Challenges in Developing Online Learning and Experimentation Algorithms in Digital Health (CHIL 2022)

Walter Dempsey

04/08/2022

Methods in Mobile Health (research group)

- Introductions, discuss format and the upcoming week research talks
- Digital Health and Micro-randomized trials (MRTs)
  - Use case: synthetic HeartSteps
- *Primary* Analysis of MRT data
  - Causal inference
- Debiased Machine Learning and MRTs

Introductions and Format

Introductions

- 30-second introductions
  - Name
  - Department
  - Degree
  - Interest level
  - 1 ‘fun’ fact

Format

- 1.5 hour meeting
  - 45 minutes on the background
  - 45 minutes on the new paper/research
- Czar of communications: Hanna Venera

What do I do?

- Plan 4 weeks ahead
  - Next week: Jieru and XXXX
  - Discussing ‘
- Czar of recruitment: Jieru Shi
- Czar of coordination: Madeline Abbott
- Czar of github management: TBD

Topic suggestions

- Topic 1: Off-policy evaluation and Markov Decision Processes
- Topic 2: Safe RL and online reinforcement learning
- Topic 2: Debiased ML and Moderation Effects
- Topic 3: MRTs and inference under a joint RL policy (i.e., partial pooling)
- Topic 4: Wearable Sensors and functional data analysis
- Topic 5: High frequency longitudinal data (EMA) and scalable Gaussian Processes
- Topic 6: Data Privacy, MRTs, and CATE

What do I do now?

- Sign up!
  - Pick a topic that most interests you
  - Find a colleague with similar interest (or ask us)
  - Email HMJ and me to set up a meeting (Jieru coordinates, Madeline to attend)
  - We will decide on the `tutorial' and` research' content
  - Meet week before with me and Madeline to go over presentation
- Propose new topics and 'recruit a colleague'

Mobile health and micro-randomized trials

Mobile Health Interventions

- Pull intervention
  - Static content that can be accessed *whenever the user finds it necessary*
  - E.g., Mindfulness or guided meditation, exercise tips, low salt food alternative lookup
- Push intervention
  - App component that is *triggered by the system itself*
  - No need for user involvement
  - E.g., prompt to encourage self-monitoring, prompt to encourage physical activity, prompt to use the low salt alternative component.

HeartSteps (PI Klasnja)

- Develop a mobile activity coach for individuals who are at risk of coronoary artery disease
- Include right combination of pull components with
- Push components *delivered at the right times* to encourage activity throughout the day

HeartSteps V1: Evening Planning

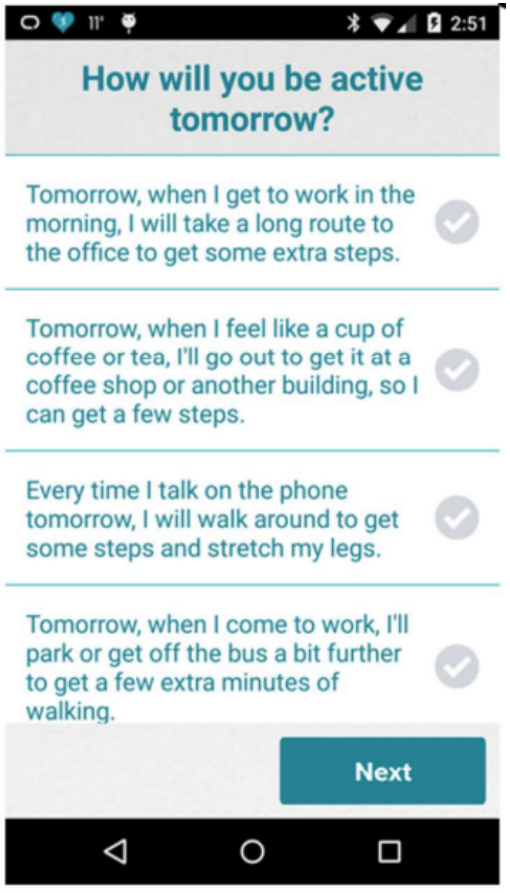


Figure 1: Event Planning Example

HeartSteps V1: Activity Suggestion

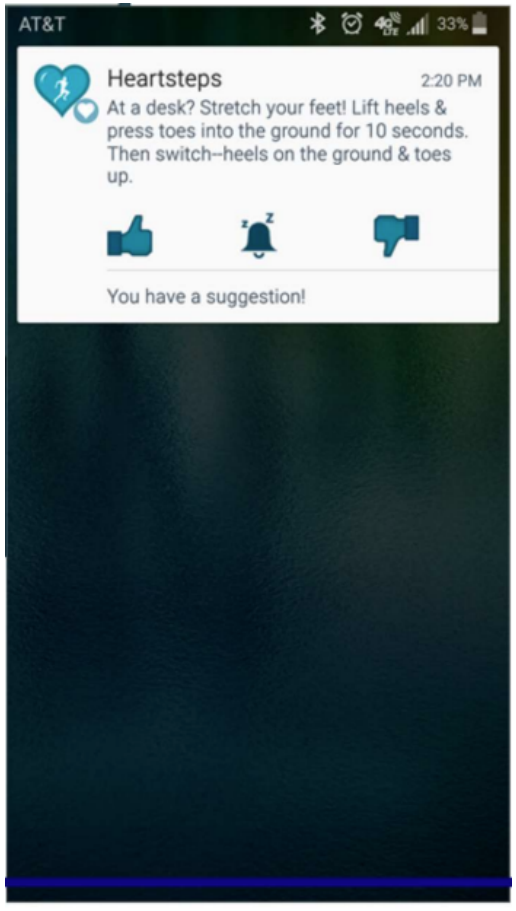


Figure 2: Activity Suggestion Example

Questions to improve the activity suggestions

- Do tailored activity suggestions have an effect at all?
- Does the effect of suggestions change over time?
- When should we send suggestions for optimal effect?

HeartSteps

- Three iterative studies
  - 42 day micro-randomized study
    - Sedentary individuals
  - 90 day + 270 micro-randomized study
    - Personalized randomization scheme
    - People with Stage 1 hypertension
- Synthetic HeartSteps
  - Build from the 90 day MRT study.
  - Randomizations set similar to MRT for illustrative purposes.

Micro-randomized trial

- Micro-randomization = each user is randomized many times = sequential experimentation
- Randomization may use online predictions as well as reinforcement learning
- Probabilistic budgets on number of treatment pushes to manage treatment burden

Data from wearable devices

- On each individual we observe the sequence

$$(O_1, A_1, Y_2, O_2, A_2, Y_3, \ldots, O_t, A_t, Y_{t+1}, \ldots, O_T, A_T, Y_{T+1})$$

- t: Decision point
- O_t: Observations at t^{th} decision point
- A_t: treatment push at t^{th} decision point
- Y_{t+1}: Proximal outcome (e.g., reward, utility, cost)

Micro-randomized trial elements:

- Decision points, t (Times at which an intervention can be provided)
  - Regular intervals in time (e.g., every 10 minutes)
  - At user demand
- **HeartSteps**: Approximately every 2-2.5 hours at user-specific times labeled as monring, mid-day, mid-afternoon, early evening, after dinner

Micro-randomized trial elements:

- Treatments A_t:
- Types of treatments that can be provided at a decision point, t
- **HeartSteps**: tailored activity suggestion (yes/no)

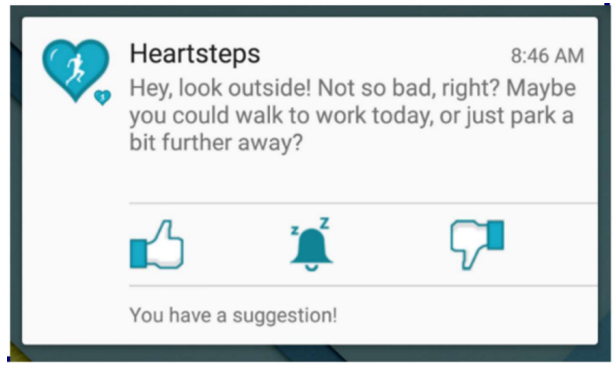


Figure 3: Activity Suggestion Example

Micro-randomized trial elements

- Randomization:
  - A stochastic strategy for selecting among the treatments at each decision point
  - The probabilistic distribution of A_t
- **HeartSteps**: push or do not push tailed suggestion
  - $P(A_t = 1) = 0.6$ (push)
  - $P(A_t = 0) = 0.4$ (do not push)
- Alternatives:
  - Use predictions in combination with an online algorithm to uniformly spread recommendations across *risk times* throughout a day
  - Use reinforcement learning to personalize the treatment recommendations

Micro-randomized trial elements:

5. Proximal outcome: Y_{t+1}

- Mediators that are thought to be critical to achieving a longer term clinical health outcome such as improved heart health
- **HeartSteps**: Activity (step count) over next 30 minutes
- Question: how can we use AI to determine the length of time over which the proximal outcome is measured?

Availability

- Interventions, A_t, can only be delivered at a decision point if the decision point is *available* for the user
  - O_t includes $I_t = 1$ if available, and $I_t = 0$ if not
- Availability is known pre-decision point, i.e., pre-treatment
- Availability is not adherence, nor is it the same as interruptibility, receptivity

Why Micro-randomization?

- Randomization (+ representative sample) is a gold standard in providing data to assess causal effects
- Sequential randomizations will enhance replicability of data analyses (decision rule development)

Experimentation for continuous improvement

- “Iterative nature of experimentation” (RA Fisher & George Box)
- " At Google, experimentation is practically a mantra; we evaluate almost every change that potentially affects what our users experience." (4 Google scientists)
- “Online experiments are widely used to compare specific design alternatives, but they can also be used to produce generalizable knowledge and inform strategic decision making. Doing so often requires sophisticated experimental designs, iterative refinement, and careful logging and analysis” (3 Facebook scientists)

Synthetic HeartSteps Data

- An MRT simulator based on Heartsteps V2 has been built in R and is available here
- This talk and the batch synthetic data are available here
- *ID*: Numeric id taking values between 1-40
- *Day*: Day-in-study (numeric)
- *Decision time*: Numeric indicator of indicator of decision time per day (1-5)
- *Dosage/burden*: Pre-defined function of past pushes (walking + anti-sedentary messages)
- *Engagement Indicator*: Binary indicator of whether the number of screens encountered in app from prior day from 12am to 11:59pm is greater than the 40% quantile of the screens collected.

Synthetic HeartSteps Data

- *Temperature*: Temperature (In Celsius degree) at the current location
- *Location*: 1 if at a location other than home or work; 0 if at home or work (pre-specified)
- *Variation Indicator*: Indicator of higher recent variation in step counts than median long-term variation
- *Pre-treatment Steps*: Log-transformed steps 30 mins prior to the current decision time from the tracker; $\log(y + 0.5)$.
- *Square root of steps yesterday*: The square root of step counts from the tracker collected from 12am to 11:59 pm

Synthetic HeartSteps Data

id	day	decision.time	dosage	engagement	other.location
1	1	1	0.00	1	1
1	1	2	0.00	1	1
1	1	3	0.00	0	1
1	1	4	1.00	1	1
1	1	5	1.95	0	1

Synthetic HeartSteps Data

decision.time	variation	temperature	logpresteps	sqrt.totalsteps	prior.anti
1	0	0.64	0.66	0.49	0
2	0	0.77	-0.69	0.49	1
3	1	0.81	0.49	0.49	0
4	0	0.83	0.51	0.49	0
5	0	0.73	0.66	0.49	0

Synthetic HeartSteps Data

decision.time	MRT_avails	MRT_probs	MRT_action	MRT_reward
1	0	0.5	0	5.49
2	1	0.5	0	-0.69
3	1	0.5	1	6.01
4	0	0.5	1	-0.69
5	0	0.5	0	6.08

Causal excursions and primary analysis methods

Primary Analysis of MRT Data:

- Why consider conducting simple, interpretable, primary analyses?
  - In clinical and commercial settings
- Understand why causal excursion effects are useful for continual learning

Questions to Improve Synthetic HeartSteps Activity Suggestions

- Do tailored activity suggestions ahve an effect at all?
- Do less and more burdensome activity suggestions work equally well?
- Does the effect of suggestions change of time?
- When should we send suggestions for optimal effect?

Micro-randomized trial

- How to justify the experimtntal trial costs in clinical research setting
  - Address a question that can be stated across disciplinary boundaries and be able to provide guarantees
  - Design trial so that a variety of further interesting questions can be answered
- **Large number of stakeholders with differing data needs/interests**

Conceptual model: Primary analysis

- Data analysts want to fit a series of increasingly complex models:
- **Primary analysis**

$$Y_{t+1}\text{`} \sim' \underbrace{\alpha_0 + \alpha_1^\top Z_t}_{\text{reduce noise}} + \underbrace{\beta_0 A_t}_{\text{causal main effect}}$$

- Z_t: summary formed from t and past/present observations
  - 'Control Variables'
- β_0: is the effect, marginal over all observed and unobserved variables, of the activity suggestion on subsequent activity

Conceptual model: Secondary analysis

- **Secondary analysis**

$$Y_{t+1}{}' \sim' \underbrace{\alpha_0 + \alpha_1^\top Z_t}_{\text{reduce noise}} + \underbrace{\beta_0 A_t + \beta_1 A_t S_t}_{\text{causal moderation effect}}$$

- Z_t: control variables
- S_t: potential moderator (e.g., current level of engagement)
- $\beta_0 + \beta_1$: is the effect when an individual is engaged ($S_t = 1$), marginal over all observed and unobserved variables, of the activity suggestion on subsequent activity.

Scientific goal

- Analytic methods that are consistent with the scientific understanding of the meaning of β coefficients
  - Make the scientist developing the intervention behavioral scientist
- Analytic methods that require minimal additional assumptions
  - Don't require strong modeling assumptions that may not hold
- Causal inference challenges:
  - Time-varying treatment $(A_t, t = 1, \ldots, T)$
  - 'Independent' variables: Z_t, S_t, I_t all may be affected by prior treatment
- Robustly facilitate noise reduction via use of controls,$\sim Z_t$.

Causal effects

- Use potential outcomes to define the effect
- The effects we define are *causal excursions*
  - Contrasts at a given time, averaging over prior randomizations
  - A 1-step excursion (send suggestion vs don't send suggestion) from the underlying stochastic policy.

Potential outcomes

- $\bar{A}_t = (\bar{A}_1, \bar{A}_2, \ldots, \bar{A}_t)$
  - $\bar{a}_t$: realizations of treatments
- $Y_{t+1}(\bar{a}_{t-1})$ is the potential proximal outcome
- $I_{t+1}(\bar{a}_{t-1})$ is the potential *available for treatment* indicator
- $H_t(\bar{a}_{t-1})$ is the potential history vector
  - $S_t(\bar{a}_{t-1})$ is a vector of features from the potential history.

Fundamental problem of causal inference

- Individual level causal effect (binary A_t):

$$Y_{t+1}(\bar{A}_{t-1}, 1) - Y_{t+1}(\bar{A}_{t-1}, 0)$$

- Averaging over underlying stochastic policy
- Excursion contrast of treat $A_t = 1$ vs don't treat $A_t = 0$
- Fundamental problem:
  - Data alone cannot be used to predict or estimate the effect

Solution to the fundamental problem of causal inference

- Instead of estimating/predicting individual level causal effects,

$$Y_{t+1}(\bar{A}_{t-1}, 1) - Y_{t+1}(\bar{A}_{t-1}, 0)$$

  estimate expectations or conditional expectations of these effects:

$$E\left[Y_{t+1}(\bar{A}_{t-1}, 1) - Y_{t+1}(\bar{A}_{t-1}, 0)\right]$$

- See here for a great resource on causal inference (especially observational studies)

Marginal & Causal effects

- The **excursion effect** at decision point t:

$$E\left[Y_{t+1}(\bar{A}_{t-1}, 1) - Y_{t+1}(\bar{A}_{t-1}, 0) \mid I_t(\bar{A}_{t-1}) = 1, S_t(\bar{A}_{t-1})\right]$$

- Effect is conditional on availability; only concerns the subpopulation of users available at decision time t
- Effect is marginal over any $Y_s, s \leq t$ and $A_s, s < t$ not in $S_t(\bar{A}_{t-1})$

Switching from potential outcomes to observables:

The **excursion effect** can be expressed in terms of conditional expectations on *observable data* (under some assumptions):

$$E\left[E[Y_{t+1}|H_t, A_t = 1] - E[Y_{t+1} \mid H_t, A_t = 0] \mid I_t = 1, S_t\right] = S_t^\top \beta$$

- H_t: observation history up to time t (excluding treatment)
- S_t is a vector of data summaries and time t
- I_t indicator of availability
- Treatment effect model: we aim to conduct inference about β

An analysis method: Weighted and centered least squares estimation

- Simple, interpretable, method for complex data
- Enables unbiased inference for a causal, overall, treatment effect (the β's)
- Inference for treatment effect is not biased by how we use controls, Z_t, to reduce the noise variance in Y_{t+1}

Steps in estimation

- Estimate probabilities: $\tilde{p}_t(s) \in (0, 1)$
- Form weights

$$W_t = \left(\frac{\tilde{p}_t(S_t)}{p_t(H_t)}\right)^{A_t} \left(\frac{1 - \tilde{p}_t(S_t)}{1 - p_t(H_t)}\right)^{1-A_t}$$

- Center treatments: $A_t \to A_t - \tilde{p}_t(S_t)$
- Minimize

$$E_n \left[\sum_{t=1}^{T} \left(Y_{t+1} - Z_t^\top \alpha - (A_t - \tilde{p}_t(S_t))S_t^\top \beta\right)^2 I_t W_t\right]$$

where E_n is the empirical distribution over individuals.

Key points

- We **do not** need to assume

$$E[Y_{t+1}|A_t, I_t = 1, Z_t] = Z_t^\top \alpha + (A_t - \tilde{p}_t(S_t))S_t^\top \beta$$

- This would assume the conditional mean of Y_{t+1} is correctly specified which is a much stronger assumption!
- We only assume the **treatment effect** is correctly specified

$$E\left[E[Y_{t+1}|H_t, A_t = 1] - E[Y_{t+1} \mid H_t, A_t = 0] \mid I_t = 1, S_t\right] = S_t^\top \beta$$

- Then under moment conditions, $\hat{\beta} \to \beta_0$

Gains from randomization

- Causal inference for a marginal treatment effect
- Inference on treatment effect is robust to working model:

$$E[Y_{t+1}|I_t = 1, H_t] \approx Z_t^\top \alpha$$

- $Z_t \subseteq H_t$
- Contrast to literature on partially linear, single index models and varying coefficient models

Choice of weights

- Choice of $\tilde{p}_t(S_t)$ determines the marginalization over time under model misspecification of treatment effect
- Example: $S_t = 1$, $\tilde{p}_t(S_t) = \tilde{p}$. Then $\hat{\beta}$ is an estimator of

$$\frac{\sum_{t=1}^{T} E[I_t]\beta_t}{\sum_{t=1}^{T} E[I_t]}$$

where $\beta(t)$ is the causal excursion effect at decision point t

Synthetic HS case study

HeartSteps Schema

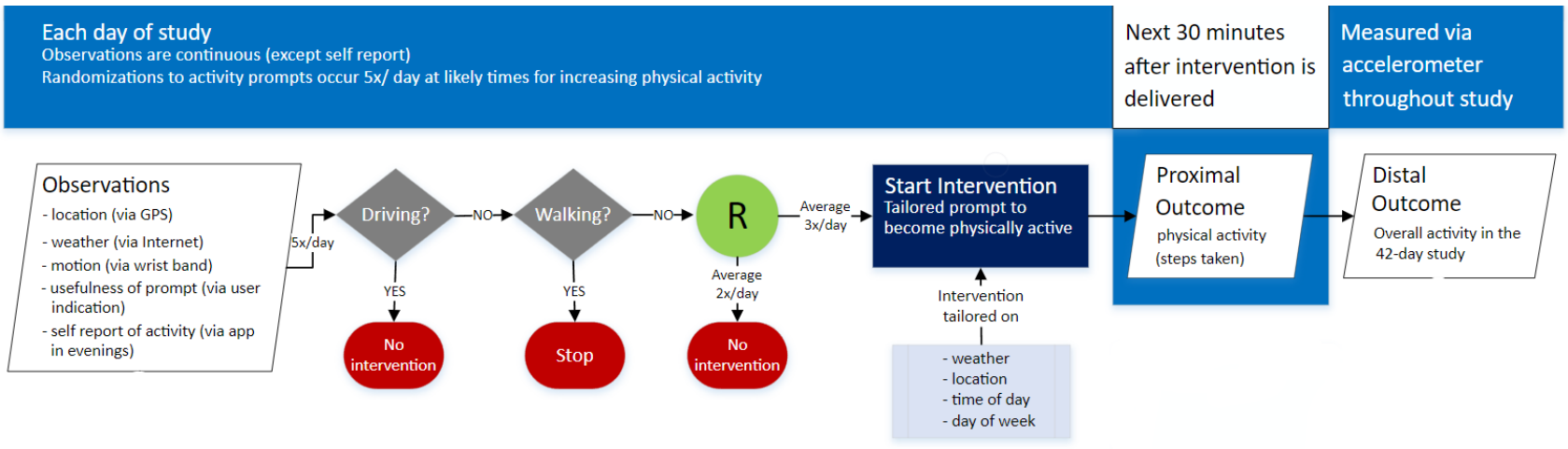


Figure 4: HS Schema

Synthetic HeartSteps

One each of the $n = 40$ participants:

- Activity suggestion, A_t:
  - Provide a tailored activity suggestion with probability 0.5
  - Do nothing (probability 0.5)
- 5 times per day $\cdot 90$ days $= 450$

Synthetic HeartSteps: Center actions

```
HS_MRT_data$MRT_action_c = HS_MRT_data$MRT_action -
  HS_MRT_data$MRT_probs
```

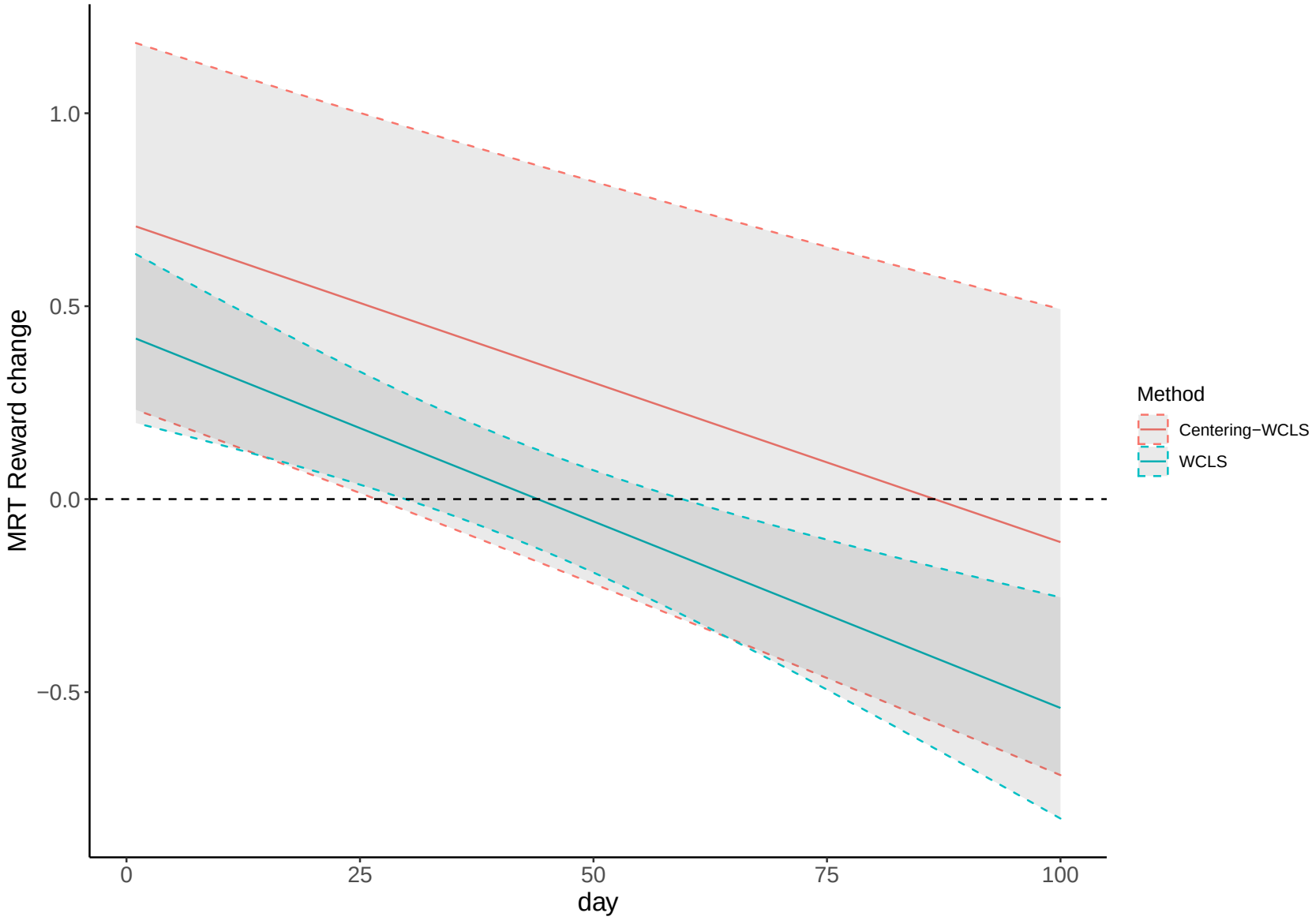
Synthetic Heartsteps:

term	estimate	std.error	statistic	p.value
(Intercept)	3.72	0.0682	2971.71	0.000
MRT_action_c	0.43	0.1097	15.07	0.000
day	0.00	0.0013	0.89	0.346
MRT_action_c:day	-0.01	0.0022	19.93	0.000

Synthetic Heartsteps:

term	estimate	std.error	statistic	p.value
(Intercept)	2.82	0.1150	602.48	0.000
MRT_action_c	0.72	0.2325	9.46	0.002
day	0.00	0.0010	9.96	0.002
logpresteps	3.24	0.0387	7019.33	0.000
engagement	0.10	0.0413	5.54	0.019
dosage	0.00	0.0116	0.00	0.950
MRT_action_c:day	-0.01	0.0015	30.45	0.000
MRT_action_c:logpresteps	-0.49	0.0767	41.37	0.000
MRT_action_c:engagement	0.38	0.0970	15.22	0.000
MRT_action_c:dosage	-0.03	0.0264	1.34	0.247

Synthetic HeartSteps:



Synthetic Heartsteps:

term	estimate	std.error	statistic	p.value
(Intercept)	2.85	0.1295	485.09	0.000
MRT_action_c	0.96	0.2773	12.10	0.001
day	0.00	0.0010	10.60	0.001
logpresteps	3.24	0.0392	6838.97	0.000
engagement	0.09	0.0407	5.43	0.020
dosage	0.00	0.0116	0.00	0.981
other.location	-0.03	0.0726	0.15	0.694
MRT_action_c:day	-0.01	0.0017	37.67	0.000
MRT_action_c:logpresteps	-0.49	0.0771	39.62	0.000
MRT_action_c:engagement	0.37	0.0981	14.07	0.000
MRT_action_c:dosage	-0.03	0.0260	1.03	0.309
MRT_action_c:other.location	-0.24	0.1107	4.80	0.028

Initial conclusions

- The data indicates a causal effect of activity suggestions on step count in the succeeding 30 minutes
  - The effect is positive for the first 20 days and then becomes negative
  - Steps in prior 30 minutes negatively impacts the effect of the suggestion
  - Engagement strongly positively impacts the effect of the suggestion
  - Being at home/work has a strong positive impact on the effect of the suggestion
  - Dosage does not have a strong effect
- Updating the stochastic policy to depend on prior steps, engagement, and other location may improve proximal outcomes further

Next Steps for intervention development

Next Steps for intervention development

- We have answered the following questions
  - Is the intervention effective, on average? Does the average effectiveness vary over time in study?
  - Does the effectiveness interact with time-varying contexts?
- Next question

How do we use this knowledge to improve the intervention?

- Use the analysis result to form a **warm-start policy** for online reinforcement-learning algorithm
- Include the identified contexts as state information in the RL algorithm

Warm-Start Policy + Bandit Algorithms

- A simple idea: use global fit to form data-informed prior for action selection strategy within a Thompson sampling algorithm
- Model mean *reward* (proximal outcome)
  $E[R_{t+1}|S_t = s, A_t = a] = r(s, a)$
  - Observe context features, s
  - Action selection strategy selects treatment, a
  - Observe reward,$\sim R$
  - *Learning algorithm updates parameters in mean/distribution of reward
  - Repeat!

Linear, "Thompson Smapling'' Bandit

- Gaussian linear model with mean reward

$$E[R_{t+1}|S_t = s, A_t = a] = r(s, a) = \eta^\top f(s, a)$$

- Initial η parameters in mean reward with a prior distribution (*warm-start!*)
- Update posterior distribution of η given (S_t, A_t, R_{t+1})
  - In this simple case, posterior is Gaussian
- Select next treatment a with probability equal to posterior probability that treatment a has highest mean reward.

Synthetic HeartSteps V2

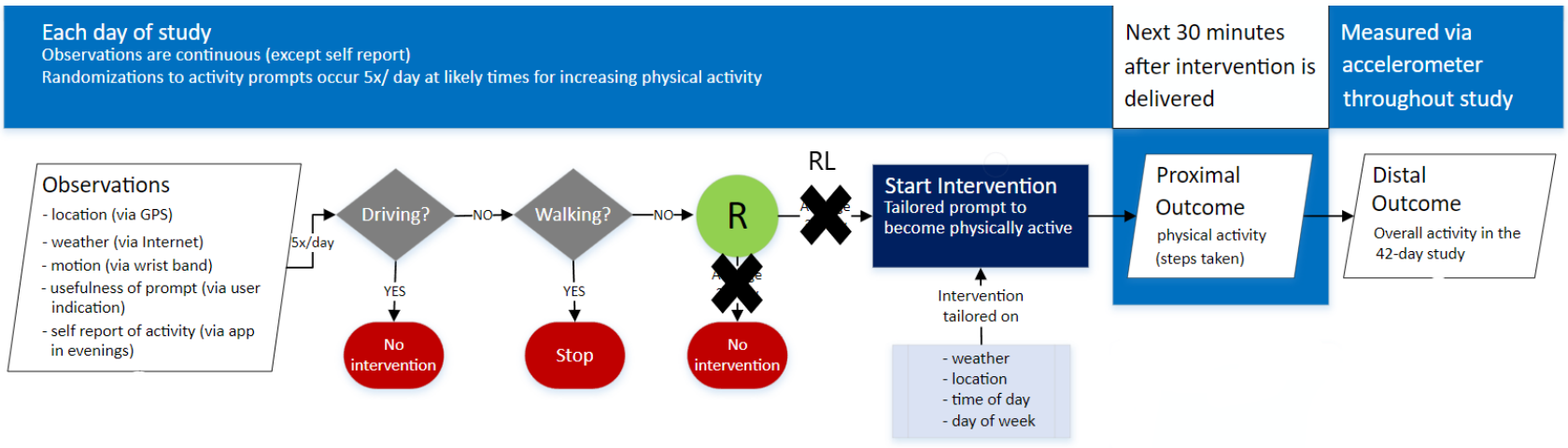
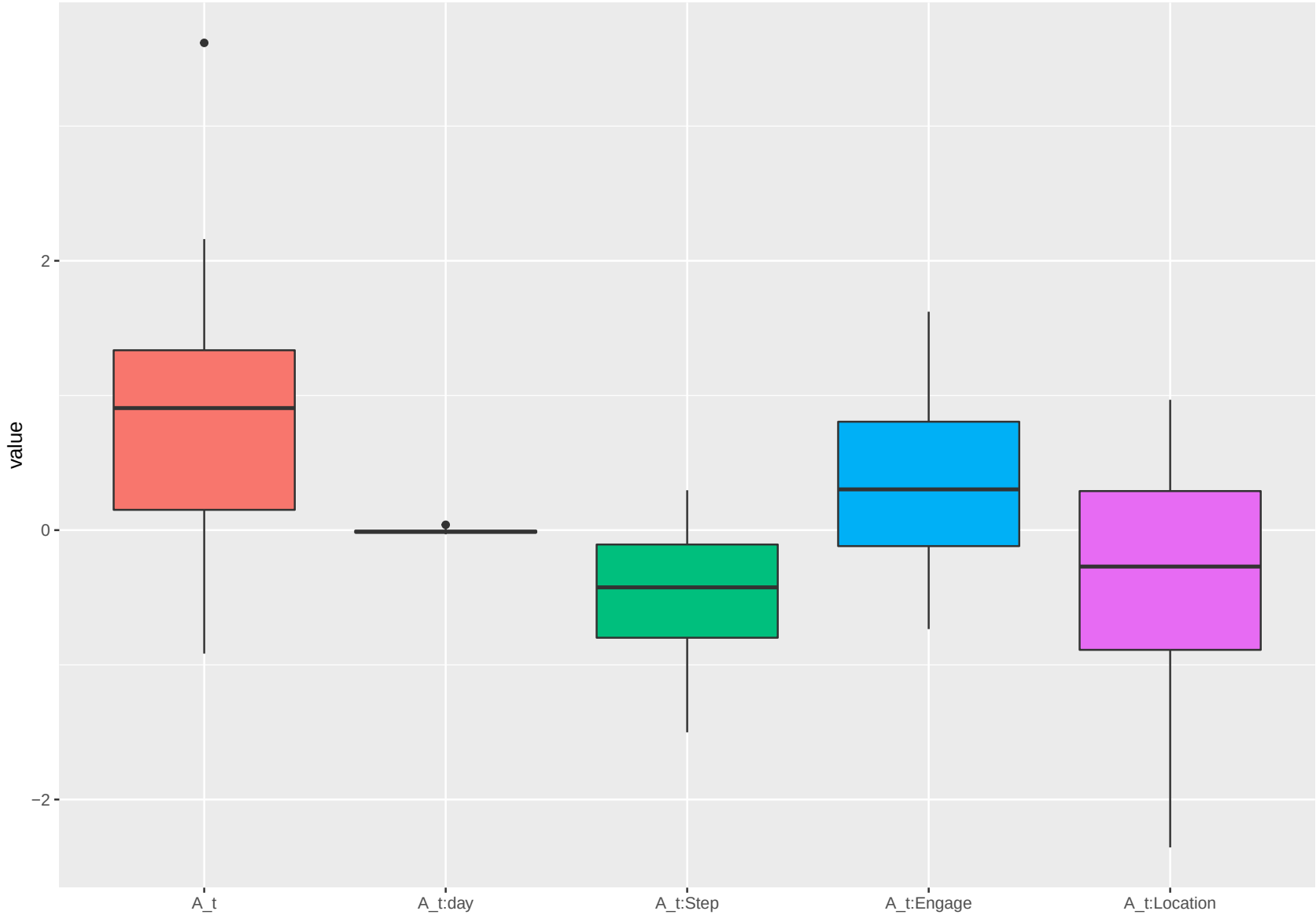


Figure 5: HS Randomization Schema

Warm-Start Policy

term	estimate	std.error	statistic	p.value
(Intercept)	2.85	0.0954	29.87	0.000
MRT_action_c	0.75	0.1908	3.93	0.000
day	0.00	0.0010	-3.18	0.001
logpresteps	3.24	0.0401	80.78	0.000
engagement	0.09	0.0489	1.92	0.054
other.location	-0.03	0.0688	-0.42	0.677
MRT_action_c:day	-0.01	0.0021	-5.35	0.000
MRT_action_c:logpresteps	-0.48	0.0802	-5.97	0.000
MRT_action_c:engagement	0.36	0.0979	3.71	0.000
MRT_action_c:other.location	-0.26	0.1377	-1.88	0.061

Assessing individual-level heterogeneity in effects



Summary

- Conceptual model (domain science)
- Micro-randomized trial
- Causal excursion effect and exploratory data analysis
- Improved treatment policy
- Confirmatory Trial

Debiased Machine Learning

DML for ATE

1. Divide up the data into Q evenly sized folds
2. Fit $\hat{m}_a = E[Y|X, A = a]$ and $\hat{e} = E[A|X]$
3. Use $\hat{m}_a$ and $\hat{e}$ to define score functions:

$$\hat{m}(X,1) - \hat{m}(X,0) + \frac{A(Y - \hat{m}(X,1))}{\hat{e}(X)} - \frac{(1-A)(Y - \hat{m}(X,0))}{1 - \hat{e}(X)} - \theta$$

DML + MRT

DML + MRT

```
## Warning in summary.glm(object): observations with zero weight not us
## calculating dispersion
```

	Estimate	Std.err	Wald	Pr(>|W|)
(Intercept)	-0.1928047	0.0617901	9.736382	0.0018066
MRT_action	0.4687127	0.0753517	38.692508	0.0000000
day	0.0040237	0.0010975	13.441471	0.0002461
MRT_action:day	-0.0098642	0.0013248	55.436377	0.0000000

```
## Warning in summary.glm(object): observations with zero weight not us
## calculating dispersion
```

	Estimate	Std.err	Wald	Pr(>|W|)
(Intercept)	2.8525256	0.0553947	2651.68918	0.0000000
MRT_action_c	0.4560119	0.0801226	32.39234	0.0000000
day	-0.0031156	0.0009272	11.29197	0.0007784
logpresteps	3.2348948	0.0391158	6839.36303	0.0000000
MRT_action_c:day	-0.0096163	0.0013552	50.35353	0.0000000

Additional resources

- Centered and weighted least squares [Boruvka et al., 2018]
- Delayed effect [Boruvka et al. 2018, Qian et al., 2020]
- Stratified MRT [Dempsey et al. 2020]
- Excursion effect with binary outcome [Qian et al., 2020]
- Sample size calculator (continuous)
- Sample size calculator (binary)
- Excursion effect with cluster-level heterogeneity [Dempsey et al. 2021]
- Increasing efficiency of causal excursion effect estimation via orthogonalization [Dempsey et al. 2022]