Proceedings of CONF-CDS 2026 Symposium: Data-Centric AI Security: Securing Models, Learning Agents, and Autonomous Systems
DOI: 10.54254/2755-2721/2026.CH33825

A Review of Contextual Bandits for Just-in-Time Adaptive Interventions in Mobile Health

Chenyuan Xu

Faculty of Arts and Social Sciences, University of Sydney, Sydney, Australia
thefoolsparkle@gmail.com

Abstract. Mobile health interventions are now common, but many still run on fixed schedules even though people's needs can change a lot within the same day, which can lead to missed opportunities for timely support and reduced effectiveness of the interventions. This limitation is one reason JITAI research has become important. A JITAI is not only about what support should be given, but also about when it should be given and withheld. This review examines that problem through the lens of contextual bandits. It explains how JITAIs are structured, why they require dynamic behavior models, and how repeated intervention decisions can be studied and formalized in mobile health. HeartSteps and Oralytics are used to show how these ideas look in practice. This study finds that contextual bandits offer a workable framework for JITAIs in mobile health, but their use is still limited by small samples, heterogeneity, changing responsivity, intervention burden, availability, and the difficulty of defining reward well.

Keywords: mobile health, just-in-time adaptive interventions, contextual bandits, personalization, digital interventions

1. Introduction

Mobile health makes it possible to deliver support in daily life rather than only in clinics or short study settings. But that does not automatically make interventions adaptive. In many cases, the timing and content of support are still decided in advance and remain fairly fixed later on. That approach can work to some extent, but it misses a basic feature of everyday behavior: the same person may be motivated and available at one moment, then tired, busy, or unreceptive a few hours later. This is the setting in which just-in-time adaptive interventions, or JITAIs, became important [1].

What changed in this area was not only the use of phones or wearable devices, but also the way the intervention problem itself was framed. Instead of treating support as a fixed plan, JITAI research gradually moved toward repeated decisions about whether support should be delivered, what kind of support should be used, and which information should guide those choices [1, 2]. This shift also raised a harder question: if support can be delivered many times, what kind of evidence is needed to judge one decision rather than another? Work on microrandomized trials pushed that question into the center of JITAI development [3]. Contextual bandits then became relevant because

DOI: 10.54254/2755-2721/2026.CH33825

they provided a clearer way to write these repeated, context-based intervention choices in mobile health [4].

This review focuses on this shift by examining how JITAIs are modeled in mobile health, why contextual bandits fit this setting, and what practical difficulties arise in real systems. The aim is to look at one narrower line of work and explain its main logic more clearly, providing researchers in this field with clear references regarding the adaptive intervention decision-making mechanism.

2. Conceptual foundations of JITAIs in mHealth

JITAIs were proposed because support needs do not appear in a fixed way, so support itself cannot always be delivered in a fixed way either. This is particularly critical in mobile health, where support is delivered in real-life settings rather than in controlled clinical environments. A person may need encouragement at one moment and find the same message unnecessary or irritating at another. Because of this, JITAIs focus on giving the right type and amount of support at the right time according to a person's changing state and context.

A useful starting point is the six-component framework proposed by Nahum-Shani et al. [1]. These components are distal outcomes, proximal outcomes, decision points, intervention options, tailoring variables, and decision rules. Distal outcome refers to the longer-term goal of the intervention, while proximal outcomes refer to shorter-term changes that help show whether one's support decision is moving in a useful direction. Decision points are the moments when support may be considered, intervention options are the possible actions, tailoring variables are the information used to guide decisions, and decision rules describe how these elements are linked. The point of this framework is not just to define terms. It shows what actually has to be decided in a JITAI.

Still, the framework is not enough if it is read too statically. Building health behavior models for JITAIs requires attention to how states, needs, and responses change over time [2]. The issue is not only which variables matter, but also how they change across situations and how intervention itself may later affect engagement or fatigue. This matters because JITAIs are not simply about sending more support. In some situations, not intervening may be the better choice. So dynamic behavior modeling is needed to explain when support may help, when it may become burdensome, and why the same intervention can work differently across different moments and different users [1, 2].

3. From JITAIs to contextual bandits

JITAIs become easier to understand once they are treated as repeated choices rather than fixed treatment plans. The question is no longer only whether an intervention works in general, but whether support should be given at a particular moment, what kind of support is available then, and what short-term change can be used to judge whether that choice helped [1, 2].

Microrandomized trials matter here because they generate evidence at the level where these choices are actually made. Unlike a traditional randomized controlled trial, which usually assigns treatment at a much coarser level, an MRT repeatedly randomizes intervention options at many relevant decision points. That makes it possible to study what happens when support is delivered at one moment rather than another, and under which conditions a specific intervention component has a proximal effect. MRTs also highlight a practical issue that becomes hard to ignore in mobile health: not every decision point is truly available for intervention [3].

This is also why contextual bandits became relevant. Tailoring variables can be treated as context, intervention options as actions, proximal outcomes as rewards, and decision rules as the basis of a policy. The mapping is useful because it turns a health intervention framework into a clearer repeated decision problem. Still, the translation is not straightforward in mHealth settings, where safety, missing data, variable collection costs, and interpretability all matter [4].

4. Core contextual bandit methods for mHealth JITAIs

In mHealth, the hard part is often not the intervention alone, but everything already happening in daily life before the intervention is delivered. A person's activity level, mood, schedule, burden, and receptivity may already shape the observed outcome at a decision point. Because of this, the key question is often not simply which action gives the highest total reward, but whether a particular intervention adds anything useful compared with doing nothing [5]. This is one reason standard contextual bandit formulations are often not enough for JITAIs.

Action Centered Contextual Bandits are useful because they treat that question directly. The baseline part of the outcome can be highly complex because it reflects many features of everyday life, while the treatment effect may still be modeled in a simpler and more stable way. That is why the zero action, or do-nothing option, matters so much in this setting. In mHealth, it is not just an empty comparison case. It is the reference point for asking whether an intervention actually helps in the current context [5].

The method also matters because it takes intervention burden seriously . In real systems, actions cannot be chosen freely without considering how often support is delivered. If support is too rare, the intervention may become ineffective. If it is too frequent, users may become fatigued and less responsive. Thus, action probabilities are not merely a technical detail. They are an integral part of the intervention design problem.

A different problem appears once personalization is taken seriously: users are clearly not the same, but each user also provides only limited data. IntelligentPooling deals with this difficulty by avoiding both extremes. It neither assumes that all users are identical nor posits that each user has enough data for a fully separate model. Instead, it learns how each user deviates from the overall population pattern, while also allowing responsivity to change over time [6]. In that sense, Action Centered Contextual Bandits and IntelligentPooling are solving different parts of the same problem. mHealth JITAIs need bandit methods, but not in their simplest textbook form [5, 6].

5. Representative applications and research platforms

HeartSteps serves as a useful example because it renders several abstract issues much more concrete. At a basic level, the system repeatedly faces a simple question: should it send an activity suggestion now, or not? But the value of the example is not just that there are repeated decisions. It also shows why "do nothing" is a real action rather than an empty baseline. A message may help in one moment and be unnecessary in another. The case also makes it easier to see why short-term reward is not the whole story. In HeartSteps, dosage matters because too much intervention can gradually become part of the problem rather than part of the solution [7].

The HeartSteps II Protocol is useful for a different reason. It should not be treated as a results paper proving effectiveness. What it really shows is what a longer JITAI research platform looks like when intervention components, micro-randomization, and theory development are all being studied together over time [8]. That matters here because mobile health bandit research does not happen in isolation. It usually sits inside a larger intervention and study design.

Oralytics makes a different set of problems visible. Here the issue is not only when to intervene, but also how reward is defined in the first place and how to judge whether an online learning system is workable before and after deployment. The related papers show this from several angles: one

DOI: 10.54254/2755-2721/2026.CH33825

focuses on balancing oral self-care quality and user burden in reward design, while others discuss pre-implementation comparison, engineering constraints, and what can actually be learned in a deployed trial [9-11]. These cases matter because they show where abstract bandit ideas stop being abstract and turn into practical design choices.

6. Practical challenges in real-world mHealth bandits

Even when contextual bandits provide a useful framework for JITAIs, real mobile health settings remain difficult because the method is always shaped by practical constraints. One major issue is heterogeneity under limited data. Different users may respond very differently to the same intervention, yet each user often contributes only a small amount of data. This makes fully individualized learning unstable, but complete pooling may ignore important differences. IntelligentPooling is useful here because it was designed for exactly this tension between personalization and data efficiency [6].

A second challenge is non-stationarity. In mHealth, users do not respond to support in a fixed way over time. An intervention that proves effective early may diminish in its impact later, and repeated prompts may gradually reduce receptivity. Thus, the learning problem extends beyond choosing actions from context; it also requires recognizing that responsivity itself may change [6].

Burden and availability create another difficulty. At many decision points, support may not be feasible because the user is busy, asleep, driving, or otherwise unavailable. Even when support is technically possible, it may still be unwise if repeated intervention creates fatigue. In practice, this means that action choice in mHealth is constrained by human burden rather than determined only by predicted reward. Real systems therefore need to treat "do nothing" as a serious option and must consider how often intervention can be delivered without damaging engagement [5, 7].

Reward design is also harder than it first appears. In oral self-care work, reward is not something that can simply be taken as given, because short-term improvement may come at the cost of future burden. Trella et al. designed a quality reward to balance desirable health behavior and user burden, and used a simulation test bed to tune reward hyperparameters before deployment [10]. This shows that in mHealth, reward design is part of method design rather than a minor technical detail.

Finally, deployment requires more than choosing an algorithm that looks good in theory. Pre-implementation work has emphasized personalization, computability, and stability as practical design criteria, while later Oralytics deployment work showed that engineering interruptions, fallback methods, and limited decision points can strongly affect what is actually feasible [9, 11]. What becomes clear here is that mHealth bandits are judged not only by learning efficiency, but also by whether they remain usable and realistic in everyday intervention settings.

7. Conclusion

This review argues that contextual bandits are useful for JITAIs not because they solve the problem completely, but because they provide a workable middle ground for mobile health interventions that require repeated decisions in changing situations. Across the papers discussed here, one point recurs consistently: JITAIs have to be treated as dynamic intervention problems prior to their translation into bandit-style methods. In that sense, contextual bandits matter because they match an important part of the intervention structure. However, many practical difficulties persist. Small samples, heterogeneity, changing responsivity, burden, availability, and reward design all make the methods harder to use in real settings. This is why work in this area keeps returning to issues such as do-nothing actions, pooling, simulation, and deployment constraints rather than only proposing more formal algorithms.

This review also has clear limitations. It focuses on a relatively small set of representative studies and does not try to cover the full range of digital health or reinforcement learning research. Neural contextual bandits were only discussed briefly. Even so, the papers reviewed here suggest that future progress will depend less on sounding more advanced in general terms, and more on keeping the methods interpretable, realistic, and usable in actual intervention settings.

This review mainly focuses on contextual bandits for JITAIs in mobile health, but it is still worth briefly mentioning neural extensions. They become relevant when the relationship between context and reward is too complex for simpler models. A representative example is NeuralUCB, which combines neural representation learning with upper-confidence-bound exploration. Still, this is a general contextual bandit method rather than a mobile-health-specific study. So in this review, neural contextual bandits are better treated as a possible next step than as part of the core argument.

References

[1] Nahum-Shani, I., Smith, S. N., Spring, B. J., Collins, L. M., Witkiewitz, K., Tewari, A., & Murphy, S. A. (2018). Just-in-time adaptive interventions (JITAIs) in mobile health: Key components and design principles for ongoing health behavior support. Annals of Behavioral Medicine, 52(6), 446-462. https: //doi.org/10.1007/s12160-016-9830-8

[2] Nahum-Shani, I., Hekler, E. B., & Spruijt-Metz, D. (2015). Building health behavior models to guide the development of just-in-time adaptive interventions: A pragmatic framework. Health Psychology, 34(Suppl), 1209-1219. https: //doi.org/10.1037/hea0000306

[3] Klasnja, P., Hekler, E. B., Shiffman, S., Boruvka, A., Almirall, D., Tewari, A., & Murphy, S. A. (2015). Microrandomized trials: An experimental design for developing just-in-time adaptive interventions. Health Psychology, 34(Suppl), 1220-1228. https: //doi.org/10.1037/hea0000305

[4] Tewari, A., & Murphy, S. A. (2017). From ads to interventions: Contextual bandits in mobile health. In J. M. Rehg, S. A. Murphy, & S. Kumar (Eds.), Mobile health: Sensors, analytic methods, and applications (pp. 495-517). Cham, Switzerland: Springer. https: //doi.org/10.1007/978-3-319-51394-2_25

[5] Greenewald, K., Tewari, A., Klasnja, P., & Murphy, S. A. (2017). Action centered contextual bandits. Advances in Neural Information Processing Systems, 30, 5973-5981.

[6] Tomkins, S., Liao, P., Klasnja, P., & Murphy, S. (2021). IntelligentPooling: Practical Thompson sampling for mHealth. Machine Learning, 110(7), 2117-2156. https: //doi.org/10.1007/s10994-021-05995-8

[7] Liao, P., Greenewald, K., Klasnja, P., & Murphy, S. A. (2020). Personalized HeartSteps: A reinforcement learning algorithm for optimizing physical activity. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 4(1), 1-22. https: //doi.org/10.1145/3381007

[8] Spruijt-Metz, D., Marlin, B. M., Pavel, M., Rivera, D. E., Hekler, E., De La Torre, S., et al. (2022). Advancing behavioral intervention and theory development for mobile health: The HeartSteps II protocol. International Journal of Environmental Research and Public Health, 19(4), 2267. https: //doi.org/10.3390/ijerph19042267

[9] Trella, A. L., Zhang, K. W., Nahum-Shani, I., Shetty, V., Doshi-Velez, F., & Murphy, S. A. (2022). Designing reinforcement learning algorithms for digital interventions: Pre-implementation guidelines. Algorithms, 15(8), 255. https: //doi.org/10.3390/a15080255

[10] Trella, A. L., Zhang, K. W., Nahum-Shani, I., Shetty, V., Doshi-Velez, F., & Murphy, S. A. (2023). Reward design for an online reinforcement learning algorithm supporting oral self-care. Proceedings of the AAAI Conference on Artificial Intelligence, 37(13), 15724-15730. https: //doi.org/10.1609/aaai.v37i13.26866

[11] Trella, A. L., Zhang, K. W., Jajal, H., Nahum-Shani, I., Shetty, V., Doshi-Velez, F., & Murphy, S. A. (2025). A deployed online reinforcement learning algorithm in an oral health clinical trial. Proceedings of the AAAI Conference on Artificial Intelligence, 39(28), 28792-28800. https: //doi.org/10.1609/aaai.v39i28.35143