Artificial intelligence decision-making in mobile health

Marianne Menictas*, Mashfiqui Rabbi*, Predrag Klasnja and **Susan Murphy** (Harvard University, the University of Michigan and the Kaiser Permanente Washington Health Research Institute, USA)

* These authors contributed equally to this article.

It is likely that you or someone you know is affected by a chronic health condition. For example, a staggering six in 10 adults in the USA are currently suffering from a chronic disease (National Center for Chronic Disease Prevention and Health Promotion, 2019). Unfortunately, chronic conditions are not treatable overnight, but they can often be improved by regular incorporation of preventative behaviours (e.g., taking medication, healthy sleeping habits, being physically active, healthy eating, etc.). However, due to the many contingencies that arise in our lives, regular incorporation of healthy behaviours is difficult, and often when we need help in enacting these behaviours, support from clinical professionals is not available.

Smartphones and wearables, however, can provide support anytime, anywhere, just-in-time, by delivering interventions to assist us in enacting healthy behaviours. One of the simplest examples is a text message or notification via our smartphone reminding us to take our medication during the next hour. But a natural question is how we can use large volumes of complex and high-frequency mobile sensor data, user feedback as well as behavioural science to determine when and where it is most helpful in order to provide supportive interventions. Artificial intelligence (AI) provides tools to effectively use these data to address these 'when' and 'where' questions. AI is defined as intelligence demonstrated by machines, such as learning and problem-solving. In mobile health, this means learning from data in order to solve the problem of when and where to deliver interventions to a user. In fact, from a conceptual point of view, AI in mobile health is more accurately an abbreviation for augmented intelligence as the goal of mobile health is to augment human capabilities to enhance self-management of healthy behaviours.

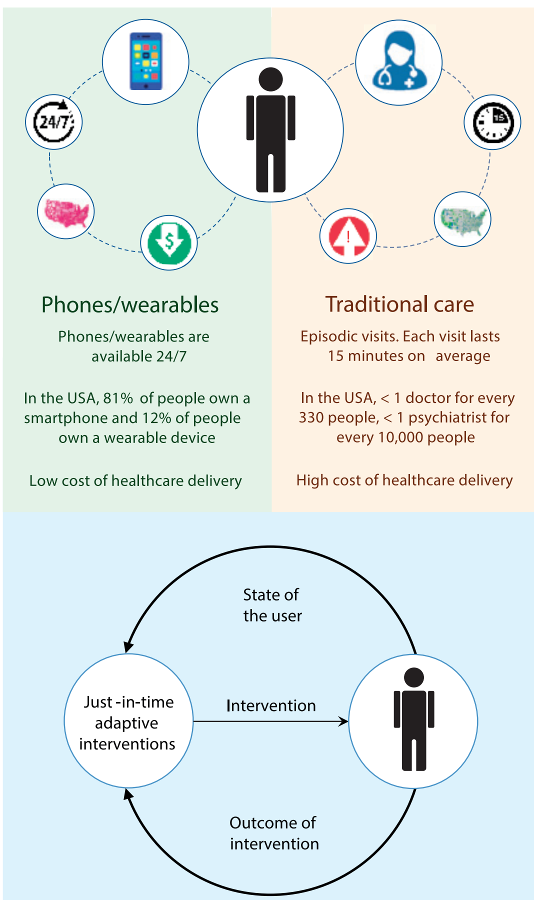

Figure 1. Top panel: comparison between AI decision-making in mobile health and traditional delivery of healthcare (e.g., visits to a clinic). **Bottom panel:** just-in-time adaptive interventions are the next generation of mobile healthcare delivery that is automated, scalable, evidence-driven and inexpensive.

Just-in-time adaptive interventions

AI decision-making in mobile health is most commonly found in just-in-time adaptive interventions (JITAIs). These are mobile health interventions that contain decision rules that take a user's current state as input in order to decide when and what type of intervention to deliver in order to support healthy behaviours. The state of the current user can include data describing the user's internal context such as stress and motivation, external context such as weather, location and so on, as well as information about the user's past behaviour and intervention response. JITAIs use information about the user's state to determine times when the user needs support and when he/she is receptive (i.e., just-in-time).

Furthermore, the decision rules adapt the type of intervention (if any) that the user would most benefit from in the current state.

Consider for example HeartSteps, a JITAI application that aims to help sedentary people reduce inactive behaviour and increase opportunistic physical activity. Paired with a wearable activity sensor, HeartSteps tracks information about the user, such as steps taken per minute, activity, location, weather, time of day and day of the week. Using this dynamically evolving data, HeartSteps' decision rules deliver tailored activity suggestions that are intended to be actionable in the user's current context.

A critical question is how to construct decision rules that deliver the right intervention for each user in their context. One approach is to use behaviour change theory, which describes mechanisms for influencing behaviour change and can be used to narrow down the contexts where specific interventions have a high likelihood of being effective. However, behavioural theory often lacks the granular precision required for JITAI delivery. For example, in the context of HeartSteps, whilst the theory suggests that cues to action work better when they are highly actionable, it provides little guidance on exactly when a cue to action should be delivered (e.g., should the system intervene if the user is currently inactive but was active recently?). Nor is the theory precise in suggesting alternative actions to be taken if messaging fails to be effective or users have found it difficult to be responsive.

AI techniques like reinforcement learning (RL) overcome this lack of precision by learning optimal decision rules from users' data as well as users' responses to the interventions. Importantly, RL provides a continuous learning paradigm that can automatically readjust the decision rules as a user's responsivity evolves. RL can adapt to long-term positive factors such as building habits or negative factors such as habituation—an aspect that is crucial to the success of any technology that aims to help people achieve and maintain healthy behaviours.

A formal introduction to RL is beyond the scope of this article, but in short, an RL algorithm learns which 'actions' to take by continually interacting with an 'environment'. In mobile health, the environment is the user and his/her state (location, mood, recent activity and so on) and the actions are different types of interventions or whether to deliver an intervention at all. The RL algorithm is used to determine which actions are best provided in a given environment in order to maximize a longer-term outcome. This outcome might be, for example, the user's average daily step count over a year or the time to a weight gain of more than 10 lbs. In recent years, RL has experienced rapid development and has been applied successfully to problems in robotic control, elevator scheduling and games such as backgammon, checkers and Go.

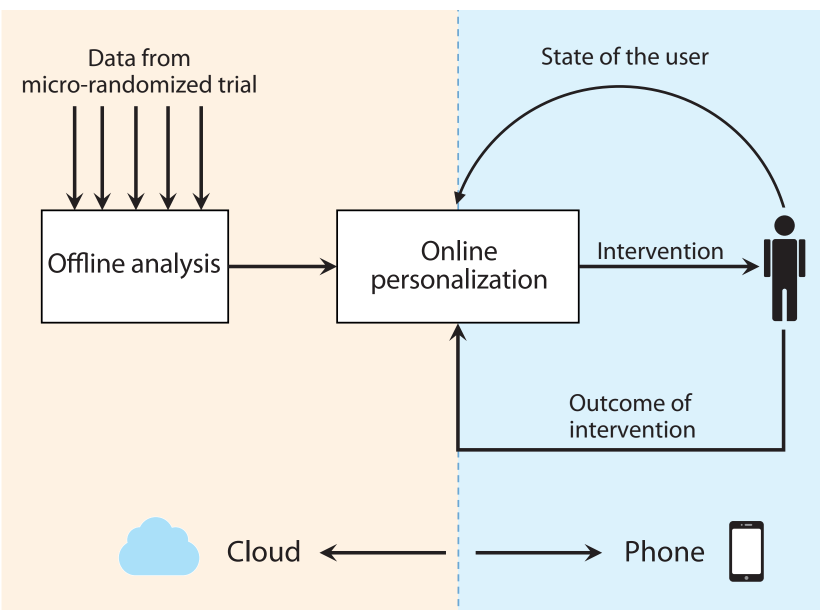

Figure 2. Two-step best practices procedure involving offline analysis and online personalization to deliver effective JITAIs.

How to use AI to construct a JITAI

Unlike the data-rich-simulated worlds of robots and games, constructing an RL algorithm to learn decision rules for an effective JITAI is fundamentally different due to the complexity of human behaviour, the limited amount of time each user will continue to use a poorly performing mobile health application as well as the limited number of users with any particular chronic disorder. In contrast, the computer program, AlphaGo, which beat the 18-time Go world champion, initially learned on 30 million moves from games played by expert human players before it played against itself thousands of times in a simulator. No such prior data or simulation environment exists that can mimic human behavioural responses under repeated health interventions. Nor is it feasible to go to the exorbitant expense required to recruit millions of people in order to conduct experiments to learn a decision rule.

In order to provide effective interventions using the RL framework, one must search for the best possible decision rule to inform the delivery of a JITAI. Ideally, these decision rules will learn the right delivery time, place and content for an intervention. To do so, we outline the following best practices procedure where we 1) collect data and conduct offline analysis and 2) create an algorithm to adaptively learn and relearn a better decision rule online as users experience the mobile health intervention.

Step 1: data collection and offline analysis

The future of AI decision-making in mobile health promises timely, personalized and effective interventions within a brief period of use. To deliver on this promise, a preliminary decision rule must be constructed prior to use online. One approach is to employ offline analyses with existing data.

Ideally, such data provides observations of user behaviour after an intervention is delivered in a given context. Specifically, each user record would consist of: (i) the current state of the user, (ii) the type of intervention delivered within the current state and (iii) the near-time outcome.

A micro-randomized trial (MRT) is an increasingly popular method (The Methodology Center, 2019) of collecting data for offline analysis. In an MRT, users may be randomized 100s or 1000s of times to different intervention options over the course of a study. For example, HeartSteps users are randomized five times a day, over 42 days, between a contextually tailored activity suggestion and no suggestion.

Once the MRT has concluded, an offline analysis can be used to construct a decision rule. The RL field has developed a variety of methods for constructing the decision rules (decision rules go by the term 'policy' in the RL literature). Variants of a very popular RL method called 'Q-learning' are often used. In Q-learning, an algorithm is used to learn a Q-function where Q can be interpreted as assessing the *relative quality of the different actions in each possible context*. Other approaches use the randomization probabilities from the MRT.

Upon completion of the offline analysis, we have an initial decision rule that works well, on average, and can be used as a starting point to rapidly personalize to each new user via online learning.

Step 2: online personalized learning

The decision rules obtained during offline analysis can be a good starting point for effective JITAIs, however, new users may be different from the users on whom the rules were learned. Thus, we aim to personalize the decision rules to the new users. RL methods can be used to personalize the prior decision rules to each user online, i.e., as the user experiences the intervention.

A key method of personalizing the decision rule online is to use an 'exploit-explore' strategy, where exploitation refers to delivering interventions that appear best given the current decision rule, and exploration refers to delivering different interventions to gather new information. Whilst using only exploration may be good for learning and relearning the decision rules, we do not want to burden users with ineffective interventions. Similarly, only exploitation is suboptimal, as this corresponds to relying on the past decision rules and never investigating whether the user's responsivity to the possible interventions has changed. Thus, a balance

between exploration and exploitation during online personalization is imperative.

The RL literature provides various algorithms for efficiently handling the exploit–explore trade-off. However, the efficacy of these methods is problem dependent. For example, some methodologies can better adapt to non-stationary behaviour (e.g., a user changes jobs unbeknownst to the mobile app), whilst others may focus on better adaptation to long-term effects (e.g., people habituating to interventions). This literature is vast, and the interested reader is directed to Sutton and Barto (2018, see further reading).

An open challenge within online personalization is that the set of interventions may not immediately reflect a user's preferences or values. For example, they may prefer specific foods (e.g., vegan diet or non-sugar-rich diet for users with diabetes), or may prefer to go to the gym but could not do so recently because of a busy schedule. Since the RL algorithm will remember interventions to which the current user (and often others) were most responsive, it may not detect changes in specific user preferences and values immediately. One way to mitigate this problem is to include 'users in the loop', that is, those who can exhibit agency over the RL algorithm. In this framework, users are given autonomy over selecting and rating specific interventions or intervention categories to provide the algorithm with information about their preferences and needs. For example, they might vote a 'thumbs up' or 'thumbs down' on an activity suggestion or may select a specific category of activity about which they want to receive notifications. Principled methods to incorporate user feedback and agency into an online personalization algorithm, however, remains an open research area.

Impact, challenges and the future

As mobile data collection technologies (e.g., wearable sensors, self-input) are maturing, the current big challenge for mobile health is to provide meaningful interventions afforded by these technologies. Just-in-time interventions can encourage regular engagement in healthy behaviours in order to prevent, as well as effectively contribute to, the management of chronic diseases. Thanks to AI, JITAIs hold the potential to be widely accessible, evidence-based, personalized and inexpensive.

Despite such great promise, achieving this goal remains challenging. While RL seems to fit seamlessly with the goals of mobile health, it requires large amounts of data. In this article, we've discussed a systematic two-step approach designed to iteratively develop and improve JITAIs based on RL algorithms.

The recognition of the benefits made possible from more accurate and effective JITAIs will pave the way to improve awareness for funding, the generation of more data and the development of better RL algorithms suited for mobile health. The next decade will probably see highly effective JITAIs made available for major preventable chronic diseases. Wherever you are, as long as you have access to a smart device, these interventions will be accessible and provide quality support even at times when clinical professionals are unavailable or unreachable, transforming healthcare delivery as we know it.

This research was partially supported by U.S. National Institutes of Health grants R01AA23187 (NIH/NIAAA), P50DA039838 (NIH/NIDA), U54EB020404 (NIH/NIBIB), U01CA229437 (NIH/NCI) and R01HL125440 (NIH/NHLBI). ■

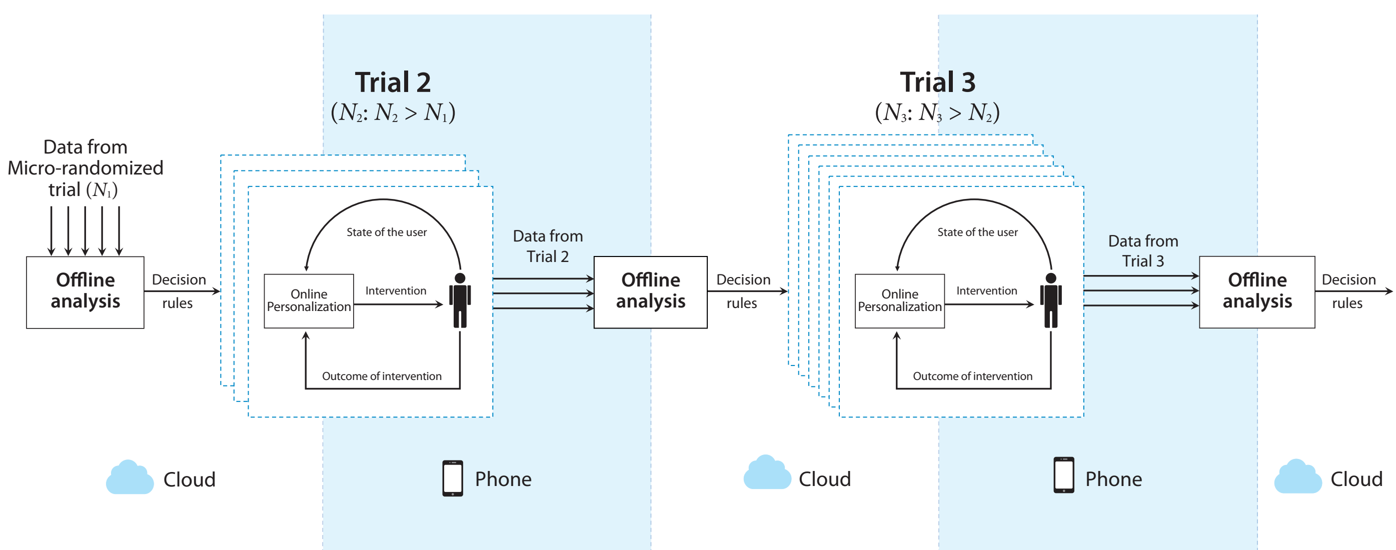

Figure 3. A path to gradually improve JITAIs requires more data and better algorithms over time. This can be achieved by iterating between offline analysis and online personalization.

Marianne Menictas is a postdoctoral fellow in the Department of Statistics at Harvard University. Marianne is currently working on computationally efficient methods for tuning models used to inform sequential decision-making in mobile health. She received her PhD in Mathematical Statistics from the University of Technology Sydney. During her postgraduate studies, Marianne focused on developing fast algorithms catered to fitting and inference for large datasets using Bayesian hierarchical models. Email: marianne_menictas@fas.harvard.edu

Mashfiqui Rabbi is a postdoctoral fellow in the Statistics Department at Harvard University. His research focuses on modelling human behaviour from phone data (e.g., sensors, self-input) and then issuing interventions that are personalized and contextualized to users' needs, daily routine and their environment. Mashfiqui holds a PhD in Information Science from Cornell University. His PhD research created the MyBehavior app, which is the first app to combine mobile data and AI to provide personalized interventions. Email: mrabbi@fas.harvard.edu

Predrag 'Pedja' Klasnja is an Assistant Professor in the School of Information at the University of Michigan and a Scientific Investigator at the Kaiser Permanente Washington Health Research Institute. He studies how mobile technologies can help individuals make lifestyle changes to improve their health. He is particularly interested in the design and evaluation of just-in-time adaptive interventions, technologies that continuously adapt to provide optimal support to individuals as their needs and circumstances change. He holds a PhD in Information Science from the University of Washington. Email: klasnja@umich.edu

Susan Murphy is a Professor of Statistics, Computer Science and Radcliffe Alumnae Professor, Harvard University. Dr Murphy's lab develops data analysis methods and experimental designs to improve real-time sequential decision-making in mobile health. She developed the micro-randomized trial for use in constructing mobile health interventions; this trial design is in use across a broad range of health-related areas. Dr Murphy is a member of the US National Academy of Sciences and the US National Academy of Medicine. In 2013 she was awarded a MacArthur Fellowship for her work on experimental designs to inform sequential decision-making. Email: samurphy@fas.harvard.edu

Further reading

- Araya, D. (2019) 3 things you need to know about augmented intelligence. Forbes, 22 January
- Ernst, D., Geurts, P. and Wehenkel, L. (2005) Tree-based batch mode reinforcement learning. J. Mach. Learn. Res. **6**, 503–556
- Klasnja, P., Hekler, E.B., Shiffman, S. et al. (2015) Microrandomized trials: an experimental design for developing just-in-time adaptive interventions. Health Psychol. **34**, 1220-1228
- Klasnja, P., Smith, S., Seewald, N.J. et al. (2018) Efficacy of contextually tailored suggestions for physical activity: a micro-randomized optimization trial of HeartSteps. Ann. Behav. Med. **53**, 573–582
- Lattimore, T. and Szepesvári, C. (2018) Bandit algorithms. Preprint. https://tor-lattimore.com/downloads/book/book.pdf (Accessed 9 October)
- Liao, P., Greenewald K., Klasnja, P. and Murphy, S.A. Personalized HeartSteps: a reinforcement learning algorithm for optimizing physical activity. Submitted, arXiv:1909.03539
- Nahum-Shani, I., Smith, S.N., Spring, B.J. et al. (2017) Just-in-time adaptive interventions (JITAIs) in mobile health: key components and design principles for ongoing health behavior support. Ann. Behav. Med. **52**, 446–462
- National Center for Chronic Disease Prevention and Health Promotion (2019). www.cdc.gov
- Raghu, A., Gottesman, O., Liu, Y. et al. (2018) Behaviour policy estimation in off-policy policy evaluation: calibration matters. arXiv preprint arXiv:1807.01066
- Russell, S.J. and Norvig, P. (2016) Artificial intelligence: a modern approach, Pearson Education Limited, Malaysia
- Silver, D. and Hassabis, D. (2016) AlphaGo: mastering the ancient game of Go with machine learning. Google Research Blog, 27 January
- Sutton, R.S. and Barto, A.G. (2018) Reinforcement learning: an introduction. MIT Press, USA
- The Methodology Center (2019) Micro-randomized trials. www.methodology.psu.edu/ra/adap-inter/mrt-projects/

