Estimating time-varying causal effect moderation in mobile health with binary outcomes

Tianchen Qian
Joint work with Daniel Almirall, Predrag Klasnja, Hyesun Yoo, Susan Murphy

Department of Statistics
Harvard University

February 18, 2019

BariFit MRT

- A micro-randomized trial (MRT) to promote weight maintenance among people who received bariatric surgery.
- Data collected from:
  - Fitbit tracker (step count)
  - user self-report (weight, calories intake)
- mHealth intervention components:
  - daily step goals
  - actionable activity suggestions
  - reminders to track food intake
  - ...
- This talk: assess the effect of (reminders to track food intake)

Data in an MRT

- On each individual: $O_1, A_1, Y_2, \ldots, O_T, A_T, Y_{T+1}$.
- t: decision point.
- A_t: treatment indicator at decision point t.
- O_t: observation accrued between decision point $t-1$ and decision point t.
- History $H_t = (O_1, A_1, Y_2, \ldots, O_t)$: information accrued prior to decision point t.

Decision Points t

- Times at which a treatment might be provided
- Times that the treatment is likely to be beneficial
- BariFit: food track reminder may be sent every morning. $t = 1, 2, \ldots, 112$ (112 days)

Treatment indicator A_t

- Whether a treatment is provided at decision point t
- (What type of treatment)
- Here we assume binary ($A_t \in \{0, 1\}$)
- Randomization probability $p_t(H_t) := P(A_t = 1 \mid H_t)$
- BariFit: whether a text message of food track reminder is sent. $p_t(H_t) = 0.5$.

Proximal outcome Y_{t+1}

- Outcome measured after decision point t (assumed to be binary here)
- Something that the treatment is directly targeting
- BariFit: whether the individual completes food log on that day
- Note the subscript!

Observation O_t

- Observation accrued between decision point $t-1$ and decision point t.
- O_1 includes baseline variables.
- BariFit: Fitbit tracker (step count)
  user self-report (e.g., weekly weight)
  baseline variables (e.g., age, gender)

Availability I_t

- Treatment A_t can only be delivered at a decision point if an individual is available.
- Available: $I_t = 1$; unavailable: $I_t = 0$. $I_t \in O_t$.
- Safety and ethical consideration: e.g., an individual is unavailable for a physical activity suggestion message if she is driving.
- Treatment effect is defined conditional on availability. (later)
- BariFit: for food track reminder, individuals are always available.
- Availability is different from adherence!

Conceptual models

- Data: $O_1, A_1, Y_2, \ldots, O_T, A_T, Y_{T+1}$
- $H_t = (O_1, A_1, Y_2, \ldots, O_t)$
- Usually data analysts fit a series of models

$$\begin{aligned} Y_{t+1} &\text{ `} \sim \text{' } g(H_t)^T \alpha + \boxed{\beta_0 A_t}, \\ Y_{t+1} &\text{ `} \sim \text{' } g(H_t)^T \alpha + \boxed{\beta_0 A_t + \beta_1 A_t S_t}, \\ &\ldots \end{aligned}$$

- $g(H_t)$: summaries from H_t; "control variables"
- S_t: potential moderators (e.g., day in the study)
- β_0, β_1: quantities of interest
- ' $\sim$ ': e.g., logit or log for binary Y

Goal

- Develop statistical methods to model and estimate the treatment effect
- Be consistent with the scientific understanding of the β coefficients
- Use control variables $g(H_t)$ for noise reduction in a robust way

Potential outcomes

- To mathematize the problem, we use potential outcomes notation (e.g., Rubin (1974))
- Define $\bar{a}_t = (a_1, \ldots, a_t)$ where $a_1, \ldots, a_t \in \{0, 1\}$
- $O_t(\bar{a}_{t-1})$: O_t that would have been observed if individual received treatment history $\bar{a}_{t-1}$.
- Similarly, $Y_{t+1}(\bar{a}_t)$, $H_t(\bar{a}_{t-1})$

Causal excursion effect

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

Causal excursion effect

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

Causal excursion effect

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

- Contrasting two excursions: following $\bar{A}_{t-1}$, then receive treatment ($A_t = 1$) vs. no treatment ($A_t = 0$) at time t.

Causal excursion effect

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

- Contrasting two excursions: following $\bar{A}_{t-1}$, then receive treatment ($A_t = 1$) vs. no treatment ($A_t = 0$) at time t.
- $S_t(\bar{A}_{t-1}) \subset H_t(\bar{A}_{t-1})$: a vector of summary variables chosen from $H_t(\bar{A}_{t-1})$.
- Effect is marginal over all variables in $H_t(\bar{A}_{t-1})$ that are not in $S_t(\bar{A}_{t-1})$

Causal excursion effect

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

- Contrasting two excursions: following $\bar{A}_{t-1}$, then receive treatment ($A_t = 1$) vs. no treatment ($A_t = 0$) at time t.
- $S_t(\bar{A}_{t-1}) \subset H_t(\bar{A}_{t-1})$: a vector of summary variables chosen from $H_t(\bar{A}_{t-1})$.
- Effect is marginal over all variables in $H_t(\bar{A}_{t-1})$ that are not in $S_t(\bar{A}_{t-1})$
- Conditional on being available: $I_t(\bar{A}_{t-1}) = 1$.

Causal excursion effect

$$\log \frac{E\{Y_{t+1}(\bar{A}_{t-1}, 1) \mid S_t(\bar{A}_{t-1}), I_t(\bar{A}_{t-1}) = 1\}}{E\{Y_{t+1}(\bar{A}_{t-1}, 0) \mid S_t(\bar{A}_{t-1}), I_t(\bar{A}_{t-1}) = 1\}}$$

- Contrasting two excursions: following $\bar{A}_{t-1}$, then receive treatment ($A_t = 1$) vs. no treatment ($A_t = 0$) at time t.
- $S_t(\bar{A}_{t-1}) \subset H_t(\bar{A}_{t-1})$: a vector of summary variables chosen from $H_t(\bar{A}_{t-1})$.
- Effect is marginal over all variables in $H_t(\bar{A}_{t-1})$ that are not in $S_t(\bar{A}_{t-1})$
- Conditional on being available: $I_t(\bar{A}_{t-1}) = 1$.

Examples

- $S_t(\bar{A}_{t-1}) = 1$: average treatment effect

$$\log \frac{E\{Y_{t+1}(\bar{A}_{t-1}, 1) \mid I_t(\bar{A}_{t-1}) = 1\}}{E\{Y_{t+1}(\bar{A}_{t-1}, 0) \mid I_t(\bar{A}_{t-1}) = 1\}}$$

- $S_t(\bar{A}_{t-1}) = (1, \text{ day in study})$

$$\log \frac{E\{Y_{t+1}(\bar{A}_{t-1}, 1) \mid \text{day}_t, I_t(\bar{A}_{t-1}) = 1\}}{E\{Y_{t+1}(\bar{A}_{t-1}, 0) \mid \text{day}_t, I_t(\bar{A}_{t-1}) = 1\}}$$

Identifiability assumptions

Assumption (consistency)

The observed data equals the potential outcome under observed treatment assignment: $O_t = O_t(\bar{A}_{t-1})$ for every t.

Identifiability assumptions

Assumption (consistency)

The observed data equals the potential outcome under observed treatment assignment: $O_t = O_t(\bar{A}_{t-1})$ for every t.

Assumption (positivity)

For every t, for every possible history H_t with $I_t = 1$, $P(A_t = a \mid H_t, I_t = 1) > 0$ for $a \in \{0, 1\}$.

Identifiability assumptions

Assumption (consistency)

The observed data equals the potential outcome under observed treatment assignment: $O_t = O_t(\bar{A}_{t-1})$ for every t.

Assumption (positivity)

For every t, for every possible history H_t with $I_t = 1$, $P(A_t = a \mid H_t, I_t = 1) > 0$ for $a \in \{0, 1\}$.

Assumption (sequential ignorability)

For every t, the potential outcomes $\{O_{t+1}(\bar{a}_t), A_{t+1}(\bar{a}_t), \ldots, O_{T+1}(\bar{a}_T) : \bar{a}_T \in \{0,1\}^{\otimes T}\}$ are independent of A_t conditional on H_t.

Identifiability

Under assumptions on previous slide,

$$\log \frac{E\{Y_{t+1}(\bar{A}_{t-1},1) \mid S_t(\bar{A}_{t-1}), I_t(\bar{A}_{t-1})=1\}}{E\{Y_{t+1}(\bar{A}_{t-1},0) \mid S_t(\bar{A}_{t-1}), I_t(\bar{A}_{t-1})=1\}}$$

$$= \log \frac{E\{E(Y_{t+1} \mid A_t = 1, H_t) | S_t, I_t = 1\}}{E\{E(Y_{t+1} \mid A_t = 0, H_t) | S_t, I_t = 1\}}$$

$$= \log \frac{E\left\{ \frac{\mathbb{1}(A_t=1)Y_{t+1}}{p_t(H_t)} \middle| S_t, I_t = 1 \right\}}{E\left\{ \frac{\mathbb{1}(A_t=0)Y_{t+1}}{1-p_t(H_t)} \middle| S_t, I_t = 1 \right\}}$$

Outline

1. Introduction
2. Special case: conditional on H_t
3. A simple and robust estimator
4. Simulation study
5. Analysis of BariFit
6. Extension: proximal outcome defined over a duration
7. Summary

Special case: conditional on H_t

Suppose for all t,

$$\log \frac{E(Y_{t+1} \mid A_t = 1, H_t, I_t = 1)}{E(Y_{t+1} \mid A_t = 0, H_t, I_t = 1)} = S_t^T \beta$$

holds for some $S_t \subset H_t$ and some parameter β.

Special case: conditional on H_t

$$\log \frac{E(Y_{t+1} \mid A_t = 1, H_t, I_t = 1)}{E(Y_{t+1} \mid A_t = 0, H_t, I_t = 1)} = S_t^T \beta$$

- working model $\exp\{g(H_t)^T \alpha\}$

Special case: conditional on H_t

$$\log \frac{E(Y_{t+1} \mid A_t = 1, H_t, I_t = 1)}{E(Y_{t+1} \mid A_t = 0, H_t, I_t = 1)} = S_t^T \beta$$

- working model $\exp\{g(H_t)^T \alpha\}$
- If working model is correct, then

$$E(Y_{t+1} \mid H_t, I_t = 1) = e^{g(H_t)^T \alpha + A_t S_t^T \beta}, \qquad (1)$$

and one can use GEE to estimate α and β.

Special case: conditional on H_t

$$\log \frac{E(Y_{t+1} \mid A_t = 1, H_t, I_t = 1)}{E(Y_{t+1} \mid A_t = 0, H_t, I_t = 1)} = S_t^T \beta$$

- working model $\exp\{g(H_t)^T \alpha\}$
- If working model is correct, then

$$E(Y_{t+1} \mid H_t, I_t = 1) = e^{g(H_t)^T \alpha + A_t S_t^T \beta}, \quad (1)$$

  and one can use GEE to estimate α and β.
- However, (1) is required to guarantee the consistency of GEE.

Semiparametric model

$$\log \frac{E(Y_{t+1} \mid A_t = 1, H_t, I_t = 1)}{E(Y_{t+1} \mid A_t = 0, H_t, I_t = 1)} = S_t^T \beta$$

$$E(Y_{t+1} \mid A_t = 0, H_t, I_t = 1) = e^{g(H_t)^T \alpha}$$

Semiparametric model

- Assume this (parametric part)

$$\log \frac{E(Y_{t+1} \mid A_t = 1, H_t, I_t = 1)}{E(Y_{t+1} \mid A_t = 0, H_t, I_t = 1)} = S_t^T \beta$$

$$E(Y_{t+1} \mid A_t = 0, H_t, I_t = 1) = e^{g(H_t)^T \alpha}$$

- Don't assume this; this becomes nonparametric

Semiparametric model

- Assume this (parametric part)

$$\log \frac{E(Y_{t+1} \mid A_t = 1, H_t, I_t = 1)}{E(Y_{t+1} \mid A_t = 0, H_t, I_t = 1)} = S_t^T \beta$$

$$E(Y_{t+1} \mid A_t = 0, H_t, I_t = 1) = e^{g(H_t)^T \alpha}$$

- Don't assume this; this becomes nonparametric

- "semi-parametric" model; Newey (1990), Tsiatis (2007)
- Robins (1994), structural nested mean model

Semiparametric estimator

- The following estimator for β, derived based on Robins (1994), is semiparametric locally efficient:

$$\mathbb{P}_n \sum_{t=1}^{T} I_t e^{-A_t S_t^T \beta} (Y_{t+1} - e^{g(H_t)^T \alpha + A_t S_t^T \beta}) V_t \begin{bmatrix} g(H_t) \\ (A_t - p_t(H_t)) S_t \end{bmatrix} = 0$$

$$\Longrightarrow (\hat{\alpha}, \hat{\beta})$$

- Robust: $\hat{\beta}$ is consistent for β with any choice of control variables $g(H_t)$
- "Locally efficient": $\hat{\beta}$ has the smallest asymptotic variance (among all semiparametric regular and asymptotically linear estimators) if $e^{g(H_t)^T \alpha}$ is a correct model for $E(Y_{t+1} \mid H_t, A_t = 0, I_t = 1)$.

$$V_t := \frac{e^{S_t^T \beta}}{e^{S_t^T \beta}\{1 - e^{g(H_t)^T \alpha}\} p_t(H_t) + \{1 - e^{g(H_t)^T \alpha + S_t^T \beta}\}(1 - p_t(H_t))}.$$

Intuition for robustness

$$\mathbb{P}_n \sum_{t=1}^{T} I_t \, e^{-A_t S_t^T \beta} \left(Y_{t+1} - e^{g(H_t)^T \alpha + A_t S_t^T \beta} \right) V_t \left[\frac{g(H_t)}{(A_t - p_t(H_t)) \, S_t} \right] = 0$$

Intuition for robustness

$$\mathbb{P}_n \sum_{t=1}^{T} I_t \, e^{-A_t S_t^T \beta}\left(Y_{t+1} - e^{g(H_t)^T \alpha + A_t S_t^T \beta}\right) V_t \begin{bmatrix} g(H_t) \\ (A_t - p_t(H_t)) \, S_t \end{bmatrix} = 0$$

- $e^{-A_t S_t^T \beta} Y_{t+1}$: "blipped-down" outcome

$$E(e^{-A_t S_t^T \beta} Y_{t+1} \mid H_t, A_t) = E\{Y_{t+1}(\bar{A}_{t-1}, 0) \mid H_t, A_t\}$$

Intuition for robustness

$$\mathbb{P}_n \sum_{t=1}^{T} I_t \, e^{-A_t S_t^T \beta}\left(Y_{t+1} - e^{g(H_t)^T \alpha + A_t S_t^T \beta}\right) V_t \left[\frac{g(H_t)}{(A_t - p_t(H_t)) \, S_t} \right] = 0$$

- $e^{-A_t S_t^T \beta} Y_{t+1}$: "blipped-down" outcome

$$E(e^{-A_t S_t^T \beta} Y_{t+1} \mid H_t, A_t) = E\{Y_{t+1}(\bar{A}_{t-1}, 0) \mid H_t, A_t\}$$

- $e^{g(H_t)^T \alpha}$: a function of H_t

Intuition for robustness

$$\mathbb{P}_n \sum_{t=1}^{T} I_t \, e^{-A_t S_t^T \beta} \left(Y_{t+1} - e^{g(H_t)^T \alpha + A_t S_t^T \beta} \right) V_t \begin{bmatrix} g(H_t) \\ (A_t - p_t(H_t)) \, S_t \end{bmatrix} = 0$$

- $e^{-A_t S_t^T \beta} Y_{t+1}$: “blipped-down” outcome

$$E(e^{-A_t S_t^T \beta} Y_{t+1} \mid H_t, A_t) = E\{ Y_{t+1}(\bar{A}_{t-1}, 0) \mid H_t, A_t \}$$

- $e^{g(H_t)^T \alpha}$: a function of H_t
- $A_t - p_t(H_t)$: centered treatment assignment

Intuition for robustness

$$\mathbb{P}_n \sum_{t=1}^{T} I_t \, e^{-A_t S_t^T \beta}\left(Y_{t+1} - e^{g(H_t)^T \alpha + A_t S_t^T \beta}\right) V_t \left[\frac{g(H_t)}{(A_t - p_t(H_t)) \, S_t} \right] = 0$$

- $e^{-A_t S_t^T \beta} Y_{t+1}$: "blipped-down" outcome

$$E(e^{-A_t S_t^T \beta} Y_{t+1} \mid H_t, A_t) = E\{Y_{t+1}(\bar{A}_{t-1}, 0) \mid H_t, A_t\}$$

- $e^{g(H_t)^T \alpha}$: a function of H_t
- $A_t - p_t(H_t)$: centered treatment assignment
- $\Longrightarrow$ The blue term and the red term are orthogonal to each other (with any $g(H_t)$).
- $\Longrightarrow$ robustness

Treatment effect of interest

- Special case considered so far: fully conditional on H_t

$$\log \frac{E(Y_{t+1} \mid A_t = 1, H_t, I_t = 1)}{E(Y_{t+1} \mid A_t = 0, H_t, I_t = 1)} = S_t^T \beta$$

- What makes more scientific sense:
  marginal over variables in H_t but not in S_t

$$\log \frac{E\{E(Y_{t+1} \mid A_t = 1, H_t) | S_t, I_t = 1\}}{E\{E(Y_{t+1} \mid A_t = 0, H_t) | S_t, I_t = 1\}} = S_t^T \beta$$

Outline

1. Introduction
2. Special case: conditional on H_t
3. **A simple and robust estimator**
4. Simulation study
5. Analysis of BariFit
6. Extension: proximal outcome defined over a duration
7. Summary

A simple and robust estimator for marginalized effect

- Control variables: $\exp\{g(H_t)^T \alpha\}$

$$\mathbb{P}_n \sum_{t=1}^{T} I_t \, e^{-A_t S_t^T \beta} \left(Y_{t+1} - e^{g(H_t)^T \alpha + A_t S_t^T \beta} \right) V_t \begin{bmatrix} g(H_t) \\ (A_t - p_t(H_t)) \, S_t \end{bmatrix} = 0$$

- Because the model assumption is now on the marginalized treatment effect, the blue term and the red term are no longer orthogonal.

Introduction

Conditional on H_t

Estimator

Simulation

BariFit

Extension

Summary

References

A simple and robust estimator for marginalized effect

- Control variables: $\exp\{g(H_t)^T\alpha\}$

$$\mathbb{P}_n \sum_{t=1}^{T} I_t \, e^{-A_t S_t^T \beta} (Y_{t+1} - e^{g(H_t)^T \alpha + A_t S_t^T \beta}) \qquad \begin{bmatrix} g(H_t) \\ \qquad S_t \end{bmatrix} = 0$$

- Choose $\tilde{p}_t(s) \in (0, 1)$
- 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 treatment: $A_t \to (A_t - \tilde{p}_t(S_t))$

A simple and robust estimator for marginalized effect

- Control variables: $\exp\{g(H_t)^T \alpha\}$

$$\mathbb{P}_n \sum_{t=1}^{T} I_t \, e^{-A_t S_t^T \beta}\left(Y_{t+1} - e^{g(H_t)^T \alpha + A_t S_t^T \beta}\right) W_t \left[\frac{g(H_t)}{(A_t - \tilde{p}_t(S_t)) \, S_t} \right] = 0$$

- Choose $\tilde{p}_t(s) \in (0, 1)$
- 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 treatment: $A_t \to (A_t - \tilde{p}_t(S_t))$
- W_t and $\tilde{p}_t(S_t)$ make the blue term and the red term orthogonal to each other.
- Boruvka et al. (2018)

A simple and robust estimator for marginalized effect

- Suppose $(\hat{\alpha}, \hat{\beta})$ solve the estimating equation:

$$\mathbb{P}_n \sum_{t=1}^{T} I_t e^{-A_t S_t^T \beta} (Y_{t+1} - e^{g(H_t)^T \alpha + A_t S_t^T \beta}) W_t \begin{bmatrix} g(H_t) \\ (A_t - \tilde{p}_t(S_t)) S_t \end{bmatrix} = 0$$

- Under moment conditions, $\hat{\beta}$ is consistent for β and is asymptoticaly normal if

$$\log \frac{E\{E(Y_{t+1} \mid A_t = 1, H_t) | S_t, I_t = 1\}}{E\{E(Y_{t+1} \mid A_t = 0, H_t) | S_t, I_t = 1\}} = S_t^T \beta$$

- Robustness: consistency of $\hat{\beta}$ doesn't require $e^{g(H_t)^T \alpha}$ to be a correct model for $E(Y_{t+1} \mid A_t = 0, H_t, I_t = 1)$

Choice of $\tilde{p}_t$

- Choice of $\tilde{p}_t(S_t)$ determines marginalization over time under model misspecification of treatment effect.
- For example, if $S_t = 1$, $\tilde{p}_t(S_t) = \tilde{p}$ for some $\tilde{p} \in (0, 1)$, then $\hat{\beta}$ converges to

$$\beta' = \log \frac{\sum_{t=1}^{T} E\{E(Y_{t+1} \mid H_t, A_t = 1) \mid I_t = 1\}}{\sum_{t=1}^{T} E\{E(Y_{t+1} \mid H_t, A_t = 0) \mid I_t = 1\}},$$

Outline

1. Introduction
2. Special case: conditional on H_t
3. A simple and robust estimator
4. **Simulation study**
5. Analysis of BariFit
6. Extension: proximal outcome defined over a duration
7. Summary

Simulation: generative model

- $E(Y_{t+1} \mid H_t, A_t) = f(Z_t) \exp\{A_t(0.1 + 0.3Z_t)\}$
- Covariate Z_t: takes value from $0, 1, 2$ with equal probability
- $f(Z_t) = 0.2\mathbb{1}(Z_t = 0) + 0.5\mathbb{1}(Z_t = 1) + 0.4\mathbb{1}(Z_t = 2)$
- $P(A_t = 1 \mid H_t) = 0.2$
- $I_t = 1$: always available

True treatment effect

$$\log \frac{E(Y_{t+1} \mid A_t = 1, H_t, I_t = 1)}{E(Y_{t+1} \mid A_t = 0, H_t, I_t = 1)} = 0.1 + 0.3Z_t$$

$$\log \frac{E\{E(Y_{t+1} \mid A_t = 1, H_t) | I_t = 1\}}{E\{E(Y_{t+1} \mid A_t = 0, H_t) | I_t = 1\}} = 0.477$$

So if we let $S_t = 1$ in the analysis model, the semiparametric locally efficient estimator would be inconsistent, and the robust estimator would be consistent.

Result

Estimator	Sample size	Bias	SD	RMSE	CP
robust	30	0.000	0.072	0.072	0.96
	50	-0.001	0.058	0.058	0.94
	100	0.002	0.041	0.041	0.94
locally efficient	30	**0.047**	0.070	0.084	0.94
	50	**0.047**	0.057	0.073	**0.89**
	100	**0.051**	0.040	0.064	**0.79**
GEE	30	**0.040**	0.068	0.080	**0.92**
	50	**0.040**	0.055	0.068	**0.88**
	100	**0.043**	0.039	0.058	**0.78**

* SD: standard deviation. RMSE: root mean squared error.
CP: 95% confidence interva coverage probability.

Outline

1. Introduction
2. Special case: conditional on H_t
3. A simple and robust estimator
4. Simulation study
5. **Analysis of BariFit**
6. Extension: proximal outcome defined over a duration
7. Summary

BariFit food track reminder

- $n = 45$ participants:
- 112 days = 112 decision points
- A_t: food track reminder is sent as text message with probability 0.5 every morning
- Y_{t+1}: binary indicator of whether the individual completes food log on that day ($Y_{t+1} = 1$ if logged calories > 0)

Estimated effect

$$\log E(Y_{t+1}) \text{ `}\sim\text{' } g(H_t)^T \alpha + \beta_0 A_t$$

- Control variables $g(H_t)$: day in study, gender, food log completion on previous day
- Estimation result for β_0:

Method	Estimate	SE	95% CI	p-value
robust	0.014	0.021	(-0.028, 0.056)	0.50
locally efficient	0.011	0.014	(-0.017, 0.039)	0.44

* SE: standard error. CI: confidence interval.

Initial conclusion

- The data indicates that there is no detectable effect of the food track reminder text message on the food log completion of that day.
- For the next iteration of BariFit...
  - Implement the reminder as part of a native app (instead of text messages) — to improve effectiveness
  - Or, combine it with other text messages (such as daily step goal) that are sent to the individuals in the morning — to reduce burden

Outline

1. Introduction
2. Special case: conditional on H_t
3. A simple and robust estimator
4. Simulation study
5. Analysis of BariFit
6. **Extension: proximal outcome defined over a duration**
7. Summary

Proximal outcome defined over a duration of time

- Sometimes the proximal outcome is measured over a duration of time during which other treatments may occur.
- On each individual: $O_1, A_1, \ldots, O_T, A_T, O_{T+1}$.
- Proximal outcome $Y_{t+\Delta}$, is a known function of the individual's data within a subsequent window of length Δ; i.e., $Y_{t+\Delta} = y(O_{t+1}, A_{t+1}, \ldots, O_{t+\Delta-1}, A_{t+\Delta-1}, O_{t+\Delta})$ for some known function $y(\cdot)$.
- Previously, $Y_{t+1} = y(O_{t+1})$.

Causal excursion effect

Let $\bar{0}$ be a vector of length $\Delta - 1$.

$$\log \frac{E\{Y_{t+\Delta}(\bar{A}_{t-1}, 1, \bar{0}) \mid S_t(\bar{A}_{t-1}), I_t(\bar{A}_{t-1}) = 1\}}{E\{Y_{t+\Delta}(\bar{A}_{t-1}, 0, \bar{0}) \mid S_t(\bar{A}_{t-1}), I_t(\bar{A}_{t-1}) = 1\}} = S_t^T \beta$$

Estimating equation for β:

$$\mathbb{P}_n \sum_{t=1}^{T} I_t e^{-A_t S_t^T \beta} (Y_{t+1} - e^{g(H_t)^T \alpha + A_t S_t^T \beta}) \tilde{W}_t \begin{bmatrix} g(H_t) \\ (A_t - \tilde{p}_t(S_t)) S_t \end{bmatrix} = 0,$$

where

$$\tilde{W}_t = W_t \times \prod_{j=t+1}^{t+\Delta-1} \frac{\mathbb{1}(A_j = 0)}{1 - p_j(H_j)}$$

Summary

- Definition of causal excursion effect for binary outcome
- A semiparametric locally efficient estimator for the effect conditional on history observed up to that time point, H_t
- A simple and robust estimator for the effect marginalized over all but a small subset S_t of H_t
- An extension to settings where the proximal outcome is defined over a duration of time during which other treatments may occur
- An analysis of marginal effect of food track reminder in BariFit MRT

Reference

Boruvka, A., D. Almirall, K. Witkiewitz, and S. A. Murphy (2018). Assessing time-varying causal effect moderation in mobile health. *Journal of the American Statistical Association 113*(523), 1112–1121.

Newey, W. K. (1990). Semiparametric efficiency bounds. *Journal of applied econometrics 5*(2), 99–135.

Robins, J. M. (1994). Correcting for non-compliance in randomized trials using structural nested mean models. *Communications in Statistics-Theory and methods 23*(8), 2379–2412.

Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of educational Psychology 66*(5), 688.

Tsiatis, A. (2007). *Semiparametric theory and missing data*. Springer Science & Business Media.