The Micro-Randomized Trial for Developing Mobile Health Interventions

Ashley Walton[1], Linda Collins[2], Lori Crosby[3], Predrag Klasnja[4], Inbal Nahum-Shani[5], Mashfiqui Rabbi[1], Maureen Walton[6], Susan Murphy[1]

[1]Harvard University, Boston, MA. [2]Penn State University, State College, PA. [3]Cincinnati Children's Hospital Medical Center, Cincinnati, OH. [4]Kaiser Permanente, Seattle, WA. [5]Institute for Social Research, University of Michigan, Ann Arbor, MI. [6]University of Michigan, Ann Arbor, MI.

Mobile Health (mHealth) technologies provide opportunities to capture aspects of behavior in everyday contexts, and extend the ability to provide treatments anytime and anywhere. The micro-randomized trial (MRT) provides data that can be used to collect evidence concerning when and in which contexts mHealth intervention components are most effective, and whether the effects deteriorate with time or past dosage. We use three completed MRTs to review the elements of an MRT, considerations important to designing an MRT, and discuss future directions.

WHAT IS AN MRT?

- MRTs can be used to decide which intervention components to retain for future study or inclusion in the mobile health intervention package.
- In classical factorial designs participants are randomized to one of the intervention options at baseline, and receive that same intervention option for the duration of the trial.
- An MRT involves sequential randomization in that participants are repeatedly randomized to receive different intervention options hundreds or even thousands of times over the course of the experiment[1].
- This repeated randomization facilitates the collection of empirical evidence regarding whether time-varying information about an individual's current context and day of the study moderates the effects of a given component.
- mHealth intervention components that interrupt individuals as they go about their life can be burdensome and potentially have iatrogenic effects[2]. Therefore it is critical to understand whether a particular component is effective, and in which contexts, when constructing an mHealth intervention.

ELEMENTS OF AN MRT

1. **INTERVENTION COMPONENT.** Anything that can be separated out for experimentation.
2. **INTERVENTION OPTIONS.** Alternative versions of an intervention component.
3. **DISTAL OUTCOME.** Long-term, clinical or health related outcome.
4. **PROXIMAL OUTCOME.** A near-term, measurable outcome, through which the intervention components aim to impact the distal outcome.
5. **DECISION POINTS.** Pre-determined times at which it might be useful to deliver an intervention component.
6. **AVAILABILITY.** Contexts where the scientific team decides, a priori, that it is unsafe or scientifically inappropriate to deliver an intervention.
7. **RANDOMIZATION PROBABILITIES.** Pre-specified probabilities that a participant will receive each intervention option if they are available at a decision point, primarily determined by considerations of participant burden.
8. **COVARIATES.** Self-report measures or sensor data, taken at baseline or observed throughout the course of the study.

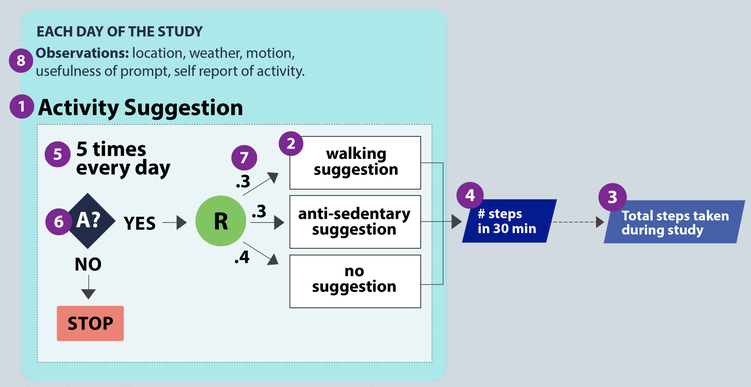

Figure 1. Example of MRT schematic for the activity suggestion component from HeartSteps.

HEARTSTEPS

Physical activity is known to decrease the risk of several health complications, yet only one in five adults in the U.S. meet the guidelines for the number of minutes of physical activity recommended per week[3]. Individuals can still experience health benefits if the required minutes are spread out across several days, and broken into more frequent but smaller amounts of time[4]. The goal of HeartSteps is to develop an intervention to increase overall levels of physical activity in sedentary adults by supporting opportunistic physical activity[5], in which brief periods of movement or exercise are incorporated into individuals' daily routines. HeartSteps Version 1 (v1), was a six-week MRT in which the intervention development team aimed to investigate whether contextually tailored activity suggestions, as well as support for planning how to be active, would increase participants' overall physical activity.

INTERVENTION COMPONENTS

1. Contextually tailored activity suggestion. Push notifications sent to participants' smartphones providing a suggestion for how to be active in the current moment, tailored to the participant's current location, weather conditions, time of day, and day of the week. The intervention options were: **(A)** a message with a walking activity that took 2-5 minutes to complete (see *Figure 2*), **(B)** an anti-sedentary activity (brief movements) that took 1-2 minutes to complete, or **(C)** no message.

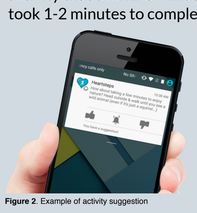

Figure 2. Example of activity suggestion

DECISION POINTS. 5 decision points every day: before morning commute, lunch time, mid-afternoon, after evening commute, and after dinner.

AVAILABILITY. Participants were unavailable when sensors on the phone indicated that they might be operating a vehicle, or were currently physically active. Participants could also hit "snooze" and turn off the activity notifications for 1, 2, 4 or 8 hours.

RANDOMIZATION. Participants were randomized with a 0.3 probability to receive a **(A)** contextually tailored walking activity, 0.3 probability of receiving an **(B)** anti-sedentary activity, and a 0.4 probability of receiving **(C)** no suggestion.

PROXIMAL OUTCOME. Total number of steps taken in the 30 minutes following randomization.

2. Daily activity planning. Participants could receive smartphone notifications in the evening that assisted them in planning for how they would be active the next day. Participants could either receive **(A)** a prompt asking participants to select a plan from a pre-defined list (structured planning), **(B)** a prompt asking participants to type their plan into a text box (unstructured planning), or **(C)** no prompt.

DECISION POINTS. One decision point every evening at a time specified individually by each participant at the beginning of the study.

AVAILABILITY. Participants were always assumed to be available at the evening decision point and were provided 1 hour to complete the planning after the prompt's delivery.

RANDOMIZATION. Participants were randomized with a 0.25 probability to receive structured planning, 0.25 probability to receive the unstructured planning, and 0.5 probability of receiving no planning.

PROXIMAL OUTCOME. Total number of steps taken on the subsequent day.

SCIENTIFIC QUESTIONS

- Does pushing the contextually tailored activity suggestion increase physical activity in the 30 minutes after the suggestion is delivered, compared to no suggestion?
- If so, does the effect of the contextually tailored activity suggestion change with time (day in study)?
- Does pushing a daily activity planning prompt increased physical activity the following day compared to no prompt?
- Is a structured, lower-burden daily planning prompt more effective at increasing physical activity the following day compared to an unstructured, higher-burden daily planning prompt?

FUTURE DIRECTIONS

- The most effective treatment for one individual in one context may be different than what is effective for another individual, and what is effective may change over time.
- The goal of HeartSteps v2 is to employ online training algorithms for assigning treatment in order to develop continually learning, personalized mHealth interventions.
- The online algorithm addresses the between and within subject heterogeneity by using this heterogeneity to improve the policy that assigns treatment to each individual, by detecting when an individual is responding differently than an "average" person to a treatment and if they respond differently in different contexts.

HIGHLIGHTS & MRT SCHEMATIC

Defining the time interval for proximal outcomes.

- Given a particular intervention component, a researcher needs to determine the appropriate interval of time after delivering that component in which to observe the component's effect on the proximal outcome.
- If the time interval is too short, then the measure of the proximal outcome will not capture the immediate effects of the intervention component.
- If this interval is too long, then the measure of the proximal outcome may include too much random noise.
- In the case of HeartSteps v1, the above considerations along with domain expertise was sufficient to determine the duration of 30 minutes.
- In some cases previous research may lack the specificity necessary to inform decisions regarding the timescales of components' proximal outcomes.
- Currently data analytic methods are being developed to inform the choice of duration over which the proximal outcome should be observed.

Defining Availability.

- HeartSteps highlights how researchers must identify conditions where components cannot be delivered in order to prevent the intervention from endangering participants, as well as prevent participants from becoming frustrated and ignoring the intervention.
- This helps ensure that intervention components may only be delivered in contexts where it is safe and have the most potential to be effective.

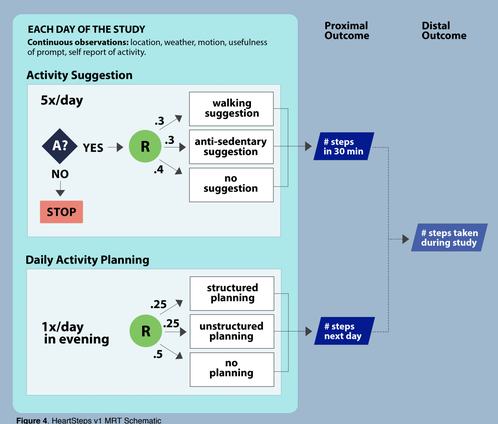

Figure 4. HeartSteps v1 MRT Schematic

SARA

Little is understood about the temporal processes that underlie the development of substance use disorders in adolescents. Mobile technology has the potential to be a powerful data collection tool to better understand these processes, as most adolescents own mobile phones that can be used to capture information continuously throughout their everyday lives. The Substance Abuse Research Assistant (SARA) is a mobile application being developed to collect data about the correlates of substance use that includes several engagement strategies designed to sustain rates of self-report over time[6]. SARA highlights how MRTs can be used to investigate engagement components, instead of/in addition to treatment components, as well as the careful considerations involved in selecting the components of an mHealth intervention to be experimented on as part of an MRT.

INTERVENTION COMPONENTS

1. Reciprocity notification. A push notification could be delivered 2 hours before the data collection period, containing an inspirational message in the form of youth-appropriate song lyrics or a celebrity quote. The intervention options were: **(A)** a reciprocity notification or **(B)** no notification.

DECISION POINTS. Once a day at 4 pm.

AVAILABILITY. No availability constraints.

RANDOMIZATION. Participants were randomized with a 0.5 probability to either **(A)** a reciprocity notification, or **(B)** no notification.

PROXIMAL OUTCOME. Whether or not participants completed the survey or active tasks on that *same* day.

2. Push reminder notification. Participants received a push notification every day at the beginning of the data collection period. They could either receive **(A)** a simple reminder or **(B)** a reminder that contained a persuasive message.

DECISION POINTS. Once a day at 6 pm.

AVAILABILITY. No availability constraints.

RANDOMIZATION. Participants were randomized with a 0.5 probability to receive **(A)** a simple reminder notification, and a 0.5 probability of receiving **(B)** a reminder notification with a persuasive message.

PROXIMAL OUTCOME. Whether or not participants completed the survey or active tasks on that *same* day.

3. Post-survey reinforcement. After completing the survey, participants could receive a notification containing a reward in the form of a meme or gif. Participations could either receive **(A)** a post-survey reinforcement, or **(B)** no reinforcement.

DECISION POINTS. Immediately after participants completed daily survey.

AVAILABILITY. Participants were unavailable if they hadn't completed the survey.

RANDOMIZATION. Participants were randomized with a 0.5 probability to receive post-survey reinforcement, and a 0.5 probability of receiving no reinforcement.

PROXIMAL OUTCOME. Whether or not participants completed the survey or active tasks on that *same* day.

4. Post-active task reinforcement. After completing the active task participants could receive a notification containing a life insight, which were visualizations of their self-reported data from the past seven days. There were two intervention options. Participants could either receive **(A)** a life insight, or **(B)** no life insight.

DECISION POINTS. Immediately after participants completed the active tasks.

AVAILABILITY. Participants were unavailable if they hadn't completed the active tasks.

RANDOMIZATION. Participants were randomized with a 0.5 probability to receive **(A)** a post-active task reinforcement, and a 0.5 probability of receiving **(B)** no reinforcement.

PROXIMAL OUTCOME. Whether or not participants completed the survey or active tasks on that *same* day.

SCIENTIFIC QUESTIONS

- Does providing an inspirational message before 2 hours before data collection result in increased completion of the daily survey and active tasks on that *same day* compared to no inspirational message?
- Does delivering a push reminder notification at the beginning of the data collection period with a persuasive message result in increased completion of the daily survey and active tasks on that *same day* compared to a push reminder notification without a persuasive message?
- Among individuals who complete the survey, does providing a reward in the form of a meme or gif increase survey completion on the *next day* compared to not providing a reward?
- Among individuals who complete the active task, does providing a life insight increase survey and active task completion on the *next day* compared to not providing a life insight?

HIGHLIGHTS & MRT SCHEMATIC

MRTs with engagement components.

- MRTs can be used to investigate engagement components.
- A researcher developing an mHealth intervention, or planning to conduct mHealth longitudinal data collection should consider including engagement components. An MRT should be used if there are open questions concerning the proximal effectiveness of engagement components.
- In the case of SARA, facilitation of self-report was intended to provide data regarding substance use and relevant contextual factors (e.g. mood, amount of free time) that might be used to develop treatment components for future mHealth interventions.

Selecting components for an MRT.

- The SARA mobile application contained several different engagement strategies, only a subset of which were randomized as part of the SARA MRT. For example SARA included a growing virtual aquarium that was progressively filled with fish as participants earned points for completing data collection. Participants could earn additional treasures for streaks in self-reporting and graduate to different "levels" where the aquarium tank turned into a sea environment.
- The fish and rewards for the aquarium were not included as part of the MRT because there were not scientific questions to motivate randomization. They also did not involve push notifications and therefore there was low risk of them interrupting or burdening participants.
- Selection of components to be included as part of an MRT should be motivated by the scientific questions that need to be addressed in order to determine whether they should be included as part of the intervention package.

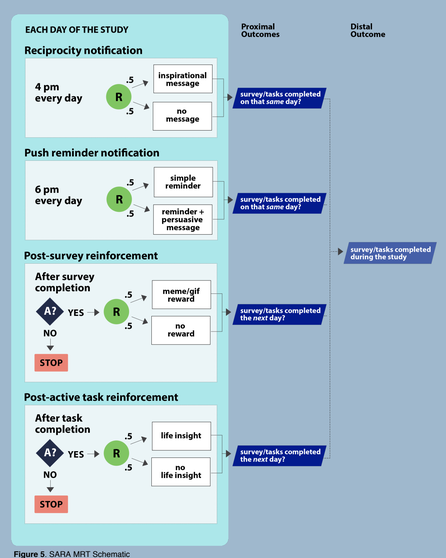

Figure 5. SARA MRT Schematic

BariFit

Individuals who have undergone bariatric surgery are required to monitor their weight and dietary intake while they progress through the different stages of recovery. Regular exercise is also recommended, where epidemiological studies indicate that higher levels of physical activity post-bariatric surgery are associated with better long-term outcomes[7]. Engaging in these multiple health behaviors and adjusting to these lifestyle changes can be challenging for someone as they are attempting to recover from surgery. BariFit is an mHealth intervention being developed by scientists at Kaiser Permanente to improve their ability to support the successful transition of these individuals to a routine where they can achieve and maintain weight loss after surgery.

INTERVENTION COMPONENTS

1. Adaptive step goal. A text message delivered every morning that provided a goal for the number of steps for participants to achieve that day. It had two different factors that each had two different intervention options. The first factor was the "set goals" factor, where participants either received **(1A)** step goals that represented the 60th percentile of their daily step counts in the previous ten days (out of the daily step counts for each of the previous ten days, which was the sixth highest), or **(1B)** goals that were variable percentiles of their step counts in the previous ten days. The second factor was the "rest days" factor, where participants were randomized to one of two options for the duration of the study: **(2A)** to have a rest day from the step goals once a week, or **(2B)** no rest days.

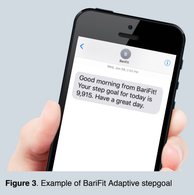

Figure 3. Example of BariFit Adaptive stepgoal

DECISION POINTS. One decision point, prior to the study.

AVAILABILITY. No availability constraints.

RANDOMIZATION. Participants were randomized with a 0.5 probability to receive **(1A)** the 60th percentile step goals, and a 0.5 probability to receive **(1B)** the variable percentile goals for the duration of the study. Also participants had a 0.5 probability of receiving **(2A)** a weekly rest day, and 0.5 probability of receiving **(2B)** no rest days for the duration of the study.

PROXIMAL OUTCOME. Number of steps taken that same day.

2. Contextually tailored activity suggestion. Text messages with suggestions for how to be active, tailored to the time of day, day of the week, and current weather conditions.

DECISION POINTS. 5 decision points every day: before morning commute, lunch time, mid-afternoon, after evening commute, and after dinner.

AVAILABILITY. Participants were unavailable for treatment if they were sensed as driving or already walking.

RANDOMIZATION. Participants had a 0.3 probability of receiving **(A)** an activity suggestion, and a 0.7 probability of receiving **(B)** nothing.

PROXIMAL OUTCOME. Total number of steps taken in the 30 minutes following randomization.

3. Reminder to track food. Text message participants could receive that reminded them to use the mobile application provided as part of the study to track what they ate that day.

DECISION POINTS. One decision point every morning.

AVAILABILITY. No availability considerations.

RANDOMIZATION. Participants had 0.5 probability of receiving **(A)** a reminder, and a 0.5 probability of receiving **(B)** no reminder.

PROXIMAL OUTCOME. Whether they used the application provided as part of the intervention to log what they ate that day.

SCIENTIFIC QUESTIONS

- Does delivering a text message with an activity suggestions tailored to the user's context increase physical activity in the thirty minutes after the suggestion is delivered compared to no suggestion?
- Does delivering a reminder result in self-monitoring of food intake that same day compared to no reminder?
- Does including a weekly "rest day" where participants do not receive a step goal result in growth in step count compared to no rest days?
- Does delivering a text message with an adaptive step goal where the magnitude is more variable (variable percentile goal) result in more growth in step count compared to an adaptive step goal where the magnitude is less variable (always 60th percentile goal)?

HIGHLIGHTS & MRT SCHEMATIC

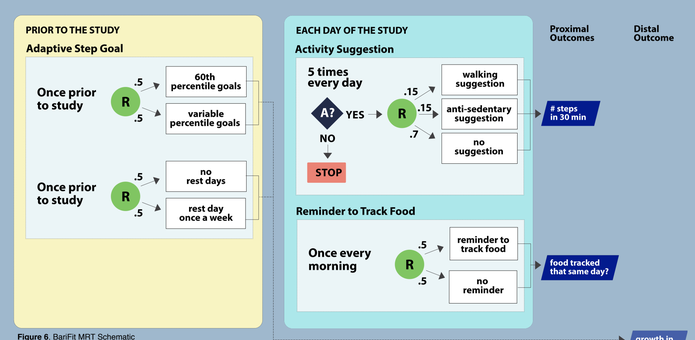

Figure 6. BariFit MRT Schematic

Sequential vs. baseline randomization.

- Determining the timescale at which each component should be randomized requires careful consideration of exactly what aspect of the intervention component is being experimented on as part of the MRT.
- The way the magnitude of the adaptive step goals varied depended on which intervention option participants were randomized to receive at the *prior to* the study (either 60th percentile or a variable percentile step goals). These two different intervention options were two different "schedules" of delivery, where they represented two different ways of varying the step goals over time.
- The manner in which the step goals varied over time did not change throughout the study; after initial randomization these schedules of delivery remained fixed. This one-time, baseline randomization was chosen in order to estimate the overall effect of using two different ways of varying the step goals over time.
- In contrast, consider the activity suggestion component where participants were randomized to one of the treatment options at five decision points each day. For this component the BariFit MRT was used to investigate what time-varying contextual factors moderated the effectiveness of the suggestions.
- The extent to which the time element of an intervention component is "unpacked" has to do with the level of granularity of the component[8]. For the activity suggestions the BariFit MRT is being used to "unpack" and understand the time element. For the adaptive step goals the time-variability is a fixed "package" like the components in a standard factorial design.

REFERENCES

1. 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, 1220–1228. **2.** Klasnja, P., Harrison, B. L., Legrand, L., Lamarca, A., Froehlich, J., & Hudson, S. E. (2008). Using Wearable Sensors and Real Time Inference to Understand Human Recall of Routine Activities. UbiComp '08, **3** C.D.C. (2014). Facts about physical activity. Retrieved from https://www.cdc.gov/physicalactivity/data/facts.htm **4.** C.D.C. (2018). How much physical activity do adults need? Retrieved from https://www.cdc.gov/physicalactivity/basics/adults/index.htm **5.**Consolvo, S., Everitt, K., Smith, I., & Landay, J. A. (2006). Design Requirements for Technologies that Encourage Physical Activity. In Proceedings of the SIGCHI conference on Human Factors in computing systems (pp. 457–466). **6.** Rabbi, M., Kotov, M. P., Cunningham, R., Bonar, E. E., Klasnja, P., Walton, M., & Murphy, S. (2018). Toward Increasing Engagement in Substance Use Data Collection : Development of the Substance Abuse Research Assistant App and Protocol for a Microrandomized Trial Using Adolescents and Emerging Adults Corresponding Author : Journal of Medical Internet Research Research Protocols, 7(7). https://doi.org/10.2196/resprot.9850 **7.** Bellicha, A., Ciangura, C., Poitou, C., Portero, P., & Oppert, J. M. (2018). Effectiveness of exercise training after bariatric surgery—a systematic literature review and meta-analysis. Obesity Reviews. **8.** Collins, L. M. (2018). Optimization of Behavioral , Biobehavioral , and Biomedical Interventions.