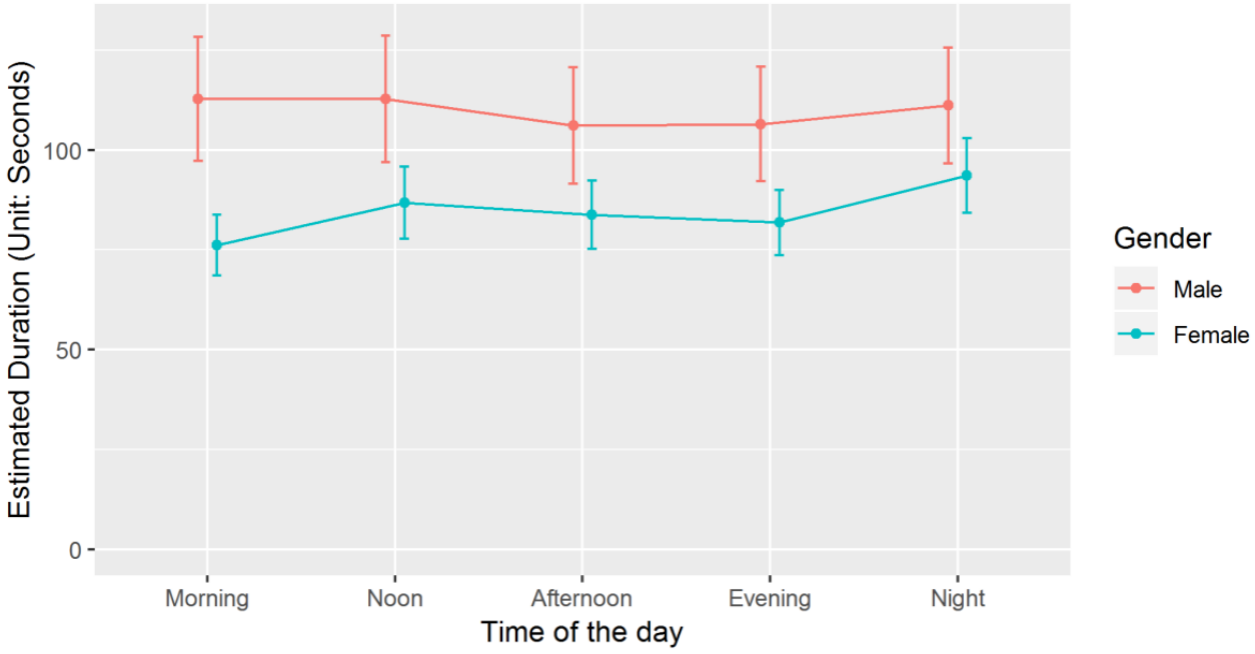
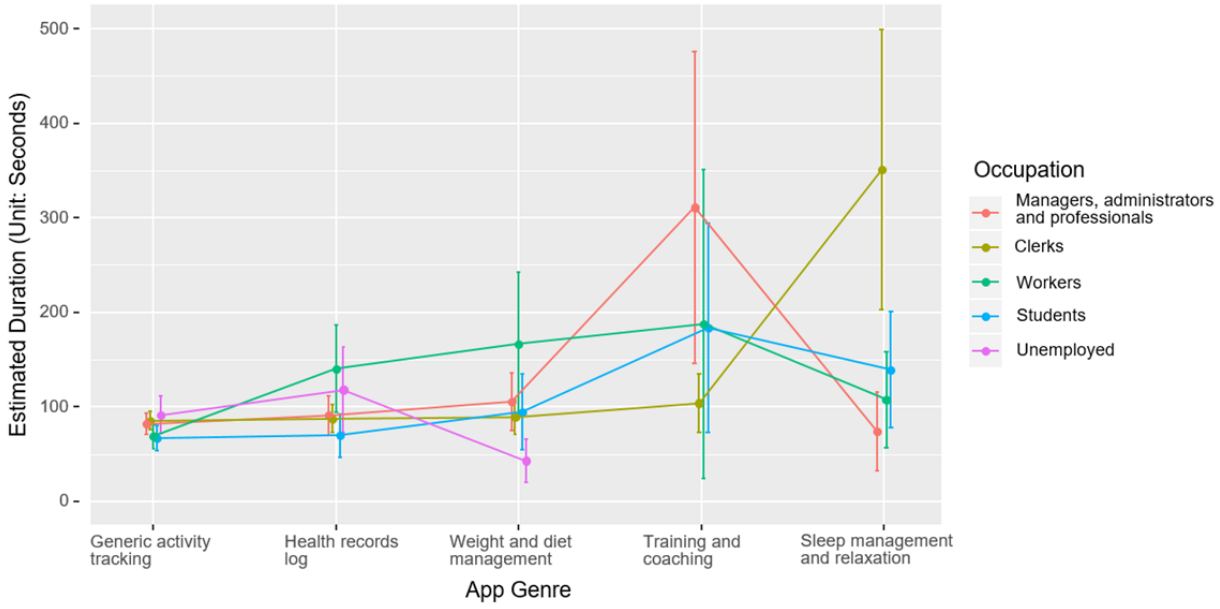


Figure 2. Interaction effect of occupation and app genre on the length of use of mobile health apps.



Discussion

Principal Findings

The first objective of this study was to assess the demographic predictors of health app adoption and use with a valid logfile dataset. Our results are consistent with previous studies that better educated and female mobile device users are more likely to adopt mHealth apps. However, we did not find the significant main effect of other demographic variables on the use of mHealth app, except for the gender difference. Second, our study aims to investigate the temporal pattern of health app use. The results indicate that the use of mHealth apps not only demonstrates a circadian rhythm but also differentiates between the weekdays and weekends. Noon and night are the peaks for the use of mHealth apps around a day. Besides, users spend longer time on mHealth apps on weekends than on weekdays. Finally, the findings show significant interaction effects between temporal features, app genres, and demographic characteristics on the use of mHealth apps. The gender difference on the use of mHealth apps is significantly greater in the morning. Users of different age groups spend significantly similar amount of time on mHealth apps at noon and night, whereas older users spend less time on mHealth apps in the morning and evening compared with the youngers. Besides, better-educated users spend significantly less time on mHealth apps in the daytime and during the weekends. We also found that demographics (gender, education, occupation, parenting status, and marital status) could differ the use of different genres of mHealth apps.

Comparison With Prior Work

Compared with prior studies, the adoption rate of mHealth apps in Hong Kong revealed in this study is quite higher than that found in Germany [10] and the United States [14] and that found in Hong Kong in 2016 [20]. Consistent with prior research [9,10], our work shows that females with better education background are more likely to adopt a health app. The consistency indicates a relatively high validity of self-report measures on health app adoption. However, our results show some discrepancy with prior studies on the use of mHealth apps. Only gender is found to play a role in explaining the mHealth app use in our study. However, we did not find the significant main effects of other demographic variables (eg, age and education) as claimed in previous studies. This is possibly because of the difference of the measurements. Prior studies measured the app use as individual's self-reported frequency of use, whereas our study aims to measure it in a more correct way, as calculating the actual duration of app use. As individuals may overestimate their health activity frequency to meet the social desirability, the validity of self-reported measure is often questioned especially in the frequency estimates [21]. Further efforts are needed to verify our results and re-examine the demographic effects in terms of the degree of health app use.

The study also finds a weekly and daily circadian rhythm of mHealth apps use. People spend longer time on health apps on the weekends than in a weekday. Besides, health app use reaches the peak at noon and night, throughout the 24-hour cycle a day. The pattern is largely matched with the temporal pattern in the use of mobile devices that mobile phone use being moderately active during the daytime but intensifying during break time and after work time [22,23]. These could be explained by the time availability theory that one of the most powerful predictors of technology use is whether the users have available time at that certain moment [24,25]. People are more likely to use mHealth apps during the weekends and the break time at noon and night because they have more spare time to engage in mobile app activities.

Finally, our findings provide evidence that users of different demographic (ie, gender, age, education, occupation, and parenting status) demonstrate different temporal patterns in the use of mHealth apps across circadian rhythm and the day of the week. First, the use difference between male and female in the morning indicates the gender role in the household: wives generally take more housework and child care responsibilities at home in the morning [26]. The unequal division of household workload responds to the gender differences of time allocation on health-related activities. Second, our result suggests that family role could also influence the temporal patterns in the use of mHealth apps for users with different parenting status. Users who have kids at home spend less time on mHealth apps during the weekends than the weekdays, which is exactly inversed with the users without kids. This is possibly because the workload for parents increases dramatically when their children come back home on the weekends. Besides, our work indicates that users of different occupations have different time schedules in the use of mHealth apps. The differences may reflect the outcomes of the social division of labor. Time in paid labor could constraint user's time allocation on other activities (eg, household time, leisure time, and health-related activity time) [27]. Thus, the employed and unemployed share distinct circadian clock patterns in the use of mHealth apps. Finally, our findings provide evidences for educational and age differences in the temporal patterns of the use of mHealth apps. The reasons are unclear but may be related to the social norms and cohort effects in health habit formation.

Limitations and Future Work

This study is subject to some limitations that can be addressed in future studies. For example, we lack the information about user's health conditions, initial motives, and behavioral outcomes in the use of mHealth apps. Research has found strong predictive powers of motivations on health-related behavior change (eg, HIV prevention and smoking cessation) [28]. People with different health conditions and motivations may also demonstrate different temporal patterns in the use of mHealth apps, which may further lead to different behavioral change outcomes. For example, people with stronger motivations may have more stable use schedules no matter the time window of the day or the day of the week, which may positively influence the health-related behavioral change. Future studies could consider combining the digital track data with survey or interview approaches to investigate the cognitive and biological mechanism underlying the health-related behavioral change in terms of the temporal patterns.

Second, although the meter in this study could passively track the activities on the devices without any interruptions to the panelists, there is concern that surveillance may have effects on user's behaviors on mobile phone. One possible assumption is that panelists may avoid doing extreme activities (eg, watching adult movies) and try to conduct more health-conducive activities to meet social desirability. However, this limitation should have minimal impacts on the validity of the results, as the efforts for panelists to take to meet the social desirability in this study are relatively high (eg, users have to participate in real training and coaching activities for several months). This impact could be rather minimal compared with that of self-reported measures because participants in survey method could easily conceal their real status and report more frequent health activities without any actual efforts.

Conclusions

By analyzing the behavioral log data of mobile devices collected from a representative panel in Hong Kong, this study explores the temporal patterns in the use of mHealth apps and examines the intertwined effects of demographic factors, temporal features, and app genres on the use of mHealth apps. Users of different demographic characteristics are found to have their own preference of the app genres and distinct schedules on app use across the circadian rhythm and the week.

Our study could contribute to the public health research and industries both theoretically and practically. First, our research adds to existing research by reporting the temporal patterns of mHealth app use. Maintaining the biological circadian rhythm is a necessity for human health. The intervention of the circadian rhythm in an individual's health-related habit could influence their health situation. Thus, understanding the temporal patterns

of individual's health app use is of great importance for health maintaining and health intervention research.

Second, our study could contribute to the promotion of health-related apps and activities. Our work shows that users with different demographic characteristics prefer different genres of mHealth apps. For example, managers, administrators, and professionals spend more time on training and coaching apps, whereas less time on sleep management and relaxation apps. Thus, our findings can assist the health promotion practitioner accurately select the appropriate target group and meet the health needs of the target audience.

Finally, our research could also contribute to the development of mHealth apps. We explored the temporal patterns in the use of mHealth apps and found that users of different demographic (ie, gender, age, education, occupation, and parenting status) have different use schedules on mHealth apps across the circadian rhythm and the week. Thus, mHealth app developers could consider more about the demographic differences in temporal patterns when they design and develop health apps to meet the customers' needs.

Acknowledgments

The study was supported in part by grant GE100028 from Michigan State University and by grant 9360120 from City University of Hong Kong.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Multilevel negative binomial regression predicting the use of mobile health apps.

[PDF File (Adobe PDF File), 119KB - mhealth_v7i5e13679_app1.pdf]

References

1. Kumar S, Nilsen WJ, Abernethy A, Atienza A, Patrick K, Pavel M, et al. Mobile health technology evaluation: the mHealth evidence workshop. Am J Prev Med 2013 Aug;45(2):228-236 [FREE Full text] [doi: 10.1016/j.amepre.2013.03.017] [Medline: 23867031]
2. Turkle S. Always-on/Always-on-you: the tethered self. Handbook Mobile Commun Stud 2008:121-138. [doi: 10.7551/mitpress/9780262113120.003.0010]
3. Silva BM, Rodrigues JJ, de la Torre Díez I, López-Coronado M, Saleem K. Mobile-health: a review of current state in 2015. J Biomed Inform 2015 Aug;56:265-272 [FREE Full text] [doi: 10.1016/j.jbi.2015.06.003] [Medline: 26071682]
4. Somers C, Grieve E, Lennon M, Bouamrane M, Mair FS, McIntosh E. Valuing mobile health: an open-ended contingent valuation survey of a National Digital Health Program. JMIR Mhealth Uhealth 2019 Jan 17;7(1):e3. [doi: 10.2196/mhealth.9990]
5. 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]
6. Marler JD, Fujii CA, Utley DS, Tesfamariam LJ, Galanko JA, Patrick H. Initial assessment of a comprehensive digital smoking cessation program that incorporates a mobile app, breath sensor, and coaching: cohort study. JMIR Mhealth Uhealth 2019 Feb 4;7(2):e12609. [doi: 10.2196/12609]
7. Langford AT, Solid CA, Scott E, Lad M, Maayan E, Williams SK, et al. Mobile phone ownership, health apps, and tablet use in us adults with a self-reported history of hypertension: cross-sectional study. JMIR Mhealth Uhealth 2019 Jan 14;7(1):e12228. [doi: 10.2196/12228]
8. Heron KE, Romano KA, Braitman AL. Mobile technology use and mHealth text message preferences: an examination of gender, racial, and ethnic differences among emerging adult college students. Mhealth 2019;5:2 [FREE Full text] [doi: 10.21037/mhealth.2019.01.01] [Medline: 30842950]
9. Ernsting C, Dombrowski SU, Oedekoven M, Kanzler M, Kuhlmey A, Gellert P. Using smartphones and health apps to change and manage health behaviors: a population-based survey. J Med Internet Res 2017 Apr 5;19(4):e101 [FREE Full text] [doi: 10.2196/jmir.6838] [Medline: 28381394]
10. Krebs P, Duncan DT. Health app use among us mobile phone owners: a national survey. JMIR Mhealth Uhealth 2015 Nov;3(4):e101 [FREE Full text] [doi: 10.2196/mhealth.4924] [Medline: 26537656]
11. del Boca FK, Darkes J. The validity of self-reports of alcohol consumption: state of the science and challenges for research. Addiction 2003 Dec;98 Suppl 2:1-12. [doi: 10.1046/j.1359-6357.2003.00586.x] [Medline: 14984237]
12. Farhud D, Aryan Z. Circadian rhythm, lifestyle and health: a narrative review. Iran J Public Health 2018 Aug;47(8):1068-1076 [FREE Full text] [Medline: 30186777]

13. Knutsson A. Health disorders of shift workers. Occup Med-C 2003;53(2):103-108. [doi: 10.1093/occmed/kqg048]
14. Fox S, Duggan M. Pew Research Center. 2012. Mobile Health 2012 URL: http://www.pewinternet.org/2012/11/08/mobile-health-2012/ [WebCite Cache ID 75o0q4IH4]
15. Central Intelligence Agency. Life expectancy at birth URL: https://www.cia.gov/library/publications/the-world-factbook/rankorder/2102rank.html [WebCite Cache ID 75o1NbAap]
16. International Telecommunication Union. ITU releases 2018 global and regional ICT estimates URL: https://www.itu.int/en/mediacentre/Pages/2018-PR40.aspx [WebCite Cache ID 75o1cdTv2]
17. Internet World Stats. Internet Usage in Asia URL: https://www.internetworldstats.com/stats3.htm [accessed 2019-01-30] [WebCite Cache ID 75o1vA1Nt]
18. Dutton WH, Rogers EM, Jun SH. Diffusion and social impacts of personal computers. Commun Res 2016 Jun 30;14(2):219-250. [doi: 10.1177/009365087014002005]
19. Hilbe J. Negative Binomial Regression. Cambridge: Cambridge University Press; 2016.
20. Shen C, Wang MP, Chu JT, Wan A, Viswanath K, Chan SS, et al. Health app possession among smartphone or tablet owners in Hong Kong: population-based survey. JMIR Mhealth Uhealth 2017 Jun 5;5(6):e77 [FREE Full text] [doi: 10.2196/mhealth.7628] [Medline: 28583905]
21. Brener ND, Billy JO, Grady WR. Assessment of factors affecting the validity of self-reported health-risk behavior among adolescents: evidence from the scientific literature. J Adolesc Heal 2003 Dec;33(6):436-457. [doi: 10.1016/S1054-139X(03)00052-1]
22. Zhu JJ, Chen H, Peng T, Liu XF, Dai H. How to measure sessions of mobile phone use? Quantification, evaluation, and applications. Mob Media Commun 2018 Feb 2;6(2):215-232. [doi: 10.1177/2050157917748351]
23. Canneyt S, Bron M, Haines A, Lalmas M. Describing Patterns and Disruptions in Large Scale Mobile App Usage Data. In: Proceedings of the 26th International Conference on World Wide Web Companion. 2017 Presented at: WWW'17 Companion; April 3-7, 2017; Perth, Australia. [doi: 10.1145/3041021.3051113]
24. Webster J, Lichty L, Phalen P. Ratings Analysis: Theory and Practice. United Kingdom: Routledge; 2009.
25. Webster J, Wakshlag J. A theory of television program choice. Comun Res 1983;10(4):430-446. [doi: 10.1177/009365083010004002]
26. Shelton B. The distribution of household tasks: does wife's employment status make a difference? J Fam Issues 1990;11(2):115-135. [doi: 10.1177/019251390011002001]
27. Gough M, Killewald A. University of Michigan Institute for Social Research, Population Studies Center. 2010. Gender , Job Loss , and Housework?: The Time Availability Hypothesis Revisited URL: https://www.psc.isr.umich.edu/pubs/abs/6410 [accessed 2019-05-08] [WebCite Cache ID 78DC01v2f]
28. Fisher W, Fisher J, Harman J. The Information-Motivation-Behavioral Skills model: a general social psychological approach to understanding and promoting health behavior. Soc Psychol Found Heal Illn 2003:82-106. [doi: 10.1002/9780470753552.ch4]

Abbreviations

ICC: intraclass correlation coefficient
LR: likelihood-ratio
mHealth: mobile health
OR: odds ratio

Edited by G Eysenbach; submitted 11.02.19; peer-reviewed by J Cho, C Shen, C Fernández, KL Mauco; comments to author 12.03.19; revised version received 15.04.19; accepted 29.04.19; published 23.05.19.

Please cite as:
Guan L, Peng TQ, Zhu JJH
Who is Tracking Health on Mobile Devices: Behavioral Logfile Analysis in Hong Kong
JMIR Mhealth Uhealth 2019;7(5):e13679
URL: http://mhealth.jmir.org/2019/5/e13679/
doi:10.2196/13679
PMID:31120429

Original Paper

Development and Implementation of a Mobile Phone–Based Prevention of Mother-To-Child Transmission of HIV Cascade Analysis Tool: Usability and Feasibility Testing in Kenya and Mozambique

Nami Kawakyu[1,2], RN, MPH; Ruth Nduati[3,4], MBChB, MPH; Khátia Munguambe[5,6], MSc, PhD; Joana Coutinho[7], BSN; Nancy Mburu[3], BS; Georgina DeCastro[7], MS; Celso Inguane[2], MPH, PhD; Andrew Zunt[8], BS; Neil Abburi[8], BS; Kenneth Sherr[2,9], MPH, PhD; Sarah Gimbel[1,2,9,10], RN, MPH, PhD

[1]Center for Global Health Nursing, University of Washington, Seattle, WA, United States
[2]Department of Global Health, University of Washington, Seattle, WA, United States
[3]Network of AIDS Researchers in East and Southern Africa, Nairobi, Kenya
[4]Department of Paediatrics and Child Health, University of Nairobi, Nairobi, Kenya
[5]University of Eduardo Mondlane, Maputo, Mozambique
[6]Manhiça Health Research Centre, Manhiça, Mozambique
[7]Health Alliance International, Beira / Chimoio, Mozambique
[8]Paul G Allen School of Computer Science and Engineering, University of Washington, Seattle, WA, United States
[9]Health Alliance International, Seattle, WA, United States
[10]Department of Family and Child Nursing, University of Washington, Seattle, WA, United States

Corresponding Author:
Nami Kawakyu, RN, MPH
Center for Global Health Nursing
University of Washington
Box 357263
Seattle, WA, 98195
United States
Phone: 1 206 221 7477
Fax: 1 206 685 9551
Email: nkawakyu@uw.edu

Abstract

Background: Prevention of mother-to-child HIV transmission (PMTCT) care cascade failures drive pediatric HIV infections in sub-Saharan Africa. As nurses' clinical and management role in PMTCT expand, decision-support tools for nurses are needed to facilitate identification of cascade inefficiencies and solutions. The mobile phone–based PMTCT cascade analysis tool (mPCAT) provides health facility staff a quick summary of the number of patients and percentage drop-off at each step of the PMCTC care cascade, as well as how many women-infant pairs would be retained if a step was optimized.

Objective: The objective of this study was to understand and improve the mPCAT's core usability factors and assess the health workers' experience with using the mPCAT.

Methods: Overall, 2 rounds of usability testing were conducted with health workers from 4 clinics and leading experts in maternal and child health in Kenya and Mozambique using videotaped think aloud assessment techniques. Semistructured group interviews gauged the understanding of mPCAT's core usability factors, based on the Nielsen Usability Framework, followed by development of cognitive demand tables describing the needed mPCAT updates. Post adaptation, feasibility was assessed in 3 high volume clinics over 12 weeks. Participants completed a 5-point Likert questionnaire designed to measure ease of use, convenience of integration into work, and future intention to use the mPCAT. Focus group discussions with nurse participants at each facility and in-depth interviews with nurse managers were also conducted to assess the acceptability, use, and recommendations for adaptations of the mPCAT.

Results: Usability testing with software engineers enabled real-time feedback to build a tool following empathic design principles. The revised mPCAT had improved navigation and simplified data entry interface, with only 1 data entry field per page. Improvements to the results page included a data visualization feature and the ability to share results through WhatsApp. Coding was simplified to enable future revisions by nontechnical staff—critical for context-specific adaptations for scale-up. Health care workers and facility managers found the tool *easy to use* (mean=4.3), used the tool *very often* (mean=4.1), and *definitely intended to continue to use* the tool (mean=4.8). Ease of use was the most common theme identified, with emphasis on how the tool readily informed system improvement decision making.

Conclusions: The mPCAT was well accepted by frontline health workers and facility managers. The collaborative process between software developer and user led to the development of a more user-friendly, context-specific tool that could be easily integrated into routine clinical practice and workflow. The mPCAT gave frontline health workers and facility managers an immediate, direct, and tangible way to use their clinical documentation and routinely reported data for decision making for their own clinical practice and facility-level improvements.

(JMIR Mhealth Uhealth 2019;7(5):e13963) doi:10.2196/13963

KEYWORDS

mHealth; quality improvement; engineering; HIV; mother to child transmission; implementation science

Introduction

Background

Gaps in the implementation of prevention of mother-to-child HIV transmission (PMTCT) services drive pediatric HIV infection in sub-Saharan Africa (SSA) [1-3]. Although PMTCT studies in SSA have been shown to reduce HIV transmission rates among breastfed infants to as low as 1% to 3% [4,5], the rates are several times higher in practice because of the high drop-out rates of HIV-positive women along the PMTCT care cascade [1-3,6-9]. Barriers to PMTCT service uptake and retention include lack of coordinated and continuous care for HIV-positive mother-baby pairs [10-12]; only 50% of infants born to HIV-positive women in Kenya and Mozambique are tested for HIV within the first 2 months of life [13,14].

To simplify the cascade and improve PMTCT services, the World Health Organization introduced Option B+ in 2013 by which HIV-positive women initiate lifelong antiretroviral therapy (ART) during pregnancy regardless of CD4 count [15]. Since then, studies have found Option B+ to be cost-effective [16-19]. A study in Malawi found that Option B+ not only prevented infant infections, but also increased the mother's 10-year survival rate by more than 4-fold compared with the standard of care [19]. There are also multiple health system challenges to implementing Option B+. The most common among them are health facility resource limitations such as shortage of staff and drugs [10,20,21]. Nurse staffing and workloads are strongly correlated with pregnant women completing steps along the complex cascade of Option B+ services [22]. Consequently, understaffing and undertraining of nurses lead to long wait times, inefficient patient flow, gaps in patient tracking [20,21], and thus, suboptimal adherence and retention of HIV-positive women [23-25]. Given that the quality of Option B+ services is dependent on specific health facility characteristics [22], interventions to improve services should be adaptable to the specific health facility's context.

The systems analysis and improvement approach (SAIA) is a 5-step process using systems engineering theory to guide facility-level staff and managers to maximize effectiveness of PMTCT services [26]. The first step is to improve the understanding of the inefficiencies by conducting a cascade analysis using the PMTCT cascade analysis tool (PCAT), which provides a systems-level view to track patient flow through the PMTCT cascade. This is followed by value stream mapping of the cascade to guide identification and prioritization of modifications addressing workflow inefficiencies. Steps 3 to 5 use continuous quality improvement (CQI) to iteratively test and modify facility-level improvement strategies. The SAIA intervention increased antiretroviral coverage by 3-fold and screening of HIV-exposed infants (HEI) by 17-fold in Kenya and Mozambique [27]. However, the PCAT's usability was inhibited by low computer availability and literacy at the health facilities [28]. As such, use of the PCAT was led by study nurses rather than facility personnel. Subsequently, the SAIA team developed and beta tested a mobile phone–based app of the PCAT (mPCAT) to facilitate its independent use by frontline facility staff.

Mobile phones have become widely available in SSA, with 73% of the population estimated to have a mobile cellular subscription [29]. Mobile phone systems for surveillance and data collection have been found to improve data quality [30-32] and be more time- and cost-efficient [30,32-35] than the traditional pen-and-paper method. The accessibility and literacy of mobile phones in SSA makes it an advantageous platform for scale-up and provides an opportunity for facility-level health care staff to use it for data-driven decision making.

Objectives

The objective of this paper was to report the process of adapting the beta-tested version of the mPCAT app through usability testing in Kenya and Mozambique and share findings from feasibility testing conducted to examine the acceptability and fit of the mPCAT as part of the broader SAIA intervention.

Methods

The usability and feasibility testing were conducted with nurses, health facility support staff, facility managers, and PMTCT experts between March and December 2017 in Kenya and Mozambique to understand and improve the mPCAT's core

usability factors and assess health workers' experience with using the mPCAT.

Setting

Sites for usability and feasibility testing were high-volume public facilities that provided a full range of PMTCT services and selected for inclusion based on experience with Option B+ service delivery and experience with SAIA, including the Microsoft Excel-based PCAT. A total of 2 sites in Kenya, one in the capital city of Nairobi and the other in the coastal city of Mombasa, participated in usability testing. In Mozambique, 2 sites, one in the capital city of Maputo and the other in Chimoio, the capital of Maniça Province in Central Mozambique, participated in usability testing. For feasibility testing, 2 sites in Nairobi and 1 site in Maputo were selected. The 2 sites in Kenya were selected as there was particularly high enthusiasm in Kenya in using the mPCAT tool, especially among health managers. Human resources for health are much more constrained in central Mozambique; thus, feasibility testing in more than 1 facility was not possible.

Kenya

Kenya adopted Option B+ in 2014. An estimated 76% of HIV-positive pregnant women in Kenya received ARTs for PMTCT in 2017 [14]. Transmission rates are estimated to be 5% at 6 weeks [36] and between 8% to 15% when including the breastfeeding period [9,36]. Only 51% of HEIs receive virological testing within the first 2 months of life [13].

Mozambique

Although Mozambique has made impressive achievements in increasing PMTCT coverage to 95%, it is challenged by high drop-out rates ranging from 32% loss to follow-up in Northern Mozambique [23] to only 5% to 30% of women returning for pharmacy refills at 90 days in Central Mozambique [20]. Owing to this high drop-out, mother-to-child HIV transmission rates double from 3% at 6 weeks to 6% at the end of the breastfeeding period [36]. Only half of HEIs receive virological testing within the first 2 months of life [13].

Intervention

The objective of the mPCAT is to facilitate facility-level staff and manager's rapid, independent, and quantitative tracking of patient flows through the PMTCT cascade. Staff enter routinely collected facility data and the tool calculates the number and proportion of women and children flowing through each step of the PMTCT cascade, broken down into the flow from antenatal care (ANC) through birth and the subsequent postpartum period. The tool calculates the number lost at each step and estimates the additional number of women and HEIs who would complete all steps of the PMTCT cascade, if each step was optimized. The results are reviewed by the health facility staff to prioritize improvements by cascade steps to optimize overall PMTCT services using CQI tools as part of the SAIA.

Usability Testing

Usability testing was conducted between March and June 2017 to examine how well the mPCAT functions and assess user experiences of the app's learnability, efficiency, memorability, error recovery, and satisfaction [37]. Usability testing is an important part of designing effective mobile health (mHealth) tools that are easy to use [37], increasing the likelihood for adoption of mHealth projects [32,33,38].

Participants

A total of 12 nurses were purposively selected from study sites to ensure representation across the PMTCT care cascade (ANC, maternity, and children at-risk care). Furthermore, 11 leading experts in maternal and child health in Kenya and Mozambique, including those from the Ministry of Health, University of Eduardo Mondlane, and University of Nairobi, were also purposively selected for usability testing. The sample size was based on previously published development and feasibility studies testing mobile phone technology to promote behavior change [39,40].

Data Collection

At the start of usability testing, participants were reoriented to the SAIA and introduced to the mPCAT by 2 study facilitators. Software developers were also present during usability testing to observe the participants' use of the mPCAT app. Participants were then given instructions to perform a series of sample tasks during videotaped think aloud exercises [37], where participants were asked to give their impression of the task, including what was difficult, what questions users had, and any feedback on improvements. Tasks included downloading the app, navigating through the app, and entering and interpreting data. At the end of each meeting, in-depth semistructured group interviews were conducted by 2 data collectors to explore what users found difficult with navigation, in entering data, interpreting data, and their preferences for the appearance of the app, such as its color scheme, graphics, and labels. Nielsen Usability Framework of learnability efficiency, memorability, error recovery, and satisfaction [37] guided the development of the think aloud exercises, as well as the interview guide.

Data Analysis

A codebook was developed through iterative coding of the videotaped think aloud exercises in Microsoft Excel guided by Nielsen Usability Framework. This codebook was then applied and revised through coding of in-depth group interview transcripts. Coding was completed independently and concurrently by 2 staff members in Kenya and 2 staff members in Mozambique. Coding was discussed with the principal investigator, who acted as a facilitator and tiebreaker. Codes were organized into usability themes and then incorporated into a cognitive demands table, a method to synthesize data such that the findings can be directly applied to inform product improvements [41]. The cognitive demands table headings used in the analysis were mPCAT task, description of difficult cognitive element, reason for difficulty, missing knowledge or skills affecting task, and potential cues or strategies that can be used to ease the task. The findings from the usability test informed updates to the mPCAT app, which was used for feasibility testing.

Feasibility Testing

Feasibility testing was conducted from September to December 2017 to assess whether the adapted mPCAT could be

successfully used by health care workers engaged in Option B+ service delivery as part of their system analysis and improvement efforts. The feasibility domains of interest included domains defined by Bowen et al as acceptability, implementation, and adaptation [42]. Feasibility studies are critical for the development of mHealth interventions to ensure smooth integration into health systems and successful scale-up [32,33].

Participants

Participants for feasibility testing were purposively selected from study sites based on prior experience working in Option B+ and representation across the PMTCT care cascade (ANC, maternity, children at-risk care). A total of 16 nurses, nurse managers, and other health facility support staff, such as counselors working in Option B+ services, participated in the questionnaires for feasibility testing. Furthermore, 10 nurses and health facility support staff participated in the focus group discussions (FGDs) and 3 nurse managers, 1 each from the study sites, were selected to participate in the in-depth interviews. The sample size was determined based on previously published feasibility studies testing mobile phone technology to promote behavior change [39,40].

Data Collection

Feasibility testing included a baseline visit by 2 study staff to reorient participants to the SAIA and introduce participants to the mPCAT and follow-up visits at 2, 4, 8, and 12 weeks post baseline to monitor and support implementation. The final visit included a 5-point Likert questionnaire designed to measure ease of use, convenience and integration into work, and future intention to use the mPCAT. FGDs with nurse participants at each facility and in-depth interviews with nurse managers were also conducted by 2 study staff to assess the acceptability, use, and recommendations for adaptions of the mPCAT.

Data Analysis

Descriptive statistics of the questionnaire data was calculated using R Studio (version 1.0.153). A codebook was developed through iterative coding of transcripts from FGDs and in-depth interviews using ATLAS.ti (version 8). Codes were guided by Bowen et al's feasibility framework [42]. One study staff and the principal investigator concurrently and independently coded 2 transcripts and then discussed and revised the codebook. This codebook was then applied to the remaining 4 transcripts and revised and reapplied to all transcripts until no revisions to the codebook was necessary. Once coding was finalized, codes were grouped into major themes guided by the feasibility framework and exemplary quotations were identified to represent the themes.

Ethical Approval for Usability and Feasibility Testing

The study procedures were reviewed and ethics approval obtained from the Institutional Review Board at the University of Washington (number 51026). The ethics review board of Kenyatta National Hospital and University of Nairobi in Kenya (P221/03/2016), as well as the ethics review board of the University of Eduardo Mondlane in Mozambique (CIBS FM&HCM/49/2016) approved the study. Furthermore, administrative approval was obtained from the Mozambican Ministry of Health. Written consent was obtained from all participants. No unique identifiers were collected, and the importance of maintaining confidentiality was emphasized during training of data collectors and at the start of group interviews and FGDs. Study data were stored in password-protected files that were only available to study staff for the purpose of data analysis.

Results

Usability Testing

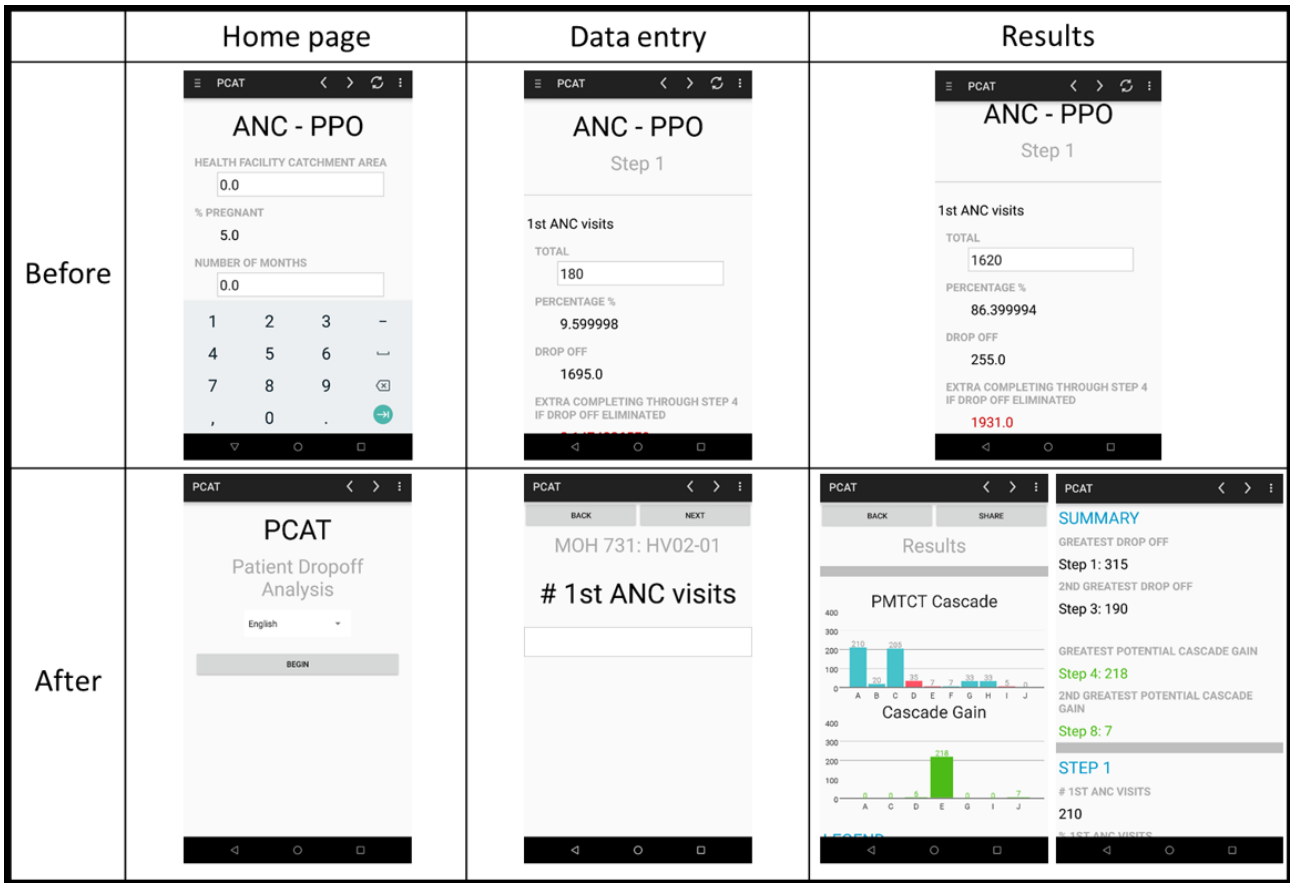
Data Entry Interface of the Mobile Phone–Based Prevention of Mother-To-Child HIV Transmission Cascade Analysis Tool: Strengths, Challenges, and Improvements

Participants described 2 main challenges with data entry in mPCAT. First, participants had difficulty understanding what data should be entered into the mPCAT. This was partly because of the differences in terminologies used in the national data collection forms and the initial mPCAT. For example, in Kenya, *previously identified HIV+* is reported in the national forms as *known positive on entry to ANC*.

To address this, terminologies used for data entry labels were revised to match national forms. In addition, toggling functionality was introduced in the Kenya version so that users could toggle over the labels and they could be shown which form and data variable point should be used to input data. Data points are not enumerated on national forms in Mozambique; therefore, this feature was not added for the Mozambique version.

Participants also had trouble with navigating the mPCAT app. The original version of the app opened directly to the data entry page with multiple textboxes without an explanation or a home screen to orient the unfamiliar user. Participants also found it challenging to scroll up and down a page to enter multiple data entry fields on 1 page and expressed a preference for scrolling left to right. Many participants had difficulty finding the *Next* button and took many tries to advance to the next screen. The *Back* option was similarly difficult for participants to find. As a result, most users thought that there was no option beyond starting over, if they accidentally advanced to the next screen or made a data entry error in a previous screen. Another design challenge identified in the initial app was that, because there was no *Save* feature, users needed to complete data entry in 1 sitting without interruption. In usability testing, participants were also observed skipping data entry fields, which led to errors in the results.

Figure 1. Screenshots of home page, data entry page, and results page before and after mobile phone–based prevention of mother-to-child HIV transmission cascade analysis tool usability testing. ANC: antenatal care; MOH: Ministry of Health; PCAT: prevention of mother-to-child HIV transmission cascade analysis tool; PMTCT: prevention of mother-to-child HIV transmission; PPO: prophylaxis.



To address these issues, a home page was added to orient the user to the app (Figure 1). Navigation was changed so that each page only had 1 data entry field and users had to answer the item before being able to advance to the next page (Figure 1). A clearly visible *Back* button at the top of the screen was added along with an adjacent *Next* button. The option to save and exit was also added so that users could return at a later time to enter data or retrieve reports. The large font size in the first version of the mPCAT was well received by users, many of who needed but did not wear glasses, and preserved.

Data Visualization and Summary Results Interface of the Mobile Phone–Based Prevention of Mother-To-Child HIV Transmission Cascade Analysis Tool: Strengths, Challenges, and Improvements

Participants had difficulty viewing and interpreting the numeric results displayed on the same page as the data entry page. On some pages with multiple data entry fields, users had to scroll down to view the results. As the results were displayed at each step of the PMTCT cascade, users would have to navigate between each page to compare the results, making it difficult to identify the greatest drop-off and optimization point. To address this, a summary results page with 3 component parts were added. One was a data visualization feature to graphically present the results and another, a summary of where the greatest drop-off in the cascade occurs and where the greatest potential is for gains. Third, the numeric results were color coded where red indicated losses and green indicated gains.

Participants also requested a feature on the app so that the results could be exported to print or shared with others. As such, a feature to export the results was added, including the ability to share the results through WhatsApp. Though the results could also be shared by email, participants expressed a preference for sharing results through WhatsApp, as it was a commonly used tool by nurses and other health care staff.

Feasibility Study

Acceptability of the Mobile Phone–Based Prevention of Mother-To-Child HIV Transmission Cascade Analysis Tool Among Users

Acceptability is the extent to which those delivering and receiving the strategy find it appropriate and satisfying [42]. Overall, users found the mPCAT app easy to use (mean=4.19, SD 0.98), reported to use it regularly during the study period (mean=4.00, SD 1.16), and intended to use the app in the future (mean=4.75, SD 0.58; Table 1). On average, responses were 1.1 points higher among participants in Mozambique than in Kenya.

Table 1. The mean (SD) of participant acceptance of mobile phone–based prevention of mother-to-child HIV transmission cascade analysis tool (N=16).

Participant acceptance	Kenya (n=10), mean (SD)	Mozambique (n=6), mean (SD)	Total (N=16), mean (SD)
Ease of use	3.70 (0.95)	5.00 (0)	4.19 (0.98)
Regular use	3.40 (1.08)	5.00 (0)	4.00 (1.16)
Intent to use	4.60 (0.70)	5.00 (0)	4.75 (0.58)

A common theme expressed by participants during FGDs and in-depth interviews (IDIs) was how easy the mPCAT app was to use, both for entering data and in interpreting the results. With regard to ease of data entry, one participant explained:

> *The mPCAT is a very easy app, it is not complicated and the questions are easy to understand. It was really easy using the mPCAT.* [FGD, Staff E]

With regard to ease of data interpretation, one participant expressed how the new data visualization feature added to the mPCAT after usability testing improved data interpretation:

> *The previous mPCAT that we were using, it was not as easy as the one we are using now. Because the other one isn't giving us a bar graph, but the one that we're using now, it is able to give you the conclusion in the form of a graph. You are able to see easily, this is where we need to improve, this is what we are doing good.* [FGD, Staff L]

Participants emphasized how the mPCAT made data interpretation easy and accessible, and how this facilitated data-driven decision making:

> *With the mPCAT tool, you were able to see your gaps immediately. And you are able to do your interventions immediately. So, it's a tool that I would recommend because it gives you instant, instant results.* [IDI, Manager A]

One example of a gap identified by users of the mPCAT was the low rate of CD4 testing among HIV-positive women. The mPCAT flagged that the facility was systematically missing CD4 testing of HIV-positive women, leading the facility to track women with missed CD4 tests and increase awareness among providers of this service delivery gap.

Implementation of Mobile Phone–Based Prevention of Mother-To-Child HIV Transmission Cascade Analysis Tool in Health Facilities

Implementation is the extent to which a strategy can be delivered in a defined but uncontrolled context [42]. All 3 study facilities reported using the mPCAT and there was consensus among staff that leadership, such as the facility-in-charge, was supportive of implementation. There was variability among FGD participants on the perceived effect of mPCAT on workload. One participant responsible for data analysis and reporting expressed that it decreased their workload by streamlining workflow. Participants in another facility described how using the mPCAT and making system changes increased their workload as once inefficiencies were identified, strategies needed to be adapted to rectify the challenges. Participants in a third facility found that the mPCAT neither increased nor decreased their workload, as the initial learning phase increased the workload but the improvements resulted in more efficient services, thereby reducing the workload.

Participants expressed that mPCAT, as part of the SAIA intervention, helped improve communication and team building across different services such as ANC and laboratory. As one participant explained:

> *There has been a great improvement between the teams. We have jointly identified the problems and interventions.* [FGD, Staff I]

Another participant elaborated:

> *It really improved our communication. Before, we didn't speak as a team. We only discussed these issues rarely, but after [starting SAIA] we discussed our results monthly.* [FGD, Staff H]

Integration of Mobile Phone–Based Prevention of Mother-To-Child HIV Transmission Cascade Analysis Tool Into the Existing Health System

Integration is the extent to which a strategy can be integrated within an existing system [42]. A key recommendation for future implementation and successful integration of the mPCAT as part of SAIA was to orient and train all those who are part of the PMTCT care cascade on mPCAT and SAIA and ensure that there is clear communication to participants on expectations related to facility-level system improvements. FGD participants from 1 facility explained that although all those who were aware of the intervention were supportive of making changes to workflow and patient flow, it was challenging to get involvement and acceptance of proposed changes among those who were not trained on the mPCAT and SAIA.

Discussion

Principal Findings

Improvements to the mPCAT's interface, including changes to its appearance, navigation, and input fields, were made during usability testing. Acceptability and adoption of the mPCAT was high, with users describing the app easy to use, both for entering and interpreting data. All 3 facilities in the feasibility study used the mPCAT to inform system improvement decision making in delivering Option B+ and reported that this process increased communication across service teams at the health facility.

Implications for Practice, Scale-Up, and Future Research

Usability testing revealed that a simplified interface would be easier for first-time users to enter the data without errors. This adaptation integrated the important user-centered design principle of universal design. That is, the product should be accessible to as many people as possible, with particular

attention to the user's level of technological literacy [43]. In this case, as most of the study participants had relatively low technological literacy, having a single data entry field per page, which had to be filled before being able to advance to the next page, assisted in focusing the user's attention and minimizing data entry errors, especially missing data entry fields.

Several important implications for successful mHealth app development and implementation were identified through the usability and feasibility testing. The presence of software developers during usability testing was very helpful in developing a more usable app, allowing them to see firsthand the priorities and challenges of end users and understand the reality of their work environments. Although the participants were able to provide feedback on what was confusing or difficult about the app, they were often unsure of how the app could be changed to be more usable. Having software developers help facilitate the usability test meant that they were able to offer several options for changes or add features that the participants could respond to, options that participants often would not have known to suggest themselves. Similarly, software developers were able to observe users and identify issues that users might not have flagged themselves. For example, developers observed that participants often skipped data entry inputs, which would lead to errors in the results. Hence, the adapted app was revised such that each item would need to be answered before a user could advance to the next item, improving error recovery. In addition, software developers were able to observe and immediately troubleshoot bugs and error messages that users encountered, rather than expecting nontechnical users to describe the issue from a distance. This collaborative process between users and software developers was critical in the development of a more user-centered, empathic design [44,45]. It also allowed for the development of an app that was locally contextualized, an important factor in the acceptance and successful adoption of mHealth projects [32,33].

Software updates to mPCAT also included simplifying coding so that minor changes could be updated by nontechnical persons. In the early version of the mPCAT app, coding for data entry items and results were scattered throughout the code, making it difficult for any nondeveloper to find the codes to change a data entry item. This was addressed by reformatting the code to pull all questions and formulas from a static JavaScript Object Notation (JSON) file, creating a Web-based graphical user interface that allows someone to provide a list of data entry fields and formulas that converts back to a JSON format that the app requires. Thus, future changes such as changing the data entry questions from English to another language, adding new items to the cascade, and altering formulas used to calculate results can be done by a nondeveloper.

Simplified coding and use of open-access software are 2 strategies that can ensure ease of adaptation of the mPCAT to local contexts such as changing the language of the app (eg, from English to Portuguese) or revising data entry fields to match national data collection forms. Currently, the original PCAT in Microsoft Excel has been or is in the process of being adapted to various other chronic care cascades in SSA, such as pediatric HIV (Kenya), hypertension (Mozambique), major mental health disorders (Mozambique), family planning uptake for people living with HIV (Kenya), preexposure prophylaxis uptake by at-risk youth (Kenya). It is anticipated that other app versions of the cascade analysis tool will be developed in the future. The mPCAT for PMTCT services is also being used in the SAIA-scale trial (R01MH113435) in 1 province in Mozambique, serving a population of approximately 2 million inhabitants.

Software applications, including mHealth apps, require continuous updating. National data collection forms change, health information systems change (eg, District Health Information System 2 [DHIS2]), phone technologies change, and therefore, apps need to constantly evolve to adapt to these changes. Changes to keep mHealth apps relevant require long-term funding, a challenge, given limited funding opportunities for long-term mHealth projects [32]. This is further challenged with common scenarios where mHealth apps are free and have limited advertising potential to sustain its own funding. In addition, the limited availability of a skilled electronic health and mHealth workforce in SSA [32,46-48] is also a critical gap that needs investment to ensure the continued success and sustainability of mHealth in strengthening health systems and improving health outcomes. Similarly, given the high turnover of health workers in SSA [49], ongoing staff training on the mPCAT, both independently and within the broader context of systems analysis and improvement approaches such as SAIA, is needed.

By harnessing the availability and power of mobile phones, frontline health workers were able to immediately, directly, and tangibly use their clinical documentation and routinely reported data to inform improvements to their own practice and workflow in their health facility. It is not uncommon in SSA, where there is a critical health workforce shortage [49], for health workers to dedicate significant time on data collection, entry, and reporting without active engagement in data use [50,51]. In addition, few mHealth technologies currently exist to support health workers' use of data to inform decision making about HIV services [52]. The mPCAT tool and findings from this study substantially contribute to minimizing this critical gap between frontline health worker and independent data use for decision making.

Limitations

There are limitations to this study. First, care should be taken when generalizing these findings to health settings beyond the study sites. The number of participants was small and although participation across PMTCT service delivery points (ANC, maternity, and children at-risk care) was secured, it was not representative of all health care providers who offer Option B+ services in Kenya and Mozambique. However, the purpose and benefit of feasibility studies is to assess how an intervention or strategy is implemented in uncontrolled, real-world settings, to inform future effectiveness trials, implementation, and scale-up [32,35,42]. Second, self-reported measurement of mHealth app use can be biased; a more objective measure would be to collect usage data on frequency and duration from the app itself, a feature to consider for future app updates. However, during FGDs and IDIs, participants described specific examples of how the mPCAT was used to inform system improvements that

triangulated the self-reported use data. Third, although this study followed up 3 months post-implementation with users, sustained use of the mPCAT post-study period was not evaluated. However, sustained use is being studied as part of a trial evaluating the effectiveness of SAIA (R01MH113435), of which mPCAT is an integral part, when administered and managed by the Ministry of Health. Through this study, the mPCAT is being used by health workers in all health facilities in Maniça Province with PMTCT services.

Conclusions

The collaborative process between software developer and user led to the development of an optimally usable and feasible mPCAT tool that was easy to use and easy to integrate into work routines. The mPCAT allows health workers and managers to quickly, clearly, and autonomously identify system inefficiencies in the delivery of Option B+ services and make data-driven decisions for system improvements. Findings from this study are critical in informing the development of a future controlled trial to determine the effectiveness of the mPCAT in reducing PMTCT dropout rates and contribute to the elimination of new pediatric HIV infections in SSA.

Acknowledgments

This work was supported through a grant from the United States National Institute of Health and National Institute of Allergy and Infectious Diseases (Usability and feasibility of a phone-based decision support tool for Option B+; R21AI124399 [Principal investigator: SG]) and through the *University of Washington/Fred Hutch Center for AIDS Research, a National Institute of Health* –funded program under award number AI027757. The authors would also like to thank the health workers and managers who participated in the study in both Kenya and Mozambique.

Conflicts of Interest

None declared.

References

1. Barker PM, Mphatswe W, Rollins N. Antiretroviral drugs in the cupboard are not enough: the impact of health systems' performance on mother-to-child transmission of HIV. J Acquir Immune Defic Syndr 2011 Feb 01;56(2):e45-e48. [doi: 10.1097/QAI.0b013e3181fdbf20] [Medline: 21084998]
2. Bhardwaj S, Carter B, Aarons GA, Chi BH. Implementation research for the prevention of mother-to-child HIV transmission in Sub-Saharan Africa: existing evidence, current gaps, and new opportunities. Curr HIV/AIDS Rep 2015 Jun;12(2):246-255 [FREE Full text] [doi: 10.1007/s11904-015-0260-1] [Medline: 25877252]
3. Doherty K, Ciaranello A. What is needed to eliminate new pediatric HIV infections: the contribution of model-based analyses. Curr Opin HIV AIDS 2013 Sep;8(5):457-466 [FREE Full text] [doi: 10.1097/COH.0b013e328362db0d] [Medline: 23743788]
4. Shapiro RL, Hughes MD, Ogwu A, Kitch D, Lockman S, Moffat C, et al. Antiretroviral regimens in pregnancy and breast-feeding in Botswana. N Engl J Med 2010 Jun 17;362(24):2282-2294 [FREE Full text] [doi: 10.1056/NEJMoa0907736] [Medline: 20554983]
5. Chasela CS, Hudgens MG, Jamieson DJ, Kayira D, Hosseinipour MC, Kourtis AP, BAN Study Group. Maternal or infant antiretroviral drugs to reduce HIV-1 transmission. N Engl J Med 2010 Jun 17;362(24):2271-2281 [FREE Full text] [doi: 10.1056/NEJMoa0911486] [Medline: 20554982]
6. Bonawitz R, Sambambi K, Vickery M, Hamomba L, Hammond E, Duncan J, et al. Identifying gaps in prevention of mother to child transmission of HIV: a case series of HIV-positive infants in Zambia. Pediatr Infect Dis J 2016 Dec;35(7):772-776. [doi: 10.1097/INF.0000000000001155] [Medline: 27093163]
7. Ezeanolue EE, Powell BJ, Patel D, Olutola A, Obiefune M, Dakum P, Nigeria Implementation Science Alliance. Identifying and prioritizing implementation barriers, gaps, and strategies through the Nigeria implementation Science alliance: getting to zero in the prevention of mother-to-child transmission of HIV. J Acquir Immune Defic Syndr 2016 Dec 01;72 Suppl 2:S161-S166 [FREE Full text] [doi: 10.1097/QAI.0000000000001066] [Medline: 27355504]
8. Braun M, Kabue MM, McCollum ED, Ahmed S, Kim M, Aertker L, et al. Inadequate coordination of maternal and infant HIV services detrimentally affects early infant diagnosis outcomes in Lilongwe, Malawi. J Acquir Immune Defic Syndr 2011 Apr 15;56(5):e122-e128 [FREE Full text] [doi: 10.1097/QAI.0b013e31820a7f2f] [Medline: 21224736]
9. Sirengo M, Muthoni L, Kellogg TA, Kim AA, Katana A, Mwanyumba S, KAIS Study Group. Mother-to-child transmission of HIV in Kenya: results from a nationally representative study. J Acquir Immune Defic Syndr 2014 May 01;66 Suppl 1:S66-S74 [FREE Full text] [doi: 10.1097/QAI.0000000000000115] [Medline: 24732822]
10. Gourlay A, Birdthistle I, Mburu G, Iorpenda K, Wringe A. Barriers and facilitating factors to the uptake of antiretroviral drugs for prevention of mother-to-child transmission of HIV in sub-Saharan Africa: a systematic review. J Int AIDS Soc 2013 Jul 19;16:18588 [FREE Full text] [doi: 10.7448/IAS.16.1.18588] [Medline: 23870277]

11. Mugasha C, Kigozi J, Kiragga A, Muganzi A, Sewankambo N, Coutinho A, et al. Intra-facility linkage of HIV-positive mothers and HIV-exposed babies into HIV chronic care: rural and urban experience in a resource limited setting. PLoS One 2014;9(12):e115171 [FREE Full text] [doi: 10.1371/journal.pone.0115171] [Medline: 25546453]
12. Adetokunboh OO, Oluwasanu M. Eliminating mother-to-child transmission of the human immunodeficiency virus in sub-Saharan Africa: the journey so far and what remains to be done. J Infect Public Health 2016;9(4):396-407 [FREE Full text] [doi: 10.1016/j.jiph.2015.06.010] [Medline: 26194038]
13. UNAIDS.: AIDSINFO; 2017. Country factsheets: Mozambique URL: http://www.unaids.org/en/regionscountries/countries/mozambique [WebCite Cache ID 788Qo0MyJ]
14. UNAIDS.: AIDSINFO; 2017. Country factsheets: Kenya URL: http://www.unaids.org/en/regionscountries/countries/kenya [WebCite Cache ID 788RXeHu9]
15. World Health Organization. World Health Organization. Geneva, Switzerland: World Health Organization; 2012. Programmatic update: use of antiretroviral drugs for treating pregnant women and preventing HIV infection in infants: executive summary URL: https://apps.who.int/iris/handle/10665/70892 [WebCite Cache ID 788RtvUjp]
16. Tweya H, Keiser O, Haas AD, Tenthani L, Phiri S, Egger M, et al. Comparative cost-effectiveness of Option B+ for prevention of mother-to-child transmission of HIV in Malawi. AIDS 2016 Mar 27;30(6):953-962 [FREE Full text] [doi: 10.1097/QAD.0000000000001009] [Medline: 26691682]
17. Gopalappa C, Stover J, Shaffer N, Mahy M. The costs and benefits of Option B+ for the prevention of mother-to-child transmission of HIV. AIDS 2014 Jan;28(Suppl 1):S5-14. [doi: 10.1097/QAD.0000000000000083] [Medline: 24468947]
18. Karnon J, Orji N. Option B+ for the prevention of mother-to-child transmission of HIV infection in developing countries: a review of published cost-effectiveness analyses. Health Policy Plan 2016 Oct;31(8):1133-1141. [doi: 10.1093/heapol/czw025] [Medline: 27016949]
19. Fasawe O, Avila C, Shaffer N, Schouten E, Chimbwandira F, Hoos D, et al. Cost-effectiveness analysis of Option B+ for HIV prevention and treatment of mothers and children in Malawi. PLoS One 2013;8(3):e57778 [FREE Full text] [doi: 10.1371/journal.pone.0057778] [Medline: 23554867]
20. Napúa M, Pfeiffer JT, Chale F, Hoek R, Manuel J, Michel C, et al. Option B+ in Mozambique: formative research findings for the design of a facility-level clustered randomized controlled trial to improve ART retention in antenatal care. J Acquir Immune Defic Syndr 2016 Dec 01;72(Suppl 2):S181-S188 [FREE Full text] [doi: 10.1097/QAI.0000000000001061] [Medline: 27355507]
21. Helova A, Akama E, Bukusi EA, Musoke P, Nalwa WZ, Odeny TA, et al. Health facility challenges to the provision of Option B+ in western Kenya: a qualitative study. Health Policy Plan 2017 Dec 01;32(2):283-291 [FREE Full text] [doi: 10.1093/heapol/czw122] [Medline: 28207061]
22. Rustagi AS, Gimbel S, Nduati R, Cuembelo MD, Wasserheit JN, Farquhar C, with input from the SAIA Study Team. Health facility factors and quality of services to prevent mother-to-child HIV transmission in Côte d'Ivoire, Kenya, and Mozambique. Int J STD AIDS 2017 Dec;28(8):788-799 [FREE Full text] [doi: 10.1177/0956462416668766] [Medline: 27590913]
23. Llenas-García J, Wikman-Jorgensen P, Hobbins M, Mussa MA, Ehmer J, Keiser O, SolidarMedIeDEA-Southern Africa. Retention in care of HIV-infected pregnant and lactating women starting ART under Option B+ in rural Mozambique. Trop Med Int Health 2016 Aug;21(8):1003-1012 [FREE Full text] [doi: 10.1111/tmi.12728] [Medline: 27208807]
24. Haas AD, Tenthani L, Msukwa MT, Tal K, Jahn A, Gadabu OJ, et al. Retention in care during the first 3 years of antiretroviral therapy for women in Malawi's option B+ programme: an observational cohort study. Lancet HIV 2016 Apr;3(4):e175-e182 [FREE Full text] [doi: 10.1016/S2352-3018(16)00008-4] [Medline: 27036993]
25. Woelk GB, Ndatimana D, Behan S, Mukaminega M, Nyirabahizi E, Hoffman HJ, et al. Retention of mothers and infants in the prevention of mother-to-child transmission of HIV programme is associated with individual and facility-level factors in Rwanda. J Int AIDS Soc 2016;19(5 Suppl 4):20837 [FREE Full text] [doi: 10.7448/IAS.19.5.20837] [Medline: 27443268]
26. Sherr K, Gimbel S, Rustagi A, Nduati R, Cuembelo F, Farquhar C, With input from the SAIA Study Team. Systems analysis and improvement to optimize pMTCT (SAIA): a cluster randomized trial. Implement Sci 2014 May 08;9:55 [FREE Full text] [doi: 10.1186/1748-5908-9-55] [Medline: 24885976]
27. Rustagi AS, Gimbel S, Nduati R, Cuembelo MD, Wasserheit JN, Farquhar C, With input from the SAIA Study Team. Implementation and operational research: impact of a systems engineering intervention on PMTCT service delivery in Côte d'Ivoire, Kenya, Mozambique: a cluster randomized trial. J Acquir Immune Defic Syndr 2016 Dec 01;72(3):e68-e76 [FREE Full text] [doi: 10.1097/QAI.0000000000001023] [Medline: 27082507]
28. Gimbel S, Rustagi AS, Robinson J, Kouyate S, Coutinho J, Nduati R, With input from the SAIA study team. Evaluation of a systems analysis and improvement approach to optimize prevention of mother-to-child transmission of HIV using the consolidated framework for implementation research. J Acquir Immune Defic Syndr 2016 Dec 01;72(Suppl 2):S108-S116 [FREE Full text] [doi: 10.1097/QAI.0000000000001055] [Medline: 27355497]
29. The World Bank. The World Bank.: The World Bank; 2016. Mobile cellular subscriptions URL: https://data.worldbank.org/indicator/IT.CEL.SETS.P2?locations=ZG [WebCite Cache ID 788SLJNQ1]
30. -. Health worker mHealth utilization: a systematic review. Comput Inform Nurs 2016 May;34(5):240. [doi: 10.1097/01.NCN.0000484013.73650.48] [Medline: 27153228]

31. Hall CS, Fottrell E, Wilkinson S, Byass P. Assessing the impact of mHealth interventions in low- and middle-income countries--what has been shown to work? Glob Health Action 2014;7:25606 [FREE Full text] [Medline: 25361730]
32. Aranda-Jan CB, Mohutsiwa-Dibe N, Loukanova S. Systematic review on what works, what does not work and why of implementation of mobile health (mHealth) projects in Africa. BMC Public Health 2014;14:188 [FREE Full text] [doi: 10.1186/1471-2458-14-188] [Medline: 24555733]
33. Sondaal SF, Browne JL, Amoakoh-Coleman M, Borgstein A, Miltenburg AS, Verwijs M, et al. Assessing the effect of mHealth interventions in improving maternal and neonatal care in low- and middle-income countries: a systematic review. PLoS One 2016;11(5):e0154664 [FREE Full text] [doi: 10.1371/journal.pone.0154664] [Medline: 27144393]
34. Brinkel J, Krämer A, Krumkamp R, May J, Fobil J. Mobile phone-based mHealth approaches for public health surveillance in sub-Saharan Africa: a systematic review. Int J Environ Res Public Health 2014 Nov;11(11):11559-11582 [FREE Full text] [doi: 10.3390/ijerph111111559] [Medline: 25396767]
35. Amoakoh-Coleman M, Borgstein AB, Sondaal SF, Grobbee DE, Miltenburg AS, Verwijs M, et al. Effectiveness of mHealth interventions targeting health care workers to improve pregnancy outcomes in low- and middle-income countries: a systematic review. J Med Internet Res 2016 Aug 19;18(8):e226 [FREE Full text] [doi: 10.2196/jmir.5533] [Medline: 27543152]
36. Joint United Nations Programme on HIV/AIDS. UNAIDS. Geneva, Switzerland: UNAIDS; 2016. On the Fast-Track to an AIDS-Free Generation: The Incredible Journey of the Global PlanTowards the Elimination of New HIV Infections Among Children by 2015 and Keeping their Mothers Alive URL: http://www.unaids.org/sites/default/files/media_asset/GlobalPlan2016_en.pdf [WebCite Cache ID 788SZzoEU]
37. Nielsen J. Usability Engineering. Mountain View, California: Academic Press Inc; 1993.
38. Gagnon M, Ngangue P, Payne-Gagnon J, Desmartis M. m-Health adoption by healthcare professionals: a systematic review. J Am Med Inform Assoc 2016 Jan;23(1):212-220. [doi: 10.1093/jamia/ocv052] [Medline: 26078410]
39. Collins R, Kashdan T, Gollnisch G. The feasibility of using cellular phones to collect ecological momentary assessment data: application to alcohol consumption. Exp Clin Psychopharmacol 2003 Feb;11(1):73-78. [Medline: 12622345]
40. Robinson E, Higgs S, Daley A, Jolly K, Lycett D, Lewis A, et al. Development and feasibility testing of a smart phone based attentive eating intervention. BMC Public Health 2013 Jul 09;13:639 [FREE Full text] [doi: 10.1186/1471-2458-13-639] [Medline: 23837771]
41. Militello LG, Hutton RJ. Applied cognitive task analysis (ACTA): a practitioner's toolkit for understanding cognitive task demands. Ergonomics 1998 Nov;41(11):1618-1641. [doi: 10.1080/001401398186108] [Medline: 9819578]
42. Bowen DJ, Kreuter M, Spring B, Cofta-Woerpel L, Linnan L, Weiner D, et al. How we design feasibility studies. Am J Prev Med 2009 May;36(5):452-457 [FREE Full text] [doi: 10.1016/j.amepre.2009.02.002] [Medline: 19362699]
43. Punchoojit L, Hongwarittorrn N. Usability studies on mobile user interface design patterns: a systematic literature review. Advances in Human-Computer Interaction 2017;2017:1-22. [doi: 10.1155/2017/6787504]
44. Leonard D, Rayport JF. Spark innovation through empathic design. Harv Bus Rev 1997;75(6):102-113. [Medline: 10174792]
45. Thomas J, McDonagh D. Empathic design: research strategies. Australas Med J 2013;6(1):1-6 [FREE Full text] [doi: 10.4066/AMJ.2013.1575] [Medline: 23423953]
46. World Health Organization, Regional Office for Africa. World Health Organization. Malabo, Equatorial Guinea: World Health Organization, Regional Office for Africa; 2010. Ehealth Solutions in the African Region Current Context and Perspectives (Document AFR/RC60/5) URL: https://apps.who.int/iris/bitstream/handle/10665/19931/AFR-RC60-R3.pdf?sequence=1&isAllowed=y
47. Gerber T, Olazabal V, Brown K, Pablos-Mendez A. An agenda for action on global e-health. Health Aff (Millwood) 2010 Feb;29(2):233-236 [FREE Full text] [doi: 10.1377/hlthaff.2009.0934] [Medline: 20348066]
48. Mechael P, Batavia H, Kaonga N, Searle S, Kwan A, Goldberg A, et al. mHealth Alliance and Columbia University.: mHealth Alliance and Columbia University; 2010. Barriers and Gaps Affecting mHealth in Low and Middle Income Countries: Policy White Paper URL: http://www.globalproblems-globalsolutions-files.org/pdfs/mHealth_Barriers_White_Paper.pdf [WebCite Cache ID 788XyoomN]
49. Mbemba GI, Gagnon M, Hamelin-Brabant L. Factors influencing recruitment and retention of healthcare workers in rural and remote areas in developed and developing countries: an overview. J Public Health Afr 2016 Dec 31;7(2):565 [FREE Full text] [doi: 10.4081/jphia.2016.565] [Medline: 28299160]
50. Okaisu EM, Kalikwani F, Wanyana G, Coetzee M. Improving the quality of nursing documentation: an action research project. Curationis 2014 Dec 03;37(2):E1-11. [doi: 10.4102/curationis.v37i2.1251] [Medline: 26864179]
51. Gimbel S, Chilundo B, Kenworthy N, Inguane C, Citrin D, Chapman R, et al. Donor data vacuuming: audit culture and the use of data in global heath partnerships. Med Anthropol Theory 2018 May 15;5(2):99. [doi: 10.17157/mat.5.2.537]
52. Gimbel S, Kawakyu N, Dau H, Unger J. A missing link: HIV-/AIDS-related mHealth interventions for health workers in low- and middle-income countries. Curr HIV/AIDS Rep 2018 Dec;15(6):414-422. [doi: 10.1007/s11904-018-0416-x] [Medline: 30259258]

Abbreviations

ANC: antenatal care

ART: antiretroviral therapy
CQI: continuous quality improvement
FGD: focus group discussion
HEI: HIV-Exposed Infants
IDIs: In-depth interviews
JSON: JavaScript Object Notation
mHealth: mobile health
mPCAT: mobile phone–based prevention of mother-to-child HIV transmission cascade analysis tool
PCAT: prevention of mother-to-child HIV transmission cascade analysis tool
PMTCT: prevention of mother-to-child HIV transmission
SAIA: system analysis and improvement approach
SSA: sub-Saharan Africa

Edited by G Eysenbach; submitted 08.03.19; peer-reviewed by O Adetokunboh, J Gass; comments to author 03.04.19; revised version received 20.04.19; accepted 23.04.19; published 13.05.19.

Please cite as:
Kawakyu N, Nduati R, Munguambe K, Coutinho J, Mburu N, DeCastro G, Inguane C, Zunt A, Abburi N, Sherr K, Gimbel S
Development and Implementation of a Mobile Phone–Based Prevention of Mother-To-Child Transmission of HIV Cascade Analysis Tool: Usability and Feasibility Testing in Kenya and Mozambique
JMIR Mhealth Uhealth 2019;7(5):e13963
URL: http://mhealth.jmir.org/2019/5/e13963/
doi:10.2196/13963
PMID:31094351

Original Paper

Automatic Identification of Physical Activity Type and Duration by Wearable Activity Trackers: A Validation Study

Diana Dorn[1], MS; Jessica Gorzelitz[1], MS; Ronald Gangnon[2], PhD; David Bell[1], PhD; Kelli Koltyn[1], PhD; Lisa Cadmus-Bertram[1], PhD

[1]Department of Kinesiology, University of Wisconsin-Madison, Madison, WI, United States

[2]Department of Population Health Sciences, University of Wisconsin-Madison, Madison, WI, United States

Corresponding Author:
Lisa Cadmus-Bertram, PhD
Department of Kinesiology
University of Wisconsin - Madison
2000 Observatory Drive
Madison, WI, 53562
United States
Phone: 1 608 265 5946
Email: cadmusbertra@wisc.edu

Abstract

Background: Activity trackers are now ubiquitous in certain populations, with potential applications for health promotion and monitoring and chronic disease management. Understanding the accuracy of this technology is critical to the appropriate and productive use of wearables in health research. Although other peer-reviewed validations have examined other features (eg, steps and heart rate), no published studies to date have addressed the accuracy of automatic activity type detection and duration accuracy in wearable trackers.

Objective: The aim of this study was to examine the ability of 4 commercially available wearable activity trackers (Fitbits Flex 2, Fitbit Alta HR, Fitbit Charge 2, and Garmin Vívosmart HR), in a controlled setting, to correctly and automatically identify the type and duration of the physical activity being performed.

Methods: A total of 8 activity types, including walking and running (on both a treadmill and outdoors), a run embedded in walking bouts, elliptical use, outdoor biking, and pool lap swimming, were tested by 28 to 34 healthy adult participants (69 total participants who participated in some to all activity types). Actual activity type and duration were recorded by study personnel and compared with tracker data using descriptive statistics and mean absolute percent error (MAPE).

Results: The proportion of trials in which the activity type was correctly identified was 93% to 97% (depending on the tracker) for treadmill walking, 93% to 100% for treadmill running, 36% to 62% for treadmill running when preceded and followed by a walk, 97% to 100% for outdoor walking, 100% for outdoor running, 3% to 97% for using an elliptical, 44% to 97% for biking, and 87.5% for swimming. When activities were correctly identified, the MAPE of the detected duration versus the actual activity duration was between 7% and 7.9% for treadmill walking, 8.7% and 144.8% for treadmill running, 23.6% and 28.9% for treadmill running when preceded and followed by a walk, 4.9% and 11.8% for outdoor walking, 5.6% and 9.6% for outdoor running, 9.7% and 13% for using an elliptical, 9.5% and 17.7% for biking, and was 26.9% for swimming.

Conclusions: In a controlled setting, wearable activity trackers provide accurate recognition of the type of some common physical activities, especially outdoor walking and running and walking on a treadmill. The accuracy of measurement of activity duration varied considerably by activity type and tracker model and was poor for complex sets of activity, such as a run embedded within 2 walking segments.

(JMIR Mhealth Uhealth 2019;7(5):e13547) doi:10.2196/13547

KEYWORDS

fitness trackers; exercise; accelerometry; data accuracy

Introduction

Background

Adequate physical activity participation is one of the most important behaviors people can adopt to maintain their health and well-being. Physical activity reduces the risk of several major chronic diseases [1] and early mortality [2], reduces health risks associated with overweight and obesity [3], and improves psychological outcomes, including mood and energy [4]. For adults, the 2018 American federal aerobic physical activity guidelines prescribe ≥150 min per week of moderate-intensity activity or ≥75 min of vigorous-intensity physical activity or an equivalent combination of the two [5]. Results from US national surveillance estimates show suboptimal rates of activity nationwide and highlight the importance of promoting physical activity participation to increase the overall health of the population [6].

Consumer-grade activity trackers are one tool that may help individuals increase and monitor their physical activity participation. These devices are available to consumers at a relatively low cost, with approximately 14 million Fitbits alone sold in 2018 [7] and 120 million devices projected to be sold by 2019 [8]. Such trackers have been shown to support increased physical activity participation in adults [9,10] and are suitable for incorporation into clinical research and health promotion interventions [9,11,12]. Naturally, the extent of the utility of these devices depends on the accuracy of their features. Peer-reviewed independent research is critical to understanding and validating these technologies, particularly as manufacturers typically neither provide comprehensive results of internal studies nor release their proprietary classification algorithms. Previous validation studies have reported high correlations between device step counts and the criterion [13,14] and a general underestimation of energy expenditure as compared with criterion measurements [15], and heart rate validation studies have shown that wearable devices are more accurate during rest than during moderate exercise [16-18].

With the evolution of physical activity algorithmic-based classification, manufacturers are developing beyond-the-manual logging of activities (eg, a user tagging an activity as *running* or *cycling*) toward use of pattern recognition algorithms to automatically detect certain activity types. Automatic activity recognition—called SmartTrack on Fitbit trackers [19] and Move IQ on Garmin trackers [20]—allows a wearable activity tracker to recognize and classify specific activity types, without input from the user. If accurate, this feature could enhance health-based research by lending insight into what types of activities participants are performing and what could aid researchers in tracking adherence to physical activity guidelines and fidelity to interventions.

Objectives

Given the importance of understanding the accuracy of trackers' new automatic activity detection capabilities, and the lack of current validation studies available on this feature, the objective of this study was to explore the accuracy of this feature on 4 commercially available wearable activity trackers in a controlled setting: the Fitbit Flex 2, the Fitbit Charge 2, the Fitbit Alta HR (Fitbit), and the Garmin Vívosmart HR (Garmin). Specifically, this study aimed to determine (1) the accuracy of the activity-type identification and (2) the accuracy of the measured duration of activity in a controlled setting.

Methods

Participants

This study protocol was approved by the University of Wisconsin-Madison Health Sciences Institutional Review Board. Written informed consent was obtained for all participants. Participants were adults aged from 18 to 50 years, who were able to perform moderate- and vigorous-intensity physical activity and were free from specific health limitations, including heart attack, angioplasty or heart surgery in the previous 3 months, active chest pain, shortness of breath, fainting, angina pectoris, current pregnancy, or any physical disability that would preclude use of a treadmill, an elliptical, or a bicycle. Participants were excluded from swimming if they reported an inability to swim front crawl (ie, freestyle) for 15 min. Participants were also excluded if their measured resting systolic blood pressure (via Riester ri-champion N automatic cuff, Riester) was > 180 mmHg or resting diastolic blood pressure was > 110 mmHg.

Study Design

The study comprised 4 activity modules (Table 1) designed to reflect some of the most common types of moderate-to-vigorous intensity physical activity with the intention of evaluating 34 trials per module on the basis of previous validation studies with small sample sizes [15]. Power analyses were not conducted because of a lack of previously published (or *a priori)* identified effect sizes. Participants self-selected which module or modules they wished to complete. Participants could complete more than 1 module if desired, but each module was scheduled on a different day to avoid excessively long activity participation in a single day by each participant. Participation in multiple modules was not expected to influence results.

Table 1. Description and duration of activities in each of the experimental modules.

Module and activities	Activity duration
Module A[a]	
Treadmill walk (2-4.5 mph)	15 min
Treadmill run (>4.5 mph)	15 min
Embedded run (walk, run, walk)[b]	25 min (5-min walk, 15-min run, 5-min walk)
Module B[a]	
Outdoor walk	15 min
Outdoor run	15 min
Elliptical	15 min
Module C	
Bike	15 min
Module D	
Swim	15 min

[a]Activities in this module separated by 10 min of rest.

[b]These activities were completed at the same miles per hour (mph) designation as the previous activities. The purpose of the embedded run was to test the trackers' ability to detect a run with walking before and after.

Activity Modules

Activity Module A was completed on a treadmill (Trackmaster and Precor Inc) and comprised 15 min of walking (at speeds of 2 to 4.5 mph), 10 min of stationary rest, 15 min of running (>4.5 mph), 10 min of stationary rest, and a 15-min run embedded in 5-min segments of walking to simulate a warm up and cool down before and after a run (walk-run-walk). The purpose of the treadmill run embedded in walking bouts was to test the trackers' ability to detect the running bout when preceded and followed by bouts of walking. This activity is referred to as the *embedded run*. *Activity Module B* comprised walking outdoors on a continuous path for 15 min, 10 min of stationary rest, running on a path for 15 min, 10 min of stationary rest, and using an elliptical trainer with moving arm handles (Precor Inc) for 15 min. Participants were instructed to use the moving arm handles on the elliptical machine for the activity's duration. *Activity Module C* comprised 15 min of outdoor cycling on a continuous path (referred to as *bike* or *biking*). *Activity Module D* comprised 15 min of swimming in a 25-yard, 8-lane indoor pool. Participants began wearing devices following initialization in the laboratory; then they walked to the treadmill, pool, outdoor walking path, or other activity location (<5 min). Following each activity within the module, participants began the bout of sedentary rest. No sedentary rest was observed after the last activity in each module. Directly following each activity, participants reported a rating of perceived exertion (RPE) using the Borg RPE scale, which has high validity and reliability for reporting relative exercise intensity [21,22], to assess the intensity at which the trackers might successfully detect an activity and duration. Participants removed the trackers immediately after the last activity in each module.

Devices

The 4 devices used in this study were the Fitbit Flex 2, the Fitbit Charge 2, the Fitbit Alta HR (Fitbit), and the Garmin Vívosmart HR (Garmin). These devices were chosen because at the time of the study, they were the only ones that offered the automatic activity detection feature. Participants wore all 4 devices concurrently, 2 on each wrist, and placement was randomized—although previous validation work [16] demonstrated that tracker accuracy does not differ by wrist location. For Module D, only the Fitbit Flex 2 was worn as the others do not detect swimming. Following informed consent, participants completed a brief health history form and a demographics questionnaire, and height and weight were measured once using a stadiometer (Health o meter, Welch Allyn). All trackers were initialized using their respective Fitbit or Garmin apps, using each participant's actual height, weight, sex, and age. The same 4 devices were reinitialized and used for each participant in the study to minimize the potential risk of interdevice heterogeneity. The Fitbit devices were programmed to detect activities lasting longer than 10 min (the shortest amount of time required for activity-type recognition by the devices).

Outcome Measures

Each activity performed was observed and recorded in conjunction with a timer on a laptop, which was used to record the duration of the activity; this was considered the criterion measure. The timer was stopped when the participant came to a complete stop at the end of the activity, and the duration was recorded. Upon completion of each module, the trackers were synced to their respective apps on a laptop, and the trackers' designation of activity type and duration were recorded from the apps. A correct identification occurred when the tracker's activity designation matched the actual activity performed. An erroneous identification occurred when the tracker's activity designation did not match the actual activity performed, and a missed identification occurred when the tracker did not designate any activity when there was an activity performed.

Statistical Analysis

All statistical analyses were completed using Statistical Analysis Software version 9.4 (SAS Institute Inc.) and Microsoft Excel version 1811 (Microsoft). Data from all users were inspected for errors before analysis. One participant's elliptical trainer data were excluded because of failure to follow protocol (the participant did not use the moving arm handles). A total of 2 participants who were unable to complete the swimming activity were excluded from analyses. Descriptive statistics were used to characterize the participant sample and the mean Borg RPE of each activity. Frequency counts were calculated to analyze the number of correct identifications and the amount of missed and erroneous identifications. Summary statistics and box plots were produced to analyze the duration of activities that were correctly identified by the trackers. Mean absolute percent error (MAPE) (1/n* | *true duration – device-based duration measure* |) was calculated for each activity and device to establish the differences between the criterion duration (duration measured with a laptop timer) and device measured duration.

Results

Participants in this study (N=69) were 61% female, had a mean age of 26.4 (SD 8.7) years, and had an average body mass index of 23.9 (SD 4.1) kg/m^2. Of the 69 participants, 36 (52%) participants chose to complete 1 module, 10 (15%) participants chose to complete 2 modules, 16 (23%) participants chose to complete 3 modules, and 7 (10%) participants chose to complete all 4 modules. Although the study was designed for 34 trials per module, the actual number of trials used in the analysis varied slightly because of infrequent but occasional inability to sync devices and retrieve data between study visits. The mean Borg RPE rating was 8.6 (SD 1.8) for the treadmill walk (corresponds to an *extremely light* to *very light* intensity); 13.6 (SD 1.9) for the treadmill run (*somewhat hard* to *hard* intensity); 14.2 (SD 1.6) for the embedded run (*somewhat hard* to *hard* intensity), 8.3 (SD 1.6) for the outdoor walk (*extremely light* to *very light* intensity), 13.7 (SD 1.4) for the outdoor run (*somewhat hard* to *hard* intensity), 13.0 (SD 1.7) for using an elliptical machine (*somewhat hard* intensity); 11.3 (SD 1.4) for outdoor biking (*light* to *somewhat hard* intensity); and 14.2 (SD 2.1) for swimming in a lap pool (*somewhat hard* to *hard* intensity).

Activity Identification

Devices were highly successful in recognizing the simple bouts of walking and running on a treadmill or outdoors, struggled to recognize the embedded run, and had more variable success in recognizing the other activities (elliptical trainer, biking outdoors, and swimming; Figure 1). Overall, the activity with the highest proportion of correct identifications was outdoor running, which was detected for 100% of trials on all devices. Outdoor walking had a rate of correct identifications between 97% and 100%, dependent on tracker brand and model. Although slightly lower, both treadmill walking and running had rates of identification above 92%. Treadmill walking was correctly detected between 93% and 97% of trials dependent on tracker brand and model, and treadmill running was correctly detected between 93% and 100% of trials. The embedded run was detected less often between 36% and 62% of trials. As we were interested in the trackers' ability to detect a run when it was preceded or followed by a short bout of walking, the correct designation for this activity was *run*.

Elliptical use was correctly detected in 91% to 93% of trials for Fitbit devices, but only 3% of trials by the Garmin device. Similarly, biking had a high rate of correct identifications by the Fitbit devices, between 94% and 97% of trials, and a lower recognition rate of 44% of trials by the Garmin device. Swimming was correctly detected for 88% of trials on the Fitbit Flex 2 device. Although we requested the use of front crawl stroke for the duration of the swim module, 5 participants switched to other strokes during the swim. When these participants were excluded from the analysis, the Fitbit Flex 2 had a correct recognition rate of 85% for 27 trials.

The types of observed misclassifications are shown in a confusion matrix (Table 2). A total of 9 misclassifications occurred for the Garmin device during the elliptical activity, the most misclassifications of any activity. A smaller number of misclassifications were observed for treadmill walking and running, outdoor walking, and swimming. The 2 misclassifications that occurred for the outdoor walking activity were misclassified for the same participant.

Figure 1. Percent correct identifications of eight activities by four wearable trackers.

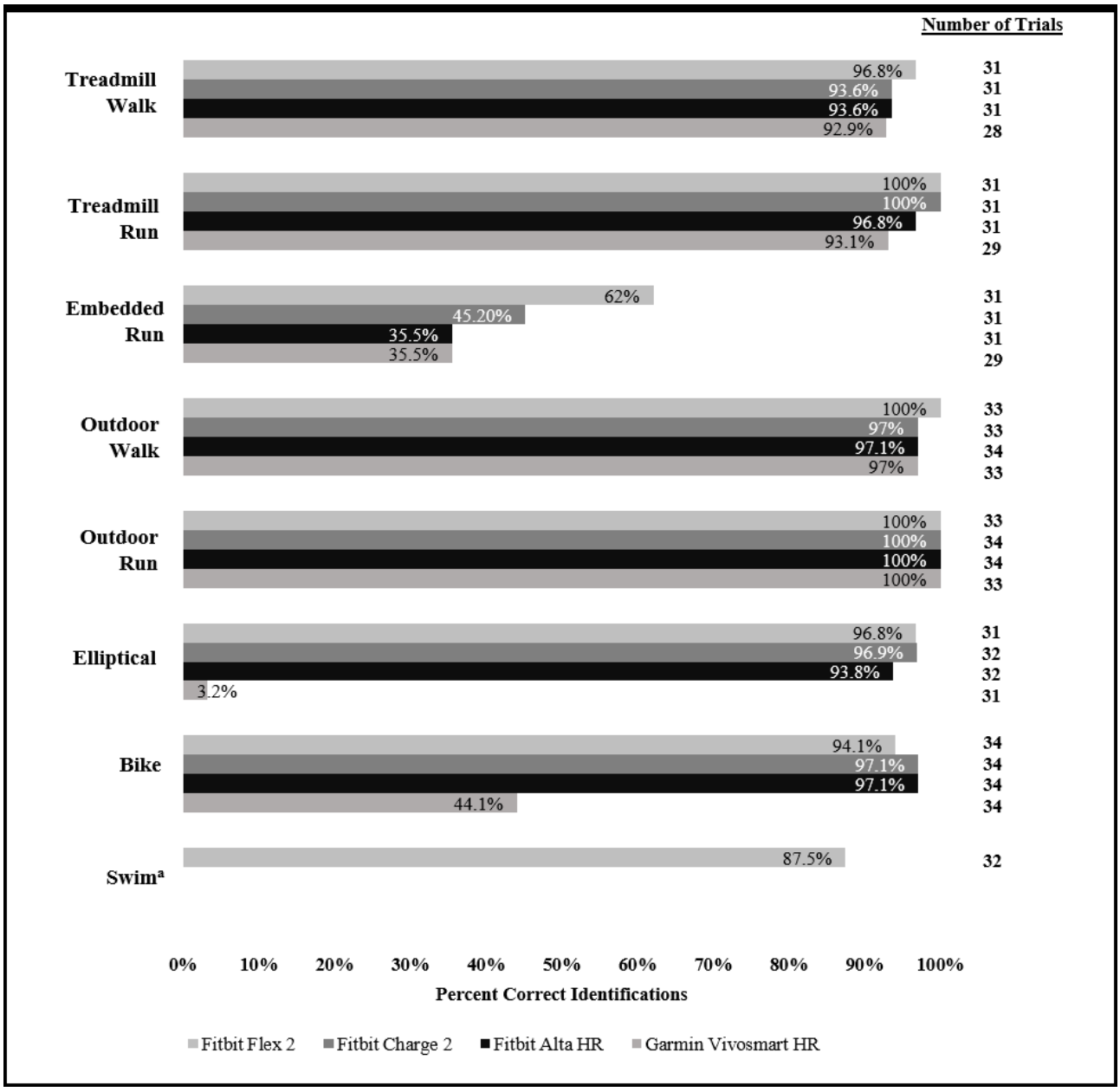


Table 2. Confusion matrix of activity identifications by actual activity type and device.

Activity and device	Identifications		Misclassified as a
	Correct	Erroneous	
Treadmill walk			
Fitbit Flex 2	30	1	Run (1)
Fitbit Charge 2	31	0	—[a]
Fitbit Alta HR	31	0	—
Garmin Vívosmart HR	28	0	—
Treadmill run			
Fitbit Flex 2	31	0	—
Fitbit Charge 2	31	0	—
Fitbit Alta HR	30	1	Aerobic workout (1)
Garmin Vívosmart HR	29	0	—
Embedded run			
Fitbit Flex 2	31	0	—
Fitbit Charge 2	31	0	—
Fitbit Alta HR	31	0	—
Garmin Vívosmart HR	29	0	—
Outdoor walk			
Fitbit Flex 2	33	0	—
Fitbit Charge 2	32	1	Elliptical (1)
Fitbit Alta HR	33	1	Elliptical (1)
Garmin Vívosmart HR	33	0	—
Outdoor run			
Fitbit Flex 2	33	0	—
Fitbit Charge 2	34	0	—
Fitbit Alta HR	34	0	—
Garmin Vívosmart HR	33	0	—
Elliptical			
Fitbit Flex 2	31	0	—
Fitbit Charge 2	32	0	—
Fitbit Alta HR	32	0	—
Garmin Vívosmart HR	22	9	Fitness (9); Other (6); Run (2)
Swim			
Fitbit Flex 2	28	4	Aerobic workout (2); Walk (2)

[a]No activity misclassifications for this device/activity combination.

Duration Accuracy

Results of the duration accuracy analysis show that detected mean duration was less than 4 min over or under the actual activity duration, except from the treadmill run for all Fitbit trackers, which was overestimated by a mean of 20 min (Figure 2). As the activity bouts were relatively short, even a few minutes of deviation in measurement of activity duration resulted in a substantial MAPE (Table 3). MAPEs were lower for all devices for treadmill walking (7% to 7.9%), outdoor walking (4.9% to 11.8%), and outdoor running (5.6% to 9.6%) than other activities, including treadmill running (8.7% to 144.8%), the embedded treadmill run preceded and followed by a walk (23.6% to 28.9%), an elliptical trainer (9.7% to 13%), biking (9.5% to 17.7%), or swimming (26.9%). Box plots for recorded duration show a spread in time for activities that were completed outdoors or in the pool, as participants needed to return to the activity start site (outdoors) or to the pool edge at the end of the activity. Thus, actual time fluctuated slightly

around 15 min. All devices estimated treadmill walking within 1 min of the actual duration.

The Garmin device underestimated treadmill running duration by approximately 1 min. The embedded run duration was overestimated by a mean of 3 min by all devices. Outdoor walking was overestimated by a mean of 1 min by all devices. Outdoor running was overestimated by a mean of 1 min by Fitbit devices and was underestimated by about 1 min by the Garmin tracker. Using an elliptical and biking were overestimated by a mean of 1 min by all devices. Swimming was overestimated by approximately 2 min by the Fitbit Flex 2. The Garmin device was the only wearable tracker to underestimate activity duration, and when activity was overestimated by other devices, it was overestimated by less than 4 min on average by all devices, except for the treadmill run, which was overestimated by a mean of 20 min.

Figure 2. Box plots of the actual duration of each activity, compared to the duration estimated by the automatic activity recognition feature on each of the four trackers.

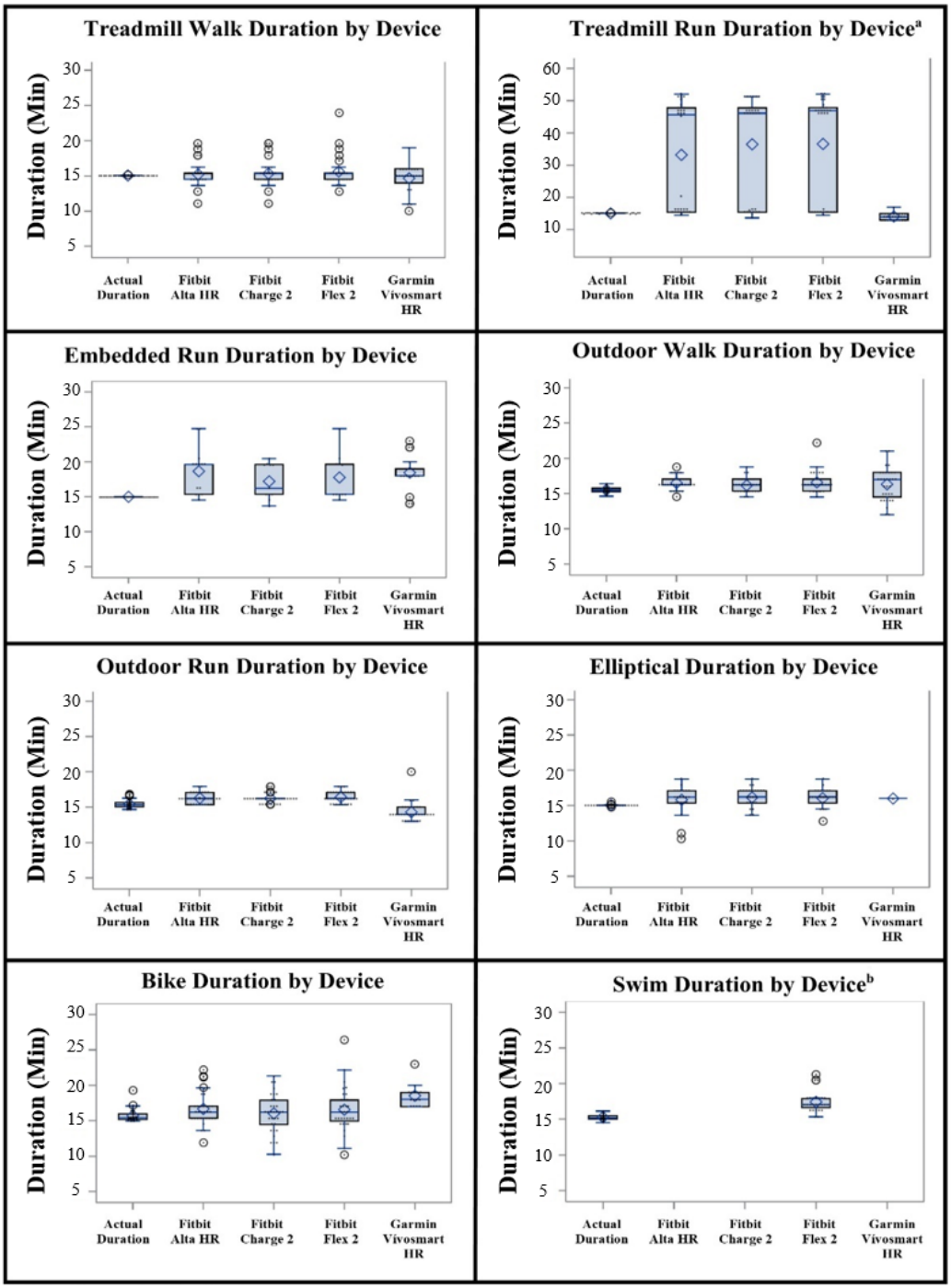


Table 3. Mean absolute percent error of duration measures by device and activity.

Activity type	Device			
	Fitbit			Garmin Vívosmart HR
	Flex 2	Charge 2	Alta HR	
Treadmill walk	7.9	7	7.6	7.7
Treadmill run	139.4	115.4	144.8	8.7
Embedded run	23.6	24.8	28.9	23.9
Outdoor walk	8.1	6.9	4.9	11.8
Outdoor run	6.6	5.6	5.6	9.6
Elliptical	9.7	10.2	10.2	13
Bike	13.3	9.5	12.9	17.7
Swim	26.9	—[a]	—	—

[a]No data acquired as Only the Fitbit Flex 2 is waterproof and can measure swimming.

Discussion

Principal Findings

The most important findings of this study are that the trackers correctly identified most of the common activity types tested in a controlled setting, especially walking and running, and that the trackers had variable success in identifying the duration of the performed activities. This study measured the accuracy of the automatic activity-type recognition and duration feature on 4 wearable activity trackers during walking, running, using an elliptical trainer, biking, and swimming. Although there are countless permutations of possible physical activity behavior (eg, combinations of intensity, duration, setting, and stops and starts), the goal of this study was to conduct a preliminary laboratory-based test of multiple common activity types. All tested trackers had better recognition of ambulatory-based activities such as walking or running as opposed to biking, using an elliptical, or swimming. This finding is consistent with previous validation studies that show wearable activity trackers generally have high correlations of step counts when compared with the criterion during ambulatory activity [23]. Although this study did not aim to validate step counts (which have already been well studied) [23], walking and running were the only activities tested that involved stepping, which may bolster the trackers' detection capability.

Outdoor walking was misclassified as using an elliptical once by both the Fitbit Charge 2 and Fitbit Alta HR. These misclassifications occurred for the same participant and may be because of the participant having an arm movement pattern that mimics an elliptical arm handle pattern during walking. The high misclassification rate of using an elliptical by the Garmin device is concerning and indicates that it is likely unsuitable for apps where participants or patients may engage in use of elliptical machines as a regular part of their exercise regimen. With regard to swimming, although the Fitbit brand recommends swimming freestyle consistently for best detection results, it also claims the Fitbit Flex 2 can measure the backstroke, breaststroke, and butterfly [24]. Although we did not aim to measure strokes other than freestyle, it should be noted that the device did correctly identify all 5 of the swim trials that experienced a switch in stroke. This is important as stroke switching represents a more natural pattern than swimming a single stroke for an entire bout of activity, suggesting the tracker may be robust for capturing mixed-stroke swimming. However, because of the very small number of trials in which multiple strokes are used, additional data are needed to confirm whether this is true.

The detected mean duration was less than 4 min over or under the actual duration of the activity for outdoor walking and running, the embedded run, using an elliptical, outdoor biking, and swimming. Of these activities, MAPEs were the lowest for outdoor walking and running and treadmill walking. The treadmill run duration for all Fitbit trackers was considerably overestimated by an average of 15 min. The MAPE for treadmill running was also noticeably larger than other activities falling between 115.4% and 144.8%. The considerable overestimation was likely because of the trackers failing to recognize the 10 min of stationary rest after the treadmill run and including some duration of the embedded run (walk, run, walk). The addition of time occurred during 16 trials detected by the Fitbit Alta HR, 20 trials detected by Fitbit Charge 2, and 20 trials detected by the Fitbit Flex 2. If confirmed by future studies, this may indicate that trackers need improvement before they are suitable for satisfactorily identifying and measuring exercise sessions comprising intermittent or combined activities. This is a particular concern as many free-living activities are performed in short bouts and adjacent to other movements; thus, our findings suggest that at least for short bouts of activity (approximately 15 min), current trackers are prone to substantial error in duration estimates.

Similar results were observed for most activities by the Fitbit devices, regarding both activity detection and duration. The 3 Fitbit devices yielded activity detection results that were at most 3.2% different from one another for any ambulatory activity, excluding the embedded run, and had identical results for the outdoor run. This similarity is also observed for the nonambulatory activities. The Fitbit devices yielded results that were, at most, 3.1% different from one another (at least 0.1%) for using an elliptical and, at most, 3% different or identical for outdoor biking. When comparing the Fitbit brand MAPEs, the

treadmill run activity shows the most variability in results with the Fitbit Alta HR and Fitbit Flex 2 differing by 29.4% and with all other activities having identical results or differing by, at most, 5.3%. It is likely that these devices use similar, if not identical, classification algorithms, as they are developed and manufactured by the same company. It is important to consider that future Fitbit devices may continue to yield similar results across devices. If the use of identical technology across devices is confirmed by the company, this may eliminate the need for testing multiple devices in a single study.

This paper provides, to our knowledge, the first data on the automatic activity-type recognition feature in wearable trackers. A major strength is the inclusion of multiple common activities in different contexts (eg, walking outdoors and on a treadmill). The use of a standardized protocol is both a strength and a limitation. As no previous data on this feature are available, we chose a standardized protocol because it provides the ability to directly observe the type and duration of activity participation. However, it is unknown how well the results of this structured experimental study would generalize to free-living settings. Individuals who use wearable activity trackers tend to wear them all day, during which a wide variety of movements and activities are performed. By contrast, in a laboratory-based protocol, the devices are worn only during the study visit and are removed after completion of the final activity in each module. This removal may aid the tracker in determining the stop time of those activities. A second limitation is that the we did not attempt to determine how the devices' accuracy might be affected by the duration, speed, or intensity of an activity. Our goal was to test multiple activity types, both indoor and outdoor, to provide a preliminary test of the activity recognition function. Although it is never possible to test all possible combinations of activity type, duration, setting, speed, and intensity, additional research will provide substantial insight on how these factors generally affect accuracy of the automatic activity recognition feature.

Our results add to the evidence base of previous validation studies that have examined the accuracy of other wearable tracker features, such as steps, which show high correlation between criterion and device [13,14], energy expenditure, which is generally underestimated when compared with the criterion [15], and heart rate, which is more accurate at rest than while exercising [16-18]. Although wearable trackers do not correctly measure all activity metrics perfectly for every trial, their ability to provide objective measures of some aspects of activity is a great advancement from even 10 years ago, with technology likely to improve rapidly with time.

Conclusions

Wearable activity trackers correctly identified the type of isolated bouts of some common physical activities, although accuracy decreased substantially for certain activities that were adjacent (or nearly so) with one another, and 1 device had very poor detection on the elliptical trainer. Future directions for research on wearable trackers include testing the recognition capabilities of more and complex permutations of activity participation, including aerobic workouts, such as kickboxing or high intensity interval training, sports such as basketball and soccer, slower walking or light-intensity activities (of particular interest for chronic disease populations and older adults), intermittent activities such as jogging interrupted by stoplights, and other free-living activity patterns at a variety of exercise intensities. Owing to the nature of the rapidly evolving wearable technology market, future directions also include testing the automatic activity detection capability in new devices as they are released to the public.

Acknowledgments

LCB's time is funded by the National Cancer Institute (1K07CA178870). The authors would like to thank the Physical Activity Epidemiology undergraduate students (Bianca Briones and Rachel Schueller) and the study participants for their contributions to this study.

Conflicts of Interest

None declared.

References

1. Lee I, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT, Lancet Physical Activity Series Working Group. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. Lancet 2012 Jul 21;380(9838):219-229 [FREE Full text] [doi: 10.1016/S0140-6736(12)61031-9] [Medline: 22818936]
2. Arem H, Moore SC, Patel A, Hartge P, Berrington DG, Visvanathan K, et al. Leisure time physical activity and mortality: a detailed pooled analysis of the dose-response relationship. JAMA Intern Med 2015 Jun;175(6):959-967 [FREE Full text] [doi: 10.1001/jamainternmed.2015.0533] [Medline: 25844730]
3. Blair SN, Brodney S. Effects of physical inactivity and obesity on morbidity and mortality: current evidence and research issues. Med Sci Sports Exerc 1999 Nov;31(11 Suppl):S646-S662. [Medline: 10593541]
4. Physical Activity Guidelines Advisory Commitee. Physical Activity Guidelines Advisory Committee report, 2008. To the Secretary of Health and Human Services. Part A: executive summary. Nutr Rev 2009 Feb;67(2):114-120. [doi: 10.1111/j.1753-4887.2008.00136.x] [Medline: 19178654]
5. Piercy KL, Troiano RP, Ballard RM, Carlson SA, Fulton JE, Galuska DA, et al. The physical activity guidelines for Americans. J Am Med Assoc 2018 Jan 20;320(19):2020-2028. [doi: 10.1001/jama.2018.14854] [Medline: 30418471]

6. Troiano RP, Berrigan D, Dodd KW, Mâsse LC, Tilert T, McDowell M. Physical activity in the United States measured by accelerometer. Med Sci Sports Exerc 2008 Jan;40(1):181-188. [doi: 10.1249/mss.0b013e31815a51b3] [Medline: 18091006]
7. Statista. 2019. Number of Fitbit devices sold worldwide from 2010 - 2018 (in 1,000s) URL: https://www.statista.com/statistics/472591/fitbit-devices-sold/ [accessed 2019-04-13] [WebCite Cache ID 77diTSBYZ]
8. Internicola D. Scientific American. 2014 Dec 22. Activity trackers get smarter at measuring your fitness URL: https://www.scientificamerican.com/article/activity-trackers-get-smarter-at-measuring-your-fitness/ [accessed 2019-04-14] [WebCite Cache ID 77dilpbOK]
9. Cadmus-Bertram LA, Marcus BH, Patterson RE, Parker BA, Morey BL. Randomized trial of a Fitbit-based physical activity intervention for women. Am J Prev Med 2015 Sep;49(3):414-418. [doi: 10.1016/j.amepre.2015.01.020] [Medline: 26071863]
10. Cadmus-Bertram L, Marcus BH, Patterson RE, Parker BA, Morey BL. Use of the Fitbit to measure adherence to a physical activity intervention among overweight or obese, postmenopausal women: self-monitoring trajectory during 16 weeks. JMIR Mhealth Uhealth 2015;3(4):e96 [FREE Full text] [doi: 10.2196/mhealth.4229] [Medline: 26586418]
11. Phillips SM, Cadmus-Bertram L, Rosenberg D, Buman MP, Lynch BM. Wearable technology and physical activity in chronic disease: opportunities and challenges. Am J Prev Med 2018 Jan;54(1):144-150. [doi: 10.1016/j.amepre.2017.08.015] [Medline: 29122356]
12. Gualtieri L, Rosenbluth S, Phillips J. Can a free wearable activity tracker change behavior? The impact of trackers on adults in a physician-led wellness group. JMIR Res Protoc 2016 Nov 30;5(4):e237 [FREE Full text] [doi: 10.2196/resprot.6534] [Medline: 27903490]
13. Diaz KM, Krupka DJ, Chang MJ, Peacock J, Ma Y, Goldsmith J, et al. Fitbit®: an accurate and reliable device for wireless physical activity tracking. Int J Cardiol 2015 Apr 15;185:138-140. [doi: 10.1016/j.ijcard.2015.03.038] [Medline: 25795203]
14. Case MA, Burwick HA, Volpp KG, Patel MS. Accuracy of smartphone applications and wearable devices for tracking physical activity data. J Am Med Assoc 2015 Feb 10;313(6):625-626. [doi: 10.1001/jama.2014.17841] [Medline: 25668268]
15. O'Driscoll R, Turicchi J, Beaulieu K, Scott S, Matu J, Deighton K, et al. How well do activity monitors estimate energy expenditure? A systematic review and meta-analysis of the validity of current technologies. Br J Sports Med 2018 Sep 07 Epub ahead of print(forthcoming). [doi: 10.1136/bjsports-2018-099643] [Medline: 30194221]
16. Cadmus-Bertram L, Gangnon R, Wirkus EJ, Thraen-Borowski KM, Gorzelitz-Liebhauser J. The accuracy of heart rate monitoring by some wrist-worn activity trackers. Ann Intern Med 2017 Apr 18;166(8):610-612. [doi: 10.7326/L16-0353] [Medline: 28395305]
17. Wang R, Blackburn G, Desai M, Phelan D, Gillinov L, Houghtaling P, et al. Accuracy of wrist-worn heart rate monitors. JAMA Cardiol 2017 Jan 01;2(1):104-106. [doi: 10.1001/jamacardio.2016.3340] [Medline: 27732703]
18. Wallen MP, Gomersall SR, Keating SE, Wisløff U, Coombes JS. Accuracy of heart rate watches: implications for weight management. PLoS One 2016;11(5):e0154420 [FREE Full text] [doi: 10.1371/journal.pone.0154420] [Medline: 27232714]
19. Fitbit. 2017. Fitbit SmartTrack - Auto Exercise Recognition URL: https://www.fitbit.com/smarttrack [accessed 2019-04-14] [WebCite Cache ID 77djOKMdF]
20. Garmin. 2017. Understanding Move IQ URL: https://support.garmin.com/en-US/?faq=zgFpy8MShkArqAxCug5wC6&searchType=noProduct&utm_source=faqSearch [accessed 2019-04-14] [WebCite Cache ID 77dkDvIQr]
21. Borg GA. Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982;14(5):377-381. [Medline: 7154893]
22. Chen M, Fan X, Moe S. Criterion-related validity of the Borg ratings of perceived exertion scale in healthy individuals: a meta-analysis. J Sports Sci 2002 Nov;20(11):873-899. [doi: 10.1080/026404102320761787] [Medline: 12430990]
23. 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]
24. Fitbit. 2019 Apr 03. How do I track swims with my Fitbit device? URL: https://help.fitbit.com/articles/en_US/Help_article/2087 [accessed 2019-04-13] [WebCite Cache ID 77dj2ESRs]

Abbreviations

MAPE: mean absolute percent error
RPE: rating of perceived exertion

Edited by G Eysenbach; submitted 30.01.19; peer-reviewed by J Sasaki, M Boukhechba, K Winfree, S Gomersall, MS Aslam, C Fincham, R Islam; comments to author 27.02.19; revised version received 13.03.19; accepted 23.03.19; published 23.05.19.

Please cite as:
Dorn D, Gorzelitz J, Gangnon R, Bell D, Koltyn K, Cadmus-Bertram L
Automatic Identification of Physical Activity Type and Duration by Wearable Activity Trackers: A Validation Study
JMIR Mhealth Uhealth 2019;7(5):e13547
URL: http://mhealth.jmir.org/2019/5/e13547/
doi:10.2196/13547
PMID:31124470

Original Paper

Social Inequity and Structural Barriers to Completion of Ecological Momentary Assessments for Young Men Who Have Sex With Men and Trans Women Living With HIV in San Francisco

Caitlin M Turner[1], MPH; Sean Arayasirikul[1], PhD; Dillon Trujillo[1], MPH; Victory Lê[1], BA; Erin C Wilson[1], MPH, DrPH

San Francisco Department of Public Health, San Francisco, CA, United States

Corresponding Author:
Caitlin M Turner, MPH
San Francisco Department of Public Health
25 Van Ness Avenue
Suite 500
San Francisco, CA, 94102
United States
Phone: 1 4155549018
Email: caitlin.turner@sfdph.org

Abstract

Background: Ecological momentary assessments (EMAs) administered via text messaging facilitate real-time data collection. With widespread cell phone access, EMAs are becoming more available to even the most disenfranchised communities, such as those living with HIV. However, structural barriers disproportionately burden young men who have sex with men (MSM) and trans women (TW) living with HIV and threaten participation in HIV research.

Objective: We aim to identify structural barriers to completing EMA text surveys nested within a digital HIV care intervention for young MSM and TW living with HIV in San Francisco.

Methods: A total of 10,800 EMA text messages were delivered daily over 90 days to 120 participants enrolled in the Health eNav intervention (2017-2018) at the San Francisco Department of Public Health. EMA surveys inquired about participants' daily affect, sexual behaviors, substance use, and treatment adherence. Survey completion was calculated after 30, 60, and 90 days of follow-up. We described characteristics of nonstarters (those who provided less than four complete responses to the first seven EMA surveys) and analyzed structural correlates of days to first weeklong or more EMA survey noncompletion using multivariable Cox proportional hazards regression. Qualitative interviews were used to evaluate the acceptability of EMA surveys.

Results: Participants completed 4384 of 10,800 (40.59%) EMA surveys. Completion of 70% or more of EMA surveys was attained by 56 of 120 participants (46.7%) at 30 days of follow-up, 40/120 (33.3%) at 60 days of follow-up, and 30/120 (25.0%) by the end of the 90-day study period. Twenty-eight participants (23.3%) were identified as nonstarters, and were more likely to be recently incarcerated (prevalence ratio [PR] 2.3, 95% CI 1.3-4.4), forego basic needs for HIV medications (PR 2.4, 95% CI 1.3-4.5), and be diagnosed with HIV in the last year (PR 2.2, 95% CI 1.1-4.1). Adjusting for nonstarters, young MSM and TW living in temporary/transitional housing (adjusted hazard ratio [aHR] 1.8, 95% CI 1.1-3.0), foregoing HIV medications to afford basic needs (aHR 1.7, 95% CI 1.1-2.7), and having less than a college education (aHR 3.5, 95% CI 1.4-9.0) had greater hazard of weeklong or more EMA survey noncompletion. Overall, there was high acceptability of the EMA surveys.

Conclusions: Although access to and use of technology is increasingly ubiquitous, this analysis demonstrates persisting gaps in EMA completion by socioeconomic factors such as incarceration, education level, housing, and competing needs for young MSM and TW living with HIV in San Francisco. Moreover, those recently diagnosed with HIV were more likely to experience an immediate drop-off in completing EMA surveys. EMAs are feasible for individuals not experiencing social inequity and structural barriers. HIV prevention technologies addressing these barriers and leveraging similar methodology may prove effective for young MSM and TW living with HIV.

(JMIR Mhealth Uhealth 2019;7(5):e13241) doi:10.2196/13241

KEYWORDS

ecological momentary assessment; EMA; MSM; trans women; people living with HIV; feasibility; acceptability

Introduction

Ecological momentary assessments (EMAs) facilitate real-time data collection and may be administered through a number of technological platforms, such as handheld devices or mobile phones. Previous studies have used the terms “daily diaries” or “electronic diaries” interchangeably with EMAs; however, in this study, we will use only EMA for consistency. A range of research projects have used EMAs to improve data accuracy and to capture nuances in affect or mood [1], substance use [1-5], and, more recently, behaviors among persons living with HIV [5-8].

With ubiquitous mobile phone access, EMA and mobile interventions are available for communities with the greatest burden of HIV. As of 2014, more than 80% of young adults owned a mobile phone for information gathering and communication purposes [9]. A number of studies found that young men who have sex with men (MSM) use mobile devices to access sexual health information and consider mobile HIV interventions to be acceptable [10-12]. In a study of young, black MSM, researchers concluded that mobile HIV interventions may increase participation in HIV prevention and treatment [13]. One study of young MSM living with HIV showed high feasibility and acceptability of daily diaries measuring affect, substance use, and other risk behaviors [8]. Another study found acceptability of a peer-administered mobile health intervention among persons living with HIV who also had a history of substance use disorders and low adherence to antiretroviral treatment [5]. However, no studies exist that similarly examine the feasibility and acceptability of EMA among trans women (TW) living with HIV.

Currently, TW are among the highest risk groups for HIV with more than 49 times the odds of HIV infection than other reproductive age groups [14]. Efforts to collect data for this community could be vastly improved by using EMAs. However, to our knowledge, no quantitative studies of EMAs or other mobile health interventions exist for TW. Just as young MSM and other HIV risk groups consistently demonstrate high EMA compliance [1,2-4,8,15], TW living with HIV may similarly respond well to such a data collection method.

However, TW and MSM living with HIV face significant structural barriers which may threaten their participation in EMA research. For this analysis, we define structural barriers as institutional or social determinants of health (eg, housing instability, incarceration, and competing needs). Consistent with literature positing social factors as fundamental causes of HIV and other adverse outcomes [16,17], we approach EMA engagement in HIV interventions through a structural lens.

Seropositivity for HIV can be both an effect and a determinant of structural vulnerability and socioeconomic crises. Housing instability [18], racial/ethnic disparities [19], and poverty [20] promote HIV infection. People living with HIV experience barriers to economic opportunities due to HIV-related stigma and financial/logistical demands of accessing HIV-related care [21,22]. Competing needs (eg, prioritizing housing over health care) may put people at increased risk for HIV infection or may be exacerbated after people contract HIV [23]. When coupled with discrimination and adversity due to sexual or gender minority status, socioeconomic crises are exponentially worse for young MSM and TW living with HIV [24-26] compared to their nonminority counterparts. In one study of morbidity and mortality among persons living with HIV, homelessness and incarceration were associated with an increased mortality risk for MSM. These factors were not significantly associated for nonsexual minority participants and were not assessed for TW [27].

Few studies examine structural barriers to EMA completion. Two EMA feasibility studies among MSM showed that men without college degrees had lower EMA completion [2,4]. Prior research demonstrated inconsistent results regarding the relationship between race/ethnicity and EMA completion: two studies cite no differences in EMA completion by race/ethnicity [4,6], and one study showed lower EMA completion among MSM of color [2]. A study of risk behaviors among young people who inject drugs found that homeless participants had lower completion of EMAs [1]. Structural barriers to EMA completion have yet to be explored among MSM and TW living with HIV.

Individual characteristics such as recent HIV diagnosis and technological behaviors may also influence EMA study participation. We hypothesized that participants who generally exhibited high mobile technology interactions (ie, higher frequencies of sending, receiving, or checking text messages) would have greater EMA completion. We also hypothesized that newly diagnosed persons may be facing a unique set of experiences—higher depression risk [28] and HIV-related stigma—which may interfere with EMA completion. Finally, prior research has shown mixed results in terms of the relationship between age and EMA completion. A study among substance-using MSM showed lower completion among older MSM [2]; conversely, another study demonstrated that EMA was feasible and acceptable among older adults living with HIV [7]. However, technological behaviors, recent HIV diagnosis, and age have yet to be explored as possible facilitators or barriers to EMA completion among young TW and MSM living with HIV.

To address these research gaps, we conducted a mixed methods study of EMA feasibility and acceptability among young MSM and TW living with HIV who were enrolled in an HIV digital care navigation intervention in San Francisco. We also assessed EMA participation by structural and individual characteristics.

Methods

Study Design and Procedures

We analyzed data from Health eNav at the San Francisco Department of Public Health from 2017 to 2018. The Institutional Review Board at the University of California, San Francisco, approved the study procedures (IRB #16-19675). Health eNav is a digital HIV care navigation intervention that uses short message service (SMS) text messaging to improve outcomes along the HIV care continuum for young MSM and TW living with HIV.

On enrollment, participants completed EMAs that monitored their daily affect, sexual behaviors, substance use, and treatment adherence. This study focuses on the EMA component of the larger study. To assess feasibility, we analyzed EMA completion, hypothesizing that engagement in the EMA portion of the study would relate to engagement in the larger intervention. Characterizing EMA completion was also a necessary step for later incorporating it as an intervention exposure variable in future analyses. Participants were compensated US $1 for each completed EMA survey for up to US $90 over the EMA portion of the study. If participants completed more than 80% of their EMA surveys, they earned a bonus of US $100. Incentives were provided in the form of a gift card.

In addition to quantitatively analyzing feasibility, we assessed acceptability by conducting semistructured, in-depth interviews with a subsample of 16 participants 12 months after enrollment. Participants were purposively sampled to obtain diversity in race/ethnicity, gender identity, and level of engagement with the intervention. Participants were provided US $75 gift cards for their time. Interviews lasted 30 to 45 minutes and took place during a time that was most convenient for the participant. The interview guide was iterated to maximize coverage of participant experiences through theoretical sampling to reach theoretical saturation [29] and to address the following research question: "What factors impacted acceptability of EMA for young people living with HIV?" The interview guide assessed the following constructs: preferences, length, and individual- and health-related effects. Interviews were audio-recorded and transcribed verbatim. Transcriptions were randomly checked for quality and accuracy against original recordings. Qualitative interview data were coded and analyzed using grounded theory [29]. Two members of the research team independently coded qualitative data, line by line, and organized codes into categories to identify specific factors that shaped acceptability of EMAs for young people living with HIV.

Participants

Eligible participants (1) identified as MSM or TW, (2) were between the ages of 18 and 34 years, (3) reported living in San Francisco, and (4) were newly diagnosed with HIV or not engaged/retained in care or not virally suppressed. Convenience sampling was used to recruit potential participants from five clinics and community-based organizations (CBOs) in San Francisco serving young people living with HIV. Posters, palm cards, and presentations (eg, in-service presentations, team huddles) were used to advertise study recruitment and to educate CBO and clinical staff. Staff referred potential participants to the study through phone/email communication and/or in-person meetings. Enrolled participants were also invited to refer peers from their social network.

Interested participants completed an in-person or telephone eligibility screening. If eligible, participants met with research staff at study offices located at the local health department to obtain informed consent and complete study enrollment activities. Overall, 171 people were screened for eligibility; of those, 140 were eligible, but 20 were lost to follow-up and did not enroll. In total, 120 participants were enrolled in the study. The majority of enrolled participants (90.8%, 109/120) were referred by clinics/CBOs and the rest (9.2%, 11/120) through peer referral. All 120 enrolled participants are included in this analysis.

Measures and Variable Selection

For the EMA component of the study, participants were sent automated SMS text messages once per day at 8:00 am, 12:00 pm, or 8:00 pm for 90 consecutive days via mSurvey [30]. They were required to respond to EMA surveys within 24 hours. Participants could receive between 17 and 31 daily EMA texts depending on their responses. For example, if a participant responded "yes" to having sex within the last 24 hours, the participant would receive follow-up questions about whether condoms were used. Had the participant responded "no," the participant would not receive subsequent questions about condom use. EMAs tended to take less than 5 minutes to complete each day. EMA completion was calculated day by day; EMA responses were considered incomplete if any or all of the 17 to 31 EMA texts received were skipped in a given day. Hypothesized barriers to EMA completion were collected via computer-assisted self-interviewing (CASI) surveys administered to participants at baseline and merged with EMA completion data.

Ecological Momentary Assessment Survey Completion

First, we computed cumulative EMA completion with the proportion of EMA surveys completed out of the total number of EMA surveys sent. We also calculated the percentage of participants who provided complete responses to 70% or more of the EMA surveys over 30 days of study follow-up. Health eNav followed participants for a relatively extended time period of 90 days; therefore, we also calculated EMA completion after 60 and 90 days.

We also characterized EMA completion in nontraditional yet clinically relevant ways. Hypothesizing that participants who failed to complete EMA surveys from the start of the study would be less likely to complete EMA surveys throughout the rest of the study, we defined "nonstarters" as those who did not complete a majority of the EMA surveys they received in the first week of the study (ie, those who provided less than four complete responses to the first seven EMA surveys sent).

We used survival analysis methods to calculate the number of study days that transpired before a participant failed to send complete responses to EMA surveys for at least seven consecutive days (ie, time to first weeklong or more EMA survey noncompletion), consistent with a previous study of EMA among substance-using MSM [2]. For this study, not completing EMA surveys for 7 or more days implied that no value was added by administering daily assessments over traditional weekly, biweekly, or monthly assessments. In addition, 1 week was a meaningful unit of time to characterize participants who may have had low completion in tandem with the overall digital navigation intervention. Classifying noncompletion in this way would allow us to later characterize participation in the overall digital HIV care intervention.

To visualize time to weeklong or more EMA survey noncompletion, we created an event plot and a Kaplan-Meier

survival curve. Nonstarters were excluded from the Kaplan-Meier survival curve because we hypothesized that they represented a distinct outcome group.

Hypothesized Barriers to Ecological Momentary Assessment Survey Completion

Factors associated with time to first weeklong or more EMA survey noncompletion were selected a priori or were hypothesized as barriers to completion. Sociodemographics, such as age (in years), race/ethnicity (black/African American, Hispanic/Latinx, other/multiple, or white), and education level (less than high school, high school or GED, and at least some college), have been shown previously to influence EMA completion, with older participants, people of color, those with less than a college education, or those with lower income having lower EMA completion [2,4,6]. We also included housing status (living with a family member, friend, or partner who rents/owns a home; living in temporary/transitional housing; experiencing homelessness; or renting/owning a home) as a possible indicator of noncompletion, given previous literature citing homelessness as a barrier to participation [1]. Recent incarceration and competing needs (eg, foregoing HIV medications to afford basic needs such as food, housing, and/or clothing and vice versa) have been shown to exacerbate health outcomes for those living with HIV [23,27] but had yet to be explored as possible barriers to EMA participation.

We hypothesized that young TW and MSM who were recently diagnosed with HIV would face more challenges completing EMA surveys due to competing needs, stigma, and identity development related to seroconversion. Finally, we hypothesized that participants who had greater interactions with technology would have higher EMA completion. Therefore, we included frequency of sending/receiving or checking text messages on a mobile phone, categorized as once a day or less, several times a day, and several times an hour or all the time.

Statistical and Qualitative Analyses

Analyses were conducted using Stata 14 software (StataCorp LP, College Station, TX, USA). Days to first weeklong or more EMA survey noncompletion, demographics, structural and socioeconomic factors, HIV diagnosis status, and technology behaviors were described for the entire sample (N=120) and bivariable Poisson binomial regression models [31] of these characteristics were analyzed for nonstarters compared to those who completed four or more EMA surveys in the first week of EMA surveys received.

Multivariable Cox proportional hazards models were used to calculate possible associations between the aforementioned sample characteristics and hazard of weeklong or more EMA survey noncompletion. Because nonstarters were analyzed as a separate group, we included nonstarter status as a covariate in multivariable survival analysis models of associations between structural barriers and noncompletion of EMA surveys. Proportional hazards assumptions were checked for each model. If any violations occurred, nonstarter status was interacted with time, and the models were rerun.

To assess acceptability of EMA, we conducted structured qualitative interviews postintervention of 16 participants. We used content analysis [32] to identify key themes, attitudes toward the EMA surveys, opinions about EMA survey length, and effect of the EMA surveys in general and on medical care.

Results

Young MSM and TW living with HIV in Health eNav were racially/ethnically diverse, with a majority identifying as black or African American (22/120, 18.3%) or as Hispanic/Latinx (38/120, 31.7%). The average age was 27.8 (SD 4.1) years. Most participants lived in temporary/transitional housing (43/120, 35.8%) or rented/owned a home (39/120, 32.5%). A majority of participants earned US $1300 or less in monthly income (90/120, 75.0%), yet over half of the sample had an associate's, technical/vocational, or bachelor's degree or higher (68/120, 56.7%). Almost one in five participants reported being incarcerated in the last 6 months. Over a quarter of the sample had competing needs, either going without HIV medications because they needed money for basic needs or going without basic needs to afford HIV medications. Most participants (82/120, 68.3%) were diagnosed with HIV over a year prior to their baseline CASI assessment. A majority of participants reported that, in general, they frequently sent, received, or checked text messages on a mobile phone (Table 1).

Table 1. Sample characteristics for all participants and by nonstarter status (ie, completed less than four of the first seven EMA surveys) from the 2017-2018 Health eNav study (N=120).

Sample characteristics	All (N=120)	Completed ≥4 of first 7 EMA surveys (n=92)	Nonstarters (n=28)	PR[a] (95% CI[b])
Days to first weeklong or more EMA survey noncompletion, mean (SD)	41.87 (35.42)	51.35 (33.10)	10.71 (23.06)	0.95 (0.92-0.99)
Demographics				
Age (years), mean (SD)	27.75 (4.07)	28.43 (5.88)	25.25 (2.22)	0.96 (0.88-1.04)
Gender identity, n (%)				
Trans woman	17 (14.2)	10 (11)	7 (25)	2.02 (1.02-4.02)
Man	103 (85.8)	82 (89)	21 (75)	Ref[c]
Race/ethnicity, n (%)				
Black or African American	22 (18.3)	16 (17)	6 (21)	1.75 (0.60-5.04)
Hispanic or Latinx	38 (31.7)	27 (29)	11 (39)	1.85 (0.72-4.79)
Other[d] or multiple	28 (23.3)	22 (24)	6 (21)	1.37 (0.47-1.03)
White	32 (26.7)	27 (29)	5 (18)	Ref
Structural factors				
Housing status, n (%)				
Lives with a family member, friend, or partner who rents/owns a home	21 (17.5)	16 (18)	5 (18)	1.86 (0.60-5.72)
Temporary/transitional housing[e]	43 (35.8)	30 (33)	13 (46)	2.36 (0.92-6.04)
Homeless or shelter	17 (14.2)	12 (13)	5 (18)	2.29 (0.76-6.93)
Rents/owns an apartment or house	39 (32.5)	34 (37)	5 (18)	Ref
Income in the last month (US$), n (%)				
$601-$1300	30 (25.0)	21 (23)	9 (32)	2.18 (0.75-6.32)
$251-$600	30 (25.0)	24 (26)	6 (21)	1.45 (0.45-4.64)
$0-$250	30 (25.0)	21 (23)	9 (32)	2.18 (0.75-6.32)
≥$1301	29 (24.2)	25 (27)	4 (14)	Ref
Education, n (%)				
High school/GED	39 (32.5)	27 (29)	12 (43)	1.9 (0.93-3.91)
Less than high school	13 (10.8)	8 (9)	5 (18)	2.38 (0.99-5.72)
Some college or more	68 (56.7)	57 (62)	11 (39)	Ref
Incarceration in the last 6 months, n (%)				
Yes	23 (19.2)	13 (14)	10 (36)	2.34 (1.25-4.39)
No	97 (80.8)	79 (86)	18 (64)	Ref
Competing needs, n (%)				
Went without HIV medications to have money for basic needs (eg, food, housing, and/or clothing)	38 (31.7)	66 (72)	16 (57)	1.62 (0.85-3.08)
Went without basic needs (eg, food, housing, and/or clothing) to have money for HIV medications	32 (26.7)	26 (28)	12 (43)	2.38 (1.28-4.45)
HIV diagnosis status, n (%)				
Diagnosed in the last year	38 (31.7)	24 (26)	14 (50)	2.16 (1.14-4.08)
Diagnosed more than one year ago	82 (68.3)	68 (74)	14 (50)	Ref
Behaviors toward technology, n (%)				
Frequency of sending and receiving text messages on a mobile phone				

Sample characteristics	All (N=120)	Completed ≥4 of first 7 EMA surveys (n=92)	Nonstarters (n=28)	PR[a] (95% CI[b])
Several times a day	39 (32.5)	32 (35)	7 (25)	0.45 (0.18-1.12)
Several times an hour or all the time	66 (55.0)	51 (55)	15 (54)	0.57 (0.26-1.22)
Once a day or less	15 (12.5)	9 (10)	6 (21)	Ref
Frequency of checking for text messages on a mobile phone				
Several times a day	42 (35.0)	35 (38)	7 (25)	0.40 (0.15-1.04)
Several times an hour or all the time	66 (55.0)	50 (54)	16 (57)	0.58 (0.26-1.29)
Once a day or less	12 (10.0)	7 (8)	5 (18)	Ref

[a]PR: crude prevalence ratio from bivariable Poisson binomial regression models comparing nonstarters to those who completed four or more of the first seven EMA surveys received.

[b]CI: confidence interval.

[c]Ref: reference group.

[d]"Other" race/ethnicity included participants who identified as American Indian or Alaska Native (n=6) or Asian (n=7).

[e]Temporary/transitional housing included participants who lived in single room occupancy hotels, motels, boarding houses, halfway houses, drug treatment centers, independent living units, domestic violence shelters, battered persons' shelters, or "safe houses."

Ecological Momentary Assessment Survey Completion

A total of 10,800 EMA surveys were sent to all 120 Health eNav participants over 90 days. Cumulatively, participants completed 4384 of 10,800 (40.59%) EMA surveys. At least 70% EMA completion was achieved by 56 of 120 participants (46.7%) by 30 days of follow-up, 40 of 120 participants (33.3%) by 60 days of follow-up, and 30 participants (25.0%) by 90 days of follow-up.

Almost one in four participants in Health eNav were nonstarters who completed less than four of the first seven EMA surveys they received. Nonstarters were more likely to have had less days to first weeklong or more EMA survey noncompletion (crude prevalence ratio [PR] 0.95, 95% confidence interval [CI] 0.92-0.99, *P*=.02), be TW (PR 2.02, 95% CI 1.02-4.02, *P*=.045), be incarcerated within the last 6 months (PR 2.34, 95% CI 1.25-4.39, *P*=.008), forego basic needs to afford HIV medications (PR 2.38, 95% CI 1.28-4.45, *P*=.006), or be diagnosed with HIV within the last year (PR 2.16, 95% CI 1.14-4.08, *P*=.02).

Days to weeklong or more EMA survey noncompletion for each participant are visualized in Figure 1. Figure 2 displays the proportion of the Health eNav sample (excluding nonstarters) who had not yet experienced a weeklong or more EMA survey noncompletion over the 90-day follow-up period.

Excluding nonstarters, the average time to weeklong or more EMA survey noncompletion was approximately 51 (SD 33) days. Of the 120 participants, 85 (70.8%) experienced a weeklong or more noncompletion; the average time to weeklong or more noncompletion was 22 (SD 20) days.

Figure 1. Event plot censoring the follow-up days that transpired before participants' first experience of a weeklong or more ecological momentary assessment (EMA) survey noncompletion from the 2017-2018 Health eNav study (N=120).

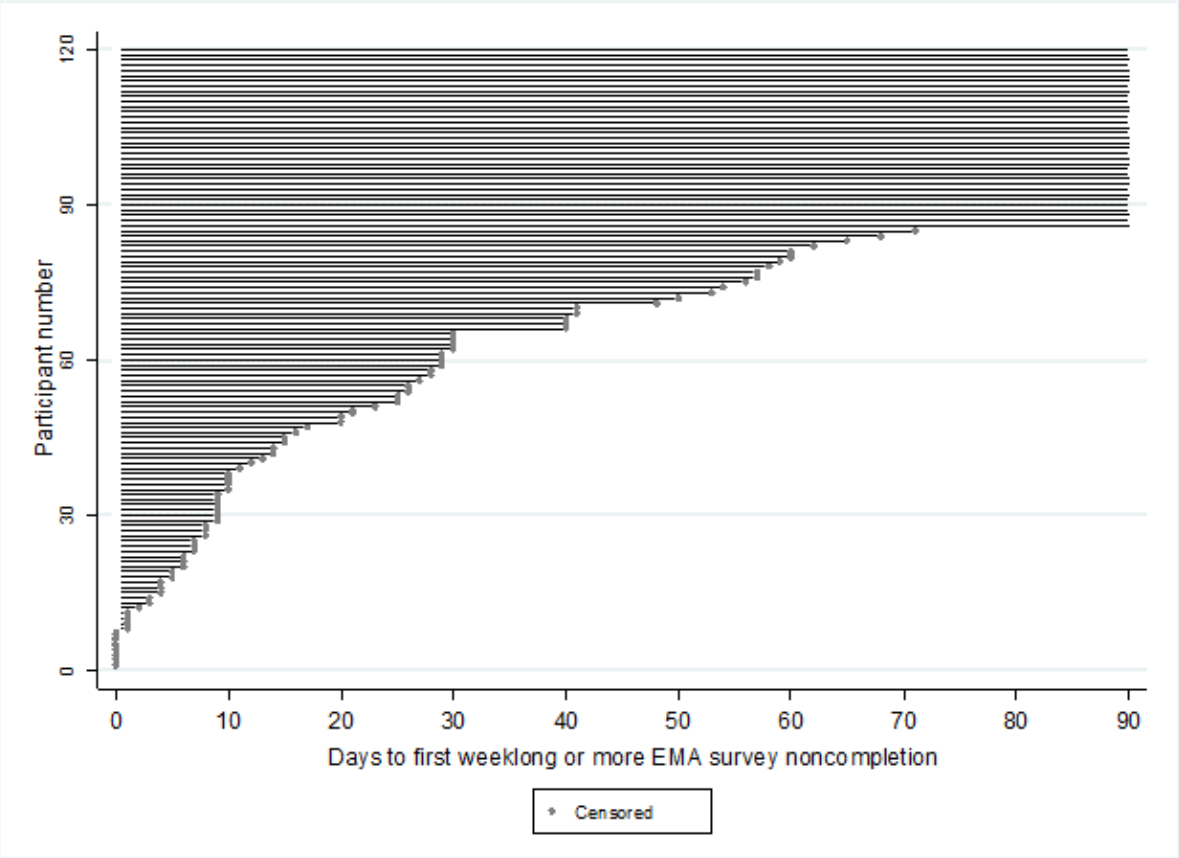
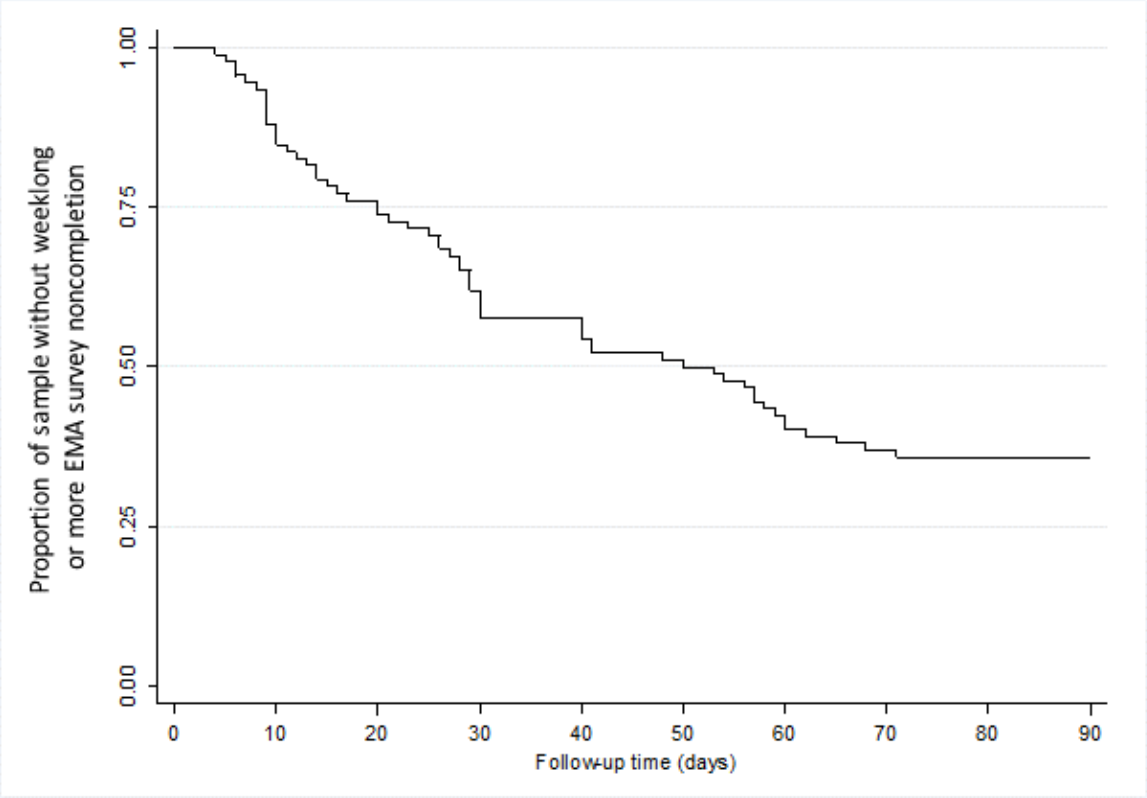


Figure 2. Kaplan-Meier survival curve visualizing the proportion of sample (excluding nonstarters, n=28) without a weeklong or more EMA survey noncompletion over the 90-day follow-up period from the 2017-2018 Health eNav study (n=92).



Hypothesized Barriers to Ecological Momentary Assessment Survey Completion

Table 2 shows results from multivariable Cox proportional hazards models.

Adjusting for nonstarter status, we found that young MSM and TW living with HIV who lived in temporary/transitional housing (eg, single room occupancy hotels, motels, boarding houses, halfway houses, drug treatment centers) had higher hazards of weeklong or more EMA survey noncompletion compared to those who rented or owned a home (adjusted hazard ratio [aHR] 1.78, 95% CI 1.06-3.01, *P*=.03). Similarly, those who had a high school education compared to those with some college education or more (aHR 1.83, 95% CI 1.16-2.89, *P*=.01) and those who went without HIV medications to afford basic needs (aHR 1.71, 95% CI 1.09-2.71, *P*=.02) had a higher adjusted hazard of weeklong or more EMA noncompletion.

After testing proportional hazards assumptions and assessing variable time dependence using Therneau and Grambsch's test of nonzero slopes [33], we concluded that nonstarter status should be interacted with time in our Cox proportional hazards models to assess the difference in hazards ratios by nonstarter status. Including the time-nonstarter status interaction, these post hoc survival analyses yielded similar conclusions as our original multivariable models for participants living in temporary/transitional housing (aHR 1.72, 95% CI 1.01-2.94, *P*=.047) and those who had a high school education (aHR 1.74, 95% CI 1.12-2.70, *P*=.01). However, competing needs (ie, foregoing HIV medications to afford food, housing, and/or clothing) was no longer significantly associated with greater hazard of weeklong or more EMA noncompletion (Table 2).

Table 2. Sample characteristics in relation to Cox proportional hazards of weeklong or more ecological momentary assessment survey noncompletion from the 2017-2018 Health eNav study (N=120).

Sample characteristics	With nonstarter status covariate, aHR[a] (95% CI[b])	With time-nonstarter status interaction, aHR (95% CI)
Demographics		
Age	1.03 (0.96-1.09)	1.03 (0.97-1.09)
Gender identity		
Trans woman	1.60 (0.85-2.99)	1.41 (0.79-2.51)
Man	Ref[c]	Ref
Race/ethnicity		
Black or African American	1.18 (0.65-2.12)	1.08 (0.58-2.01)
Hispanic or Latinx	0.75 (0.42-1.36)	0.76 (0.43-1.36)
Other or multiple	0.50 (0.23-1.09)	0.56 (0.27-1.16)
White	Ref	Ref
Structural factors		
Housing status		
Lives with a family member, friend or partner who rents/owns a home	0.87 (0.38-1.97)	1.26 (0.63-2.50)
Temporary/transitional housing	1.78 (1.06-3.01)	1.72 (1.01-2.94)
Homeless or shelter	1.49 (0.73-3.04)	1.47 (0.71-3.06)
Rents/owns an apartment or house	Ref	Ref
Income in the last month (US$)		
$601-$1300	1.16 (0.60-2.26)	1.34 (0.73-2.46)
$251-$600	1.04 (0.57-1.91)	1.18 (0.68-2.06)
$0-$250	1.13 (0.63-2.04)	1.15 (0.63-2.11)
≥$1301	Ref	Ref
Education		
High school/GED	1.83 (1.16-2.89)	1.74 (1.12-2.70)
Less than high school	1.12 (0.31-4.03)	1.55 (0.58-4.14)
Some college or more	Ref	Ref
Incarceration in the last 6 months		
Yes	1.67 (0.98-2.85)	1.44 (0.84-2.47)
No	Ref	Ref
Competing needs		
Went without HIV medications to have money for basic needs (eg, food, housing, and/or clothing)	1.71 (1.09-2.71)	1.46 (0.92-2.32)
Went without basic needs (eg, food, housing, and/or clothing) to have money for HIV medications	1.58 (0.96-2.61)	1.29 (0.78-2.15)
HIV diagnosis status		
Diagnosed in the last year	0.80 (0.46-1.39)	0.92 (0.56-1.51)
Diagnosed prior to the last year	Ref	Ref
Attitudes and behaviors toward technology		
Frequency of sending and receiving text messages on a mobile phone		
Several times a day	0.70 (0.35-1.43)	0.79 (0.38-1.67)
Several times an hour or all of the time	0.56 (0.29-1.06)	0.66 (0.35-1.25)
Once a day or less	Ref	Ref

Sample characteristics	With nonstarter status covariate, aHR[a] (95% CI[b])	With time-nonstarter status interaction, aHR (95% CI)
Frequency of checking for text messages on a mobile phone		
Several times a day	0.75 (0.33-1.69)	0.84 (0.37-1.92)
Several times an hour or all of the time	0.61 (0.27-1.35)	0.72 (0.33-1.57)
Once a day or less	Ref	Ref

[a]aHR: adjusted hazard ratio (adjusted for nonstarter status as a time-independent covariate or nonstarter status interacted with time).

[b]CI: confidence interval.

[c]Ref: reference group.

Acceptability of Ecological Momentary Assessments Among Young People Living With HIV

Overall, most participants (14/16, 86%) who were qualitatively interviewed found the EMA surveys to be acceptable. Five of 16 (31%) participants cited that the surveys were easy to complete. Two participants reported that they enjoyed having the EMA surveys as part of their daily routine. Others (4/16, 25%) reported EMA participation was motivating, citing that completing the EMA surveys proved that they had the ability to commit to something and it was an opportunity to talk about uncomfortable topics such as sex or drugs. One participant said they "felt important" when asked questions about themselves on a daily basis. Few participants (3/16, 19%) felt unfavorably about the EMA surveys. For one participant, completing EMA surveys was an added burden and responding was a low priority in the context of other stressors in their life; however, they did find the remuneration motivating. One participant simply was uninterested, and another participant felt paranoid when they were in public and received a notification to begin their EMA survey. Regardless of whether participants found the EMA surveys to be favorable, some participants (4/16, 25%) found the surveys to be repetitive or redundant.

In terms of survey length and duration, only two (13%) felt that the surveys were too long or should be offered over a shorter duration of time. Most (9/16, 56%) suggested administering the surveys over a longer period of time to make a habit out of medication adherence and other self-monitoring skills they learned from the EMA surveys. They also felt that a longer EMA duration would allow them to see more positive or negative growth and become more accustomed to having a daily routine.

Many participants felt that the EMA surveys had a positive impact on their lives. Most participants (9/16, 56%) felt that EMA surveys improved self-monitoring of behaviors/mood and offered reminders to take HIV medications. Many (6/16, 38%) reported that the EMA surveys created a designated time for self-reflection regarding mood, habits related to substance use or sexual behaviors, and medication adherence. Specifically, one participant learned to ask themselves "What am I doing?" and "What can I change?" For one participant who regularly used substances, the EMA surveys were something to look forward to. For three other participants, the EMA surveys were similar to talking to a friend or having a loved one check up on them; EMA surveys increased their engagement in social relationships and community. In some cases (3/16, 19%), EMA surveys improved moods or habits. However, three participants (19%) reported no impact of the EMA surveys on their lives. Not all participants were in a stable situation to reap benefits from the surveys, citing housing (1/16, 6%), personal crises or stress (1/16, 6%), or phone turnover (2/16, 13%) as barriers.

Several participants cited health-related benefits from the EMA surveys. Seven (44%) felt that they received more insight into their physical and mental health status or cared more about their health. Half the sample felt they had more agency in managing their health care, citing confidence in changing doctors (1/16, 6%), adhering to medical appointments (3/16, 19%), knowing their rights as a patient (1/16, 6%), and being better able to report about their health and habits during doctor's visits (1/16, 6%). One participant credited the EMA surveys for helping monitor their viral load and CD4 count.

Discussion

Young MSM and TW living with HIV in San Francisco had moderate cumulative completion of EMA surveys (4384/10,800, 40.59%). Seventy percent or higher EMA completion was achieved by approximately half of the participants by 30 days of follow-up, and a quarter of participants by the end of the study (90 days). Results from this study suggest that although young MSM and TW living with HIV have access to mobile technology and they frequently send, check, and receive texts on their mobile phone, they still face substantial structural barriers to participating in EMA surveys administered via text messaging. Participants who experienced an earlier time to weeklong or more EMA survey noncompletion, were TW, were recently incarcerated, or who had competing needs were more likely to complete less than four of the first seven EMA surveys they received. A higher hazard of weeklong or more EMA survey noncompletion was experienced by participants who lived in temporary/transitional housing, had less than a college education, or went without HIV medications to afford basic needs. However, when we specified an interaction between time and nonstarter status, the association between competing needs and hazard of EMA noncompletion was no longer statistically significant. Qualitative interviews further cemented the role of structural barriers to EMA completion, with participants who reported low acceptability of the EMA surveys implicating housing instability and mobile phone turnover. Separately, in ad hoc analyses, we found that 30.0% (36/120) of the sample experienced no structural barriers (not less than college educated and not foregoing HIV medications for basic needs and not living in temporary/transitional housing). The average time to

failure for these 36 individuals was 55.1 days, compared to 36.2 days for individuals who experienced at least one of these structural barriers (one-sided *t* test statistic: t_{118}=2.75, *P*=.007).

Recent HIV diagnosis was also associated with greater likelihood of completing less than four of the first seven EMA surveys received. It could be that some young MSM and TW recently diagnosed with HIV were triggered by reminders of their new status during the course of the intervention, which interfered with their EMA participation. Although EMA surveys were not directly implicated in the following postintervention qualitative interview, one participant mentioned that participation was difficult because they did not want to be reminded about their new HIV diagnosis every day with the digital care navigation. Young MSM and TW recently diagnosed with HIV are an especially important group to engage in HIV interventions due to heightened susceptibility to depression and subsequent poor linkage to care [28]. This study provides novel, preliminary evidence that tailoring interventions to be sensitive to diagnosis timing may be critical for intervention participation among those who are newly diagnosed.

Our findings add novelty to the prior EMA literature. The cumulative completion we observed was lower than one study of young MSM living with HIV [8]; however, this could be a consequence of having a longer follow-up period compared to traditional EMA studies that span 1 week to 60 days [4,7,8]. Moreover, instead of excluding participants who failed to comply with EMA surveys during a calibration phase, we identified nonstarters (ie, those who completed less than four of the first seven EMA surveys they received) and analyzed them as a separate group. A number of EMA feasibility studies exclude participants during the calibration phase [8] or require that participants possess mobile phones with unlimited text plans [34], which inflates completion and excludes important information about participants who may experience additional challenges to EMA completion. This limits the generalizability of results. Had we removed nonstarters, similar to how participants were removed during calibration or screening phases in other studies, EMA completion would have increased by 10%.

In addition to identifying and analyzing nonstarters, we examined EMA noncompletion of 1 week or more. This enabled us to explore hypothesized correlates of the days that transpired before a weeklong or more noncompletion of EMA surveys. This time-to-event outcome offered greater nuance than a binary (low versus high) EMA completion outcome and subsequently greater precision in detecting significant associations. In using survival analysis methods, we were also able to account for censoring from participant dropout or from participants who had not yet experienced a weeklong or more noncompletion by the end of study follow-up.

The strengths of this analysis should be considered with its limitations in mind. Because this was the first study of EMA among young MSM and TW living with HIV, our analyses were largely exploratory, and findings should be interpreted as hypothesis-generating. Variable selection was based on prior literature and hypotheses; however, there were no prior EMA studies focusing on MSM and TW living with HIV. Although we utilized novel outcome classifications in our analyses, our time-to-event outcome definition poses another limitation to our study. Choosing a time-to-event of 1 week or greater could have been too short or too long to delineate meaningful windows of EMA noncompletion. However, having 1 week or longer of EMA noncompletion was a clinically relevant time unit. Several components of the Health eNav intervention necessitated weekly check-ins. Thus, failing to complete EMA surveys for 1 week or longer could signify noncompletion with the larger digital navigation intervention. In addition, if participants experienced noncompletion of a week or more, then there would be no added value to administering daily EMA surveys in lieu of weekly CASIs. A final limitation is that some point estimates were imprecise, probably due to the small sample size. However, relative to other EMA studies among MSM that have studied up to 70 participants for 30- or 60-day periods [3,8,15,34], this study consists of a larger number of participants and uses survival analysis methods with greater precision to detect associations. Additionally, this EMA study is embedded within a larger parent study that enrolled young people living with HIV who might benefit from an intervention aiming to improve linkage, engagement, and retention in HIV care. As a result, this sample may exhibit selection bias skewed toward the hardest-to-reach young persons living with HIV.

This analysis has a number of strengths and limitations that should be considered in subsequent research. Future studies should incorporate more complex causal pathways to examine correlates of EMA completion. The outcome classifications used in this analysis (eg, identifying nonstarters and days to first weeklong or more EMA survey noncompletion) represent a few of the ways in which EMA feasibility can be assessed; future research should similarly explore other ways of mapping noncompletion patterns depending on the study design and population. There are a number of additional structural and individual barriers to EMA completion beyond the scope of this study that could be examined, particularly for other HIV risk communities. If future studies find similar barriers to EMA completion for young sexual and gender minority groups living with HIV, this will lend credence to creating interventions that best address the unique needs of these communities. Our findings suggest that EMA is feasible and acceptable for individuals not experiencing social inequity and structural barriers, and that HIV prevention technologies addressing the previously mentioned barriers (eg, housing instability, incarceration, competing needs, educational constraints, and HIV diagnosis recency) and leveraging similar methodology may prove effective for young MSM and TW living with HIV.

Relatedly, these findings suggest that EMAs may be especially sensitive to experiences of structural barriers. A potential implication of these findings may be the uptake of EMAs in public health and clinical settings to better detect and identify the onset of structural barriers and social inequity in relation to real-time monitoring of engagement and retention in care, especially among vulnerable populations. Future studies should assess the intervention potential of EMAs in combination with other interventions at the individual level and systems level. EMA as a source of real-time feedback may inform personalization of public health service and health care

utilization. Overall, these analyses lay the groundwork not only for future EMA studies among MSM and TW living with HIV, but also for characterizing EMA completion in other communities.

Acknowledgments

The authors would like to send their sincerest gratitude to the Health eNav participants and research team for making this research possible. This work was supported by the Health Resources and Services Administration (award number H97HA28895).

Conflicts of Interest

None declared.

References

1. Mackesy-Amiti ME, Boodram B. Feasibility of ecological momentary assessment to study mood and risk behavior among young people who inject drugs. Drug Alcohol Depend 2018 Dec 01;187:227-235. [doi: 10.1016/j.drugalcdep.2018.03.016] [Medline: 29684890]
2. Turner CM, Coffin P, Santos D, Huffaker S, Matheson T, Euren J, et al. Race/ethnicity, education, and age are associated with engagement in ecological momentary assessment text messaging among substance-using MSM in San Francisco. J Subst Abuse Treat 2017 Dec;75:43-48 [FREE Full text] [doi: 10.1016/j.jsat.2017.01.007] [Medline: 28237053]
3. Yang C, Linas B, Kirk G, Bollinger R, Chang L, Chander G, et al. Feasibility and acceptability of smartphone-based ecological momentary assessment of alcohol use among African American men who have sex with men in Baltimore. JMIR Mhealth Uhealth 2015;3(2):e67 [FREE Full text] [doi: 10.2196/mhealth.4344] [Medline: 26085078]
4. Rendina HJ, Ventuneac A, Mustanski B, Grov C, Parsons JT. Prospective measurement of daily health behaviors: modeling temporal patterns in missing data, sexual behavior, and substance use in an online daily diary study of gay and bisexual men. AIDS Behav 2016 Dec;20(8):1730-1743 [FREE Full text] [doi: 10.1007/s10461-016-1359-0] [Medline: 26992392]
5. Westergaard RP, Genz A, Panico K, Surkan PJ, Keruly J, Hutton HE, et al. Acceptability of a mobile health intervention to enhance HIV care coordination for patients with substance use disorders. Addict Sci Clin Pract 2017 Dec 26;12(1):11 [FREE Full text] [doi: 10.1186/s13722-017-0076-y] [Medline: 28441962]
6. Hensel DJ, Fortenberry JD, Harezlak J, Craig D. The feasibility of cell phone based electronic diaries for STI/HIV research. BMC Med Res Methodol 2012 Jun 12;12:75 [FREE Full text] [doi: 10.1186/1471-2288-12-75] [Medline: 22691189]
7. Moore RC, Kaufmann CN, Rooney AS, Moore DJ, Eyler LT, Granholm E, et al. Feasibility and acceptability of ecological momentary assessment of daily functioning among older adults with HIV. Am J Geriatr Psychiatry 2017 Aug;25(8):829-840 [FREE Full text] [doi: 10.1016/j.jagp.2016.11.019] [Medline: 28065496]
8. Cherenack EM, Wilson PA, Kreuzman AM, Price GN, Adolescent Medicine Trials Network for HIV/AIDS Interventions. The feasibility and acceptability of using technology-based daily diaries with HIV-infected young men who have sex with men: a comparison of internet and voice modalities. AIDS Behav 2016 Dec;20(8):1744-1753 [FREE Full text] [doi: 10.1007/s10461-016-1302-4] [Medline: 26837629]
9. Lenhart A, Purcell K, Smith A, Zickuhr K. Social media & mobile internet use among teens and young adults. Pew Internet & American Life Project 2010 Feb 03 [FREE Full text]
10. Kingdon MJ, Storholm ED, Halkitis PN, Jones DC, Moeller RW, Siconolfi D, et al. Targeting HIV prevention messaging to a new generation of gay, bisexual, and other young men who have sex with men. J Health Commun 2013;18(3):325-342. [doi: 10.1080/10810730.2012.727953] [Medline: 23320963]
11. Kubicek K, Carpineto J, McDavitt B, Weiss G, Iverson EF, Au C, et al. Integrating professional and folk models of HIV risk: YMSM's perceptions of high-risk sex. AIDS Educ Prev 2008 Jun;20(3):220-238 [FREE Full text] [doi: 10.1521/aeap.2008.20.3.220] [Medline: 18558819]
12. Kubicek K, Carpineto J, McDavitt B, Weiss G, Kipke MD. Use and perceptions of the internet for sexual information and partners: a study of young men who have sex with men. Arch Sex Behav 2011 Aug;40(4):803-816 [FREE Full text] [doi: 10.1007/s10508-010-9666-4] [Medline: 20809373]
13. Holloway IW, Winder TJ, Lea CH, Tan D, Boyd D, Novak D. Technology use and preferences for mobile phone-based HIV prevention and treatment among black young men who have sex with men: exploratory research. JMIR Mhealth Uhealth 2017 Dec 13;5(4):e46 [FREE Full text] [doi: 10.2196/mhealth.6436] [Medline: 28408360]
14. Baral SD, Poteat T, Strömdahl S, Wirtz AL, Guadamuz TE, Beyrer C. Worldwide burden of HIV in transgender women: a systematic review and meta-analysis. Lancet Infect Dis 2013 Mar;13(3):214-222. [doi: 10.1016/S1473-3099(12)70315-8] [Medline: 23260128]
15. Wray TB, Kahler CW, Monti PM. Using ecological momentary assessment (EMA) to study sex events among very high-risk men who have sex with men (MSM). AIDS Behav 2016 Jan 8;20(10):2231-2242. [doi: 10.1007/s10461-015-1272-y] [Medline: 26746212]

16. Poundstone KE, Strathdee SA, Celentano DD. The social epidemiology of human immunodeficiency virus/acquired immunodeficiency syndrome. Epidemiol Rev 2004;26(1):22-35. [doi: 10.1093/epirev/mxh005] [Medline: 15234945]
17. Link BG, Phelan J. Social conditions as fundamental causes of disease. J Health Soc Behav 1995:80-94. [Medline: 7560851]
18. Riley ED, Gandhi M, Hare C, Cohen J, Hwang S. Poverty, unstable housing, and HIV infection among women living in the United States. Curr HIV/AIDS Rep 2007 Dec;4(4):181-186. [Medline: 18366949]
19. Singh GK, Daus GP, Allender M, Ramey CT, Martin EK, Perry C, et al. Social determinants of health in the United States: addressing major health inequality trends for the nation, 1935-2016. Int J MCH AIDS 2017;6(2):139-164 [FREE Full text] [doi: 10.21106/ijma.236] [Medline: 29367890]
20. Latkin CA, German D, Vlahov D, Galea S. Neighborhoods and HIV: a social ecological approach to prevention and care. Am Psychol 2013;68(4):210-224 [FREE Full text] [doi: 10.1037/a0032704] [Medline: 23688089]
21. Dray-Spira R, Lert F, Marimoutou C, Bouhnik A, Obadia Y. Socio-economic conditions, health status and employment among persons living with HIV/AIDS in France in 2001. AIDS Care 2003 Dec;15(6):739-748. [doi: 10.1080/09540120310001618595] [Medline: 14617496]
22. Liu Y, Canada K, Shi K, Corrigan P. HIV-related stigma acting as predictors of unemployment of people living with HIV/AIDS. AIDS Care 2012;24(1):129-135. [doi: 10.1080/09540121.2011.596512] [Medline: 21777074]
23. Kalichman SC, Pellowski J, Kalichman MO, Cherry C, Detorio M, Caliendo AM, et al. Food insufficiency and medication adherence among people living with HIV/AIDS in urban and peri-urban settings. Prev Sci 2011 Sep;12(3):324-332. [doi: 10.1007/s11121-011-0222-9] [Medline: 21607719]
24. Pellowski JA, Kalichman SC, Matthews KA, Adler N. A pandemic of the poor: social disadvantage and the US HIV epidemic. Am Psychol 2013;68(4):197-209 [FREE Full text] [doi: 10.1037/a0032694] [Medline: 23688088]
25. Poteat T, Reisner SL, Radix A. HIV epidemics among transgender women. Curr Opin HIV AIDS 2014 Mar;9(2):168-173. [doi: 10.1097/COH.0000000000000030] [Medline: 24322537]
26. Silva-Santisteban A, Raymond HF, Salazar X, Villayzan J, Leon S, McFarland W, et al. Understanding the HIV/AIDS epidemic in transgender women of Lima, Peru: results from a sero-epidemiologic study using respondent driven sampling. AIDS Behav 2012 May;16(4):872-881. [doi: 10.1007/s10461-011-0053-5] [Medline: 21983694]
27. Bekele T, Globerman J, Watson J, Hwang SW, Hambly K, Koornstra J, Positive Spaces Healthy Places Study Team. Elevated mortality and associated social determinants of health in a community-based sample of people living with HIV in Ontario, Canada: findings from the Positive Spaces, Healthy Places (PSHP) study. AIDS Behav 2018 Jul;22(7):2214-2223. [doi: 10.1007/s10461-018-2040-6] [Medline: 29557541]
28. Bhatia R, Hartman C, Kallen MA, Graham J, Giordano TP. Persons newly diagnosed with HIV infection are at high risk for depression and poor linkage to care: results from the Steps Study. AIDS Behav 2011 Aug;15(6):1161-1170 [FREE Full text] [doi: 10.1007/s10461-010-9778-9] [Medline: 20711651]
29. Corbin J, Strauss A. Strategies for qualitative data analysis. In: Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory. Thousand Oaks, CA: Sage Publications; 2008:65-86.
30. mSurvey. URL: https://msurvey.co/ [accessed 2019-04-18] [WebCite Cache ID 77jmvbr5j]
31. Barros AJ, Hirakata VN. Alternatives for logistic regression in cross-sectional studies: an empirical comparison of models that directly estimate the prevalence ratio. BMC Med Res Methodol 2003 Oct 20;3:21 [FREE Full text] [doi: 10.1186/1471-2288-3-21] [Medline: 14567763]
32. Elo S, Kyngäs H. The qualitative content analysis process. J Adv Nurs 2008 Apr;62(1):107-115. [doi: 10.1111/j.1365-2648.2007.04569.x] [Medline: 18352969]
33. Therneau T, Grambsch P. Modeling Survival Data: Extending the Cox Model. New York: Springer-Verlag; 2000.
34. Smiley S, Elmasry H, Webb Hooper M, Niaura R, Hamilton A, Milburn N. Feasibility of ecological momentary assessment of daily sexting and substance use among young adult African American gay and bisexual men: a pilot study. JMIR Res Protoc 2017 Feb 02;6(2):e9 [FREE Full text] [doi: 10.2196/resprot.6520] [Medline: 28153816]

Abbreviations

aHR: adjusted hazard ratio
CASI: computer-assisted self-interviewing
CI: confidence interval
EMA: ecological momentary assessment
MSM: men who have sex with men
PR: prevalence ratio
TW: trans women

Edited by G Eysenbach; submitted 26.12.18; peer-reviewed by L Nguyen, B Henwood, A Rudolph; comments to author 28.01.19; revised version received 25.03.19; accepted 29.03.19; published 08.05.19.

Please cite as:
Turner CM, Arayasirikul S, Trujillo D, Lê V, Wilson EC
Social Inequity and Structural Barriers to Completion of Ecological Momentary Assessments for Young Men Who Have Sex With Men and Trans Women Living With HIV in San Francisco
JMIR Mhealth Uhealth 2019;7(5):e13241
URL: https://mhealth.jmir.org/2019/5/e13241/
doi:10.2196/13241
PMID:31066714

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org