ORIGINAL RESEARCH

A counselor in your pocket: feasibility of mobile health tailored messages to support HIV medication adherence

Paul F Cook[1]
Jane M Carrington[2]
Sarah J Schmiege[1]
Whitney Starr[3]
Blaine Reeder[1]

[1]University of Colorado College of Nursing, Aurora, CO, USA; [2]University of Arizona College of Nursing, Tucson, AZ, USA; [3]University of Colorado School of Medicine, Aurora, CO, USA

→ Video abstract

Point your SmartPhone at the code above. If you have a QR code reader the video abstract will appear. Or use:

http://youtu.be/7tbYkroHe-w

Correspondence: Paul F Cook
University of Colorado College of Nursing, Campus Box C288-04, Aurora, CO 80045, USA
Tel +1 303 724 8537
Fax +1 303 724 8560
Email paul.cook@ucdenver.edu

Purpose: Medication adherence is a major challenge in HIV treatment. New mobile technologies such as smartphones facilitate the delivery of brief tailored messages to promote adherence. However, the best approach for tailoring messages is unknown. Persons living with HIV (PLWH) might be more receptive to some messages than others based on their current psychological state.

Methods: We recruited 37 PLWH from a parent study of motivational states and adherence. Participants completed smartphone-based surveys at a random time every day for 2 weeks, then immediately received intervention or control tailored messages, depending on random assignment. After 2 weeks in the initial condition, participants received the other condition in a crossover design. Intervention messages were tailored to match PLWH's current psychological state based on five variables – control beliefs, mood, stress, coping, and social support. Control messages were tailored to create a mismatch between message framing and participants' current psychological state. We evaluated intervention feasibility based on acceptance, ease of use, and usefulness measures. We also used pilot randomized controlled trial methods to test the intervention's effect on adherence, which was measured using electronic caps that recorded pill-bottle openings.

Results: Acceptance was high based on 76% enrollment and 85% satisfaction. Participants found the hardware and software easy to use. However, attrition was high at 59%, and usefulness ratings were slightly lower. The most common complaint was boredom. Unexpectedly, there was no difference between mismatched and matched messages' effects, but each group showed a 10%–15% improvement in adherence after crossing to the opposite study condition.

Conclusion: Although smartphone-based tailored messaging was feasible and participants had clinically meaningful improvements in adherence, the mechanisms of change require further study. Possible explanations might include novelty effects, increased receptiveness to new information after habituation, or pseudotailoring, three ways in which attentional processes can affect behavior.

Keywords: adherence, communication, feasibility, HIV, technology

Introduction

Nonadherence to prescribed medications is a significant barrier to treatment of many chronic diseases, and nowhere is this more important than in antiretroviral therapy (ART) for human immunodeficiency virus (HIV) infection. Adherence is most commonly defined in terms of the percentage of prescribed doses of medication that a person actually takes.[1] Persons living with HIV (PLWH) face particularly great challenges because of the level of adherence required. Adherence of at least 80% is acceptable for most chronic diseases such as diabetes and hypertension.[2] However,

PLWH are asked to maintain adherence levels of 95% or higher because better clinical outcomes, including fewer hospitalizations, fewer opportunistic infections, and higher virologic suppression rates are associated with these higher adherence levels.[3]

Varied interventions have been tested to improve PLWH's adherence, but the success of these interventions is highly variable.[4] Tailoring messages to match patients' demographic and psychological characteristics is one strategy that may increase the efficacy of adherence interventions: Noar et al[5] found in a meta-analysis of tailored print messages that 1) tailoring messages based on a theory is more effective than tailoring on surface message features alone; 2) tailoring on four to six theory-based characteristics at once has better results than tailoring on either more than six or less than four variables; and 3) specific theory-based strategies are effective, including tailoring on variables such as perceived control (self-efficacy), social support, type of coping strategy used, and level of motivation for change. More frequent or ongoing messaging has also been recommended as a way to enhance the efficacy of behavior change interventions for PLWH.[6]

Mobile technology has grown in popularity and accessibility, especially in the minority and lower-income groups that are most heavily affected by HIV,[7] and presents exciting new opportunities for intervention. Technology-based adherence interventions can have a broad reach and are potentially more cost-effective than in-person counseling methods[8] despite having slightly weaker overall effect sizes.[6] To date, most technology-based interventions to promote healthier behavior have offered either nontailored reminders,[9] or text messages that are personalized by a human counselor using clinical judgment.[10] However, technology also makes it relatively easy for investigators to automatically tailor messages via an algorithm, using variables such as those identified by Noar et al[5] to optimize impact. Two recent studies[11,12] have moved in this direction: in both studies, participants selected a relevant barrier to adherence and then received a message that was randomly selected from a pool of appropriate content based on the selected barrier and other baseline participant characteristics. One of these intervention studies showed positive effects on adherence,[11] while the other found no relationship between tailored message satisfaction and adherence.[12]

Another appealing feature of mobile technology is the ability to tailor messages based on an individual's psychological state at a particular point in time, rather than messages tailored to context-independent baseline variables as in past research.[11,12] Momentary states – PLWH's point-in-time thoughts, emotions, and motivation in the context of their everyday lives – have been identified as predictors of their immediate behavior, and have different relationships to behavior than the same variables measured retrospectively.[13,14] For example, mood, stress, and motivation are momentary state variables that fluctuate from one point in time to another. With mobile technology it might be possible to use momentary states as a basis for tailored adherence messages. This type of "personalized counseling" approach – delivering different messages to different people at different times – is congruent with the current movement toward personalized medicine in health care.[15]

Theoretical model and tailored message development

For the current study, we developed a novel intervention in which PLWH completed daily surveys about their momentary states on a smartphone, and then immediately received a message tailored based on their survey responses. First, participants identified one of ten barriers to adherence that they considered most relevant to them. Participants were asked to identify a potential barrier regardless of whether they were currently adherent, based on the idea that even currently adherent patients can become nonadherent if barriers increase. The list of barriers was generated with HIV primary care experts, and ten base messages addressing those barriers were created. These included statements such as, "Talk to your providers about any side effects", and, "Keep taking medication even when using alcohol or other drugs". The note to Table 1 includes a complete list of barriers.

The ten base messages were designed to be as tailoring-neutral as possible: that is, they were short, factual statements presenting accurate information about ART adherence without framing the message on any theory-relevant dimensions. Using a computer algorithm, the message was then simultaneously tailored along five separate dimensions. Specific tailoring variables were selected based on a model of momentary states and behavior that was developed in our prior research[16] and that was tested in a project serving as a parent study for the current investigation (Cook, unpublished data, 2012). Within each of the five tailoring domains, two versions of a statement linked to that domain were created as part of each of the ten messages, based on relevant theories of health behavior (Table 1).

The theoretical model underlying the intervention suggests that adherence behavior is based primarily on a participant's motivational state; the tailored messages were therefore intended to work at the level of motivation by

Table 1 Message tailoring dimensions

Construct	What is assessed to determine tailoring	Difference between tailored message versions	Specific matching prediction and relevant theory of health behavior	Average construct validity rating of ten tailored messages
Control beliefs	High vs low sense of control over problems	Emphasis on reasons for change ("think about") vs actions to take ("do")	Problem-solving focus → ↑ adherence for high control beliefs; reasons for change better for low control beliefs (Prochaska and DiClemente's transtheoretical model[33])	M =3.58 (SD =0.19)
Mood	Level of positive emotional arousal (high vs low)	Focus on feared outcomes (permanent vs short-term), and use of affect-inducing words	Feared consequences → ↑ adherence for high (positive) mood; lower-fear presentation better for low (negative) mood (Leventhal's illness perception model[34])	M =3.86 (SD =0.31)
Situational stress	Level of current stress (high vs low) from multiple sources	Deep focus (content) vs surface focus (vivid images, expert quotes, personal relevance)	For high stress, surface focus → ↑ adherence; content focus better for low stress when people process information more deeply (Lazarus and Folkman's coping theory[35])	M =3.69 (SD =0.29)
Coping	Approach (domain of gains) vs avoidance (domain of losses)	Emphasis on benefits of behavior change vs costs of not changing behavior	Gain frame → ↑ adherence for high (active) coping; loss frame better for low (passive) coping (Kahneman and Tversky's prospect theory[36])	M =3.91 (SD =0.12)
Social support	Perceived current social support (+) and stigma (−)	Focus on self vs others in reasons for behavior change	Other-focus → ↑ adherence if support is seen as high; self-focus better if support is low (Bronfenbrenner's socio-ecological model[37])	M =3.72 (SD =0.26)

Notes: Construct validity ratings were made by 15 independent experts in health behavior change research, using a scale from 1= "not relevant" to 4= "very relevant and succinct". Ratings were averaged across experts, and across ten different base messages corresponding to different barriers to adherence. Experts also gave a rating of 3.84 points for the messages' overall integration and consistency, and 3.89 points for their relevance to ART adherence and clinical utility, on the same 4-point scale. The ten barriers identified through consultation with HIV primary care experts in a previous stage of development were: 1) low perceived importance of treatment, 2) lack of confidence in treatment, 3) adverse effects of treatment, 4) feeling healthy, 5) feeling sick, 6) lack of financial resources, 7) treatment complexity, 8) stigma, 9) alcohol or other drug use, and 10) forgetting.
Abbreviations: vs, versus; M, mean; SD, standard deviation; HIV, human immunodeficiency virus.

synchronizing message wording with momentary psychological states. Messages were not designed to modify PLWH's momentary states directly, but rather were presented in a way that was hypothesized to make messages more acceptable to people in specific states. For instance, when PLWH reported high levels of perceived control, they were predicted to attend more to internal motivators, eg, "Taking care of your health is something very important that you can do for yourself". By contrast, PLWH reporting lower levels of perceived control were hypothesized to respond better to external reasons for behavior change, eg, "People are counting on you to take the best care of your health that you can".

To validate final messages, we used a five-step content validity process described by Lynn:[17] 1) relevant content domains were identified by the first author (ie, control beliefs, mood, stress, coping, and social support); 2) a group of 15 experts (five or more experts are recommended) generated possible message wordings for all five tailoring dimensions; 3) individual statements were combined into usable forms by the principal investigator (PI); 4) seven experts (three or more

experts are recommended) judged the resulting statements on each of the original domains; and 5) the same experts judged the overall integration and consistency of the complete messages, their relevance to ART adherence, and their clinical utility. Results of this content validation process are also shown in Table 1. All individual items were rated above 3.5 on Lynn's 4-point scale; Lynn suggested that items with scores of 3 or higher have acceptable content validity.[17]

Design and study aims

Our primary aim in the current study was to test the feasibility of tailored messages matched to momentary states as a way to improve PLWH's ART adherence. Feasibility testing is an important step in the development of new technologies. Any new technology can be characterized as either emerging, promising, or effective in single or multiple contexts.[18] Based on the literature review presented above, although reminder messages are likely to be effective to at least a moderate degree in improving adherence, technology for systematically tailoring messages is still at the emerging stage. Necessary

further development steps include discovery, feasibility testing with end-users, and efficacy testing to determine the potential benefits of a technology-based intervention. During feasibility testing, researchers measure participants' acceptance of the new technology, their perceptions of its ease of use, and their willingness to adopt this type of technology if it were widely available. Feasibility testing is related to the treatment delivery and receipt steps of intervention fidelity assessment.[19]

In this study, we measured feasibility based on acceptance of the intervention including recruitment and attrition, and by participant self-report on perceived ease of use of the daily surveys, and perceived usefulness of the new technology. Ease of use and usefulness are orthogonal dimensions of people's experiences with technology, and both contribute to a technology's ultimate adoption in practice.[20] Ease of use is defined as whether a technology is simple and convenient, while perceived usefulness is defined as whether a technology has potential benefits.[21]

As a secondary aim, we also conducted a pilot randomized controlled trial (RCT) using an experimental crossover design to test the tailored messages' effects on adherence. Although pilot studies do not necessarily produce reliable estimates of an intervention's effects, they can be useful in identifying its more or less useful aspects and can guide revisions to make the intervention more efficacious in future full-scale RCTs.[22]

Methods
Participants
This study gained ethical approval from the Colorado Multiple Institutional Review Board and all participants provided written informed consent. Participants were 37 PLWH recruited from a larger observational parent study of momentary state predictors of ART adherence behavior, who were receiving HIV care at the Infectious Disease Group Practice at the University of Colorado Hospital, Aurora CO, USA. The study's inclusion and exclusion criteria were 1) documented HIV infection and current ART treatment based on medical records; 2) ability to speak, read, and write English; 3) age over 18 years and less than 81 years old; and 4) no current substance abuse, cognitive impairment, psychiatric or medical disorder, or other condition (as evaluated by the participant's HIV primary care clinician) that would substantially interfere with study participation. Participants were recruited during the second half of the parent study from April 2013 through May 2014, and the current, secondary study ended as planned. Not all of the 87 parent-study participants were recruited due to the more limited scope and

shorter recruitment duration of the current study; however, during the time the current study was recruiting, all PLWH who completed the parent study were invited to participate. There were no additional selection criteria beyond those originally used in the parent study. Study flow is shown in the CONSORT diagram in Figure 1.

Participants' average age was M =42.6 years (standard deviation [SD] =7.98 years), and 22% (8/37 participants) were women. Of the 37 participants, 51% were minority group members, including six Latino/Latina participants, eight African–Americans, three Native Americans, and two other non-White participants. The majority of participants (22/37 participants, or 59%) were gay, lesbian, bisexual, or transgender (GLBT), with the remaining 15 participants (41%) identifying as heterosexual. Participants had M =13.0 years of education on average (SD =2.16 years), equivalent to some college, although a few participants had either less than a high school education or a graduate degree. Participants' average level of ART adherence during the parent study was 80.0% (SD =21.8%), and just over half of the participants (19/37 participants, or 51%) had an overall adherence level during the parent study that was below the minimum recommended level of 95%.

Sample representativeness
Compared to national epidemiology, our current sample included a higher percentage of White men who have sex with men and better-educated patients but the level of diversity was high compared to the Western US region. Because of the potential for sampling bias when participants are recruited from a larger parent study, we tested for any demographic differences between PLWH who participated in this secondary tailored messaging pilot study and PLWH who did not participate (either because of refusal or because they ended the parent study before this secondary pilot study began). We found no differences between participants and nonparticipants in terms of age, sex, race/ethnicity, sexual orientation, or years of education (all P-values >0.13), suggesting an absence of sampling bias.

Baseline differences between groups
Table 2 shows participant demographics by group. As is common in small studies, there were some failures of randomization. Specifically, participants in the AB order group (matched messages first) were more likely to be women and less likely to be GLBT than participants in the BA order group (mismatched messages first). These two variables are related, because most women in the study were heterosexual, while the largest subgroup of GLBT
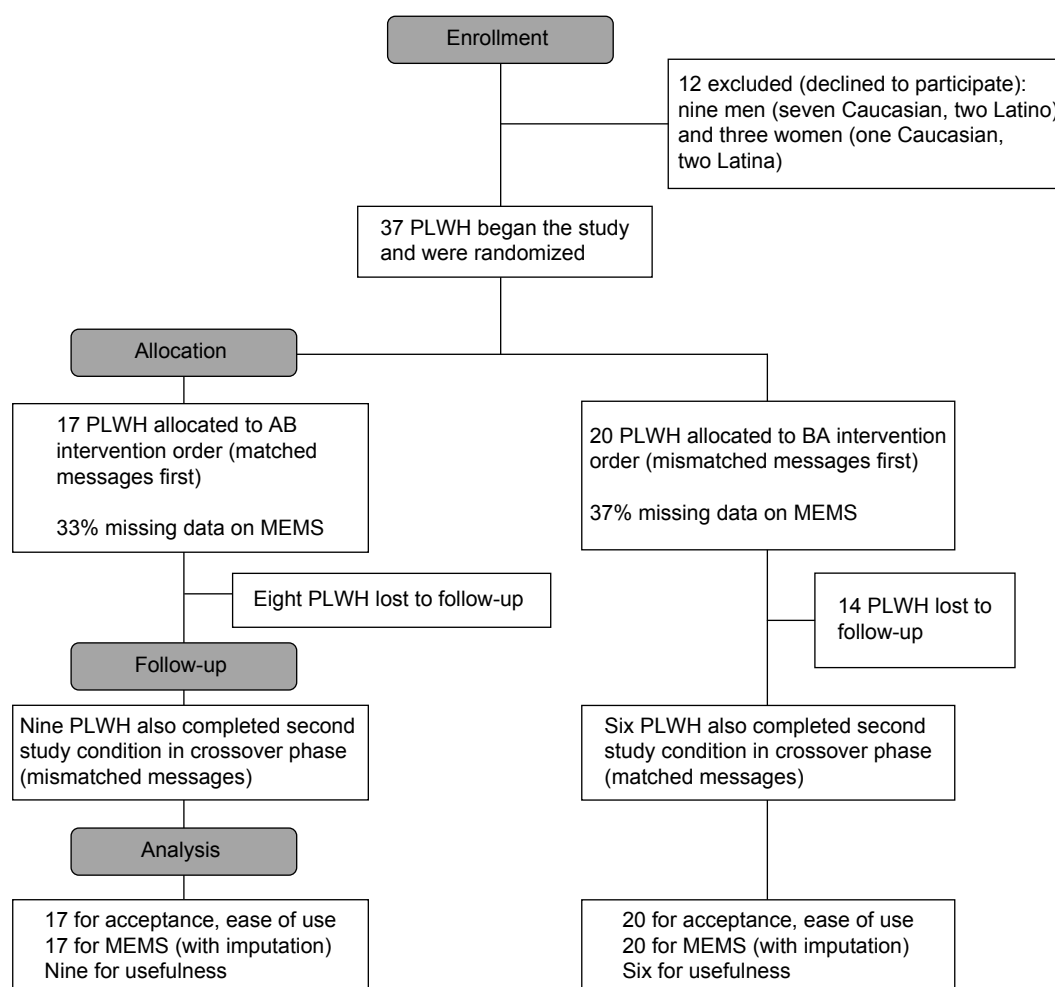
Figure 1 CONSORT diagram showing recruitment, enrollment, and retention.
Abbreviations: PLWH, persons living with HIV; HIV, human immunodeficiency virus; MEMS, medication event monitoring system; CONSORT, Consolidated Standards of Reporting Trials.

participants consisted of patients who were men who have sex with men. Sex was therefore included as a covariate in subsequent analyses, but sexual orientation was not included, because of multicollinearity between these two variables. There were no baseline differences between groups on other important covariates like age, race/ethnicity, or years of education. There was a nonsignificant but still relatively large difference in baseline ART adherence, 72% in the AB order group versus 87% in the BA order group. There was also a nonsignificant but still potentially meaningful

Table 2 Participant demographics by order of study conditions

Variable	AB condition order (matched first)	BA condition order (mismatched first)	Significant difference? Yes/no
Age, years	M =43.7 (SD =5.87)	M =41.8 (SD =9.23)	No: $t(30)$ =0.71, P=0.48, d=0.25
Sex	35% women (6/17)	10% women (2/20)	Yes: χ^2=4.59, P=0.03, ϕ=0.37
Race/ethnicity	47% non-White (8/17)	55% non-White (11/20)	No: χ^2=0.003, P=0.06, ϕ=0.01
	3 Latino/Latina	3 Latino/Latina	
	2 African–American	6 African–American	
	1 Native American	2 Native American	
	2 other ethnicity		
Sexual orientation	29% GLBT (5/17)	85% GLBT (17/20)	Yes: χ^2=10.5, P=0.001, ϕ=0.56
Years of education	M =12.9 (SD =1.45)	M =13.0 (SD =2.54)	No: $t(27)$ =0.11, P=0.92, d=0.04
Baseline MEMS adherence (%)	M =72% (SD =22%)	M =87% (SD =20%)	No: $t(25)$ =1.85, P=0.08, d=0.64

Abbreviations: MEMS, medication event monitoring system; SD, standard deviation; GLBT, gay, lesbian, bisexual, or transgender; M, mean.

difference in the number of study observations completed in each of the two conditions, as reported below in the findings about acceptance. Despite the lack of statistical significance, we included these potential confounds as covariates in our outcome analysis.

Procedures

Baseline data collection

A research assistant met individually with each participant in a consultation room at the clinic where the participant regularly received HIV care. Each participant's most recent CD4 and viral load laboratory test results, which were usually collected within the past 3 months as per standard of care, were extracted from clinic charts with the participant's authorization. As part of the parent study, each participant had already received a smartphone, had completed baseline questionnaires, and had provided demographic information. During their participation in the parent study, PLWH had also completed 3 months of daily smartphone-based surveys about their control beliefs, mood, stress, coping, social support, and motivation, and had stored their ART medication in medication event monitoring system (MEMS) pill bottles that electronically monitored bottle openings as a measure of adherence. Surveys were completed once daily, at random times determined by the software. The parent study involved assessment only; no tailored messages were delivered.

Randomization

Group assignment was generated using simple randomization by the research assistant at the time the participant completed the parent study. For the first 2 weeks, participants were randomized to receive either matched messages (AB message order) or mismatched messages (BA message order). Participants then crossed over to the other study condition (mismatched or matched) to complete surveys and to receive the other type of messages for a final 2 weeks before a concluding in-person study visit. As in the parent study, participants were paid US $25 for the in-person visit. The investigators also paid for another month of data service on the participant's smartphone so that he or she could continue to complete surveys and receive tailored messages. Participants were blind to their initial group assignment and to which intervention condition was considered the active one, although they were informed as part of the study's consent process that there would be 2 weeks of messages matched to their survey responses and 2 weeks that were mismatched. The research assistant who administered questionnaires also

generated the allocation sequence, resulting in a single-blind study.

Message delivery

During the current secondary intervention study, participants continued answering surveys on their smartphones for another 4 weeks, again once daily at times randomly determined by the software. The same survey questions were used as in the parent study. However, at this point, tailored messages were generated by an algorithm in the survey software based on the participant's responses, and were delivered immediately after completion of the daily survey. The survey was completed on a Samsung phone (Samsung Electronics Co, Ltd, Suwon, South Korea) with an Android operating system (Google, Inc, Mountain View, CA, USA) and pre-installed Apptive (Austin, TX, USA) software to deliver the daily tailored messages. The intervention message appeared as an alert on the device, which also retained all other telephone and data capabilities. Each message contained a URL to an online survey that opened in the smartphone's built-in web browser. Web-based surveys were completed online using SurveyMonkey (Palo Alto, CA, USA), and survey questions asked about the participant's current medication adherence and motivation to take ART medication; in addition, questions on five theoretically relevant aspects of participants' daily experiences were used to tailor intervention messages, as described in Table 1. The researchers pre-programmed tailored messages in SurveyMonkey using the software's decisional logic capabilities to either match or mismatch message text to participants' survey responses, depending on which study condition the participant was assigned to at that point in time.

Tailored messaging intervention

Each time an intervention was delivered, the participant was first asked to select one of ten possible barriers to adherence. The participant then received a tailored text message addressing the barrier he or she had selected. A single base message for the selected barrier was tailored based on a theoretical model of momentary state influences on behavior, as described in the "Introduction" section. A cut-off score was used for each of the momentary state subscales to classify participants as either "high" or "low" with respect to each of the five variables. Separate components of the message were then tailored based on each of the five variables. Figure 2 shows sample messages that are matched and mismatched based on a particular set of momentary states, and identifies the message components that were tailored based on each of the five measured variables. With two versions for each of

Patient description: Imagine that a participant has decided to stop taking medication, and her survey results show low control beliefs, positive mood, high stress, active coping, and high social support. The same participant would receive the following message in a "matched" week:

> **MATCHED**
>
> **IMPORTANT MESSAGE: A vacation from meds is no vacation for you!** *It's not easy to take medication every day.* TAKING ALL OF YOUR HIV MEDICATION <u>is essential to prevent getting sicker.</u> **People are counting on you to take the best care of your health that you can.** *Think about what makes it hard for you to take medication right now, and tell your health care providers.*

The same participant, with the same set of momentary states, would receive the following message instead in a "mismatched" messaging (control group) week:

> **MISMATCHED**
>
> **HIV medications work best when taken at least 95% of the time.** *You can take control of your health.* SKIPPING DOSES OR TAKING A BREAK FROM HIV MEDICATION <u>can make your HIV worse.</u> **Taking care of your health is something very important that you can do for yourself.** *Keep taking your medication! Don't stop or skip without talking to your health care providers first.*

This message is predicted to be less acceptable to participants with the same set of momentary states, who will be more likely to discount the information. The second message might still improve adherence, but is expected to have a weaker effect than the theory-matched message.

Key to tailored message components:
Bold = tailored for stress level (high stress = more dramatic or emotional, low stress = more factual)
Italics = tailored for control beliefs (high control = action focus, low control = support focus)
SMALL CAPS = tailored for coping (high coping = gain-focused, low coping = focus on avoiding loss)
<u>Underline</u> = tailored for mood (four options, nested within coping messages: positive mood = more fear-arousing message framing, negative mood = less fear-arousing message framing)
Bold italics = tailored for social support (high support = other-focused, low support = self-focused)

Figure 2 Example of tailored messaging intervention.
Abbreviation: meds, medications.

five momentary state variables for each of ten barriers, the tailoring intervention had a total of 320 possible message variations ($2^5 \times 10$) in each study condition.

Mismatched message control group

During the 2 weeks when they were assigned to the mismatched message control condition (either before or after the crossover, depending on random assignment), participants again completed daily surveys about their momentary states and barriers to adherence, but this time they received a message that was systematically tailored in the opposite way

from what we theoretically predicted would most facilitate participants' use of the message. Thus, we expected any given message wording to facilitate adherence under some psychological conditions (when it was a matched message) but not others (when it was mismatched). Because matched and mismatched messages were identical in length, delivery method, amount of tailoring, and factual content, we expected the mismatched-message condition to provide a relatively strong control for "pseudotailoring" effects that have been documented in the literature.[23] Pseudotailoring is a phenomenon in which participants respond better to a message if they believe it has been generated especially for them. Such messages may seem more salient to the participant, but do not contain "active ingredients" based on relevant theory. Mismatched messages are equivalent to matched messages in terms of their potential for pseudotailoring effects, but because they are tailored in exactly the opposite way from what theory predicts should be effective, they should have weaker effects on behavior. In the current study, we compared messages that were either completely matched or completely mismatched on five dimensions simultaneously, which did not allow us to identify the effects of individual message components, but did provide the strongest comparison possible.

Measures
Primary outcomes: feasibility of new technology
Because the primary aim of this study was to examine feasibility of the new smartphone-based tailored messaging intervention, the primary outcome measures focused on acceptance, ease of use, and perceived usefulness, which are constructs of the technology acceptance model.[21] At the end of the parent study, participants completed a self-report technology acceptance questionnaire about acceptance of the daily smartphone-based survey. By that point in time, they had been using the technology to complete daily surveys for 12 weeks, but had not received any tailored messages. Acceptance was measured with a single question: "What number best represents your attitude toward the smartphone survey tool?". Answers were measured on a 0–100 scale with reference points at 0= completely unacceptable, 50= neutral, and 100= completely acceptable. Our analysis of acceptance also considered study enrollment and attrition data as indirect behavioral evidence of acceptance.

To measure ease of use, the technology acceptance questionnaire also included a series of items developed by a nurse informaticist, (Dr. Jane Carrington) on concepts including overall ease of use, consistency, efficiency, memorability

of the items, level of language used, perceived cognitive load, format in which information was presented, errors, and overall performance of the device. These items were subjected to exploratory factor analysis using data from the parent study, and the initial pool of 24 items was reduced to a 12-item ease of use scale (Cronbach's alpha =0.96) representing general satisfaction with the technology, the survey questions, the smartphone hardware, and the tailored messages For example, one question asked, "Did using the survey feel natural to you?". Response choices ranged from "exceptionally unnatural" to "exceptionally natural".

Finally, participants completed a tailored messaging survey at the end of the intervention phase, which asked them more specific usefulness questions about the tailored-messaging intervention and their willingness to use this technology for adherence support. Sample items included, "If you had it to do over again would you still enroll in this study?" and, "Would you recommend a similar program to a friend who was taking the same medication?". One question about personalization of the survey items was dropped because it did not load with the others; Cronbach's alpha was 0.72 for the resulting seven-question scale. Two items about participants' behavioral intentions were also examined individually. Finally, three open-ended questions on this survey asked participants if any messages seemed wrong, offensive, or unhelpful, what changes to the messages they would suggest, and what other feedback participants could offer.

Manipulation check on the tailored messaging intervention

A separate question on the tailored messaging survey asked whether participants noticed any differences between the intervention and control messages. We believed that our tailoring strategy was relatively subtle and that participants would be unable to notice a major difference, which is important because participants otherwise might not have found the matched and mismatched messages to be equally credible or because noticeable differences might have led to pseudotailoring effects. Three additional questions similarly asked whether the participant noticed any difference in how personalized the messages were, in how helpful the messages were, or in how much the participant liked the messages.

Secondary outcome: effect of the intervention on adherence

For our secondary aim, we measured adherence using MEMS caps, electronic devices that record actual pill bottle openings in real time.[1] MEMS are regarded as a relatively objective measure, have low reactivity after a 6-week initial run-in

(completed during the parent study), and capture real-time data about adherence.[24]

Data analysis

For the first aim, we examined descriptive statistics and visualized data on the feasibility domains of acceptance, ease of use, and perceived usefulness. Data were relatively complete for the acceptance and ease-of-use self-report measures, and complete for the behavioral acceptance measures of enrollment and attrition, but the usefulness measure was administered only at the end of the crossover condition, resulting in a high proportion of missing data due to attrition. Descriptive analyses for Aim 1 used available observations only with no adjustment for missing data.

For the second aim, we first examined descriptive results on the manipulation check items, with the goal of testing whether participants found the matched and mismatched messages equally credible. We next tested for any between-group differences in baseline demographic characteristics that might indicate failures of randomization. Finally, we analyzed MEMS data across groups and over time using a 2×2 repeated-measures analysis of variance (ANOVA). We expected to find an interaction between initial group assignment (matched versus mismatched) and time showing that matched messages improved adherence more than mismatched messages did within each experimental group. We did not expect a main effect of condition, because all participants eventually received both interventions. The analysis controlled for participants' baseline adherence, and for the number of days of tailored messages actually received in case there were dose-response effects. No power analysis was completed for the current preliminary study, because effect size estimation rather than statistical significance testing was the primary goal; sample sizes of 30 or greater are generally considered adequate for this type of pilot study.[25] The MEMS measure had some missing data, 33% of observations in the AB group and 37% of observations in the BA group. While this amount of missing data is slightly outside the 30% level at which missing data can be imputed without bias,[26] MEMS observations appeared to be missing at random with respect to participant demographics or baseline adherence. Therefore, missing data points were handled using multiple imputation prior to analysis.

Results
Aim 1: feasibility of the tailored messaging intervention
Acceptance
The recruitment process suggested high acceptance based on the 76% enrollment rate among PLWH who were offered the

chance to participate (37/49 patients; Figure 1). This high level of acceptance must be considered in context, because participants had already been involved in a previous study and were receiving a month of smartphone data service as an incentive for participation. However, participants also rated the intervention very acceptable based on a 0–100 visual analogue scale (mean (M) =85%; SD =17.9%) with no ratings below 50%. Table 3 shows self-reported acceptance results, together with scores on the ease of use and perceived usefulness measures. In exploratory correlations with participant demographics, we found no relationship between acceptance and either participants' sex or minority race/ethnicity (all P-values >0.33), but there was a significant negative relationship between the overall acceptance rating and participants' age (r=0.66; P=0.02), meaning that older adults found the tailored messaging intervention less acceptable than younger adults did.

We examined attrition from the study (also shown in Figure 1), including the number of tailored messages actually received, as an additional behavioral measure of acceptance. All participants received at least one tailored text message. The average number of messages received by participants was higher during the 14 days they received matched messages (M=10.7 messages; SD =12.7 messages, versus M=6.0 messages; SD =8.6 messages) during the 14 days they spent in the mismatched study condition. Although this difference was not statistically significant (P=0.48), it might suggest better acceptance of the matched messages than of the mismatched messages. Additionally, 22/37 participants failed to complete the second (crossover) part of the study, an attrition rate of 59%. These participants generally failed to return for scheduled appointments and did not complete end-of-study paperwork or respond to multiple outreach attempts. No participants reported leaving the study because of difficulty using the technology. This high attrition rate should again be considered in the context of the parent study, where participants had already been completing surveys on a smartphone for 3 months at the time they enrolled in this secondary study. Therefore, participants' perceptions

of burden may have been higher in our sample than would otherwise have been the case.

Ease of use

The ease-of-use measure had a mean rating of 5.96 out of 7 possible points (SD =1.06 points), showing relatively high satisfaction with the ease of receiving and understanding text messages about adherence. The ease-of-use scale had no correlation with participants' sex or minority race/ethnicity (all P-values >0.17), but did have a significant negative correlation with participants' age (r=−0.48; P=0.03), meaning that older participants found the daily surveys harder to use.

Usefulness

On the tailored messaging survey, participants gave the intervention a rating of 2.85 points (SD =0.32 points) out of 4 possible points, indicating that it generally or mostly met their needs. Two items about participants' behavioral intentions were also examined individually. On these items, participants also indicated a relatively high level of willingness to receive this type of support message if they could do it over again (M =3.00 points; SD =0.00 points), and an even higher level of willingness to recommend the survey and messages to a friend who was taking similar medication (M =3.40 points; SD =0.55 points). On open-ended feedback items, participants said that all text messages were appropriate and acceptable. Results on these usefulness measures were limited by attrition as shown in Figure 1, because these items were not completed until the end of the study. Importantly, there were no adverse events reported during the study, suggesting little potential for harm even when messages were systematically mismatched to participants' survey responses.

Manipulation check: lack of perceived differences between study conditions

As expected, participants perceived few differences between the intervention and control conditions, with 20% saying there were no differences between the first 2 weeks and second 2 weeks of the tailored-messaging intervention, 40%

Table 3 Summary statistics for self-report feasibility measures

Variable	N	Possible range	M (SD)	Minimum	Maximum
Acceptance	14	0–100	85.3 (17.9)	49	100
Ease of use	25	1–7	5.96 (1.06)	3.24	7.00
Usefulness	5	0–4	2.79 (0.23)	2.50	3.13
Would do it again	5	0–4	3.00 (0.00)	3.00	3.00
Would recommend to a friend	5	0–4	3.40 (0.55)	3.00	4.00

Notes: N varies by analysis because the technology acceptance questionnaire was completed by participants at the end of the baseline phase for the first two measures (with some missing data on the acceptance item), and the tailored messaging survey was completed at the end of the intervention phase, with sample size affected by attrition. Results on the perceived usefulness measures are therefore less reliable.
Abbreviations: SD, standard deviation; N, number of responses; M, mean.

saying there was a small difference that was barely notice-able, and the remaining 40% saying there was a noticeable difference but not a large or important one. These findings suggest that the message framing intervention was relatively subtle, as intended. Similarly, 80% of respondents said that both sets of messages related equally well to their needs across the matched and mismatched conditions, with the remaining 20% saying that neither set of messages seemed to relate well to their needs. These responses also suggested a lack of perceived differences between message types.

When asked how much they liked the tailored messages, 60% of participants again said there was no difference between study conditions, with 20% reporting a noticeable difference and another 20% saying there was a large and important difference in how much they liked the messages. Finally, when asked how helpful the messages were, 20% said there was no difference, and 40% said there was only a small difference in helpfulness of the tailored messages across study conditions. On this item, however, 40% of participants did rate the difference in helpfulness between study conditions as noticeable. No participants chose the most extreme category stating that differences in helpfulness were large and important. Unfortunately, due to the way these questions were worded (ie, "Did you notice a difference?"), it was not possible to determine which set of messages the participants considered more acceptable or helpful. At first glance, an assumption might be that participants preferred the matched messages, but the adherence results below could lead one to question this assumption.

Aim 2: effect of the intervention on adherence

After controlling for baseline adherence level, sex, and the number of days with a study observation, there was a sta-tistically significant effect of time in the repeated-measures ANOVA (multivariate $F_{(1, 20)} =5.50$; $P=0.03$; $\eta^2=0.16$). This result indicates that both groups' adherence improved from the initially assigned messaging condition to the crossover condition, regardless of the order in which the two condi-tions were presented – AB or BA (Figure 3). The expected interaction between initial group assignment and time was nonsignificant, and the observed effect size was small (multivariate $F_{(1, 20)} =0.18$; $P=0.68$; $\eta^2=0.005$). There was also no overall difference between experimental conditions ($F_{(1, 20)} =0.17$; $P=0.90$; $\eta^2=0.0005$). The most plausible threat to validity in this analysis was attrition: because relatively few participants completed the crossover phase, it seemed possible that participants who remained were simply more adherent
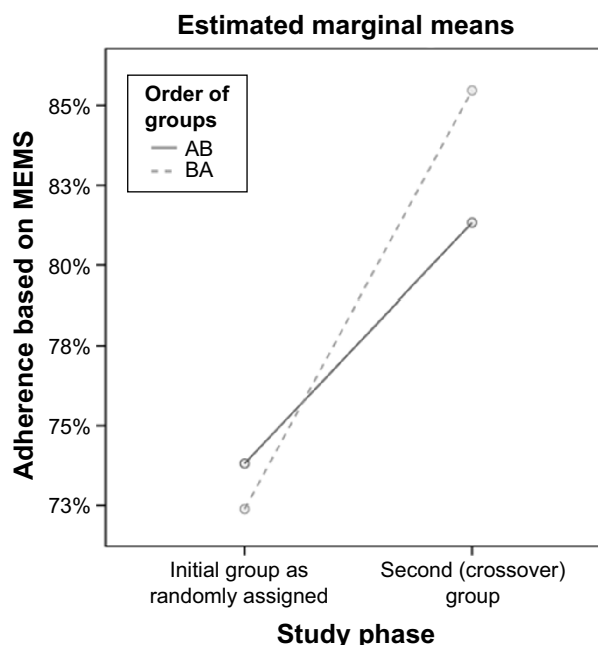

Estimated marginal means

Figure 3 Change in adherence over time.
Notes: AB order = matched, then mismatched messages; BA order = mismatched, then matched messages. Covariates in the model are evaluated at the following average values: sex =0.72 (1= male, 0= female); baseline MEMS =0.8354; number of days with observations =7.48 days.
Abbreviation: MEMS, medication event monitoring system.

overall, a nonrandom source of attrition that would call into question our use of imputation for missing data. We checked for this possibility in a sensitivity analysis that used only cases with complete data for both the intervention and crossover phases, and found an identical pattern of results: there was no difference between matched and mismatched messages, but there was an overall increase in adherence from the originally assigned intervention (regardless of which one was assigned) to the crossover phase. Additionally, there was no difference in baseline adherence between patients who remained in the study (M =81.8%) and those who dropped out prematurely (M =80.0%; $t_{(25)} =0.16$, $P=0.88$). Findings from these analy-ses gave us greater confidence in the obtained results.

Discussion

The current study was designed to evaluate feasibility of a novel text-messaging intervention to promote ART adherence based on PLWH's psychological state at the time the message was received. The intervention was feasible based on accep-tance (as shown by a 76% enrollment rate and a 0–100 visual analogue scale rating of 85%) and participants' high ratings of ease of use for the daily surveys. Perceived usefulness was slightly lower, with ratings between "a little bit" and "mostly" useful on a 4-point scale, so there appeared to be room for improvement in terms of making the intervention appealing

to PLWH. The usefulness findings were more tentative, because they were based on a questionnaire completed at the end of the study when results were affected by attrition. Nevertheless, perceived usefulness affects people's sustained use of a new technology over time,[27] so the slightly lower usefulness ratings suggested a need to make the intervention more appealing in order to maximize its chances for successful adoption in practice. On the other hand, there were no comments suggesting problems with the message content, and no adverse events even when messages were systematically mismatched to PLWH's momentary states. Participants reported occasional problems using the hardware and software; in particular, limited battery life and a tendency to miss messages when the phone was powered down were common reported problems. The main concern appeared to be that the technology could be made more exciting or engaging to promote repeated use over time, not that there was anything specifically unacceptable about the tailored messaging intervention.

The present study had a relatively high rate of attrition, and it is unknown whether the same usefulness considerations led participants to discontinue. We checked baseline adherence as a potential confound, and there was no evidence that participants' decision to leave the study was related to their overall level of medication adherence. Anecdotally, the most common complaint was simply that surveys and messages were too repetitive, that the intervention became boring over time, and that participants wanted more variety. This might relate to the fact that participants perceived few differences among tailored messages, even between the completely matched and completely mismatched message conditions. It is also possible that some participants actually had limited variety in the messages received: although there were 320 possible messages overall, if participants reported essentially the same momentary states and selected the same adherence barrier each day, they would have received the same or almost the same message each time. Alternately, participants might have abandoned the intervention as a result of alert fatigue, which could be addressed by delivering less frequent messages.

The current pilot RCT results in aim 2 showed no difference between the matched and mismatched message conditions overall, with an effect size very close to zero after controlling for covariates. This result implied that tailoring messages based on relevant theory did not have a strong effect, contrary to both our expectations and prior meta-analytic findings.[5] Only a few prior studies[11,12] have utilized tailored messages that are systematically tailored based on algorithms, but the available evidence for this type of intervention is mixed, and the current study seemed to suggest no benefit of systematic

tailoring. However, we did observe an unexpected but significant improvement in adherence at the time of the crossover between groups, regardless of whether participants crossed from intervention to control or the reverse. The improvement was equivalent to about a 15 percentage point gain in adherence, which is a potentially clinically meaningful effect. This surprising finding requires further exploration.

One possible explanation is a novelty effect; educational technology researchers have long known that simply introducing a different element can cause people to attend to their environment more carefully, regardless of what the difference entails.[28] Therefore, simply noticing that the messages had changed might have been enough to make participants attend to their content more closely and to follow advice contained in the messages. This type of effect has been documented in one other intervention using electronic tailored messaging to improve health behavior, where both the intervention and control groups showed improvement.[29] One argument against this interpretation is that participants noticed few differences between the original study condition and the crossover condition, so they might not have noticed that the second condition was novel or different compared to the first. A variation on this idea is that repetitive messages might actually lull people into a more receptive state, a principle used in clinical hypnosis.[30] Such effects might arise even if participants did not consciously notice a difference in message content. This explanation is potentially supported by the complaint that messages were boring or repetitive; although repetition of the same message reduces attention, it may also make novel or different messages more salient.[31] Pseudotailoring effects[23] are a third possible explanation if participants believed the second condition was better matched to their individual needs simply because they had been answering surveys longer by that time. Again, the fact that participants noticed few differences between study conditions might be an argument against this interpretation.

Study limitations

There were minor failures of randomization in this small-N pilot study, which we addressed in the analysis by controlling for baseline between-group differences. There might have been differences in other unmeasured variables even though random assignment was used. Missing data were another concern, and slightly exceeded the level at which imputation is bias-free; nevertheless, modern data imputation procedures are considered to provide more accurate parameter estimates than traditional case-deletion missing data strategies, even when some bias may be present.[26] We therefore imputed missing data despite this limitation, an approach supported

by the fact that MEMS values appeared to be missing at random with respect to baseline participant characteristics, including initial adherence.

Selection bias was a potential concern, because participants were recruited from a previous study and had completed 3 months of prior daily surveys; however, there were no demographic differences between PLWH who started the parent study and those who continued into the current secondary intervention pilot study. Perhaps the greatest remaining methodological limitation in the present study was the high rate of attrition between the initially assigned experimental condition and the crossover condition, which may reduce confidence in our results on usefulness and adherence. However, acceptance and ease-of-use ratings were completed at the baseline time point and were not affected by attrition, and we conducted a sensitivity analysis using complete cases to verify that improvements in adherence were not an artifact of attrition. All of these limitations could be addressed in a future full-scale RCT.

Implications for practice and research

Although a smartphone-delivered tailored messaging intervention was feasible for improving adherence in PLWH, it did not show evidence of efficacy in the way that we expected. The results demonstrated that daily tailored messages can improve adherence, but the way in which they do this is still an open question. The current intervention was designed to assess momentary states and to use information about them to affect participants' motivation, but not to modify the underlying states directly. There is some evidence that momentary states do predict adherence behavior in PLWH,[16] so these states may yet be a viable target for intervention. An alternative approach might be to design tailored messages that directly modify momentary states, rather than trying to use momentary states as a way to indirectly affect motivation. Such a strategy might lead to greater differences in message content to reduce participants' sense of boredom and to generate stronger effects.

Second, analyses currently underway from the parent study suggest that motivation does play a mediating role between other momentary states and behavior. It might be possible to retain motivation as the target for a daily text-messaging intervention, but to do so in a different way. For instance, the variables of importance, confidence, and readiness are often identified as aspects of motivation in motivational interviewing, and these are alternative dimensions that could be used to tailor daily adherence messages.[32] This strategy would allow us to retain the principle of "working at the level of motivation" that guided the current study, but to implement this principle in a different way.

A third development option might be to abandon the motivational model, but capitalize on the novelty effect produced when messages differ from one another over time. This strategy would mean providing PLWH with a wide range of interesting messages to keep their attention, with the idea that simple variations might be more useful than algorithm-driven matching of messages to participant characteristics. Such a strategy could incorporate additional features of currently available smartphone technology, such as YouTube videos, audio clips, interactive graphics, or gamification strategies like contests, badges, and leader boards. The more characteristics of a message that change at once, the weaker the ability to identify active ingredients or to control for potential pseudotailoring effects,[5] but this methodological concern might be secondary to the clinical need for a more efficacious version of the intervention. Modifying other characteristics of the intervention such as dose (message length or level of detail) and frequency (number of messages per day), or including suggestions that address practical barriers to care such as transportation, might also contribute to a stronger effect.

Conclusion

Smartphone-based tailored messages are an appealing mobile health modality that can potentially address limitations of prior adherence interventions. Tailored messages can address theoretically relevant psychological variables that are assessed in the moment, and can deliver information that is potentially more relevant to the immediate context of PLWH's daily lives. In the current study, we successfully delivered tailored messages to PLWH over a period of 4 weeks. PLWH were willing to participate in the intervention, gave it high ratings for acceptance and ease of use, and gave it moderately high ratings for perceived usefulness. These factors could likely be improved by varying message content and/or format to maintain participants' interest, strategies that might also reduce attrition. Tailored messages have the potential to improve adherence to a clinically meaningful degree, although theory-based tailoring did not appear to be the active ingredient, and the mechanisms by which smartphone-delivered messages actually affect PLWH's adherence behavior are still unclear.

Acknowledgments

and staff at the University of Colorado Hospital Infectious Disease Group Practice for their support of this study.

Disclosure

References

1. Chesney MA. The elusive gold standard: future perspectives for HIV adherence assessment and intervention. *J Acquir Immune Defic Syndr*. 2006;43(Suppl1):S149–S155.
2. Christensen AJ. *Patient Adherence to Medical Treatment Regimens*: *Bridging the Gap Between Behavioral Science and Biomedicine*. New Haven: Yale University Press; 2004.
3. Paterson DL, Swindells S, Mohr J, et al. Adherence to protease inhibitor therapy and outcomes in patients with HIV infection. *Ann Intern Med*. 2000;133(1):21–30.
4. Amico KR, Harman JJ, Johnson BT. Efficacy of antiretroviral therapy adherence interventions: a research synthesis of trials, 1996 to 2004. *J Acquir Immune Defic Syndr*. 2006;41(3):285–297.
5. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. *Psychol Bull*. 2008;133(4):673–693.
6. Noar SM, Black HG, Pierce LB. Efficacy of computer technology-based HIV prevention interventions: a meta-analysis. *AIDS*. 2009;23(1):107–115.
7. Smith A. *US Smartphone Use in 2015*. Washington, DC: Pew Research Center – Internet, Science, and Tech. Available from: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015. Accessed April 15, 2015.
8. Ybarra ML, Bull SS. Current trends in Internet- and cell phone-based HIV prevention and intervention programs. *Curr HIV/AIDS Rep*. 2007;4(4):201–207.
9. Vervloet M, van Dijk L, Santen-Reestman J, et al. SMS reminders improve adherence to oral medication in type 2 diabetes patients who are real time electronically monitored. *Int J Med Inform*. 2012;81(9):594–604.
10. Webb TL, Joseph J, Yardley L, Michie S. Using the internet to promote health behavior change: a systematic review and meta-analysis of the impact of theoretical basis, use of behavior change techniques, and mode of delivery on efficacy. *J Med Internet Res*. 2010;12(1):e4.
11. Petrie KJ, Perry K, Broadbent E, Weinman J. A text message programme designed to modify patients' illness and treatment beliefs improves self-reported adherence to asthma preventer medication. *Br J Health Psychol*. 2011;17(1):74–84.
12. Linn AJ, van Weert JCM, van Dijk L, Horne R, Smit EG. The value of nurses' tailored communication when discussing medicines: exploring the relationship between satisfaction, beliefs and adherence. *J Health Psychol*. Epub 2014 Jul 4.
13. Barta WD, Portnoy DB, Kiene SM, Tennen H, bu-Hasaballah KS, Ferrer R. A daily process investigation of alcohol-involved sexual risk behavior among economically disadvantaged problem drinkers living with HIV/AIDS. *AIDS Behav*. 2008;12(5):729–740.
14. Mustanski B. The influence of state and trait affect on HIV risk behaviors: a daily diary study of MSM. *Health Psychol*. 2007;26(5):618–626.
15. US Food and Drug Administration. *Paving the Way for Personalized Medicine*: *FDA's Role in a New Era of Medical Product Development*. Washington, DC: US Food and Drug Administration; 2013. Available from: http://www.fda.gov/downloads/ScienceResearch/SpecialTopics/PersonalizedMedicine/UCM372421.pdf. Accessed April 15, 2015.
16. Cook PF, McElwain CJ, Bradley-Springer L. Feasibility of a daily electronic survey to study prevention behavior with HIV-infected individuals. *Res Nurs Health*. 2010;33(3):221–234.
17. Lynn MR. Determination and quantification of content validity. *Nurs Res*. 1986;35(6):382–385.
18. Reeder B, Meyer E, Lazar A, Chaudhuri S, Thompson HJ, Demiris G. Framing the evidence for health smart homes and home-based consumer health technologies as a public health intervention for independent aging: a systematic review. *Int J Med Inform*. 2013;82(7):565–579.
19. Devito Dabbs A, Song MK, Hawkins R, et al. An intervention fidelity framework for technology-based behavioral interventions. *Nurs Res*. 2011;60(5):340–347.
20. Davis FD. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*. 1989;319–340.
21. Reeder B, Revere D, Olson DR, Lober WB. Perceived usefulness of a distributed community-based syndromic surveillance system: a pilot qualitative evaluation study. *BMC Res Notes*. 2011;4:187.
22. Leon AC, Davis LL, Kraemer HC. The role and interpretation of pilot studies in clinical research. *J Psychiatr Res*. 2011;45(5):626–629.
23. Weinstein ND, Rothman AJ, Sutton SR. Stage theories of health behavior: conceptual and methodological issues. *Health Psychol*. 1998;17(3):290–299.
24. Cook PF, Schmiege SJ, McClean M, Aagaard L, Kahook MY. Practical and analytic issues in the electronic assessment of adherence. *West J Nurs Res*. 2011;34(5):598–620.
25. Herzog MA. Considerations in determining sample size for pilot studies. *Res Nurs Health*. 2008;31:180–191.
26. Collins LM, Schafer JL, Kam CM. A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychol Methods*. 2001;6(4):330–351.
27. Lu HP, Gustafson DH. An empirical study of perceived usefulness and perceived ease of use on computerized support system use over time. *J Inform Manag*. 1994;14(5):317–329.
28. Clark RE. Reconsidering research on learning from media. *Rev Educ Res*. 1983;53(4):445–459.
29. Yap TL, Davis LS, Gates DM, Hemmings AB, Pan W. The effect of tailored e-mails in the workplace: part II. Increasing overall physical activity. *AAOHN J*. 2009;57(8):313–319.
30. Wright ME, Wright BA. *Clinical Practice of Hypnotherapy*. New York: Guilford Press; 1987.
31. Ayala YA, Malmierca MS. Stimulus-specific adaptation and deviance detection in the inferior colliculus. *Front Neural Circuits*. 2013;6:89.
32. Miller WR, Rollnick S. *Motivational interviewing: Helping people change*. 3rd ed. 2012. New York: Guilford Press.
33. Prochaska JO, Norcross J, DiClemente C. *Changing for Good*. New York: Avon Books; 1995.
34. Leventhal H, Diefenbach M, Leventhal EA. Illness cognition: using common sense to understand treatment adherence and affect cognition interactions. *Cognitive Therapy and Research*. 1992;16:143–163.
35. Lazarus RS, Folkman S. *Stress, Appraisal, and Coping*. New York: Springer; 1984.
36. Kahneman D, Tversky A. *Choices, values, and frames*. New York: Cambridge University Press; 2000.
37. Reifsnider E, Gallagher M, Forgione B. Using ecological models in research on health disparities. *Journal of Professional Nursing* 2005; 21(4):216–222.