SYNTHETIC DATA GENERATOR FOR ADAPTIVE INTERVENTIONS IN GLOBAL HEALTH

Aditya Rastogi, Juan Francisco Garamendi, Ana Fernández del Río, Anna Guitart, Moiz Hassan Khan, Dexian Tang & África Periáñez
benshi.ai
Barcelona, Spain
{aditya,juanfran,ana,guitart,moiz,dexian,africa}@benshi.ai

ABSTRACT

Artificial Intelligence and digital health have the potential to transform global health. However, having access to representative data to test and validate algorithms in realistic production environments is essential. We introduce *HealthSyn*, an open-source synthetic data generator of user behavior for testing reinforcement learning algorithms in the context of mobile health interventions. The generator utilizes Markov processes to generate diverse user actions, with individual user behavioral patterns that can change in reaction to personalized interventions (i.e., reminders, recommendations, and incentives). These actions are translated into actual logs using an ML-purposed data schema specific to the mobile health application functionality included with *HealthKit*, and open-source SDK. The logs can be fed to pipelines to obtain user metrics. The generated data, which is based on real-world behaviors and simulation techniques, can be used to develop, test, and evaluate, both ML algorithms in research and end-to-end operational RL-based intervention delivery frameworks.

1 INTRODUCTION

Digital tools have become essential for frontline health workers (FHWs) everywhere, particularly in resource-poor settings where they are often designed to mitigate the effects of the lack of resources or infrastructure (Marsch, 2021; Overdijkink et al., 2018). From capacity building and communication to patient management and triage assistance or medical supply delivery, there is an array of mobile health solutions to help workers provide the best care possible to their communities (Overdijkink et al., 2018; Hosny & Aerts, 2019; Wahl et al., 2018; Forero et al., 2018).

These tools provide an excellent opportunity for personalization through adaptive technologies, using the rich information about provider and patient behavior and outcomes contained in the records generated through the use of digital applications (Olaniyi et al., 2022; Guitart et al., 2021a;b; Katsaris et al., 2021). This allows us to provide additional support by adapting the content to each FHW's and patient's evolving needs and by delivering timely reminders, suggestions, or personalized incentives specifically designed to work for them at precisely that time. We thus increase the chances of success by adopting an adaptive, contextualized approach to innovation that allows for rapid deployment cycles. Big tech has excelled at this iterative model, although optimizing for engagement and monetization rather than well-being.

Statistical analysis and machine learning (ML) models play a crucial role in using the information available to understand and predict individual FHW or patient behavior in order to design the best interventions for each use case (Hosny & Aerts, 2019; Wahl et al., 2018). Reinforcement learning (RL), a subset of these methodologies, additionally provides the algorithmic framework for making decisions about when and which interventions to send (Yom-Tov et al., 2017; Forman et al., 2019; Liao et al., 2020; Wang et al., 2021; Trella et al., 2022). It allows us to fine-tune the system depending on how much weight on knowledge extraction (i.e., gathering data that will allow statistical inference with enough power) or on optimization (i.e., using all information to make the best choice for each individual even if it hinders statistical significance) we want to put.

Synthetic data is becoming increasingly important due to concerns about reproducibility and establishing baselines (Jordon et al., 2020). Healthcare data is extraordinarily sensitive, and privacy through appropriate de-identification is a crucial step (largely yet to be overcome) before any data can be shared for analysis and ML (Yoon et al., 2020; Jordon et al., 2021). Adaptive technologies involve complex feedback cycles and long-term effects, so having a way to simulate user behavior, both in the absence of interventions and in their reactions to them, becomes essential for safe development and testing before the deployment with real subjects.

We introduce *HealthSyn* (benshi.ai, 2023a), an open-source library for mobile health data generation, and describe its components and methodology. It can be used for ML model development and can be coupled to algorithmic decision-making on interventions to which the user will react, thus providing an end-to-end simulation environment for RL-based interventions in mobile health. User behavioral logs from mobile health applications are tracked using *HealthKit* (benshi.ai, 2023b), an open-source Software Developer Kit (SDK) for mobile health. It is responsible for the tracking, labeling, and organization of the logs in a data schema optimized for ML and intervention purposes. *HealthKit* tracks general in-app users' behavior (actions, sessions, etc.) and additional events organized based on the type of content (patient management, capacity building, etc.). It also delivers the interventions and tracks the user's reaction to them (e.g. whether notifications were opened, discarded, or blocked). It distinguishes between online and offline activity, which is particularly important when working with global health applications, as they are often used without an internet connection, impacting both user behavior and the interventions that will work best for them.

1.1 Our contribution

To the best of our knowledge, *HealthSyn* is the first synthetic log generation library openly available. Complemented by the open *HealthKit* SDK, responsible for most of the heavy lifting regarding data tracking, labeling, and organization, it provides a simulation environment for RL-based interventions for mobile health. Although other behavioral simulators exist (see appendix A [1]), they focus on modeling subject daily activity (e.g., sleep hours, or alcohol consumption). The simulation environment presented here generates the information as if generated by app use. It is well-suited explicitly for digital tools designed to support FHW in low- and middle-income countries.

2 Mobile Health Log Synthetic Generator Architecture

HealthSyn is made up of three components: the *environment definition* and *event generator* modules, and a collection of *external wrappers* to link those modules to *HealthKit*, data pipelines, and RL-based nudge delivery system. These components allow for the simulation of an end-to-end intervention delivery framework as depicted in Figure 1, which shows the production system (described in more detail in appendix B) shaded in orange, and how this is emulated using *HealthSyn* in green.

2.1 Environment Definition

The *environment definition* module is composed of submodules responsible for determining user behavior (both unaltered and in response to nudging) and the context (i.e., the set of traits that will be considered as defining their state) and rewards for the contextual or restless bandits from the RL-component with which *HealthSyn* can be coupled for adaptive intervention simulation.

User behavior is modeled using a Markov processes. A key component of this model is the transition probability matrix, which can be defined differently for each user or context, or be the same for the entire population. This matrix serves as a baseline for the users' behavior, and it establishes the behavior in the absence of interventions.

User response to nudges is defined using a decay model that takes into account the number of nudges received on the last d days (where d is another parameter to be set for the environment). The decay model is a linear combination of three functions, each representing different types of user behavior. These are represented in Figure 2, with parameters that can be determined through a fitting process using real-world historical data to ensure that the model accurately reflects users' actual

[1] Related works have not been included in the main paper for the sake of space, but the interested reader can find this section as an appendix.

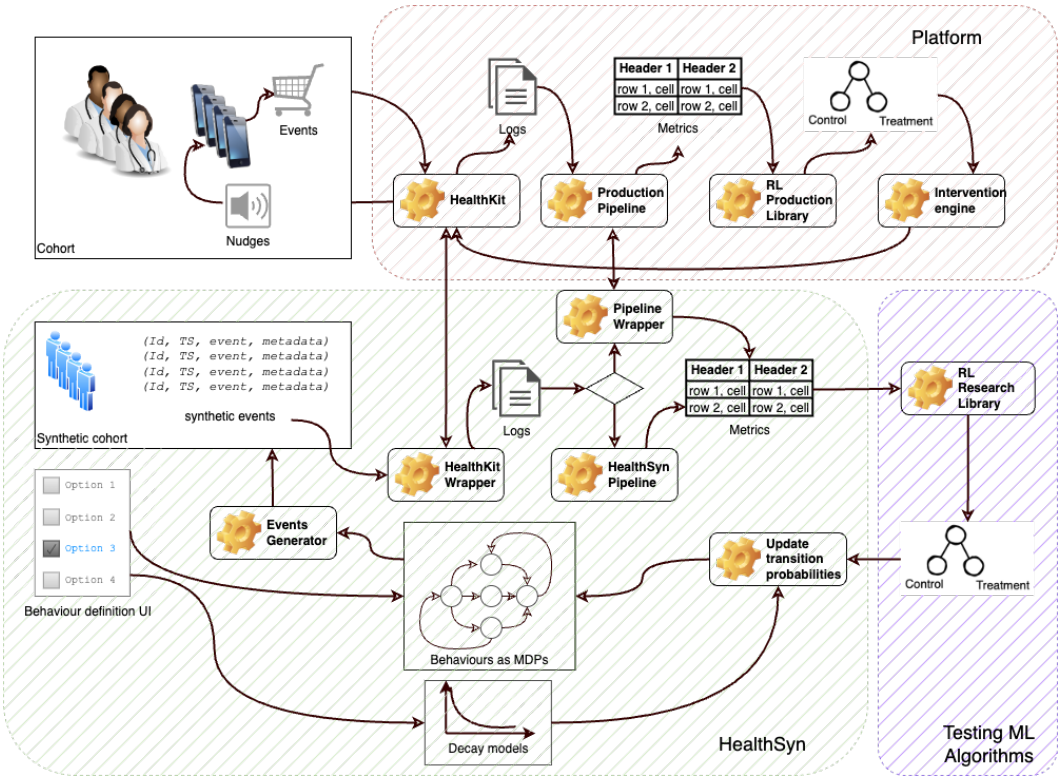

Figure 1: Overview of the production (modules involved in the delivery of interventions to real users in production settings, shaded in orange) and simulation (modules that mimic synthetic users and their reactions to interventions, shaded in green) environments. Modules included with *HealthSyn* are highlighted, and the *HealthKit* is included as the interface between users and the platform.

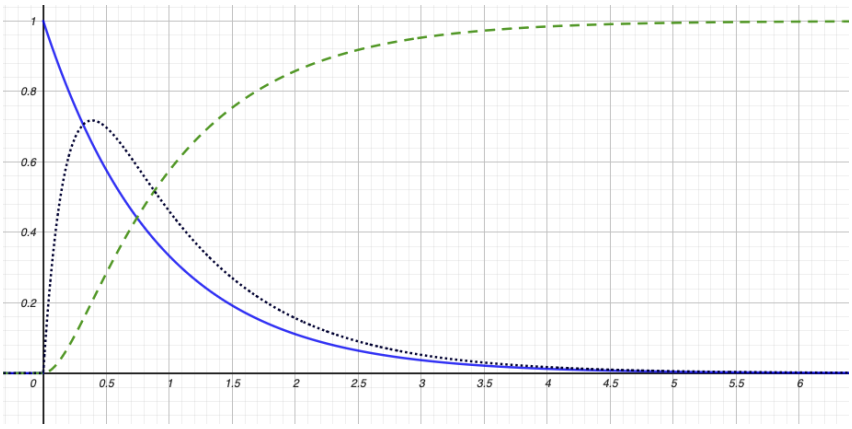

Figure 2: Decay functions used to model user reaction to nudges. The x-axis represents the number of nudges received. The solid blue line represents a response inversely proportional to the interventions (nudging has a negative effect). The black dotted line is characteristic of a positive reaction to the intervention until fatigue wears off the effect. The dashed green line represents a positive reaction that does not decay, but has decreasing returns that saturate the effect at a certain level.

behavior. One function (in solid blue) models an effect that is maximum when no interventions are received (the nudging negatively impacts users whose behavior is modeled by it exclusively). Another function (in dotted black) represents an initial positive effect that eventually whines to zero. The final function (in dashed green) models an effect that builds up with increasing returns per nudge until it saturates at a constant level. Note that the combination of these can model complex

user reactions to nudging, including long-lasting, novelty, and fatigue effects. See appendix C for a formal definition of the functions.

2.2 Synthetic In-App Activity Generation

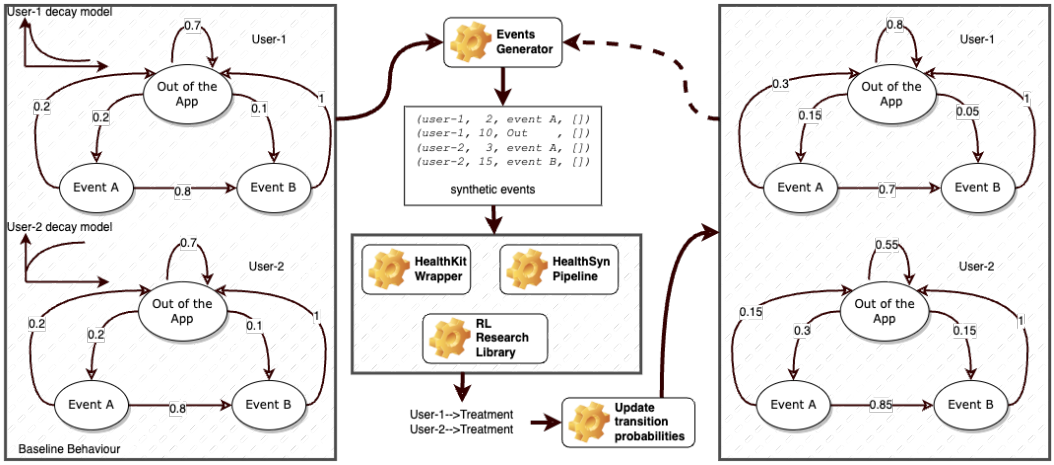

Figure 3: Overview of a single iteration in the simulation. The baseline behavior as defined in the *environment definition* and used by the *event generator* module to create a comprehensive list of in-app events or actions. These events are then translated into metrics that are used by the RL algorithm for intervention decisions, which in turn impacts user transition probabilities according to the user decay model. Dashed arrows represent input for the next iteration.

The *event generator* module generates user in-app actions or events as a collection of tuples, with each tuple representing an (id, timestamp, action, and metadata). Users are modeled as Markov processes; the generated actions are the states reached through the process.

Figure 3 illustrates the dynamic changes in the transition matrix for two users due to online interventions as the simulation progresses. In this illustration, we observe that both users start with the same baseline transition matrix and receive an intervention at the last decision point. However, the matrix evolves differently for each user based on each particular decay model. As shown in Figure 2, 'user-1' has an adverse reaction to interventions and is modeled as the blue function, while 'user-2' has a positive response and is modeled as the dashed green line. The 'action generator' module adjusts the transition matrix accordingly, with probabilities of specific actions decreasing and probabilities of transitioning to the 'Out of the app' state increasing for 'user-1' and the opposite occurring for 'user-2'. For users who do not receive an intervention, the transition probabilities will not change.

2.3 External Wrapper

This module translates the simulated user actions into actual logs and metrics. It serves as the bridge between the simulated world and the real world, and it is composed of three sub-modules. The *HealthKit Wrapper* reads and interprets the actions and then uses the SDK to generate the associated logs in the SDK schema. The *Pipeline Wrapper* reads the logs generated by the SDK wrapper, uses an existing data pipeline to compute the metrics, and saves them into a database. Finally, the *RL Wrapper* converts these metrics into context and rewards (according to the environment definition) that can be directly fed to the chosen RL algorithm. It also feeds back information on intervention decisions to the action generator module.

3 Summary and Conclusions

We have proposed a workflow that offers a systematic approach for evaluating RL algorithms for adaptive interventions in mobile health using synthetic data. It uses *HealthSyn*, an open-source library to simulate the behavior of FHWs in digital tools, and *HealthKit*, an open-source SDK which can track and label their logs in a data schema optimized for ML.

REFERENCES

Ahmed Alaa, Alex James Chan, and Mihaela van der Schaar. Generative time-series modeling with fourier flows. In *International Conference on Learning Representations*, 2021a. URL https://openreview.net/forum?id=PpshD0AXfA.

Ahmed M. Alaa and Mihaela van der Schaar. Attentive state-space modeling of disease progression. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper/2019/file/1d0932d7f57ce74d9d9931a2c6db8a06-Paper.pdf.

Ahmed M. Alaa, Boris van Breugel, Evgeny Saveliev, and Mihaela van der Schaar. How faithful is your synthetic data? sample-level metrics for evaluating and auditing generative models, February 2021b.

benshi.ai. HealthSyn. HealthSyn GitHub repository, 2023a. URL https://github.com/benshi-ai/healthsyn.

benshi.ai. HealthKit. NPM package, 2023b. URL https://www.npmjs.com/package/@benshi.ai/js-sdk.

Greg Brockman, Vicki Cheung, Ludwig Pettersson, Jonas Schneider, John Schulman, Jie Tang, and Wojciech Zaremba. Openai gym, 2016.

Roberto Forero, Shizar Nahidi, Josephine De Costa, Mohammed Mohsin, Gerry Fitzgerald, Nick Gibson, Sally McCarthy, and Patrick Aboagye-Sarfo. Application of four-dimension criteria to assess rigour of qualitative research in emergency medicine. *BMC health services research*, 18 (1):1–11, 2018.

Evan M. Forman, Stephanie G. Kerrigan, Meghan L. Butryn, Adrienne S. Juarascio, Stephanie M. Manasse, Santiago Ontañón, Diane H. Dallal, Rebecca J. Crochiere, and Danielle Moskow. Can the artificial intelligence technique of reinforcement learning use continuously-monitored digital data to optimize treatment for weight loss? *Journal of behavioral medicine*, 42:276, 4 2019. ISSN 15733521. doi: 10.1007/S10865-018-9964-1. URL /pmc/articles/PMC6524648//pmc/articles/PMC6524648/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC6524648/.

Anna Guitart, Ana Fernández del Río, África Periáñez, and Lauren Bellhouse. Midwifery Learning and Forecasting: Predicting Content Demand with User-Generated Logs, 2021a.

Anna Guitart, Afsaneh Heydari, Eniola Olaleye, Jelena Ljubicic, Ana Fernández del Río, África Periáñez, and Lauren Bellhouse. A recommendation system to enhance midwives' capacities in low-income countries. *arXiv preprint arXiv:2111.01786*, 2021b.

Ahmed Hosny and Hugo JWL Aerts. Artificial intelligence for global health. *Science*, 366(6468): 955–956, 2019.

Eugene Ie, Chih wei Hsu, Martin Mladenov, Vihan Jain, Sanmit Narvekar, Jing Wang, Rui Wu, and Craig Boutilier. Recsim: A configurable simulation platform for recommender systems, 2019.

Daniel Jarrett, Ioana Bica, and Mihaela van der Schaar. Time-series generation by contrastive imitation. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems*, volume 34, pp. 28968–28982. Curran Associates, Inc., 2021. URL https://proceedings.neurips.cc/paper/2021/file/f2b4053221961416d47d497814a8064f-Paper.pdf.

James Jordon, Jinsung Yoon, and Mihaela van der Schaar. Measuring the quality of synthetic data for use in competitions. *CoRR*, abs/1806.11345, 2018a. URL http://arxiv.org/abs/1806.11345.

James Jordon, Jinsung Yoon, and Mihaela van der Schaar. Pate-gan: Generating synthetic data with differential privacy guarantees. In *International Conference on Learning Representations*, 2018b.

James Jordon, Jinsung Yoon, and Mihaela van der Schaar. Differentially private bagging: Improved utility and cheaper privacy than subsample-and-aggregate. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper/2019/file/5dec707028b05bcbd3a1db5640f842c5-Paper.pdf.

James Jordon, Alan Wilson, and Mihaela van der Schaar. Synthetic data: Opening the data floodgates to enable faster, more directed development of machine learning methods. *CoRR*, abs/2012.04580, 2020. URL https://arxiv.org/abs/2012.04580.

James Jordon, Daniel Jarrett, Evgeny Saveliev, Jinsung Yoon, Paul Elbers, Patrick Thoral, Ari Ercole, Cheng Zhang, Danielle Belgrave, and Mihaela van der Schaar. Hide-and-seek privacy challenge: Synthetic data generation vs. patient re-identification. In Hugo Jair Escalante and Katja Hofmann (eds.), *Proceedings of the NeurIPS 2020 Competition and Demonstration Track*, volume 133 of *Proceedings of Machine Learning Research*, pp. 206–215. PMLR, 06–12 Dec 2021. URL https://proceedings.mlr.press/v133/jordon21a.html.

Iraklis Katsaris, , and Nikolas Vidakis and. Adaptive e-learning systems through learning styles: A review of the literature. *Advances in Mobile Learning Educational Research*, 1(2):124–145, 2021. doi: 10.25082/amler.2021.02.007.

Peng Liao, Kristjan Greenewald, Predrag Klasnja, and Susan Murphy. Personalized heartsteps: A reinforcement learning algorithm for optimizing physical activity. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*, 4, 3 2020. ISSN 24749567. doi: 10.1145/3381007. URL /pmc/articles/PMC8439432//pmc/articles/PMC8439432/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC8439432/.

Lisa A Marsch. Digital health data-driven approaches to understand human behavior. *Neuropsychopharmacology*, 46(1):191–196, 2021.

Van Bach Nguyen, Belaid Mohamed Karim, Bao Long Vu, Jörg Schlötterer, and Michael Granitzer. Policy learning for malaria control. *ArXiv*, abs/1910.08926, 2019.

Babaniyi Yusuf Olaniyi, Ana Fernández del Río, África Periáñez, and Lauren Bellhouse. User engagement and churn in mobile health applications. *In Proceedings of 2022 ACM KDD Workshop on Applied Data Science for Healthcare (DSHealth 2022)*, 2022.

Sanne B Overdijkink, Adeline V Velu, Ageeth N Rosman, Monique DM Van Beukering, Marjolein Kok, and Regine PM Steegers-Theunissen. The usability and effectiveness of mobile health technology–based lifestyle and medical intervention apps supporting health care during pregnancy: systematic review. *JMIR mHealth and uHealth*, 6(4):e8834, 2018.

Sekou L. Remy and Oliver Bent. A global health gym environment for rl applications. In Hugo Jair Escalante and Raia Hadsell (eds.), *Proceedings of the NeurIPS 2019 Competition and Demonstration Track*, volume 123 of *Proceedings of Machine Learning Research*, pp. 253–261. PMLR, 08–14 Dec 2020. URL https://proceedings.mlr.press/v123/remy20a.html.

Marlesson R.O. Santana, Luckeciano C. Melo, Fernando H.F. Camargo, Bruno Brandao, Anderson Soares, Renan M. Oliveira, and Sandor Caetano. Mars-gym: A gym framework to model, train, and evaluate recommender systems for marketplaces. *2020 International Conference on Data Mining Workshops (ICDMW)*, 2020-November:189–197, 11 2020. ISSN 23759259. doi: 10.1109/ICDMW51313.2020.00035.

Ashudeep Singh, Yoni Halpern, Nithum Thain, Konstantina Christakopoulou, Ed H Chi, Jilin Chen, Alex Beutel, and Google Research. Building healthy recommendation sequences for everyone: A safe reinforcement learning approach. In *ACM RecSys 2020*, 9 2020.

Anna L. Trella, Kelly W. Zhang, Inbal Nahum-Shani, Vivek Shetty, Finale Doshi-Velez, and Susan A. Murphy. Designing reinforcement learning algorithms for digital interventions: Pre-implementation guidelines. *Algorithms 2022, Vol. 15, Page 255*, 15:255, 7 2022. ISSN 1999-4893. doi: 10.3390/A15080255. URL https://www.mdpi.com/1999-4893/15/8/255/htmhttps://www.mdpi.com/1999-4893/15/8/255.

Boris van Breugel, Trent Kyono, Jeroen Berrevoets, and Mihaela van der Schaar. Decaf: Generating fair synthetic data using causally-aware generative networks. In M. Ranzato, A. Beygelzimer, Y. Dauphin, P.S. Liang, and J. Wortman Vaughan (eds.), *Advances in Neural Information Processing Systems*, volume 34, pp. 22221–22233. Curran Associates, Inc., 2021. URL https://proceedings.neurips.cc/paper/2021/file/ba9fab001f67381e56e410575874d967-Paper.pdf.

Brian Wahl, Aline Cossy-Gantner, Stefan Germann, and Nina R Schwalbe. Artificial intelligence (AI) and global health: how can AI contribute to health in resource-poor settings? *BMJ global health*, 3(4):e000798, 2018.

Shihan Wang, Chao Zhang, Ben Kröse, and Herke van Hoof. Optimizing adaptive notifications in mobile health interventions systems: Reinforcement learning from a data-driven behavioral simulator. *Journal of Medical Systems*, 45:1–8, 12 2021. ISSN 1573689X. doi: 10.1007/S10916-021-01773-0/FIGURES/4. URL https://link.springer.com/article/10.1007/s10916-021-01773-0.

Elad Yom-Tov, Guy Feraru, Mark Kozdoba, Shie Mannor, Moshe Tennenholtz, and Irit Hochberg. Encouraging physical activity in patients with diabetes: Intervention using a reinforcement learning system. *J Med Internet Res 2017;19(10):e338 https://www.jmir.org/2017/10/e338*, 19:e7994, 10 2017. ISSN 14388871. doi: 10.2196/JMIR.7994. URL https://www.jmir.org/2017/10/e338.

Jinsung Yoon, Daniel Jarrett, and Mihaela van der Schaar. Time-series generative adversarial networks. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett (eds.), *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019. URL https://proceedings.neurips.cc/paper/2019/file/c9efe5f26cd17ba6216bbe2a7d26d490-Paper.pdf.

Jinsung Yoon, Lydia N. Drumright, and Mihaela van der Schaar. Anonymization through data synthesis using generative adversarial networks (ads-gan). *IEEE Journal of Biomedical and Health Informatics*, 24(8):2378–2388, 2020. doi: 10.1109/JBHI.2020.2980262.

Lixin Zou. Data-efficient reinforcement learning for malaria control. In Zhi-Hua Zhou (ed.), *Proceedings of the Thirtieth International Joint Conference on Artificial Intelligence, IJCAI-21*, pp. 507–513. International Joint Conferences on Artificial Intelligence Organization, 8 2021. doi: 10.24963/ijcai.2021/71. URL https://doi.org/10.24963/ijcai.2021/71. Main Track.

A Appendix: Related Work

There are many recent examples of statistical and ML research around digital interventions and mobile health; good examples are Forman et al. (2019); Yom-Tov et al. (2017); Liao et al. (2020). A data science framework for the design of RL algorithms for digital interventions is presented in Trella et al. (2022). Wang et al. (2021) uses a data-driven behavioral simulator (trained using real data) to model the user's behavior and generate simulated data that can be used to train and evaluate the RL algorithm specifically for mobile health.

The van der Schaar lab has produced seminal and ground-breaking research around synthetic data generated within the framework of privacy (Jordon et al., 2018b; 2019; 2021; Yoon et al., 2020), electronic health record (EHRs), and its application for ML for healthcare (Alaa & van der Schaar, 2019; Jordon et al., 2019; 2020; 2021; Yoon et al., 2020). Their work includes methods to generate data using Generative Adversarial Networks (GANs) (Jordon et al., 2018b; Yoon et al., 2019; 2020), taking into account its time series nature (Yoon et al., 2019; Alaa et al., 2021a; Jarrett et al., 2021), evaluation methods for synthetic data generation (Jordon et al., 2018a; Alaa et al., 2021b), fairness van Breugel et al. (2021), and methods that include models of some of the physical processes underlying clinical parameters in the generation process (Alaa & van der Schaar, 2019).

The KDD Cup Challenge 2019 included a competition on policy learning for malaria control using RL (Nguyen et al., 2019; Zou, 2021), for which a global health environment for simulation of RL

applications was introduced (Remy & Bent, 2020). Understanding the adaptive intervention delivery as a recommendation problem, a simulation environment specifically designed for RL algorithm development and testing is presented in Ie et al. (2019). The MARS-Gym framework (Santana et al., 2020) provides a set of tools and environments, built on top of the OpenAI Gym library (Brockman et al., 2016), to model, train and evaluate RL-based recommender systems. In Singh et al. (2020), a simulation environment is used to test RL recommendation algorithms in the context of safe RL.

B Appendix: Adaptive Intervention Delivery Framework

The architecture for the adaptive intervention delivery framework for mobile health is as follows. The central piece is a data-centric platform that connects to the apps through the *HealthKit* SDK, which ensures all the information that is needed is tracked in a data schema optimized for data science and machine learning purposes. A use-case-specific data pipeline then turns the incoming information into actionable insights and metrics that are then readily available for visualization, and analysis, and to be consumed by (predictive and recommendation) ML models that produce additional metrics. The complete individual characterization provided by the resulting observational, predictive, and recommendation metrics is then available for the design of adaptive interventions. They are used to define the cohort of users that will be targeted, the content and scheduling of the interventions, and the state (or context, as we are limiting the discussion here to the use of contextual and restless bandits) and rewards for the RL component. The latter then uses the information to decide whether to assign subjects to treatment or control groups at each decision point in an adaptive experiment. In the case of adaptive interventions deployed in production, it decides whether to deliver or not or which of a finite set of interventions to deliver at each decision point for all users in the targeted cohort. This is schematically represented in Figure 1 by the orange shaded block labeled *platform* and the cohort of real users it is linked to.

The platform can integrate with the digital tools (and additional sources of digitalized information) through the *HealthKit* SDK, which supports data collection in an opinionated data schema optimized for data science and ML, and which is also responsible for the delivery of content and messaging interventions back to the tool. It guarantees all incoming information into the platform is adequately tracked and organized.

In the platform, incoming contextual, patient, and FHW data will be processed by use-case-specific data pipelines into useful metrics. All these metrics are then available to run statistical and ML models that will generate additional metrics. Cohorts are defined in terms of these metrics (e.g. FHWs that are predicted to underperform in their next evaluation) and constitute the basic unit of aggregate exploration and intervention. Interventions are defined and can use the available metrics to determine the content and scheduling of the interventions (e.g. content recommendation in the capacity building module as suggested by an ML model, sent in a push notification the next time they are predicted to be online). Intervention delivery is RL-based, and it is these metrics too that can be used to build the reward (e.g. number of questions answered correctly in in-app tests) and define the state for RL algorithms.

Interventions can combine clinical and behavioral elements. Clinical interventions fall into the realm of precision medicine and involve leveraging all the information available about a patient to improve quality of care and patient prognosis. Behavioral interventions aim at ensuring an adequate level of engagement (with the tool, program, and treatment) of medical care teams and patients.

Once the interventions have been developed, their effect can be measured using the experimentation platform, which includes fully randomized and adaptive (bandit-based) designs, with and without multiple assignments.

C Appendix: Mathematical Details

User behavior is defined using a decay model on the activity concerning the number of nudges $n \geq 0$ received on the last d days ($d = 5$ in our experiments). The decay model is represented by the three equations, each describing different aspects of the user behavior:

$$f(n) = a_0 e^{-k_a n} \tag{1}$$

$$g(n) = b_0 \frac{k_a \left(e^{-k_a n} - e^{-k_b n}\right)}{k_b - k_a} \tag{2}$$

$$h(n) = c_0 \frac{k_a \left(e^{-k_b n} - 1\right) - k_b \left(e^{-k_a n} - 1\right)}{k_b - k_a} \tag{3}$$

where k_a, k_b, a_0, b_0, c_0 are parameters defining the shape of the function. These parameters can be determined through a fitting process using real-world historical data to ensure that the decay model accurately reflects users' actual behavior. The decay model is then used to generate synthetic user behavior data for testing reinforcement learning algorithms in the context of e-health interventions.

Each equation models a different type of behavior: eq. (1) models the users whose maximum interaction with the app is when they do not receive intervention, eq. (2) represents users that are reactive to the interventions but are eventually overwhelmed and stop interacting. Finally, eq. (3) models users that are reactive to interventions, but at some point, they are saturated, and the activity does not increase or decrease. Figure 2 shows the shape of the three decay models.

The final behaviour $a_i(n)$ for user i is a linear combination of f, g, and h:

$$a(n) = \alpha_i f(n) + \beta_i g(n) + \gamma_i h(n) \tag{4}$$

where α_i, β_i, γ_i are parameters defining the individual characteristic of user i and can be itself random variables.

This allows us to define behavior at three levels, to which we will refer as *individual*, *contextual*, and *intervention*. By *individual*, we mean the unique behavior of a user as an individual, independent of the features defining their context. *Contextual* behavior represents the expected behavior of a user based on their context. All users with the same context will exhibit similar behavior to some extent, which can also be regulated. This is achieved using different parameters or even different functions $f(n), g(n), h(n)$ depending on the context. Finally, the user changes in in-app activity in response to an *intervention* depend on the number of nudges n received over time. It is implemented using the decay model described above, where the number of nudges n directly depends on the assigned group of the previous decision points, i.e., eq. (1)-(3) naturally models this behaviour. Note that it is also possible to define user-specific reactions to the interventions in the app's level or type of activity. This is achieved using different weighting parameters α_i, β_i, γ_i for each user i.