Journal of the American Medical Informatics Association, 28(6), 2021, 1225–1234
doi: 10.1093/jamia/ocab001
Advance Access Publication Date: 28 February 2021
Research and Applications

Research and Applications

Adaptive learning algorithms to optimize mobile applications for behavioral health: guidelines for design decisions

Caroline A. Figueroa,[1,*] Adrian Aguilera,[1,2] Bibhas Chakraborty,[3,4,5] Arghavan Modiri,[6] Jai Aggarwal,[6] Nina Deliu,[6,7] Urmimala Sarkar,[2] Joseph Jay Williams,[6] and Courtney R. Lyles [2]

[1]School of Social Welfare, University of California Berkeley, Berkeley, California, USA, [2]UCSF Center for Vulnerable Populations, Zuckerberg San Francisco General Hospital, San Francisco, California, USA, [3]Centre for Quantitative Medicine, Duke-National University of Singapore Medical School, Singapore, Singapore, [4]Department of Statistics and Applied Probability, National University of Singapore, Singapore, Singapore, [5]Department of Biostatistics and Bioinformatics, Duke University, Durham, North Carolina, USA, [6]Department of Computer Science, University of Toronto, Toronto, Canada, and [7]Department of Statistical Sciences, Sapienza University of Rome, Rome, Italy

*Corresponding Author: Caroline A. Figueroa, MD, PhD, School of Social Welfare, University of California, 102 Haviland Hall, Berkeley, CA 94720-7400, USA (carolinefigueroa8@gmail.com)

Received 7 October 2020; Editorial Decision 21 December 2020; Accepted 7 January 2021

ABSTRACT

Objective: Providing behavioral health interventions via smartphones allows these interventions to be adapted to the changing behavior, preferences, and needs of individuals. This can be achieved through reinforcement learning (RL), a sub-area of machine learning. However, many challenges could affect the effectiveness of these algorithms in the real world. We provide guidelines for decision-making.
Materials and Methods: Using thematic analysis, we describe challenges, considerations, and solutions for algorithm design decisions in a collaboration between health services researchers, clinicians, and data scientists. We use the design process of an RL algorithm for a mobile health study "DIAMANTE" for increasing physical activity in underserved patients with diabetes and depression. Over the 1.5-year project, we kept track of the research process using collaborative cloud Google Documents, Whatsapp messenger, and video teleconferencing. We discussed, categorized, and coded critical challenges. We grouped challenges to create thematic topic process domains.
Results: Nine challenges emerged, which we divided into 3 major themes: 1. Choosing the model for decision-making, including appropriate contextual and reward variables; 2. Data handling/collection, such as how to deal with missing or incorrect data in real-time; 3. Weighing the algorithm performance vs effectiveness/implementation in real-world settings.
Conclusion: The creation of effective behavioral health interventions does not depend only on final algorithm performance. Many decisions in the real world are necessary to formulate the design of problem parameters to which an algorithm is applied. Researchers must document and evaulate these considerations and decisions before and during the intervention period, to increase transparency, accountability, and reproducibility.
Trial Registration: clinicaltrials.gov, NCT03490253.

Key words: telemedicine, machine learning, algorithms, implementation science, behavioral medicine

INTRODUCTION

Mobile health applications (apps), such as smartphone and text-messaging interventions, have proven effective in eliciting beneficial health outcomes, including better mental health management,[1,2] weight loss,[3] and increased physical activity.[4–6] However, only a small percentage of users use behavior change apps over a long period of time,[7] and mobile applications have not become part of routine medical care.[8,9] Engagement is particularly low for unsupported interventions.[10]

One explanation for low retention and declining effectiveness of apps is that they are not responsive enough to users' changing needs. More recently, an increasing interest in machine learning to optimize digital behavioral health interventions has emerged. Delivery via smartphones allows personalization of these interventions to individuals' preferences and needs.[11] Using machine learning, the content of interventions can also dynamically adapt to changes in participant behavior over time to maximize outcomes.[12]

Reinforcement learning (RL), a subfield of machine learning, is a powerful method to use in various healthcare settings because it can optimize sequences of decisions.[13,14] Several studies have started using RL for optimizing the delivery of text messaging.[15,16] For example, using simulations, Piette et al showed that using RL to optimize text-messaging for medication adherence could produce a 5%–14% increase in adherence, could predict medication barriers, and detect when messages were sent too frequently.[15] RL algorithms for mobile health make predictions based on incoming participant data and use these to make decisions for individuals (eg, what message should the participant receive and when). As more information is collected over time, the algorithm improves its predictions and, hence, makes more effective decisions. A previous mobile health study elucidated that RL algorithms can learn new strategies over time to maximize physical activity.[17] The algorithm altered its decision-making strategy when participants changed their exercise behavior (eg, walked less) because the weather worsened.[17]

However, the use of RL presents multiple challenges in the real world. A systematic review on machine learning in mobile health (beyond solely RL) identified only a few randomized controlled trials (RCTs), the highest level of evidence in clinical medicine,[18] and a lack of studies in clinical practice settings.[12] Guidelines for designing and using these algorithms in clinical settings are needed. For instance, as opposed to simulation studies, clinical studies may have unforeseen difficulties (such as data errors), involve a large interdisciplinary team, and need to be executed within a limited timeframe.

Outlining the challenges and decisions to make throughout the process of algorithm development for a clinical RCT increases transparency, replicability, and likely outcomes of behavior health studies using RL. Further, identifying and solving issues related to differences in scientific strategies of disciplines will help to increase the productivity of interdisciplinary collaborations.[19]

We recently started the Diabetes and Mental Health Adaptive Notification Tracking and Evaluation (DIAMANTE) RCT: a smartphone application that uses RL to optimize physical activity text-messaging in underserved patients with diabetes and depression.[20]

We analyzed study notes of a 1.5-year design process of implementing an RL algorithm: a collaboration between computer scientists, behavioral scientists, physicians, psychologists, and statisticians. We discussed the challenges of developing these algorithms for mobile health in real-world settings and the solutions we implemented. This provides guidelines for decision-making that can be used by other clinical and healthcare researchers.

METHODS

The DIAMANTE study

DIAMANTE is a 6-month physical activity text-messaging study, which sends individuals motivational text messages to help them increase their physical activity. The DIAMANTE study is an RCT with 3 groups (uniform random [UR], RL, and a control group). Phone-pedometers passively collect daily step counts on participants' personal phones. The study recruits low-income English- and Spanish-speaking patients with depression and diabetes served in a safety net setting. In a user-centered design period, we developed our motivational messages based on a cognitive behavioral framework, Cognition, Opportunity, Motivation, and Behavior (COM-B).[21] This process included qualitative interviews, usability phases, and crowdsourcing to categorize the messages.[22] In these phases, we identified issues that may harm user engagement related to missing data due to internet connectivity problems, server errors in sending out messages, and participants' low technical skills to access the app on their phone and transmit their data.[22]

Figure 1. Shows a schematic overview of our study and algorithm. The study is registered on clinicaltrials.gov: NCT03490253 and described elsewhere.[20]

Algorithm

The DIAMANTE study is an RCT with 3 groups (UR, RL, and a control group). In the RL group, the algorithm estimates the optimal combination of message category and timing. It thus picks the combination that will likely maximize the increase in steps for every participant in the upcoming day. In the UR group, participants receive the same messages, but they are microrandomized. This means that messages from different categories and timings are delivered with equal probabilities. This allows us to assess the causal effects of intervention components and how they change over time.[23] Microrandomization is different from regular randomization because interventions (here text messages) are repeatedly (eg, daily) randomized *within* participants.

The algorithm chooses the types of messages from different categories, their frequency, and delivery time period. We assumed that our reward variable (ie, the daily change in steps), is a linear function of contextual variables, action variables, and interactions between actions and action-contextual variables (see Figure 1). The model contains contextual variables for each participant. These include time-fixed variables, such as demographics and clinical characteristics, and time-varying variables, such as day of the week and day of study. See Supplementary Material for the list of contextual variables and algorithm details.

Data sources to identify challenges and solutions

Over the 1.5 years of the project, the team kept track of the research process using a combination of Google Docs tools, Whatsapp messenger communication, and video conferencing (Zoom/Skype). Further, the team convened to discuss, categorize, and code critical challenges and solutions on numerous occasions. CAF and JJW took

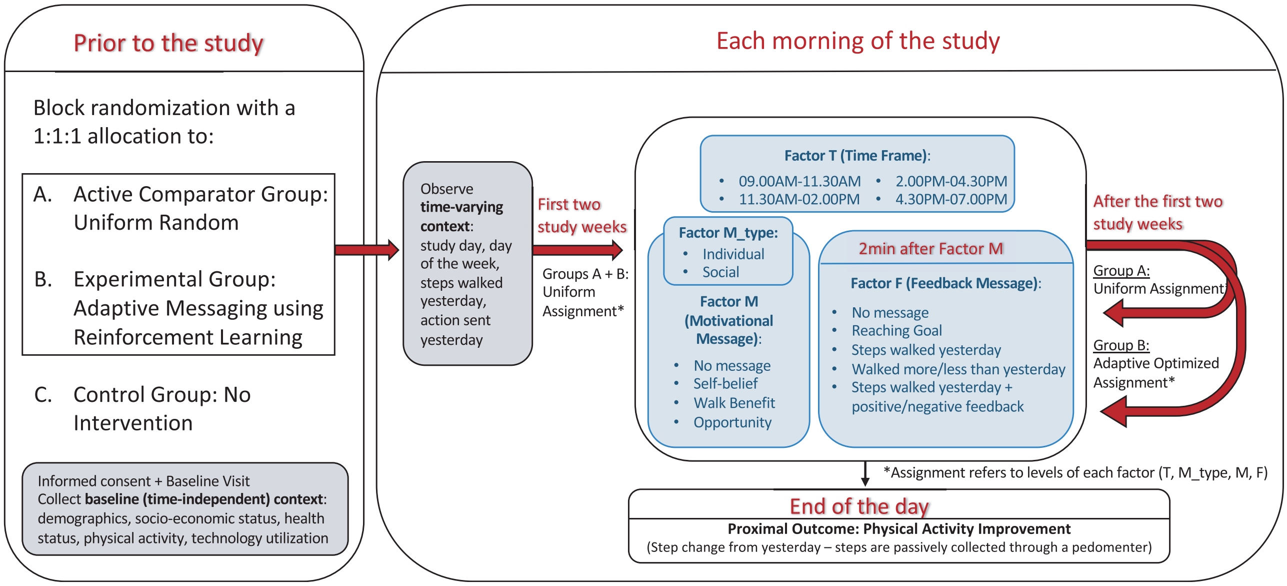

Figure 1. The DIAMANTE study is a randomized controlled trial with 3 groups (uniform random, reinforcement learning, and a control group). In the reinforcement learning group each morning, the algorithm evaluates which messages, delivered at what time period, will likely increase steps for every participant in the upcoming day. The algorithm training data consist of the historical data of all participants (contextual variables), which include time-varying variables, and select clinical/demographic data to improve prediction abilities.[20] Our action space is defined by the 5x4x4x2 DIAMANTE factorial design (5 intervention options for a "feedback" message and 4 intervention options for a "motivational" message, including the "no-message" category, 4 options for the time frame, and 2 social categories [individual or family]).

field notes and recorded audiotapes of key discussions, which were transcribed and coded.

CAF reviewed the initial list of challenges and coded the documents and recordings, which were revised by AA and CRL. We classified and grouped challenges to create thematic topic process domains.

RESULTS

Over the 1.5 year time period, the behavioral science team (Behavioral scientists, physicians and psychologist from the University of California Berkeley and San Francisco) met once per week. Additionally, AA, CAF, and RHR met weekly with the developer CK from January 2019 onwards. Starting February 2019, the data science team (computer scientists and statisticians from the University of Toronto, the National University of Singapore, and the University of Rome La Sapienza) met with a member of the behavioral science team (CAF) on a weekly basis until April 2020. JJW and ND held independent meetings with BC. JJW also held internal team meetings. Finally, the team, including CAF, JJW, JA, ND, and AM discussed the project over WhatsApp messenger.

We coded 119 pages of notes and 82 pages of exported WhatsApp conversations from February 2, 2019 to May 7, 2020. Further, we transcribed around 7 hours of data of key meetings, using an online automatic transcription service (otter.ai) combined with our own transcription where the automatic transcription failed.

Nine challenges emerged which we divided into 3 major themes: defining the model for decision-making; data handling/collection in real time; and appropriate algorithm performance vs effectiveness/implementation issues. These challenges are shown in Figure 2. We describe challenges, considerations, and solutions in more detail in Table 1.

Potential challenges during the design process and their solutions

Choosing a learning algorithm

Standard RL algorithms may perform poorly with the limited data collected in mobile health studies: treatment (here, text messages) is provided up to a few times per day. We chose algorithms for contextual multi-armed bandit (MAB) problems: a problem of deciding which arm of an experiment to try when the goal is maximizing reward from a distribution with unknown parameters. The MAB framework has been applied in diverse decision-making problems including online advertisements and has been proposed for clinical trials.[24]

We used these algorithms based on our previous work and because they might be particularly effective for mobile health.[13,25] MAB algorithms simultaneously attempt to acquire new knowledge by exploring the different intervention options (here, text messages), and optimize decisions based on acquired knowledge (eg, which text messages led to a positive reward before).[26] Each intervention option is associated with a different reward function, also depending on participants' *context*: here participant characteristics (eg, age, gender) or daily time-varying variables (eg, day of the week). Specifically, we employed Thompson sampling (TS) with Bayesian Linear Regression. TS (also called "posterior sampling" or "probability matching") is an algorithm for solving multi-armed bandit problems. Using a Bayesian approach, it tries to balance between *exploration*, choosing a strategy that would provide more information (eg, testing a new type of text message) versus *exploitation*, making the current optimal predicted decision (eg, sending the text message that was previously associated with the higher reward). TS has been widely used in real-world applications, including other mobile health studies,[16] showing promising results (also with small amounts of data).[27]

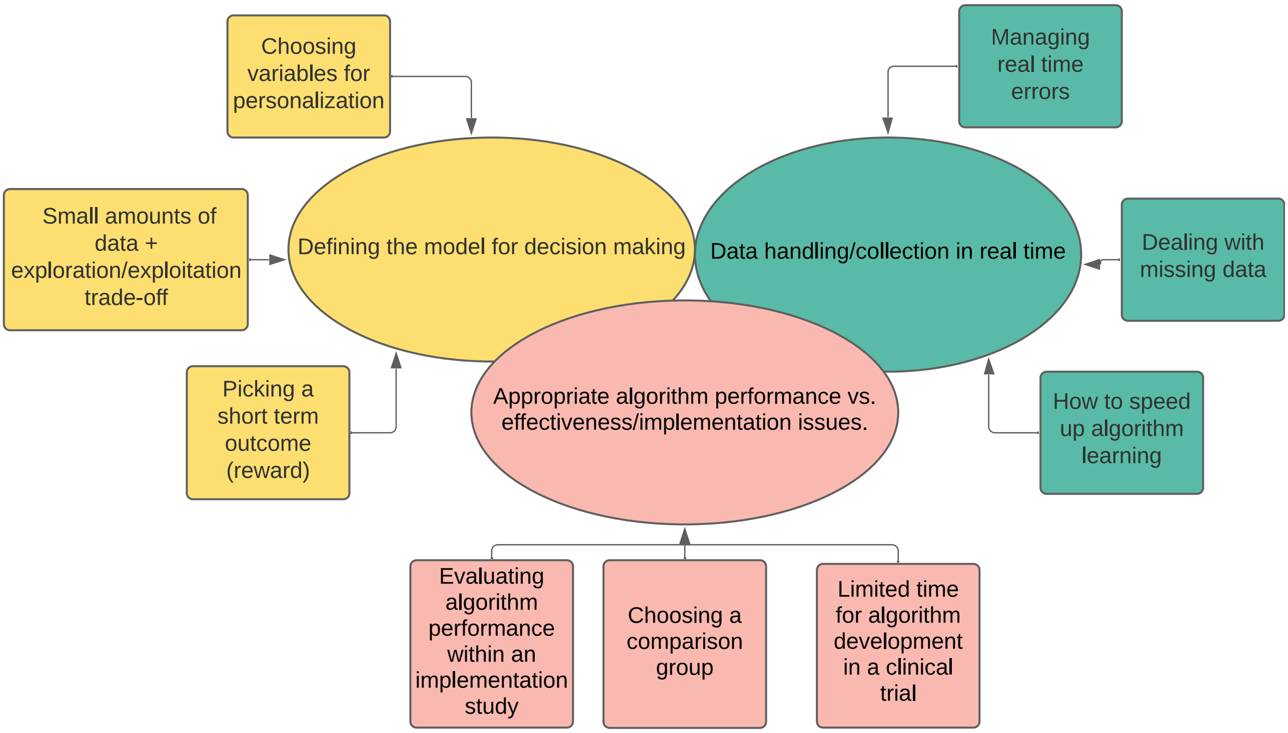

Figure 2. Challenges and considerations when using reinforcement learning in clinical mobile health studies.

MABs have been applied in education research[28] and mobile behavioral interventions, including physical activity[29,30] and sleep.[31] Using simulations, we previously showed that contextual MABs for educational technology interventions resulted in better student outcomes than noncontextual MABs or randomization.[32]

MAB algorithms have limitations. For instance, they are slow to adapt to changing circumstances due to external factors, including the weather or new illness.[13] Additionally, because they take only short-term rewards into account, they are not optimal for maximizing long-term outcomes. However, we expect that slight increases in walking from day to day may result in changes in habits[13] and, thereby, increases in overall steps over the duration of the study (6 months).

Choosing design variables

The DIAMANTE study collects a wealth of contextual variables at baseline (Supplementary Table 1), but guidelines for choosing variables to include in the model are lacking.

We decided on the following approach:

- We initially include all contextual variables that may impact physical activity behavior, based on our (incomplete) empirical knowledge, to avoid missing potential important variables. To provide regularization, we use Bayesian linear regression with a normal-inverse gamma prior to the regression coefficients (mean and variance). This keeps coefficients small and minimizes overfitting.[33]
- During the study, we will evaluate the model every 3 months to improve performance through an iterative process. This includes assessing if we should remove certain terms from the regression model (ie based on high correlations), or choose a different type of regression method (ie, LASSO-related regression, which automatically selects variables by removing predictors from the model[34]) We expect that algorithm tweaks will become less necessary as the study progresses.

Choosing a reward variable

There are no guidelines for choosing the reward variable: the model's outcome/feedback. In testing phases of our study, users indicated that receiving messages in the morning motivated them to be active later in the day, varying with their schedules. Therefore, we chose the change in steps within a 24-hour time window. This is a trade-off, as this longer time period could introduce noise, but a short time window may miss meaningful activity. The "correct" reward thus depends on the purpose of the study. Future work should evaluate algorithm performance for various types of reward variables.

Dealing with missing data

Our passive step-data collection relies on participants' built-in phone pedometers as opposed to physical activity trackers. This facilitates easier implementation but perhaps less reliable data.[35] Participants may forget to carry their phones, or lack Internet connection necessary to transmit their data, leading to faulty decision-making of the algorithm. Further, in our testing phases, participants received messages remarking they had walked 0 steps when their data was missing. These errors could lead to lower engagement and higher dropout.

To solve this issue, we took the following approach:

- We consider 0 steps faulty, since it is unlikely for a participant to walk 0 steps daily. When 0 steps are detected, participants do not receive messages.
- We code 0 steps as "NA" in the training data. This avoids potential biases in the estimation process.
- We use the last observation carried forward technique,[36] using participants' last available steps, for dealing with missing reward

Table 1. Challenges, considerations and solutions

Themes	Challenge	Considerations	Solution
Defining the model for decision making	1. Choosing a learning algorithm	We considered algorithms that maximize short-term outcomes and deal with small amounts of data. We examined different types of sampling methods used previously in mobile health studies.	We use algorithms for contextual multi-armed bandit problems (MAB) and employ Thompson Sampling (TS), successfully applied in mobile health before. This algorithm can deal with small amounts of data and address the exploration/exploitation trade-off.
	2. Choosing design variables (contextual and action variables)	We employ a 5x4x4x2 design for the action variables, and collect demographic, time-varying and clinical characteristics for personalization (contextual variables). To deal with this large number of terms, we considered: • Selecting terms based on our behavioral scientists' prior knowledge and hypotheses • Selecting terms based on prior analysis results of pilot studies • Using regression with regularization • Including all variables of possible interest	In the absence of reliable estimates at the start of the study (not enough data from pilot phases), we included all variables of interest to avoid missing important variables. To provide regularization by shrinking coefficients, we use Bayesian linear regression with a Normal-Inverse Gamma prior on the regression coefficients. During the study, we will evaluate the model once per 3 months to assess if it would be useful to remove certain terms from the regression model, and/or employ a different regularization technique.
	3. Choosing the reward variable	Physical activity studies can have many different reward variables. Possible reward variables we discussed include: • The total increase in steps within a number of hours after sending a message. • The total increase in steps from one day to the next • The total increase in steps after sending a message until sending the next message	We chose the increase in step from one day to the next (24 hour time window, from 00:00 to 23:59 pm). Our overall aim is to motivate participants to increase steps during the day, regardless of *when*. Some participants noted that the messages received early in the day motivated them to plan their activity at a later time during the day.
Data handling/collection in real time	4. Dealing with missing data	Questions we considered were: • How should the model handle missing data? • What text-messages should be sent out to participants when we register missing data?	We used the last observation carried forward (LOCF) technique for missing reward data (steps). Feedback messages will not be sent to participants when there is missing data. Instead, the system sends a message asking them to open the app (which pulls their steps). The researchers will call participants when there are >2 days of missing data.
	5. Dealing with real time algorithm errors	Dealing with data issues in real time, such as errors in execution of the script or in the algorithm training data.	Our developer sends a log of data errors to the researchers on a daily basis. Researchers use the app to identify issues in real time. The researchers check the data output every 2-3 months for missing values and micro-randomization errors.
	6. Speed up algorithm learning	To speed up the learning of the algorithm, it is recommended to have data to inform the prior distributions of the algorithm.	We took the following approach: • Conduct simulation studies using pilot data (n = 10). Based on our simulations, we confirmed that

(continued)

Table 1. continued

Themes	Challenge	Considerations	Solution
		Our study was however not designed with a period of large-scale data collection and experimentation with the algorithm before the start of the trial.	TS has a lower bias in estimation of reward with an initial uniform random policy. • All participants receive 2 weeks of uniform random policy, and then by using the data collected during this period, we update priors for parameters of TS.
Appropriate algorithm performance vs. effectiveness/implementation issues.	7. Comparison group in the context of a Randomized Controlled Trial	In the original protocol we planned to compare the RL Algorithm to a pre-specified text messaging program, that delivers brand-new messages with educational concepts. But, throughout our design process it became apparent that this is not a clean comparison, as both arms would receive different types of text-messages.	We compare the RL condition to a uniform random (UR) Policy. We removed the program with different content. Both arms will receive the same types of messages, but the UR policy picks the messages with equal probability; the RL condition will learn a decision policy. This allows us to scrutinize the benefit of RL.
	8. Evaluating algorithm performance within an implementation study	We had to find a design to optimize algorithm performance, and maximize usability in low resource settings, and to evaluate performance and usability.	We chose a three-arm (TS, UR and control) trial to balance the need of efficacy vs. effectiveness. By comparing the UR (micro-randomized) messages to the control group we can evaluate the effect of our text-messages irrespective of the use of RL. Outcomes of only the UR condition can be used to inform future RL algorithms.
	9. Limited time for algorithm development in a clinical trial	We had to discover a balance between focusing on algorithm performance versus implementability of the app. We knew we would have limited time for algorithm preparation, a slow roll-out for participant recruitment (12-15 people per month) and less data from trial participants upfront. We had to implement a variety of strategies simultaneously.	• In the preparation phase, we tested the app and algorithm in the study team, user design phases and simulations. We balanced content and structure liked by participants, with design to maximize algorithm performance to yield an externally valid approach. • We ran preliminary quality data checks on pilot data. • We analyzed the algorithm data from the study team members and a small group of participants in a pilot.

data. In simulations, we will examine if various methods of imputation for longitudinal data improve our algorithm performance.[37]

- We send participants a message asking them to open their app to automatically transmit their data.
- We contact participants by phone if we have not received their step data for > 2 days.

Addressing algorithm errors in real time

In user-testing phases, the app did not always collect steps, and messages sent by our online server were occasionally inconsistent with messages registered in our data export file.

To prevent these errors throughout the study, we took the following approach:

- The researchers receive an automated daily log of data errors (such as 0-step measurements).
- Throughout the study, we compare the data export (which messages participants received according to our analysis file) to 1) the messages logged on our online server, HealthySMS, and 2) the messages we receive ourselves, through our enrollment in the program as a continuous internal test.
- We conduct preliminary checks of the data quality every 3 months. These consist of checking for missing values, assessing microrandomization errors (eg, check the uniform assignment

in the UR group), and evaluating consistency among collected data.

Speeding up algorithm learning with limited time available

The rate of algorithm learning slows with sparse data. In our study, we only enroll approximately 10–15 participants per month—and even fewer during the COVID-19 pandemic. To speed up algorithm learning, we implemented a 2-week UR policy at the start of the study for every participant in the RL arm. UR data can be used to create informed prior distributions for the TS condition. This way, algorithm decisions are based both on prior knowledge and incoming data, which increases the speed of algorithm personalization.[38] We lacked recommendations on how long to collect uniform data before the learning phase. Based on our experience, most participants who drop out do so in the first month. Therefore, we settled on 2 weeks of UR, before switching to TS, to maximize the probability that most participants will receive RL for at least 2 weeks. In simulation studies using pilot data from our user design phases ($n = 10$), we confirmed that TS has a lower bias in estimation of reward with an initial UR policy.

Choosing a comparison group

We originally planned to compare RL to an arm with fixed content derived from the National Diabetes Prevention Program.[39] Although this approach was most comparable to existing health education interventions, we decided to cleanly evaluate the effect of RL by changing the comparison group to UR. Both arms will receive the same types of messages, but UR picks the messages with equal probability. The RL condition dynamically adapts treatment, allowing us to assess if sending messaging using a RL decision-making algorithm is superior to choosing message categories and timings at random.

Evaluating algorithm performance within an implementation study

The DIAMANTE study is a hybrid effectiveness and implementation trial, set up to ensure the effectiveness of the app in a real-world setting. Hybrid designs combine effectiveness and implementation research to reduce the time from initial concept to a working product.[40] This is important because traditional designs often result in efficacious interventions that are not effective in real-world studies, particularly for digital interventions.[41]

Our team navigated between making decisions to optimize algorithm performance and maximize usability. For instance, sending messages more than once a day with shorter reward periods may improve algorithm performance[16,17] but may also decrease user engagement.[42] Because our target users did not frequently use apps or texting, and were therefore at higher risk for dropout,[43–45] we prioritized decisions that would benefit user engagement where possible.

To further examine clinical effectiveness and implementation, we implemented a 3-arm trial design, including RL, UR, and a control group to balance the need of evaluating algorithm performance vs effectiveness and implementation of the overall intervention. Doing so, we will also be able to evaluate the effect of our text messages irrespective of the use of RL by comparing the UR (microrandomized) messaging to the control group.

Limited time for algorithm development within the context of a clinical trial

Typical RL-algorithm research involves several preparation phases, which improve algorithm performance but may take years to complete before a clinical trial. Here, we were only able to employ a preparatory phase of 9 months before the start of the RCT. Most of the limited years of study funding were dedicated to the clinical trial. Given the limited amount of time for research on algorithm performance we:

- Conducted simulation studies for debugging and defining model parameters;
- Ran preliminary quality data checks on pilot data;
- Analyzed the algorithm data from the study team members and a small group of participants in a pilot. This testing revealed crucial errors that we could fix before deploying the algorithm.

DISCUSSION

We described the challenges, decisions, and solutions of designing RL algorithms to personalize mobile health applications in real-world settings. We use the design process of a physical activity study (DIAMANTE) as an example. Qualitative analyses from 1.5 years of study notes showed that the most important decisions and challenges were related to the choice of the model and design variables for decision-making, handling missing data and algorithm errors in real time, and maintaining a balance between intervention implementation and optimal model performance. These issues need to be taken into account and should be documented during the design process of RL algorithms for clinical studies.

Approach with multiple phases

Despite the limited time to develop our RL algorithm, we employed a multiphased approach. This included simulation studies, pilot user testing, and testing of the mobile application within our own study team. We recommend that all researchers using RL or other types of machine-learning to personalize digital health interventions employ multiple phases, which can happen simultaneously, in preparation of a rollout with clinical participants. This is crucial for identifying and fixing algorithm errors. Because of the complexity of the design process, developing these models requires a multidisciplinary team with both deep technical expertise and profound knowledge of the clinical population of interest.

Further, the RL algorithm design process should not stop when the clinical study starts. Instead, decisions should continuously be evaluated, and algorithm development should work through an iterative process. Frequent data checks, automated reports about missing data, and internal testing are essential throughout the study in its entirety.

Microrandomization

Another important lesson is that an RL study must have an adequate comparison condition to disentangle the performance of the algorithm and the characteristics of the digital intervention overall on clinical outcomes. Such a comparison, as we choose here, is contrasting RL to UR, in which messages are microrandomized (ie, delivered with equal probability). This will allow us to quantify the benefit of "adaptive tailoring" using RL. We also used a period of UR for all participants to inform priors for the algorithm, with the aim to speed up learning with sparse data, which is an important consideration for mobile health. Researchers may also choose to include a period of microrandomization in order to determine decision rules[23] (eg, at what times to send messages based on participants' availability [not performed in this study]).

Increasing engagement through RL

The type and delivery frequency of text messages will adapt over time throughout the study based on data collected from participants every day. This learning algorithm aims to maximize the outcome (increases in steps) learning and updating over time based on incoming data. More commonly, digital health studies only tailor their content in a user-centered design process to the needs, wishes, and norms of a group of individuals. In this study, we tailor both the content as a whole through a user-centered design process and the text-messaging delivery on a daily basis using RL. We hypothesize that this approach will increase participant engagement and thereby will be more effective. We will assess participant engagement by examining the times they accessed the app and read the messages, how they rated the app's usability, and their qualitative opinions.[20]

The importance of developing guidelines for machine learning in mobile health

A framework article discussing the machine learning literature in health argued that the field lacks transparency, clear reporting, exploration for ethical concerns, and demonstrations of effectiveness.[46] While several studies have discussed RL algorithm performance for mobile health, (for example in simulations), few discuss all the steps needed to develop these algorithms for clinical studies. Because this is a novel field, machine learning algorithms used in applied health settings often undergo less scrutiny compared to other clinical interventions.[46] Additionally, because of the excitement around artificial intelligence (AI), some have warned that digital medicine must avoid a crisis of reproducibility like that found in other biomedical fields.[47] Recent RCT reporting guidelines for AI studies for clinical decision-making have begun to emerge.[48] Here, we provide guidelines specifically for the algorithm design part of mobile behavioral health studies.

We recommend that all studies using machine learning to optimize digital health interventions document their decision-making process and identify critical issues and challenges they encountered. This further avoids the "black box" problem of not knowing how and why algorithms are making decisions.

Issues around choosing a model for decision-making also need to be explored more. Here we chose an algorithm for contextual multi-armed bandit problems, as this algorithm may be particularly suitable for mobile health studies. There is a lack of research that compares the effectiveness of different RL models and assesses what kind of problems within mobile health they should be applied to.

Similarly, we chose changes in daily steps as our reward. Other physical activity studies using an RL algorithm made different choices. One study used 30-minute steps after participants received a motivational message (up to 5 times per day to increase short bouts of physical activity).[16] Other research chose the increase in activity since the last motivational text message.[17] Algorithm performance with various reward functions needs to be explored.

Strengths

To our knowledge, this is the first study to describe RL algorithm design decisions in the context of a multidisciplinary collaboration for mobile health clinical studies in the real world. Notably, we conduct this work in a low-income ethnic minority population for whom the greatest health disparities exist yet where novel methods are not often designed.[43,49,50] This work brings the challenge of balancing decisions that boost algorithm performance and those that maximize usability. Clinical studies with low-income populations are complex because of data errors, low-tech skills of users, working within a large interdisciplinary team, and executing studies within a limited timeframe (eg, related to funding). Many of these issues are less relevant in simulation studies or work with convenience samples. This article can be used as a framework of considerations for other interdisciplinary teams working, or wanting to work, in this space.

Limitations

Researchers working with other populations and/or health problems may encounter issues not described here. Further, we provide a framework of decision-making, but, because this is an evolving field, we cannot be certain that our choices are optimal. Detailed analyses on the final dataset, in combination with results from other studies, will provide these answers. Additionally, here we did not discuss considerations such as privacy and algorithm bias in detail, but these issues must be further explored. Further, user ratings and experiences of the content are not included as variables in the current model. Future work should focus on incorporating engagement into RL algorithms. Because user engagement can be quantified in many different ways,[51] future guidelines should also define consistent engagement measurements for studies to be comparable. Finally, here we measured steps using participants' pedometers on their personal phones to facilitate real-world implementation. The reliability of pedometers may depend on the phone model[52] and where participants carry their phone (eg, closer to the body may be more reliable).[53] Future implementation work should compare the use of wristbands and phone pedometers to measure physical activity in low-income populations.

CONCLUSION

Creating effective behavioral health interventions using RL involves many decisions beyond evaluating algorithm performance. These considerations need to be documented and evaluated before and during the intervention period to increase transparency, accountability, and replicability. As the application of machine learning into digital healthcare interventions increases, we need effective collaborations between different disciplines to do this work well in real-world settings.

FUNDING

The DIAMANTE trial is funded by an R01 to Dr Adrian Aguilera and Dr Lyles, 1R01 HS25429-01 from the Agency for Healthcare Research and Quality (AHRQ). The sponsor did not have any role in the study design; collection, management, analysis, and interpretation of data; writing of the report; the decision to submit the report for publication nor ultimate authority over any of these activities.

AUTHOR CONTRIBUTIONS

CAF reviewed the initial list of challenges and coded the documents and recordings, which were revised by AA and CRL. CAF and JJW took field notes and recorded audiotapes of key discussions, which were transcribed and coded. All authors convened to discuss, categorize, and code critical challenges and solutions on numerous occasions. CAF wrote the first draft of the article. All authors revised the manuscript for relevant scientific content and approved the final version of the manuscript.

ETHICS COMMITTEE APPROVAL

The Institutional Review Board at the University of California San Francisco approved this study (IRB: 17-22608).

SUPPLEMENTARY MATERIAL

Supplementary material is available at *Journal of the American Medical Informatics Association* online.

DATA SHARING STATEMENT

Curated technical appendices, statistical code, and anonymized data supporting the conclusions of this article are available from the authors on request.

ACKNOWLEDGMENTS

We would like to acknowledge Chris Karr who helped develop and manage the automated text-messaging service, passive data collection, and RL algorithm.

CONFLICT OF INTEREST STATEMENT

None declared.

REFERENCES

1. Firth J, Torous J, Nicholas J, Carney R, Rosenbaum S, Sarris J. Can smartphone mental health interventions reduce symptoms of anxiety? A meta-analysis of randomized controlled trials. *J Affect Disord* 2017; 218: 15–22.
2. Firth J, Torous J, Nicholas J, *et al.* The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry*2017; 16 (3): 287–98.
3. McCarroll R, Eyles H, Mhurchu CN. Effectiveness of mobile health (mHealth) interventions for promoting healthy eating in adults: a systematic review. *Prevent Med* 2017; 105: 156–68.
4. Gal R, May AM, van Overmeeren EJ, Simons M, Monninkhof EM. The effect of physical activity interventions comprising wearables and smartphone applications on physical activity: a systematic review and meta-analysis. *Sports Med Open* 2018; 4 (1): 42.
5. Murray JM, Brennan SF, French DP, Patterson CC, Kee F, Hunter RF. Effectiveness of physical activity interventions in achieving behaviour change maintenance in young and middle aged adults: a systematic review and meta-analysis. *Soc Sci Med* 2017; 192: 125–33.
6. Roberts AL, Fisher A, Smith L, Heinrich M, Potts HW. Digital health behaviour change interventions targeting physical activity and diet in cancer survivors: a systematic review and meta-analysis. *J Cancer Surviv* 2017; 11 (6): 704–19.
7. Baumel A, Muench F, Edan S, Kane JM. Objective user engagement with mental health apps: systematic search and panel-based usage analysis. *J Med Internet Res* 2019; 21 (9): e14567.
8. Lyles CR, Handley MA, Ackerman SL, *et al.* Innovative implementation studies conducted in US Safety Net Health Care Settings: a systematic review. *Am J Med Qual* 2019; 34 (3): 293–306.
9. Rowland SP, Fitzgerald JE, Holme T, Powell J, McGregor A. What is the clinical value of mHealth for patients? *NPJ Digit Med* 2020; 3 (1): 1–6.
10. Weisel KK, Fuhrmann LM, Berking M, Baumeister H, Cuijpers P, Ebert DD. Standalone smartphone apps for mental health—a systematic review and meta-analysis. *NPJ Digit Med* 2019; 2 (1): 1–10.
11. Nahum-Shani I, Smith SN, Spring BJ, *et al.* Just-in-Time Adaptive Interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. *Ann Behav Med* 2018; 52 (6): 446–62.
12. Triantafyllidis AK, Tsanas A. Applications of machine learning in real-life digital health interventions: review of the literature. *J Med Internet Res* 2019; 21 (4): e12286.
13. Rabbi M, Klasnja P, Choudhury T, Tewari A, Murphy S. Optimizing mHealth interventions with a bandit. In: *Digital Phenotyping and Mobile Sensing*. Springer; 2019: 277–91.
14. Yu C, Liu J, Nemati S. Reinforcement learning in healthcare: A survey. arXiv preprint arXiv:1908.08796; 2019.
15. Piette JD, Farris KB, Newman S, An L, Sussman J, Singh S. The potential impact of intelligent systems for mobile health self-management support: Monte Carlo simulations of text message support for medication adherence. *Ann Behav Med* 2015; 49 (1): 84–94.
16. Liao P, Greenewald K, Klasnja P, Murphy S. Personalized heartsteps: a reinforcement learning algorithm for optimizing physical activity. *Proc ACM Interact Mob Wearable Ubiquitous Technol* 2020; 4 (1): 1–22.
17. Yom-Tov E, Feraru G, Kozdoba M, Mannor S, Tennenholtz M, Hochberg I. Encouraging physical activity in patients with diabetes: intervention using a reinforcement learning system. *J Med Internet Res* 2017; 19 (10): e338.
18. Sibbald B, Roland M. Understanding controlled trials. Why are randomised controlled trials important? *BMJ* 1998; 316 (7126): 201.
19. Lach D. Challenges of interdisciplinary research: reconciling qualitative and quantitative methods for understanding human–landscape systems. *Environ Manag* 2014; 53 (1): 88–93.
20. Aguilera A, Figueroa CA, Hernandez-Ramos R, *et al.* mHealth app using machine learning to increase physical activity in diabetes and depression: clinical trial protocol for the DIAMANTE Study. *BMJ Open* 2020; 10 (8): e034723.
21. Michie S, Van Stralen MM, West R. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implement Sci* 2011; 6 (1): 42.
22. Gómez-Pathak LAA, Williams JJ, Lyles CR, *et al.* Combining user centered design and crowdsourcing to develop messaging content for a physical activity smartphone application tailored to low-income patients. *JMIR* 2020; 21177.
23. Klasnja P, Hekler EB, Shiffman S, *et al.* Microrandomized trials: an experimental design for developing just-in-time adaptive interventions. *Health Psychol* 2015; 34 (Suppl): 1220–8.
24. Agrawal S. Recent advances in multiarmed bandits for sequential decision making. *informsPubsOnline*. 2019: 167–88.
25. Tewari A, Murphy SA. From ads to interventions: contextual bandits in mobile health. In: *Mobile Health*. Springer, 2017: 495–517.
26. Greenewald K, Tewari A, Klasnja P, Murphy S. Action centered contextual bandits. Advances in neural information processing systems; 2017; 30: 5973.
27. Agrawal S, Goyal N. Analysis of Thompson sampling for the multi-armed bandit problem. In: *Conference on Learning Theory*; June 16, 2012: 1–39.
28. Williams JJ, Rafferty AN, Ang A, Tingley D, Lasecki WS, Kim J. Connecting instructors and learning scientists via collaborative dynamic experimentation. In: *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*; May 6, 2017: 3012–-8.
29. Rabbi M, Aung MH, Zhang M, Choudhury T. MyBehavior: automatic personalized health feedback from user behaviors and preferences using smartphones. In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*; September 7, 2015: 707–18.
30. Tomkins S, Liao P, Yeung S, Klasnja P, Murphy S. IntelligentPooling: Practical Thompson Sampling for mHealth; 2020. https://arxiv.org/abs/2008.01571
31. Daskalova N, Yoon J, Wang Y, Araujo C, Beltran G, Jr, Nugent N, McGeary J, Williams JJ, Huang J. SleepBandits: Guided flexible self-experiments for sleep. In: *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*; Apr 21, 2020: 1–13.
32. Shaikh H, Modiri A, Williams JJ, Rafferty AN. Balancing Student Success and Inferring Personalized Effects in Dynamic Experiments. In EDM; Jan 1, 2019.
33. Marquardt DW, Snee RD. Ridge regression in practice. *Am Statist* 1975; 29 (1): 3–20.

34. Park T, Casella G. The Bayesian lasso. *J Am Statist Assoc* 2008; 103 (482): 681–6.
35. Hurt CP, Lein DH, Jr, Smith CR, *et al.* Assessing a novel way to measure step count while walking using a custom mobile phone application. *PLoS One* 2018; 13 (11): e0206828.
36. Hamer RM, Simpson PM. Last observation carried forward versus mixed models in the analysis of psychiatric clinical trials. *Am Psychiatric Assoc* 2009.
37. Blankers M, Koeter MW, Schippers GM. Missing data approaches in eHealth research: simulation study and a tutorial for nonmathematically inclined researchers. *J Med Internet Res* 2010; 12 (5): e54.
38. Russo D, Van Roy B, Kazerouni A, Osband I, Wen Z. A tutorial on thompson sampling. arXiv preprint arXiv:1707.02038; 2017
39. Albright AL, Gregg EW. Preventing type 2 diabetes in communities across the US: the National Diabetes Prevention Program. *Am J Prevent Med* 2013; 44 (4): S346–51.
40. Curran GM, Bauer M, Mittman B, Pyne JM, Stetler C. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Med Care* 2012; 50 (3): 217–26.
41. Mohr DC, Lyon AR, Lattie EG, Reddy M, Schueller SM. Accelerating digital mental health research from early design and creation to successful implementation and sustainment. *J Med Internet Res* 2017; 19 (5): e153.
42. Eysenbach G. The law of attrition. *J Med Internet Res* 2005; 7 (1): e11.
43. Figueroa CA, DeMasi O, Hernandez-Ramos R, Aguilera A. Who benefits most from adding technology to depression treatment and how? An analysis of engagement with a texting adjunct for psychotherapy. *Telemed J E Health*. 2021; 27 (1): 39–46.
44. Avila-Garcia P, Hernandez-Ramos R, Nouri SS, *et al.* Engaging users in the design of an mHealth, text message-based intervention to increase physical activity at a safety-net health care system. *JAMIA Open* 2019; 2 (4): 489–97.
45. Nouri SS, Avila-Garcia P, Cemballi AG, Sarkar U, Aguilera A, Lyles CR. Assessing mobile phone digital literacy and engagement in user-centered design in a diverse, safety-net population: mixed methods study. *JMIR Mhealth Uhealth* 2019; 7 (8): e14250.
46. Vollmer S, Mateen BA, Bohner G, *et al.* Machine learning and artificial intelligence research for patient benefit: 20 critical questions on transparency, replicability, ethics, and effectiveness. *BMJ* 2020; 368: l6927.
47. Stupple A, Singerman D, Celi LA. The reproducibility crisis in the age of digital medicine. *NPJ Digit Med* 2019; 2 (1): 1–3.
48. Liu X, Rivera SC, Moher D, Calvert MJ, Denniston AK. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI Extension. *BMJ* 2020; 370: m3164.
49. Allen LN, Christie GP. The emergence of personalized health technology. *J Med Internet Res* 2016; 18 (5): e99.
50. Schueller SM, Hunter JF, Figueroa C, Aguilera A. Use of digital mental health for marginalized and underserved populations. *Curr Treat Options Psych* 2019; 6 (3): 243–55.
51. Short CE, DeSmet A, Woods C, *et al.* Measuring engagement in eHealth and mHealth behavior change interventions: viewpoint of methodologies. *J Med Internet Res* 2018; 20 (11): e292.
52. Höchsmann C, Knaier R, Eymann J, Hintermann J, Infanger D, Schmidt-Trucksäss A. Validity of activity trackers, smartphones, and phone applications to measure steps in various walking conditions. *Scand J Med Sci Sports* 2018; 28 (7): 1818–27.
53. Presset B, Laurenczy B, Malatesta D, Barral J. Accuracy of a smartphone pedometer application according to different speeds and mobile phone locations in a laboratory context. *J Exerc Sci Fit* 2018; 16 (2): 43–8.