ORIGINAL ARTICLE

The Potential Impact of Intelligent Systems for Mobile Health Self-Management Support: Monte Carlo Simulations of Text Message Support for Medication Adherence

**John D. Piette, PhD · Karen B. Farris, PhD ·
Sean Newman, MS · Larry An, MD ·
Jeremy Sussman, MD, MS · Satinder Singh, PhD**

Abstract

Background Mobile health (mHealth) services cannot easily adapt to users' unique needs.

Purpose We used simulations of text messaging (SMS) for improving medication adherence to demonstrate benefits of interventions using reinforcement learning (RL).

Methods We used Monte Carlo simulations to estimate the relative impact of an intervention using RL to adapt SMS adherence support messages in order to more effectively address each non-adherent patient's adherence barriers, e.g., forgetfulness versus side effect concerns. SMS messages were assumed to improve adherence only when they matched the barriers for that patient. Baseline adherence and the impact of matching messages were estimated from literature review.

RL-SMS was compared in common scenarios to simple reminders, random messages, and standard tailoring.

Results RL could produce a 5–14 % absolute improvement in adherence compared to current approaches. When adherence barriers are not accurately reported, RL can recognize which barriers are relevant for which patients. When barriers change, RL can adjust message targeting. RL can detect when messages are sent too frequently causing burnout.

Conclusions RL systems could make mHealth services more effective.

Keywords Medication adherence · Mobile health · Text messaging · Simulation models · Chronic disease management

J. D. Piette (✉) · J. Sussman
Center for Clinical Management Research, VA Ann Arbor
Healthcare System, Ann Arbor, MI, USA
e-mail: jpiette@umich.edu

J. D. Piette
Department of Health Behavior and Health Education, University of
Michigan School of Public Health, Ann Arbor, MI, USA

J. D. Piette · L. An · J. Sussman
Department of Internal Medicine, University of Michigan Medical
School, Ann Arbor, MI, USA

K. B. Farris
Department of Social and Administrative Services, College of
Pharmacy, University of Michigan, Ann Arbor, MI, USA

S. Newman · S. Singh
Division of Computer Science and Engineering, Department of
Electrical Engineering and Computer Science, College of
Engineering, University of Michigan, Ann Arbor, MI, USA

Introduction

Patient non-adherence to medications for chronic disease management is a major cause of preventable morbidity, mortality, and excess healthcare spending. As many as half of all chronically ill patients fail to take their medications as prescribed, contributing to nearly 100,000 premature deaths annually, as well as $290 billion in annual healthcare costs [1–3] and more than a 10 % increase in hospital admissions for older adults [4].

While forgetfulness is one major challenge to medication use, non-adherence often is intentional, due to patients' skepticism about the medication's potential benefits, or concerns regarding side effects or other negative consequences of long-term pharmacotherapy [5–7]. Patients often report multiple reasons for not taking their medication as prescribed [5]. Those reasons differ among patients and can change over time as patients' health beliefs are influenced by factors such as their receipt of health information or changes in their health status [8, 9].

Mobile health (mHealth) interventions such as patient text messaging (short message service or SMS) can deliver frequent prompts and health information to improve adherence [10–16]. Two large trials in Kenya have shown that SMS messages can improve antiretroviral adherence among patients with HIV/AIDS, with one of those trials demonstrating significant improvements in viral suppression [17, 18]. Studies from other countries also demonstrate that SMS messaging may improve medication adherence [13–16]. Some of the most successful text messaging adherence interventions have focused on medication reminders [19].

One well-established, evidence-based technique to ensure that patients using mHealth support services receive the information they need is to tailor messages according to the characteristics of the user [20]. Most studies of tailored health communication use comprehensive baseline assessments to determine what information will be most impactful for which patients [20], typically focusing on patients' sociodemographic characteristics, clinical history, health beliefs, and other determinants of self-care behavior. Message tailoring can significantly improve the effectiveness of health behavior change messages, including messages addressing medication underuse [21–23].

Despite its demonstrated benefits, the effectiveness of tailored communication may be limited if the assessments used to tailor information fail to address key determinants of users' behavior, or if valid patient information cannot be collected due to psychometric limitations in the measures or biases in patient reports. For some behaviors, such as medication adherence, accurate data are often lacking about what causes patients to fall short of self-care goals [24]. Patients' need for health information may change over time as they master skills or develop new concerns, and repeated tailoring to accommodate those changes is challenging. Finally, studies have shown repeatedly that patients can become desensitized to health communication [25–28], and current approaches to tailored adherence support typically cannot detect patient burnout or adapt message content and frequency so that services remain engaging and effective. The next generation of chronic disease behavior change services will need new strategies for building on the successes of tailored messaging based on up-to-date evidence regarding the content and mode of communication that are most effective for each individual.

Reinforcement learning (RL) is a field of artificial intelligence that presents new opportunities to automatically adapt mobile health communication based on feedback from patients about the impact of different types of messages with respect to attaining a given health or behavioral goal. RL algorithms applied to mobile health communication for adherence support can use information about the effectiveness of prior messages addressing each patient's reasons for non-adherence while continuing to strategically explore patients' responses to messages addressing other potential adherence

challenges that have yet to be identified. RL has been applied extensively in robotics, control systems, and resource allocation tasks [29, 30], and RL algorithms are common in digital commerce, including web advertisement; online news article selection; or product recommendation by companies such as Google [31], Yahoo [32], Amazon, and Netflix. Applications of RL to human-centered tasks related to behavior change are rare [33–35], and to our knowledge, only a few published papers have described the use of RL in mobile health disease management support [36].

The purpose of this study was to demonstrate the ways in which RL could be used to adapt the content and frequency of SMS messages to promote adherence among individuals taking a medication regularly for a chronic medical problem. To that end, we conducted Monte Carlo simulations estimating the impact of an RL-driven SMS adherence support tool relative to other common approaches to designing mobile health messaging systems, such as focusing specifically on patient reminders or tailoring communication using a baseline survey. Monte Carlo simulations allow potential impacts of a healthcare intervention or policy to be evaluated without the costs of service development and deployment, patient enrollment, and outcome measurement in a real-world trial [37–39]. A strength of simulations is that intervention effects can be estimated in the context of differing scenarios (e.g., whether or not patients misreport their adherence barriers, making standard tailoring more difficult) as well as in the context of random variability in key parameters, such as patients' daily adherence behavior. Through the simulations presented here, we estimated the relative performance of RL compared to three common approaches to mobile health messaging and in the context of real-world challenges to the effectiveness of SMS behavior change services, including a change in patients' adherence barriers over time and decreased intervention effectiveness when some patients become desensitized to messages that are sent too frequently.

Methods

Components of a Reinforcement Learning System for SMS Adherence Support

An RL system is composed of four functional units: (1) a set of action choices (in this case, the types of SMS messages addressing various potential reasons for non-adherence); (2) a means of obtaining feedback or "reinforcement" about the outcome of interest following each action choice, i.e., the "reward" in RL terminology; in the case of medication use, the reward can be defined as an ongoing measure of medication adherence during successive intervals; (3) Additional information about the "state" or context within which action choices are made, such as patients' characteristics affecting

adherence as well as information about patients' prior history of interaction with the system; and (4) an RL computational agent (the RL "engine") that adapts probabilistic action choices based on the state, action, and reward data that are accumulated over time. Each of these components is described below. Additional technical details about RL are included in the Electronic Supplementary Material (ESM).

Action Choices Action choices represent the set of potential interventions available to the RL system at each time point. For the simulations described here, we defined action choices as the choice to send an SMS message addressing one of three possible reasons for medication non-adherence. Specifically, we reviewed the literature [6, 40, 41] and identified three common barriers to appropriate medication use: (1) patients' beliefs about their disease and its implications for their health and well-being, e.g., beliefs that their illness is not serious or that their risk for complications is minimal; (2) patients' beliefs about their medication, e.g., that the medication cannot change their long-term course of illness or that it will cause side effects; (3) patients' need for strategies to address forgetfulness in taking their medication. In the simulations, the RL system was designed to "learn" for each patient which message type or action was most effective in promoting adherence; and (4) in the third simulation below, examining the impact of message fatigue, a null message (i.e., not sending any message), was considered a fourth action choice so that the system could learn the frequency of messaging that was most effective for each patient.

Adaption Based on Feedback RL systems adapt by seeking to optimize a numerical outcome or "reward." Operationally, an SMS medication adherence support system using RL would need a "sensor" for collecting data about patients' medication taking following each message sent to the patient as well as a mathematical model for translating those adherence data into a score that the system uses as its reward to evaluate the impact of those action choices. A sensor could be something simple, such as an SMS system with which patients could repeatedly self-report their adherence using a numerical rating scale. For the simulations presented here, we assumed that the RL system would be informed about patients' adherence using daily pill bottle openings provided automatically via a medication electronic monitoring system (MEMS) [42–45].

State Information When making action choices, some RL algorithms can take into account baseline information about each patient as well as information gathered during the history of interaction between the patient and the system. Such "state" information summarizes what is important to action selection. State parameters can be thought of as "effect modifiers" that influence the expected impact of a given action on patients' adherence. In SMS intervention design, state information could include information such as changes in adherence after sending previous messages, adherence barriers or health beliefs reported at baseline, or the patient's demographic characteristics. RL systems learn the best mathematical function relating state variables with probabilistic action choices. By learning an action-selection policy from prior interactions with users and applying that policy to future interactions with patients who share similar characteristics as recorded in their vector of state information, RL systems can learn more quickly how best to target messages so as to maximize the impact on the "reward."

The RL Engine There are several types of RL algorithms that could be used to process reward data and drive SMS adherence support message-selection choices [29, 46–48]. Regardless of this choice, all RL systems must address two fundamental problems: (1) balancing exploration of relatively untested action choices versus exploitation of information about choices that have already been shown to be impactful and (2) accounting for delayed effects of a given action choice on patients' future adherence behavior. This latter problem is particularly relevant in mobile health services supporting behaviors such as medication adherence, since information from SMS adherence messages hopefully affects patients' adherence not only the day the message was received but also the patient's long-term ability to self-manage their chronic disease. One subset of RL approaches (known as "bandit" algorithms) assumes that an SMS message sent today influences only the patient's likelihood of adherence that same day, i.e., the message will have no residual impact on their adherence in the future. Other algorithms (e.g., algorithms based on partially observable Markov decision processes or POMDP's) estimate a parametric model for patients' underlying pattern of behavior and use that knowledge to account for long-term effects. These differences between algorithms can contribute to the speed with which an RL system can effectively adapt to each patient's unique needs (i.e., the number of patients and patient interactions needed to optimize message choices) as well as the maximum impact that the system may have across patients and over time. Parametric models can be slow to learn, while algorithms that learn quickly may not have the greatest long-term effect. In the simulations that follow, we used a contextual bandit algorithm, LinUCB (see ESM for more information), to illustrate the way in which an RL-supported SMS adherence program can adapt quickly to each patient's needs and have a relatively important impact even in the absence of a parametric model of patients' adherence behavior [49].

Monte Carlo Simulations

To evaluate the performance of RL-based adherence messaging relative to other approaches, we conducted three Monte

Carlo simulations. For each simulation, we assumed that patients' adherence behavior was determined by the extent to which they experienced three adherence barriers, represented by the three "$\alpha_{(barrier)}$" terms in the following formula:

$$P(\text{adherence}) = \alpha_{(disease)} \times \alpha_{(medicine)} \times \alpha_{(remember)}$$

where $P(\text{adherence})$ is the probability that a patient takes his or her medication correctly each day, $\alpha_{(disease)}$ represents the extent to which the patient believes that his or her illness is important to treat (ranging from values of 0 for not at all important to 1 for extremely important), $\alpha_{(medicine)}$ represents the extent to which the patient is concerned about the negative consequences of his or her medication use (ranging from 0 for extremely concerned to 1 for not at all concerned), and $\alpha_{(remember)}$ represents the likelihood that the patient remembers to take his or her medication on a given day (ranging from 0 for will not remember to 1 for will definitely remember).

For example, if a patient is only 80 % convinced about the importance of treating his or her illness (i.e., $\alpha_{(disease)}=0.80$), the patient is only 70 % concerned about taking the medication (i.e., $\alpha_{(medicine)}=0.70$), and if the patient has only a 90 % probability of remembering to take the medication assuming he or she intends to do so (i.e., $\alpha_{(remember)}=0.90$), then that patient's probability of taking medication on a given day would be as follows:

$$(1): P(\text{adherence}) = 0.80 \times 0.70 \times 0.90 = 0.504$$

In the simulations presented here, we assumed that the patient sample included 20 patients whose only adherence barrier was a doubt about the disease's severity, 20 patients whose only barrier was a concern about the safety and efficacy of the medication, and 20 patients who both had doubts about disease severity as well as problems with forgetfulness. Based on a literature review [4, 5, 8, 9, 50, 51], we assumed that the value of each $\alpha_{(barrier)}$ in our simulations was 0.65. Studies for medications treating hypercholesterolemia, osteoarthritis, cardiovascular disease, and type 2 diabetes report adherence rates ranging from 50 to 72 % [3, 51]. Other studies including conditions such as bladder disorders and glaucoma have found adherence rates below 40 % [2]. Given that all three $\alpha_{(barrier)}$ terms=0.65, the baseline adherence rate for the overall sample of simulated patients was in this range, i.e., 57 % or the average of 65 % for the 20 patients with only doubts about their disease's severity, 65 % for the 20 patients with only medication belief concerns, and 42.25 % for the 20 patients with both disease belief as well as forgetfulness problems. Patients' actual adherence on a given day was assumed to vary around their own average adherence level, according to random variation in the relevant $\alpha_{(barrier)}$ terms, with each day's

value for each patient drawn from a Gaussian distribution with a mean of 0.65 and a standard deviation of 0.30.

SMS messages were assumed to impact adherence only if they addressed the patient's underlying reasons for non-adherence as expressed by that person's three $\alpha_{(barrier)}$ values. For example, we assumed that a reminder message would have no impact if the patient in fact was not forgetful but rather was non-adherent because of unaddressed concerns about the medication's safety. When a given SMS message did address one of the patient's reasons for non-adherence, the magnitude of the SMS message effect or "$\beta_{(match)}$" was calculated as a multiplicative improvement in the relevant $\alpha_{(barrier)}$. For example, if $\beta_{(match)}=0.6$ and a patient only had medication concerns with $\alpha_{(medicine)}=.65$, then their $\alpha_{(medicine)}$ would change to the following:

$$(2): \alpha_{(medicine)} = 0.65 + 0.6 \times (1.0-0.65) = 0.86$$

In the absence of an empirical evidence for differences in the effect size of SMS messages addressing each of the three adherence barriers, we assumed that the effect of a matching SMS message was the same for all three. The magnitude of $\beta_{(match)}$ used in the simulations presented below was based on the improvement in adherence reported in SMS adherence trials, i.e., 12 to 16 % [17, 21, 50, 51]. For example, Pop-Eleches et al. [17] found that weekly short reminders could improve adherence by up to 13 % over the control group after 48 weeks, and Petrie et al. [21] showed that text messages tailored to patient's underlying adherence barriers could improve adherence up to 15 % over a control group. To ensure that our simulations generated results that were comparable to these outcomes, we assumed that each day, the actual value of $\beta_{(match)}$ for each patient was selected at random from a normal distribution with a center of 0.7 and a standard deviation of 0.3.

Given these assumptions, the goal of the RL-supported messaging system was to increase the intervention's effectiveness over time via an increased likelihood of a match between the type of message sent to each patient on each day and the patient's adherence barriers. RL sought to increase the match success rate by modifying the probability of sending a message addressing each of the three barriers, based on feedback about that patient's adherence following prior choices by the RL engine as to which message type would be most effective. More details about the characteristics of each simulation are presented in Table 1 and ESM. The results of the simulations are presented graphically in Figs. 1, 2, and 3 as changes in the proportion of patients who took their medication each day over the course of a 180-day intervention period. Estimates for each day of adherence support were averaged over 100 independent runs, with key parameters being drawn from Gaussian distributions, as noted above.

Table 1 Summary of simulation experiments

	Simulation 1	Simulation 2	Simulation 3
Aim	Estimate the relative performance of an RL agent assuming that a third of the sample does not accurately report their non-adherence barrier at baseline	Estimate the relative performance of an RL agent assuming that a third of patients' underlying barrier to adherence changes during the course of intervention	Estimate the relative performance of an RL agent assuming that patients "tune out" messages that come too frequently
Action choices	1. Send a disease belief message 2. Send a medication belief message 3. Send a remembering strategy message	1. Send a disease belief message 2. Send a medication belief message 3. Send a remembering strategy message	1. Send a disease belief message 2. Send a medication belief message 3. Send a remembering strategy message 4. Send no message to learn about the impact of message frequency
State variables	How many times the patient took his/her medication following the last 5 times a message of each type was sent	How many times the patient took his/her medication following the last 5 times a message of each type was sent	1. How many times the patient took his/her medication following the last 5 times a message of each type was sent 2. A binary indicator for whether a message had been sent on each of the previous 2 days
Adherence after 180 days[a]	RL: 78 % (0.5 %) Tailored: 73 % (0.6 %) Random: 67 % (0.5 %) Reminders: 66 % (0.6 %)	RL: 78 % (0.6 %) Tailored: 73 % (0.6 %) Random: 68 % (0.6 %) Reminders: 66 % (0.5)	RL: 70 % (0.6 %) Tailored: 57 % (0.6 %) Random: 57 % (0.6 %) Reminders: 56 % (0.7 %)

For each experiment, we assumed that the population consisted of 60 patients including the following; 20 patients whose only barrier to medication adherence was concerns about the safety and efficacy of their medications, 20 patients with doubts about their disease severity, and 20 patients who had both concerns about their medications as well as problems remembering to take them

RL reinforcement learning

[a] The baseline adherence rate for the overall sample was assumed to be 57 %, based on a literature review. Numbers in parentheses are standard deviations

Fig. 1 Adherence rates achieved using four different message-selection strategies when a third of the population does not accurately self-report their non-adherence barrier
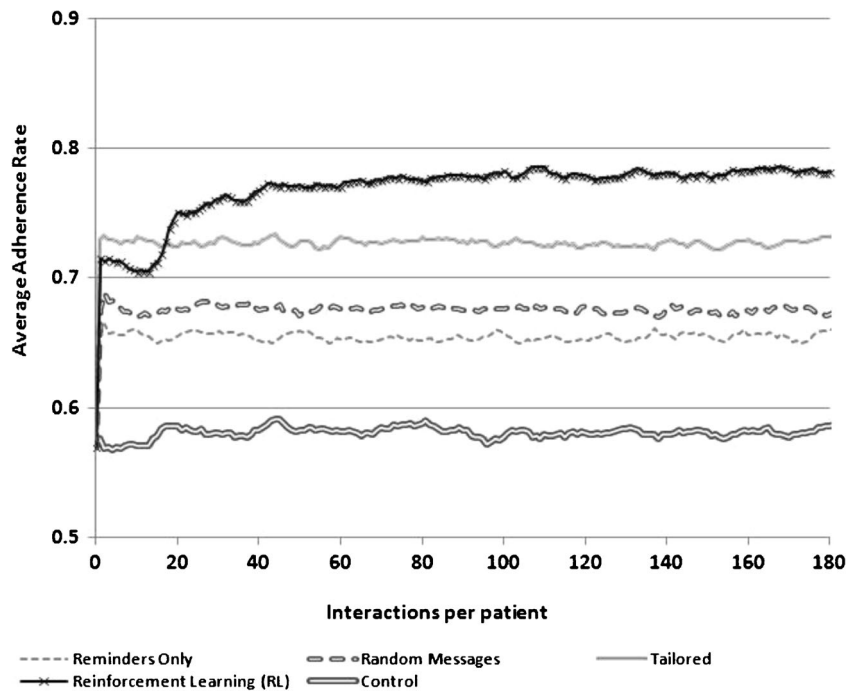
We compared the behavior of the RL-SMS adherence support service to three common alternatives to delivering SMS messaging interventions: (1) SMS messages serving only as medication adherence reminders, (2) SMS messages addressing each of the three potential barriers to adherence but sent randomly so that each patient received all message types on random days, and (3) SMS messages tailored to patients' adherence barriers as reported in a baseline survey. The RL algorithm initially weighted its choice of messages according to the baseline information about the patient's barriers—

similar to the tailored approach. However, unlike tailoring, message choices for RL were selected from a probability distribution in which messages addressing other potential adherence barriers were selected with a small probability, so that RL could "explore" these choices, learn from subsequent adherence feedback, and adjust message targeting accordingly.

The relative efficacy of the RL-based messaging system was evaluated in three scenarios, representing possible situations that could limit the efficacy of current approaches: (1) a

Fig. 2 Adherence rates achieved using four different message-selection strategies when a third of the population change their non-adherence barrier on day 90
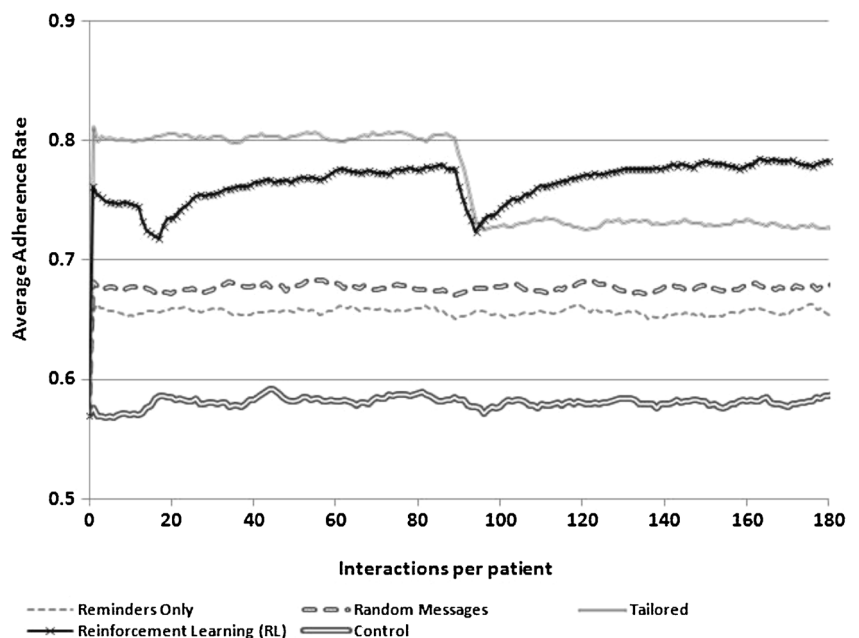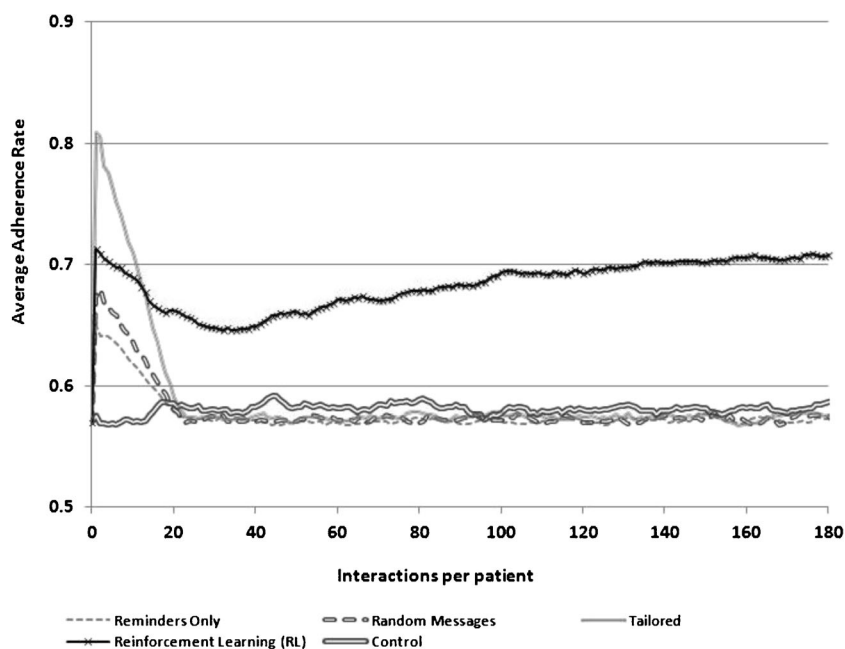
Fig. 3 Adherence rates achieved using four different message-selection strategies when there is message fatigue

scenario in which some patients' actual barriers to medication adherence are not accurately measured at baseline, (2) a scenario in which some patients' reasons for non-adherence change during the course of intervention, and (3) a scenario in which patients experiences SMS message fatigue and begin to tune-out messages that come too frequently. *In the first scenario*, we assumed that patients who were actually concerned about their medication's safety instead reported at baseline that their problem was forgetfulness. *In the second scenario*, we examined what would happen when patients who originally had doubts about their medication's safety instead developed doubts about their disease severity roughly halfway through the intervention period. In each of these two scenarios, we assumed that patients received SMS messages daily.

In the third scenario, we assumed that patients experienced message burnout when they received messages too frequently and that "too frequently" meant that they received a message after having received a message on each of the two previous days. For patients experiencing message burnout, we used a multiplicative fatigue factor that diminished the daily effect of a well-targeted message (recall that mismatched messages were assumed to have no effect). Specifically, we assumed that the effect of a given SMS message matching the patient's adherence barrier decreased 5 % each day that the patient was fatigued. For patients in the three non-RL comparison groups, we assumed that the SMS system continued to send messages daily despite message fatigue. In contrast, in the RL-supported program, the system was designed to "learn" automatically to adapt to message burnout by evaluating the impact of sending one of the three message types versus sending no message on a given day to prevent message fatigue. Each day that the

patient was not fatigued, the impact of a well-matched message increased 10 % until it reached its maximum level. See the ESM for more details.

Results

Simulation #1: Adaptation to Misreported Reasons for Non-Adherence

As shown in Fig. 1, in the initial 10 days of messaging, a tailored system resulted in higher rates of improvement in medication adherence than the alternative messaging approaches. During that initial period, the RL system (lacking experience from which to learn) began by randomly sending some messages to patients and therefore matched their adherence support needs less frequently than the tailored approach. However, with greater experience, the RL system was able to target messages effectively according to their impact on each patient's actual adherence, i.e., regardless of the patient's reported reasons for non-adherence at baseline. In contrast, the tailored approach continued to send mismatched messages to simulated patients who reported forgetfulness at baseline but who actually had concerns about their medication's safety. As a consequence, the overall improvement in adherence for the tailored system was less than for the RL system at 180 days.

The simple reminders performed worse than all other messaging strategies because only one third of the simulated sample had forgetfulness as an adherence problem, and those patients also had concerns about their disease severity.

Random messaging performed slightly better than simple reminders, since each message had at least a one third chance of being relevant for a particular patient. For simulated patients with both disease belief and memory barriers, each random message had a two thirds chance of being relevant. At the end of 180 days, the estimated daily adherence rate was 78 % for the RL messaging system, 73 % for the tailored system, 67 % for random messaging, and 66 % for simple reminders.

Simulation #2: Changes Over Time in Patients' Reasons for Non-Adherence

When assuming that simulated patients did in fact accurately report their reasons for non-adherence at baseline, the tailored approach, because it was immediately able to match adherence support messages to patients' needs, exceeded all other approaches in the initial period of intervention and continued to do so for the first 90 days (Fig. 2). During that period, the RL approach progressively improved relative to random messaging and simple reminders but still performed worse than tailored messages. However, when simulated patients who originally had concerns about their medication's safety developed doubts about their disease severity halfway through the intervention period, the tailored approach experienced a significant drop in effectiveness for users who thereafter had a mismatch between their message tailoring and their actual need for adherence support. The RL approach also experienced a drop in effectiveness at 90 days, since "lessons learned" in the first 90 days led the system to send messages incorrectly to patients whose adherence barriers changed. However, as shown in Fig. 2, the RL system was able to adapt to the change in patients' adherence barriers and began to outperform the tailored approach as well as the other two approaches around day 100. At the end of 180 days, expected daily adherence rates were 78 % for the RL system, 73 % for tailored messaging, 68 % for random messages, and 66 % for simple reminders.

Simulation #3: The Impact of Message Fatigue

As shown in Fig. 3, message fatigue led to diminishing intervention impact among patients receiving tailored messages, random messages, or simple reminders. This is because these messaging strategies were unable to adapt to the fact that when fatigued, the impact of well-targeted messages was progressively less. This resulted in a steady drop in the adherence rate for the three comparison groups, until the adherence rate in each group reached the baseline rate. In contrast, the RL messaging strategy was able to learn that the highest adherence rate for these simulated patients was achieved by repeating the pattern of sending two consecutive well-targeted messages followed by not sending a message on the third day.

Initially, the adherence rate in the RL messaging group declined as the RL system learned how patients responded (and in the process, elicited fatigue in some patients). However, by 20 interactions per patient, the adherence rate in the RL group started rising. At the end of 180 days, the expected daily adherence rate was 70 % for the RL system while the comparison groups returned to their baseline adherence rate of 57 %.

Discussion

We used Monte Carlo simulations and information about the effectiveness of adherence support services from prior trials to estimate the impact of an RL-based SMS adherence support tool compared to other common approaches to SMS adherence messaging. Our results indicate that RL is able to learn to personalize messages to individuals based on their true underlying causes for non-adherence, adapt to changes in the underlying causes for non-adherence over time, and adapt the frequency of messaging so as to avoid message fatigue. The RL system was able to outperform current approaches to designing SMS adherence interventions given reasonable assumptions regarding each approach's effectiveness when SMS messages matched patients' adherence barriers and assuming a relatively small sample of patients. These results affirm our underlying hypothesis that RL-based systems can overcome some of the limitations of current approaches to mHealth service design including approaches that tailor health communication via fixed information collected about patients' health beliefs at baseline. RL systems learn from increasing experience. As a consequence, systems will become increasingly effective as the number of patients and patient interactions grows. Because mHealth interventions can be expensive to create but less expensive to maintain, they become more efficient in the context of larger sample sizes, and therefore, RL may be ideally suited to situations in which mHealth programs are likely to be deployed.

It is important to emphasize that the scenarios explored in the current simulation trials are common in chronic disease management. For example, it is well known that patients often misreport their adherence barriers [19], and time-pressed clinicians may not be able to administer extensive tailoring surveys at program entry. For both of these reasons, RL services that require little tailoring information at baseline are an attractive alternative. Moreover, changes in patients' health status and adherence barriers are the hallmark of chronic illness, and researchers have had difficulty identifying services that can continue to be engaging and relevant in order to maintain health benefits over the long-term. Because RL-based behavior change services continually adapt according to what is working for each patient, they represent a more

patient-centered and potentially more effective alternative to current approaches.

"Message fatigue" is a major problem in chronic disease management [25–28]. Fatigue has not been addressed adequately because tolerance for frequent messages likely varies across patients, and addressing fatigue means balancing that problem with the need to make interventions as intensive as possible. Unlike less flexible approaches to messaging, RL can use feedback from patients to develop schedules that meet each individual's unique needs and preferences for information.

Our simulations leave open a number of challenging questions. Consistent with the underlying motivation for an adaptive approach, generalizations about the impact of RL behavioral interventions relative to standard messaging strategies are likely to be of limited relevance for a given application, due to the diversity of: the clinical problems for which an adaptive intervention might be useful, the characteristics of patients with whom these interventions could be applied, and the action choices that could be available to the RL system. Multiple empirical studies will be needed addressing a wide range of health behavior challenges to fully understand where and to what degree RL systems can improve behaviors over-and-above current, deterministic interventions.

Nevertheless, some general conclusions can be drawn about the characteristics of RL-based interventions that are likely to lead to a more successful adaptation to patients' needs and therefore greater effectiveness. Action choices need to be defined so that they represent distinct and meaningful alternatives that behavioral theory suggests may have varying, measurable impacts on outcomes across patients and over time. To the extent possible, the "reward" signal should be reliable, tightly linked to the action choices made by the RL system, and reported frequently enough to allow for rapid and successful intervention adaptation. Here, we presented a hypothetical example in which reward data consisted of daily adherence information available from a MEMS cap. Of course, MEMS cap information is rarely available in real-world settings and advances in passive sensor technologies may expand the number of situations in which RL-based interventions are successful in promoting behavior change. Even without those advances, researchers should continue to identify instances in which patients' self-reports can be used to guide adaptation by RL systems, since reliable behavioral reports that are correlated in predictable ways with the outcome of interest may be adequate, even if those self-reports are biased relative to the actual behavior, e.g., over-reporting of medication taking.

The mathematics needed to know how to "power" clinical trials for RL-based systems is underdeveloped relative to what is known for more standard fixed arms trials. The use of simulations, as in this paper, is one approach to determining how many patients are needed to learn a good RL policy. The benefit of these simulations depends on the availability of good mathematical models of patient behavior in response to RL system actions as well as good models of noise in the reward signal. If such models are available, then simulations can be used to fine tune the parameters of the RL algorithm (and more generally select from among competing RL algorithms) as well as determine how many interactions with patients are needed for the system to effectively learn what actions are most effective for each user. Further simulation-based and theoretical work is needed to articulate these important relationships as they apply to real-world applications, such as SMS adherence support.

Some of the parameters we used were based on real data (e.g., the baseline adherence rates and the expected impact of a message addressing a patient's adherence barriers), while other parameters such as the distribution of adherence barriers in the population were more arbitrary. To some extent, the results of our simulations reflect these choices. For example, a third of the sample in the current simulations had both concerns about their medication as well as problems remembering to take it; if those patients only had problems with forgetfulness, then random messages (with a third of those messages being reminders) would have had roughly the same impact on adherence in the overall sample as messaging exclusively targeting forgetfulness. If baseline adherence rates were significantly higher than the 57 % estimate we used, RL would have less opportunity to improve over standard tailoring. In contrast, if baseline adherence rates were much lower, then the differences between RL and simple tailoring would be larger than that shown in Figs. 1, 2, and 3. Auxiliary analyses that we conducted varying these assumptions as well as others such as the proportion of patients who misreport their adherence barriers demonstrate that the *relative* performance of RL compared to current mHealth messaging strategies would be largely unaffected by these differences. Nevertheless, the true benefits of an adaptive, patient-centered approach will be situation-specific, and more research should be pursued to understand where and how to best apply this new approach to enhancing health behavior change.

Studies have shown repeatedly that improved adherence is linked with improved health. For example, investigators have reported that hypertensive patients who are "non-adherent" have a 3.8-fold risk of stroke-related death in 2 years [52], 43 % of "high adherence" patients with hypertension achieve blood pressure control compared to 33 % of those with "medium adherence" [53], and "lower adherers" to statins have a 25 % increased risk of death compared to "high adherers" [54]. The precise functional form of the relationship between adherence and health is not well described, and it is difficult to generalize based on these simulations and available literature about specific health benefits that could be achieved through a more adaptive approach to behavioral intervention design. Nevertheless, it is fair to say that the absolute improvements in adherence shown

here relative to tailored messaging are modest, suggesting that the details of intervention design as well as the population in which that intervention is evaluated will be crucial determinants of effectiveness in real-world studies of RL-based adherence support systems. For example, auxiliary analyses indicate that if the baseline adherence rate were only 32 %, a third of the population did not accurately report their adherence barrier, and if a third developed a new adherence barrier, the difference in adherence with RL versus standard tailoring would be more than twice that presented here.

In summary, this simulation study demonstrated the possible benefits of RL in mHealth communication for the improvement of medication adherence. We found that under many real-world circumstances, RL could be more effective than tailored messages, because RL systems can learn the needs of each individual patient and how those needs change with time. RL systems also may provide a solution to patient burnout, by adapting the frequency of messages to each user so that they meet the person's needs and preferences. RL-based mHealth interventions are a promising example of how artificial intelligence can make healthcare more patient-centered.

References

1. Nehi. Thinking outside the pillbox: A system-wide approach to improving patient medication adherence for chronic disease, 2009.

2. Osterberg L, Blaschke T. Adherence to medication. *N Eng J Med*. 2005; 353: 487-497.

3. Elliott RA, Shinogle JA, Peele P, et al. Understanding medication compliance and persistence from an economics perspective. *Value Health*. 2008; 11: 600.

4. Vermeire E, Hearnshaw H, Van Royen P, et al. Patient adherence to treatment: Three decades of research. A comprehensive review. *J Clin Pharm Ther*. 2001; 26: 331-342.

5. McHorney CA, Spain CV. Frequency of and reasons for medication non-fulfillment and non-persistence among American adults with chronic disease in 2008. *Health Expect*. 2011; 14: 307-320.

6. McHorney CA, Gadkari AS. Individual patients hold different beliefs to prescription medications to which they persist vs nonpersist and persist vs nonfulfill. *Patient Prefer Adherence*. 2010; 4: 187-195.

7. Horne R, Weinman J. Patients' beliefs about prescribed medicines and their role in adherence to treatment in chronic physical illness. *J Psychosom Res*. 1999; 47: 555-567.

8. Unni E, Farris KB. Determinants of different types of medication non-adherence in cholesterol lowering and asthma maintenance medications: A theoretical approach. *Patient Educ Couns*. 2011; 83: 382.

9. Unni EJ, Farris KB. Unintentional non-adherence and beliefs in medicines in older adults. *Patient Educ Couns*. 2010; 83: 265-268.

10. Fjeldsoe B, Marshall AL, Miller YD. Behavior change interventions delivered by mobile telephone short-messaging service. *Am J Prev Med*. 2009; 36: 165-173.

11. Krishna S, Boren SA, Balas EA. Healthcare via cell phones: A systematic review. *Telemed J E Health*. 2009; 15: 231-240.

12. Free C, Phillips G, Galli L, 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: e1001362+.

13. Armstrong AW, Watson AJ, Makredes M, et al. Text-message reminders to improve sunscreen use: A randomized, controlled trial using electronic monitoring. *Arch Dermatol*. 2009; 145: 1230.

14. Shetty AS, Chamukuttan S, Nanditha A, et al. Reinforcement of adherence to prescription recommendations in Asian Indian diabetes patients using short message service (SMS)–a pilot study. *J Assoc Physicians India*. 2011; 59: 711.

15. 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: 594-604.

16. Wei J, Hollin I, Kachnowski S. A review of the use of mobile phone text messaging in clinical and healthy behavior interventions. *J Telemed Telecare*. 2011; 17: 41-48.

17. Pop-Eleches C, Thirumurthy H, Habyarimana JP, et al. Mobile phone technologies improve adherence to antiretroviral treatment in a resource-limited setting: A randomized controlled trial of text message reminders. *AIDS*. 2011; 25: 825-834.

18. Horvath T, Azman H, Kennedy GE, Rutherford GW. Mobile phone text messaging for promoting adherence to antiretroviral therapy in patients with HIV infection (review). The Cochrane Collaboration Library: www.thecochranelibrary.com 2012.

19. Misono AS, Cutrona SL, Choudhry NK, et al. Healthcare information technology interventions to improve cardiovascular and diabetes medication adherence. *Am J Manag Care*. 2010; 16: SP82-SP92.

20. Hawkins RP, Kreuter M, Resnicow K, et al. Understanding tailoring in communicating about health. *Health Educ Res*. 2008; 23: 454-466.

21. Petrie KJ, Perry K, Broadbent E, et al. A text message programme designed to modify patients' illness and treatment beliefs improves self-reported adherence to asthma preventer medication. *Br J Health Psychol*. 2012; 17: 74-84.

22. Woolford SJ, Clark SJ, Strecher VJ, et al. Tailored mobile phone text messages as an adjunct to obesity treatment for adolescents. *J Telemed Telecare*. 2010; 16: 458-461.

23. Resnicow K, Davis R, Zhang N, et al. Tailoring a fruit and vegetable intervention on ethnic identity: Results of a randomized study. *Health Psychol*. 2009; 28: 394-403.

24. Garber MC, Nau DP, Erickson SR, et al. The concordance of self-report with other measures of medication adherence: A summary of the literature. *Med Care*. 2004; 42: 649-652.

25. Boker A, Feetham HJ, Armstron A, et al. Do automated text messages increase adherence to acne therapy? Results of a randomized, controlled trial. *J Am Acad Dermatol*. 2012; 67: 1136-1142.

26. Hardy H, Kumar V, Doros G, et al. Randomized controlled trial of a personalized cellular phone reminder system to enhance adherence to antiretroviral therapy. *AIDS Patient Care STDS*. 2011; 25: 153-161.

27. Furberg RD, Uhrig JD, Bann CM, et al. Technical implementation of a multi-component, text messaging-based intervention for persons living with HIV. *J Med Internet Res*. 2012; 1: e17+.

28. Hanauer DA, Wentzell K, Laffel N, et al. Computerized automated reminder diabetes system (cards): E-mail and SMS cell phone text messaging reminders to support diabetes management. *Diabetes Technol Ther*. 2009; 11: 99-106.

29. Singh S, Bertsekas D. Reinforcement learning for dynamic channel allocation in cellular telephone systems. *Adv Neural Inf Process Syst*. 1997; 9: 974-980.

30. Abbeel P, Coates A, Quigley M. An application of reinforcement learning to aerobatic helicopter flight. In: Adv Neural Inf Process Syst 19 2007:87+.

31. Panait L, Tuyls K. Theoretical advantages of lenient q-learners: An evolutionary game theoretic perspective. In: AAMAS 2007.

32. Lihong L, Chu W, Langford J, Schapire RE. A contextual bandit approach to personalized news article recommendation. WWW 2010:661-670.

33. Singh S, Litman D, Kearns M, et al. Optimizing dialogue management with reinforcement learning: Experiments with the NJFun system. *J Artif Intell Res*. 2002; 16: 105-133.

34. Rudary M, Singh S, Pollack M. Adaptive cognitive orthotics: Combining reinforcement learning and constraint-based temporal reasoning. Proc 21 Int Conf Mach Learn 2004:719-726.

35. Boger J, Hoey J, Poupart P, et al. A planning system based on Markov decision processes to guide people with dementia through activities of daily living. *IEEE Trans Inf Technol Biomed*. 2006; 10: 323-333.

36. Murphy S. Customizing treatment to the patient: Adaptive treatment strategies. *Drug Alcohol Depend*. 2007; 88: S1-S3.

37. Lesosky M, McGeer A, Simor A, et al. Effect of patterns of transferring patients among healthcare institutions on rates of nosocomial methicillin-resistant *Staphylococcus aureus* transmission: A Monte Carlo simulation. *Inf Control Hosp Epidemiol*. 2011; 32: 136-147.

38. Beggs CB, Shepherd SJ, Kerr KG. How does healthcare worker hand hygiene behaviour impact upon the transmission of MRSA between patients? An analysis using a Monte Carlo model. *BMC Infect Dis*. 2009; 9: 1-9.

39. Benneyan JC. An introduction to computer simulation in healthcare: Patient wait case study. *J Soc Health Syst*. 1997; 5: 1-15.

40. McHorney CA. The adherence estimator: A brief, proximal screener for patient propensity to adhere to prescription medications for chronic disease. *Curr Med Res Opin*. 2009; 25: 215-238.

41. Zhao X, Villagran MM, Kreps GL, McHorney C. Gain versus loss framing in adherence-promoting communication targeting patients with chronic diseases: The moderating effect of individual time perspective. Health Commun 2011:1-11.

42. Mengden T, Vetter H, Tousset E, et al. Management of patients with uncontrolled arterial hypertension – the role of electronic compliance monitoring, 24-h ambulatory blood pressure monitoring and Candesartan/HCT. *BMC Cardiovasc Disord*. 2006; 6: 36.

43. Rudd P, Ahmed S, Zachary V, et al. Improved compliance measures: Applications in an ambulatory hypertensive drug trial. *Clin Pharmacol Ther*. 1990; 48: 676-685.

44. Schroeder K, Fahey T, Hay AD, et al. Adherence to antihypertensive medication assessed by self-report was associated with electronic monitoring compliance. *J Clin Epidemiol*. 2006; 59: 650-651.

45. Vrijens B, Vincze G, Kristanto P, et al. Adherence to prescribed antihypertensive drug treatments: Longitudinal study of electronically compiled dosing histories. *BMJ (Clin Res ed)*. 2008; 336: 1114-1117.

46. Singh S, Litman DJ, Kearns M, Walker M. Optimizing dialogue management with reinforcement learning: Experiments with the NJFun system. J Artif Intell Res 2000:1-28.

47. Bartow AG, Sutton RS. *Reinforcement learning: An introduction*. New York: A Bradford Book; 1998.

48. Brafman RI, Tennenholtz M. R-max - a general polynomial time algorithm for near-optimal reinforcement learning. *J Mach Learn Res*. 2002; 3: 213-231.

49. Li L, Chu W, Langford J, Schapire RE. A contextual-bandit approach to personalized news article recommendation. WWW 2010; Apr 26-30:661-670.

50. Haynes RB, Ackloo E, Sahota N, McDonald HP, Yao X. Interventions for enhancing medication adherence. Cochrane Database Syst Rev (Online) 2008:CD000011

51. Yeaw J, Benner JS, Walt JG, et al. Comparing adherence and persistence across 6 chronic medication classes. *J Manag Care Pharm*. 2009; 15: 728-740.

52. Herttua K, Taba AG, Martikainen P, et al. Adherence to antihypertensive therapy prior to the first presentation of stroke in hypertensive adults: Population-based study. *Eur Heart J*. 2013; 34: 2933-2939.

53. Bramley TJ, Gerbino PP, Nightengale BS, et al. Relationship of blood pressure control to adherence with antihypertensive monotherapy in 13 managed care organizations. *J Manag Care Pharm*. 2006; 12: 239-245.

54. Rasmussen JN, Chong A, Alter DA. Relationship between adherence to evidence-based pharmacotherapy and long-term mortality after acute myocardial infarction. *JAMA*. 2007; 297: 177-186.