Revision response for paper "Estimating Causal Excursion Effects for Binary Outcomes Using Per-Decision Inverse Probability Weighting"

Yihan Bao, Lauren Bell, Elizabeth Williamson, Claire Garnett, Tianchen Qian

November 27, 2022

Response to Associate Editor

Thank you for your comments. We have revised the paper carefully to address your comments. Your original comments are in italic and the reply is in normal font. Quotes from the revised version of the manuscript are indented and in a smaller font, and the text added during revision are colored in blue.

Major comments

1. *I struggled to understand the causal parameters targeted in this paper. The defined potential outcomes and casual effects seem incorrect to me. This issue must be remedied. More specifically.*

    - *I was confused by Equation (1), because I think $\bar{a}_t$ should represent a fixed treatment sequence, and not a random one (see also my specific comment on this issue). Therefore, how come the history includes A_1,$A_2(a_1)$,$A_3(\bar{a}_2)$,...? Shouldnt the treatment values be just the fixed treatment sequences a_1,a_2,a_3...,a_{t-1}? Similarly, according to the current definitions in the manuscript, the left hand side of Equation (2) are random variables (functions of the random variables $\bar{A}_{t-1}$ and as indicated by the authors later on this page, these are stochastic processes). This is also relevant to Assumption 1. This means that, for example, the fully marginal excursion effect targeted in the data example, depends on the distribution of treatment assignment, right?*

      Reply. Thanks for your comment. You are correct about the potential outcome for the history, and we corrected it in the paper as follows:

The potential outcome of H_t under $\bar{a}_{t-1}$ is

$$H_t(\bar{a}_{t-1}) = \{X_1, a_1, X_2(a_1), a_2, X_3(\bar{a}_2), a_3, \ldots, X_{t-1}(\bar{a}_{t-2}), a_{t-1}, X_t(\bar{a}_{t-1})\}. \quad (1)$$

You are also correct that the fully marginal excursion effect indeed depends on the distribution of the treatment assignments, and this is a feature of the causal excursion effects. We added the following paragraphs in Section 2.2 where we introduce the causal excursion effect to elaborate on this point:

> The causal excursion effect (3) is commonly used in the primary and secondary analyses for MRTs (Nahum-Shani et al., 2021; Klasnja et al., 2021). It differs from most other causal effects under time-varying treatments such as marginal structural models and structural nested mean models (Robins, 1994; Robins et al., 2000). Instead of contrasting fixed treatment trajectories, (3) contrasts between two stochastic treatment trajectories, $(\bar{A}_{t-1}, 1, \bar{0}_{\Delta-1})$ and $(\bar{A}_{t-1}, 0, \bar{0}_{\Delta-1})$. That is, (3) is a causal contrast between two excursions from the treatment policy: following the treatment policy up to time $t-1$ and then deviating from the policy to assign $A_t = 1$ and no treatment for the next $\Delta - 1$ decision points, and following the treatment policy up to time $t-1$ and then deviating from the policy to assign $A_t = 0$ and no treatment for the next $\Delta - 1$ decision points. (3) is conditional on $I_t(\bar{A}_{t-1}) = 1$ and $S_t(\bar{A}_{t-1})$, meaning that it concerns the subpopulation who, after following the treatment policy up to time $t-1$, are eligible to be randomized at t and within a stratum defined by $S_t(\bar{A}_{t-1})$. In practice we often impose a model on how $\text{CEE}\{S_t(\bar{A}_{t-1})\}$ depends on $S_t(\bar{A}_{t-1})$ to pull across S_t-strata.
>
> $\text{CEE}\{S_t(\bar{A}_{t-1})\}$ depends on the treatment policy in the MRT, because the variables in $H_t(\bar{A}_{t-1}) \setminus S_t(\bar{A}_{t-1})$ are marginalized over in the conditional expectations in (3). Related marginalization ideas were considered by Robins (2004); Neugebauer et al. (2007). Such dependence is scientifically desirable for two reasons. First, the causal excursion effect approximates the treatment effect in a real-world implementation. A good MRT policy would alreay have incorporated implementation considerations such as burden and feasibility through the choice of randomization probability and availability criteria, so a feasible policy will not deviate too far from the MRT policy. Second, the causal excursion effect indicates the effective deviations from the MRT policy and how it might be improved. A fully marginal effect (by setting $S_t(\bar{A}_{t-1}) = 1$ in (3)) indicates whether a treatment is worth further investigation, and an effect modification analysis (by setting $S_t(\bar{A}_{t-1})$ to be certain time-varying covariates in (3)) indicates whether the MRT policy should be modified to depend on time-varying covariates. In addition, the dependence of $\text{CEE}\{S_t(\bar{A}_{t-1})\}$ on the MRT policy resembles the primary analysis in factorial designs and allows one to design trials with a higher power to detect meaningful effects. For a more comprehensive discussion on causal excursion effect and its use in MRT, see Boruvka et al. (2018), Qian et al. (2021), and Qian et al. (2022).

- *It seems to me that the causal effect is well-defined only among those eligible under certain treatment decision process? if one compares the mean outcomes under two different treatment processes, the mean is over different groups. More gen-*

erally, because $S_t(\bar{A}_{t-1})$ is a random variable, does that mean that the population for which the causal effect is defined is random and depends on the previous treatments' assignment mechanism? What does that mean about the practical utility of such an estimand?

Reply. Thanks for your comment.

Regarding your first question, you are correct that the causal effect is only meaningful among available individuals (those with $I_t(\bar{A}_{t-1}) = 1$). And indeed, for different treatment policies (not different observed treatment trajectories), the causal excursion effect can be over different groups. More generally, the causal excursion effect depends on the distribution of the treatments (a.k.a. the treatment policy in the MRT), and this point is elaborated in Section 2.2; see our reply to your previous comment.

Regarding your second question, $S_t(\bar{A}_{t-1})$ is indeed random, and conditional on $S_t(\bar{A}_{t-1})$ means that we are focusing on the following subpopulation: those who will be in the same $S_t(\bar{A}_{t-1})$-stratum when followed the MRT policy from time 1 to time $t-1$. We usually impose a model on the causal excursion effect, which implies that we pool across S_t-strata. You are correct that the $S_t(\bar{A}_{t-1})$-strata can depend on different treatment policies. We believe this is still a practical estimand, because of the following (added text in Section 2.2):

> Such dependence is scientifically desirable for two reasons. First, the causal excursion effect approximates the treatment effect in a real-world implementation. A good MRT policy would alreay have incorporated implementation considerations such as burden and feasibility through the choice of randomization probability and availability criteria, so a feasible policy will not deviate too far from the MRT policy. Second, the causal excursion effect indicates the effective deviations from the MRT policy and how it might be improved. A fully marginal effect (by setting $S_t(\bar{A}_{t-1}) = 1$ in (3)) indicates whether a treatment is worth further investigation, and an effect modification analysis (by setting $S_t(\bar{A}_{t-1})$ to be certain time-varying covariates in (3)) indicates whether the MRT policy should be modified to depend on time-varying covariates.

- *The authors should explain the connection between their work and existing research, for example, works by Robins, Hernan and colleagues on marginal and nested structural models, and other relevant literature for repeated treatment scenarios. If the outcome is always defined on some horizon, why the problem cannot be embedded in a time-to-event model?*

Reply. Thanks for your comment.

Regarding your first point, we have now discussed the difference between our work and the marginal structural models and structural nested models, in that these

models consider contrasts between fixed treatment trajectories whereas our model contrasts between stochastic treatment trajectories:

> It differs from most other causal effects under time-varying treatments such as marginal structural models and structural nested mean models (Robins, 1994; Robins et al., 2000). Instead of contrasting fixed treatment trajectories, (**??**) contrasts between two stochastic treatment trajectories, $(\bar{A}_{t-1},1,\bar{0}_{\Delta-1})$ and $(\bar{A}_{t-1},0,\bar{0}_{\Delta-1})$.

We also made the connection to two more recent literature that have similar ideas of marginalizing over a subset of the history:

> ...the variables in $H_t(\bar{A}_{t-1}) \setminus S_t(\bar{A}_{t-1})$ are marginalized over in the conditional expectations in (**??**). Related marginalization ideas were considered by Robins (2004); Neugebauer et al. (2007).

And we directed the readers to a list of literature, including a comprehensive comparison of various causal estimands under time-varying treatments by Guo et al. (2021):

> See Boruvka et al. (2018), Qian et al. (2021), and Qian et al. (2022) for more discussion on causal excursion effect and its use in MRT. See Guo et al. (2021) for a comprehensive comparison of various causal estimands under time-varying treatments.

We also added paragraphs to discuss the connections between our per-decision IPW idea and related efficiency improvement techniques in the literature. The following is added to Section 3:

> Related techniques of reducing the number of terms in the product of IP weights to improve efficiency were developed for other settings. In off-policy reinforcement learning, the per-decision importance sampling technique was proposed to construct IP weights related to ours by leveraging the fact that earlier outcomes cannot depend on later treatments (Precup, 2000). In Cox marginal structural model for survival analysis, to determine the contribution of a subject to a risk-set at time t, a stabilized weight was proposed which involves IP weights not through the end of the follow-up but only up till time t (Robins et al., 2000). Those settings are analogous to our setting with $T = 1$, i.e., when focusing on a single proximal outcome that decomposes into a sequence of sub-outcomes. Our contribution is to extend the per-decision IPW technique for estimating causal excursion effects on longitudinal proximal outcomes. In Supporting Information [TQ: to fill], we provide a more detailed explication on how our method connects to and differs from the literature.

And additional details are included in Supporting Information [TQ: to fill].

Regarding your second point, "why the problem cannot be embedded in a time-to-event model?" [TQ: depends on the quantity of interest.]

- *As pointed out by one of the reviewers, the notion of maximum property of proximal outcome seems to be identical to (discrete) time-to-event outcome, which*

can be studied using standard tools I think. What do the authors gain by their presentation?

Reply. Thanks for your comment. [TQ: Again a survival question. Our weight is indeed the same as the stablized weight in cox-MSM(?). What we gained is a new estimator for causal excursion effect, which was not addressed in the literature. Therefore, we should be clear about the connection of the maximum property with relevant notions in discrete-time survival setting (e.g., cite the education survival paper) after we introduce the maximum property.]

2. *Should β be β_t? or do the authors assume that β is the same for all t? or study the causal eect at single t only? am not sure if S_t can be of the same length for all t? In practice, do the authors recommend re-running the analysis for each t or putting structure as a function of certain time-varying variables as they do in the example.*

Reply. Thanks for your comment. We assume that β is the same for all t. We now clarified this by revising equation (5) to

$$\text{CEE}(S_t) = S_t^T \beta \quad \text{for all } 1 \le t \le T.$$

S_t is of the same length for all t, and we clarified this using two examples in Section 2.2 to:

> $S_t(\bar{A}_{t-1})$ is a length-p vector of summary variables from $H_t(\bar{A}_{t-1})$, usually chosen by the domain scientist based on the scientific question of interest regarding effect modification. In the Drink Less MRT example, we will first set $S_t(\bar{A}_{t-1}) = 1$ to obtain the fully marginal effect (i.e., averaged over all participants and all days in the study) of a push notification on the app use in the subsequent 72-hour window. We will then set $S_t(\bar{A}_{t-1}) = (1, \text{Day}_t)$ to assess how the effect is moderated by days in the study.

3. *Theorem 1: First, the authors state in the preamble to the Theorem that it provides proof of consistency. This does not seem to appear in the statement covered by the Theorem as it only pertains to convergence in distribution. The proof in the Web Appendix does show that the expectation of the estimating equation is zero under the true value of β, β^*. But some (what I believe to be minor techniques) parts of the argument are missing. This is also needed for the variance estimator to be consistent.*

Reply. Thanks for your comment. We revised the theorem statement to explicitly say that $\hat{\beta}$ is consistent:

> Under regularity conditions, $\hat{\beta}$ is consistent for β^* and $\sqrt{n}(\hat{\beta} - \beta^*)$ is asymptotically normal...

[TQ: To use some uniform convergence result to show that the variance estimator is consistent. Need to read Carrie's proof, first.]

4. *Could one extend Equation (5) Theorem 1 to any parametric model? I think this generalization should be rather simple, shouldnt it?*

   Reply. Thanks for your comment. Equation (5) and Theorem 1 can can be easily generalized to any transformations of S_t, by replacing the model $\text{CEE}(S_t) = S_t^T\beta$ with $\text{CEE}(S_t) = f_t(S_t)^T\beta$ for some function f_t and replacing S_t by $f_t(S_t)$ in the estimating quation. We now have this sentence after Equation (5):

   > Equation (**??**) can also accommodate flexible time-varying effects by including basis functions of t in S_t.

   We also added the following to the Discussion Section:

   > Equation (**??**) assumes that the causal excursion effect is linear in S_t. This can be immediately generalized to transformations of S_t by replacing S_t with $f_t(S_t)$ for arbitrary fixed function f_t.

Specific comments

1. *Page 4: This a bit nit-picking I admit, but I believe the potential outcomes exist without formally seeing them as an instantiation of a random variable. Feel free to ignore this comment if you disagree, but note its relevance to Equation (3) as mentioned above.*

   Reply. Thanks for your comment.

2. *Page 4: "Formally suppose $Y_{t,\Delta} = y(H_{t+\Delta})$ for some function $y(\cdot)$." I think this statement should be made clearer. Do the authors mean that the proximal outcome is defined as specific outcome that can be written as a function of the the history and Δ? What do the authors mean by "some function"?*

   Reply. Thanks for your comment. We meant that our method allows $Y_{t,\Delta}$ to be any transformation or summary of $H_{t+\Delta}$, i.e., everything observed before $A_{t+\Delta}$. We revised as follows to improve clarify:

   > Formally, suppose $Y_{t,\Delta}$ is a function of $H_{t+\Delta}$; i.e., the definition of the proximal outcome for decision point t can depend on anything observed up to and including $X_{t+\Delta}$.

3. *Page 7: I think it should be "provide**s** information".*

   Reply. Thanks for your comment. We corrected the typo.

4. *Page 9: "out method" ⇒ "our method".*

   Reply. Thanks for your comment. We corrected the typo.

5. *Page 10: Please clarify what Y_{it} stands for? I thought the definition of Y was always dependent on Δ.*

   Reply. Thanks for your comment. We think you are referring to $Y_{it,\Delta}$ in $\epsilon_{it}(\alpha,\beta) = Y_{it,\Delta} - e^{g(H_{it})^T\alpha + A_{it}S_{it}^T\beta}$. This is the $Y_{t,\Delta}$ for the i-th individual. We added the following text to clarify:

   > The $\epsilon_{it}(\alpha,\beta)$ is a residual term defined as $\epsilon_{it}(\alpha,\beta) = Y_{it,\Delta} - e^{g(H_{it})^T\alpha + A_{it}S_{it}^T\beta}$, where $Y_{it,\Delta}$ is the proximal outcome $Y_{t,\Delta}$ for the i-th individual.

6. *Page 12: The authors state that $\tilde{p}_{it}(S_{it})$ can depend arbitrarily as along it at most depends on S_{it}. Then they give common choices. I think it will be useful to explain how it practice one should decide on these values. This also seems to me very similar to stabilized weights for marginal structural models.*

   Reply. Thanks for your comment. You are correct that it is very similar to stabilized weights for marginal structural models (MSM); the main difference is that in MSM the numerator of the stabilized weight can only depend on the baseline effect modifiers in the causal effect model, and here $\tilde{p}_{it}(S_{it})$ can only depend on S_{it}, which is the time-varying effect modifiers in the causal excursion effect model. We revised to make this connection to the MSM literature and clarified how in practice one should decide on the value of $\tilde{p}_{it}(S_{it})$:

   > The M_{it} is defined as $M_{it} = \left\{\frac{\tilde{p}_t(S_{it})}{p_t(H_{it})}\right\}^{A_{it}} \left\{\frac{1-\tilde{p}_t(S_{it})}{1-p_t(H_{it})}\right\}^{1-A_{it}}$, which is similar to the stabilized IPW (Robins et al., 2000) and is employed by other estimators for causal excursion effects (Boruvka et al., 2018; Qian et al., 2021). Here $\tilde{p}_t(S_{it}) \in (0,1)$ is similar to the numerator of a stabilized weight and can be chosen to improve efficiency by making $\tilde{p}_t(S_{it})$ as close to $p_t(H_{it})$ as possible. For example, if $p_t(H_{it})$ is a constant or depends only on S_{it}, one can simply set $\tilde{p}_t(S_{it}) = p_t(H_{it})$ which makes $M_{it} \equiv 1$; if $p_t(H_{it})$ depends on variables in $H_{it} \setminus S_{it}$, one should not set $\tilde{p}_t(S_{it}) = p_t(H_{it})$ but can set $\tilde{p}_t(S_{it})$ to be the prediction from a logistic regression with A_{it} being the response and S_{it} being the predictor.

7. *Theorem 1: I dont think β^* was defined.*

   Reply. Thanks for your comment. We added the following to Theorem 1 to clarify:

   > Let β^* denote the true value of β corresponding to the data generating distribution P_0.

Response to Reviewer 1

Thank you for your comments. We have revised the paper carefully to address your comments. Your original comments are in italic and the reply is in normal font. Quotes from the revised version of the manuscript are indented and in a smaller font.

1. *For the proximal outcome defined by equation (3) in Section 2.3, I wonder how is it different from the corresponding proximal outcome defined by Qian et al (2021)? If I am not wrong, they should be the same? If the authors can clarify their differences, which will help to compare both the existing and proposed methods more clearly?*

   Reply. Thanks for your comment. The general definition of the proximal outcome in this paper is the same as Qian et al. (2021) in the following sense: in both papers, $Y_{t,\Delta}$ is a function of $H_{t+\Delta}$. Our methodology contribution is that for a special case of such proximal outcomes (one that satisfies the maximum property, Equation (3) in our paper), we proposed a more efficient estimator than Qian et al. (2021). To clarify, we added the following in Section 2.2 where we define the proximal outcome:

   > We consider MRTs where the proximal outcome takes binary value and is defined over a window of length Δ (of decision points) following each decision point; the same proximal outcome was considered by Qian et al. (2021).

2. *In the last two paragraphs of Section 2.3, i.e., "Equation (3) holds if the binary outcome is whether an event of interest occurs between t and $t+\Delta$ and such events are 'instantaneous'." and "If, however, Y is defined by whether an event is sustained for a period of time, then (3) may not hold.", one limitation is that the outcome Y has to be instantaneous rather than sustained. I wonder if there are other limitations for the proposed method, perhaps they can be summarized in the Discussion Section?*

   Reply. Thanks for your comment. [TQ: To explain that this is not really a limitation (by pointing to the appendix). To respond to this point after talking about the choice of reference policy, etc. In particular, can discuss when this new method will not lead to substantial efficiency gain.]

3. *My main concern is the estimating equation U defined in (6). Though the proof of its zero expected value is given in Supplementary Materials, it will be better if the authors can also give the full derivation of this equation. In other words, if the manuscript can present how to build or create this estimating equation based on Robins (1994)?*

   Reply. Thanks for your comment. [TQ: In the paper we already briefly commented on this. To write an appendix section to describe the EIF (Qian 2021), which was

derived using the same technique as Robins (1994). Then, describe how our estimating equation is constructed by adapting each of the terms in the EIF. This will have to be ad hoc, though.]

4. *If I am not wrong, should the factor $e^{-S_{it}^T\beta}$ be replaced by the factor $^{-A_tS_{it}^T\beta}$ in equation (6)?*

   Reply. Thanks for noticing this. You are correct and we corrected the typo.

5. *"The M_{it} is a marginalization weight also employed by other estimators for causal excursion effects (Boruvka et al., 2018; Qian et al., 2021) and is defined as...", it is not so clear why the factor M_{it} is important to be included in equation (6)? Does this factor have value of one most of the time? Is it a good idea to demonstrate how ignoring M_{it} in the estimating equation will affect the parameter estimates in the simulation or application sections?*

   Reply. Thanks for your comment. M_{it} is similar to the stabilized inverse probability weight in marginal structural models, and it's purpose is to allow the estimating equation to estimate the marginalize causal effect. You are correct that when the true randomization probability $p_t(H_{it})$ is a constant, one can (and should) set $\tilde{p}_t(S_{it}) = p_t(H_{it})$ so that $M_{it} = 1$. However, if $p_t(H_{it})$ depends on variables in $H_{it} \setminus S_{it}$, then one cannot make $M_{it} = 1$ because $\tilde{p}_t(S_{it})$ can only depend on S_{it}. Regarding your suggestion to "demonstrate how ignoring M_{it} in the estimating equation will affect the parameter estimates in the simulation", a similar simulation was done in Boruvka et al. (2018, Table 2) showing that if one sets $M_{it} = 1$ when one shouldn't, the estimator will be inconsistent.

   We revised our paper as follows to clarify the role of M_{it} and when it can / cannot be set to 1:

   > $M_{it} = \left\{ \frac{\tilde{p}_t(S_{it})}{p_t(H_{it})} \right\}^{A_{it}} \left\{ \frac{1-\tilde{p}_t(S_{it})}{1-p_t(H_{it})} \right\}^{1-A_{it}}$, which is similar to the stabilized IPW (Robins et al., 2000) and is employed by other estimators for causal excursion effects (Boruvka et al., 2018; Qian et al., 2021). Here $\tilde{p}_t(S_{it}) \in (0,1)$ is similar to the numerator of a stabilized weight and can be chosen to improve efficiency by making $\tilde{p}_t(S_{it})$ as close to $p_t(H_{it})$ as possible. For example, if $p_t(H_{it})$ is a constant or depends only on S_{it}, one can simply set $\tilde{p}_t(S_{it}) = p_t(H_{it})$ which makes $M_{it} \equiv 1$; if $p_t(H_{it})$ depends on variables in $H_{it} \setminus S_{it}$, one should not set $\tilde{p}_t(S_{it}) = p_t(H_{it})$ but can set $\tilde{p}_t(S_{it})$ to be the prediction from a logistic regression with A_{it} being the response and S_{it} being the predictor.

6. *The pd-EMME is developed in Section 3, where explains how it differs from the existing EMME. Why not this section also gives a brief description of the GEE approach, which*

is covered in Section 4, e.g., how does the estimating equation of GEE differ from equation (6)? Why does the GEE approach give inconsistent β estimates when the nuisance model $g(H_t)$ is mis-specified, can it be explained by words and proved mathematically in the Appendix? Though the we can observe the estimating equation of GEE from the R codes, it is not clear why the weight of $\{1-e^{g(H_{it})^T\alpha+A_{it}S_{it}^T\beta}\}^{-1}$, i.e. `weight <- (1 - exp_AXdm_beta * exp_Zdm_alpha)^(-1)`*, is used?*

Reply. Thanks for your comment. We added a section in Appendix (Section E) to explicate the particular estimating equation of GEE we used, which is a GEE for binary outcome with log link. That the consistency of GEE estimator requires the marginal mean model $E(Y_{t,\Delta} \mid H_t, A_t)$ is a standard result in Z-estimator theory; in essence, one requires the residual ϵ_{it} to have mean 0 conditional on H_{it}, A_{it}. The weight $\{1-e^{g(H_{it})^T\alpha+A_{it}S_{it}^T\beta}\}^{-1}$ comes from the product of the derivative of the mean model (which involves $e^{g(H_{it})^T\alpha+A_{it}S_{it}^T\beta}$) and the inverse variance (the variance equals $e^{g(H_{it})^T\alpha+A_{it}S_{it}^T\beta}\{1-e^{g(H_{it})^T\alpha+A_{it}S_{it}^T\beta}\}$). Please see Appendix Section E for our added paragraphs. We also added the following to the main paper, and added the prefix"log-linear" when referring to GEE in the main paper:

> See Supporting Information E for the exact form of the log-linear GEEs we used.

7. *In Section D of the Supplementary Material, equation (D.2) is defined by $1-\prod_{j=1}^{\Delta}P(R_{t+j}(\bar{A}_{t-1},1,\bar{0})=0|H_t)$ for $j=1,\ldots,\Delta$, and equation (D.3), i.e., $E(Y_{t,\Delta}(\bar{A}_{t-1},1,\bar{0})|H_t)$ (shouldnt be (D.7)?) is defined by $1-\phi_1(Z_t)\prod_{j=1}^{\Delta-1}E(\phi_0(Z_{t+j}))$ for $j=1,\ldots,\Delta-1$. Similarly, the equation of $E(Y_{t,\Delta}(\bar{A}_{t-1},0,\bar{0})|H_t)$ should be (D.4)? The last equation should be (D.5) instead of (D.3)? I am confuse why the equations (D.2) and (D.3) are not consistent, i.e., why is $j=1,\ldots,\Delta-1$ instead of $j=1,\ldots,\Delta$ in equation (D.3) and $\phi_1()$ is included in (D.3) but not in (D.2)? If I understand correctly, in equation (D.2), $P(R_{t+1}(\bar{A_{t-1}},1,\bar{0})=0|H_t)$ can be considered as $\phi_1()$, the rest terms of the products are from $j=1$ to $j=\Delta-1$?*

Reply. We apologies for the confusion/typos and thank you for checking our derivation so carefully. We have greatly revised and expanded Appendix D to clarify the proof and correct typos. Please refer to Appendix D for detail. To answer your specific questions, indeed we misnumbered the equations and these are now fixed. The use of $j=1,\ldots,\Delta-1$ instead of $j=1,\ldots,\Delta$ in the product of now equations (D.7) and (D.8) is because the Δ terms are factored as a single $\phi_1(Z_t)$ (or $\phi_0(Z_t)$) term and $\Delta-1$ of $3/C\cdot 0.5^{1/\Delta}$ terms, so the total number of terms in the product is always Δ.

Response to Reviewer 2

Thank you for your comments. We have revised the paper carefully to address your comments. Your original comments are in italic and the reply is in normal font. Quotes from the revised version of the manuscript are indented and in a smaller font.

1. ***Novel contribution needs stronger justification***

   The core contribution of this manuscript is to recognize that when $Y_{t,\Delta}$ is the maximum of a sequence of binary variables, the product of $\Delta-1$ weights in the CEE criterion can be reduced to a smaller product, leading to efficiency gains. Similar IPW tricks exist in the literature. In off-policy learning, for example, if we collect under a policy $\pi(a_t \mid s_t)$ and wish to evaluate a distal outcome dened as the cumulative sum of variables $Y = \sum_{t=1}^{T} Y_{t+1}$ under a dierent policy $\mu(a_t \mid s_t)$, then the IPW approach leads to

$$\mathbb{E}[W_{1:T} Y_t] = \mathbb{E}\left[W_{1:T} \sum_{t=1}^{T} Y_{t+1}\right] = \mathbb{E}\left[\sum_{t=1}^{T} W_{1:t} Y_{t+1}\right]$$

   where $W_{1:t} = \prod_{s=1}^{t} \frac{\pi(A_t|S_t)}{\mu(A_t|S_t)}$. The argument relies on the fact that Y_{t+1} is fixed given H_{t+1} and $E[W_{t+1:T} \mid H_{t+1}] = 1$. The current paper should connect with the relevant literature on efficiency gain techniques such as those above. Implicitly for binary outcomes like those considered in this manuscript, it is even more clear since $Y_{t,\Delta} | Y_{t,\Delta-s} = 1$ is deterministically one.

   Reply. [TQ: To acknowledge this literature early on in the introduction. This actually means that we can generalize this to continuous outcome settings, too, and we should briefly mention this in the discussion.]

2. Δ ***choice and reference distribution***

   The authors emphasize the case where the reference distribution for the future treatments is equivalently equal to $\bar{0}_{\Delta-1}$. The selection of the reference distribution depends on the scientific question. So, in general, the estimator will not "effectively discard" (page 2) the data. Can the authors comment on the benefits of this approach when we consider other reference distributions? For example, under Boruvka et. al. (2018), the weights are equivalently 1 in the future and so this issue does not arise.

   To that end, the choice of Δ and reference distribution are often chosen by the scientist to balance these efficiency/power concerns. That is, Δ is chosen to assess effects over proximal windows that are not too long relative to the reference distribution. I think for Δ large, the reference distribution would change, i.e., don't treat for some short window and then follow the MRT. I think the proposed approach would extend naturally to such

settings. My question is whether the authors can characterize the benefit over these decision choices. Again, Boruvka et. al. (2018), comes to mind where $W_{t,\Delta} = 1$ and so there is no benefit of the proposed approach. So it seems to be a function of distance of the reference distribution from the randomization probability.

Reply. Thanks for your comment.

We agree that the reference distribution should be selected based on the scientific question. In Web Appendix H we now present a generalization of our method to other reference distributions that are in between the reference distribution we considered in the main paper (where the future treatments are set to $\bar{0}_{\Delta-1}$) and the reference distribution in Boruvka et al. (2018) (where the future treatments will follow the MRT policy). We also included additional simulation studies in Web Appendix H [TQ: Yihan to complete the simulation studies] to show how the efficiency gain by our method (compared to the EMEE method by Qian et al. (2021)) depends on the distance of the reference distribution from the MRT policy. Indeed, the efficiency gain is larger when the distance of the reference distribution from the MRT policy is greater, and there is no efficiency gain if the reference distribution is the MRT policy (as in Boruvka et al. (2018)).

As you said, the choice of Δ and reference distribution are often chosen by the scientists. In addition to the efficiency/power concerns, another critical consideration is what reference distribution is meaningful scientifically. For example, in the Sense2Stop MRT where the proximal outcome window is 120 minutes and randomization occurs every 1 minute, due to burden considerations it is of scientific interest to consider a reference distribution where all the future treatments are set to 0, because it was deemed too burdensome for the user if an additional treatment is sent within 120 minutes of a previous treatment. There are other settings where the reference distribution would be chosen to be the MRT policy like Boruvka et al. (2018). The dependence of the causal excursion effect on the reference distribution is an important topic, and future work on the statistical and robustness considerations about the reference distribution is needed.

3. ***Time-to-event outcome and underlying scientific question***

The maximum property seems to suggest $Y_{t,\Delta}$ is really a stand-in for an interval-censored time-to-event. If so, do we still want to model on the log-relative risk scale? Or is there a better scale for the 'time-to-engagement' proximal outcome, such as the mean restricted survival time? What is learned by modelling such an outcome as a binary variable rather than a time-to-event outcome that may be suitably censored?

To that end, the authors suggest the scientific question concerns time until engagement (up to 72 hours). This reads to me as a survival time proximal outcome. So isn't the main efficiency loss from treating the outcome like a binary outcome, rather than a (potentially) censored survival outcome? That is, can the authors comment on the issues of dichotomization that individuals such as Frank Harrell and Stephen Senn have discussed at length in the design and analysis of clinical trials.

If the researchers were interested in the delayed effect of intervention on engagement three days later, then that is not a cumulative outcome and we cannot rely on the method presented. Is that correct? If so, my question is whether $Y_{t,\Delta=3}$ and the subsequent analysis can provide information to answer questions about the effect on R_{t+3} directly? Perhaps you want to look at this as a function of Δ? Are the conclusions sensitive to which approach we take? My point here is that can we rely on the methodology proposed to help us more broadly with scientific questions about binary outcomes?

Reply. Thanks for your comment and for pointing us to the literature on the issues of dichotomization by individuals including Frank Harrell and Stephen Senn.

To your first paragraph on whether a binary outcome or a time-to-event outcome should be used: We agree with you that which outcome is suitable depends on the scientific question. An analysis on a binary outcome is about *whether* an event occurs (within a prespecified time window), and an analysis on a time-to-even outcome is about *how soon* an event occurs. For many mobile health applications, the intervention aims to nudge an individual so that they start to adopt certain healthy behavior that they wouldn't do otherwise. Example scenarios include using activity suggestions to encourage individuals to walk (HeartSteps study, Klasnja et al., 2015), using mindfulness practice reminders to reduce stress (Sense2Stop study, Battalio et al., 2021), and using push notifications to encourage app-based self-monitoring (Drink Less study, Bell et al., 2020). In these examples, it matters the most that an individual takes the walk, practices mindfulness, or completes their drinking log in the app; it doesn't matters as much whether they take the walk or practice mindfulness within 10 minutes or 30 minutes after receiving the notification, or whether they log their drinking behavior immediately after a drink or in the following day. Therefore, for these applications, we argue that a binary outcome suits the scientific question better. There are also examples where a time-to-event outcome is more scientifically meaningful. An example is a mobile intervention on problem anger management that one of the authors (Qian) is working on: once problem anger is detected and intervention delivered, we want the intervention to bring down the anger level as soon as possible.

There are two additional caveats to this distinction between binary and time-to-event

outcomes. The first is that in order for the binary outcome (of whether an event occurs within a prespecified time window) to make sense, the time window must be carefully chosen and reflect scientific considerations. The second is that it is often helpful to complement an analysis on one outcome type with an analysis on the other outcome type to add insights to how an intervention works. In practice, these considerations should be adequately discussed with the scientific team to ensure that the statistical formulation reflects the scientific question of interest.

We will address the second paragraphs of your comment by summarizing the relevant points on dichotomization by Stephen Senn and other individuals, and then providing our perspective.

The first point by Stephen Senn is that dichotimizing loses efficiency compared to analyzing the original survival outcome. While we agree with this statement, we believe the choice between a time-to-event outcome and a binary outcome would primarily based on the scientific question, as discussed above.

The second point by Stephen Senn is that one needs to be careful in interpreting the causal effect on a binary outcome. Sander Greeland has a related point, that "any increase in average risk or response seen in one group over another may stem from a small improvement in everyone or a much larger improvement in a few". We completely agree with this point. Indeed, our causal excursion effect on the log relative risk scale should not simply be interpreted as the proportion of individuals whose behavior are changed by the intervention. This relates to the third point by Stephen Senn below.

The third point by Stephen Senn is that to demystify whether the causal effect on a binary outcome is due to a small improvement in everyone or a much larger improvement in a few, one should do repeated experiments. We agree that in order to realize the potential of personalized medicine, one needs not only know whether an intervention works when averaging over the population, but also how it works (e.g., how the treatment interacts with effect modifiers or how the effect differs across individuals). Indeed, in addition to the marginal causal excursion effect analysis (by setting $S_t = \emptyset$ in our model), our method allows for effect modification analysis (by setting a user-specified non-empty S_t), and such analyses help with answering the previous questions. Analyses with even finer granuality such as person-specific analyses may also be conducted using MRT data.

To your third paragraph: The notation system in the main paper doesn't handle the problem you described. For this problem, one could use the EMEE estimator by Qian et al. (2021). For the sake of this discussion, however, one could alternatively use the generalized version of our proposed estimator, described in Web Appendix A, to

handle this problem. In particular, for each decision point t, one would define a double-subscript version of the R variables: $R_{t,t+1} \equiv 0$, $R_{t,t+2} \equiv 0$, and $R_{t,t+3}$ as whether the individual engaged on the third day. Then, defining $Y_{t,\Delta=3} = \max(R_{t,t+1}, R_{t,t+2}, R_{t,t+3})$ as the engagement on the third day, we could use the proposed method to estimate the causal effect on the engagement on the third day. Admittedly, using our method in this case will yield exactly the same result as Qian et al. (2021) without any efficiency improvement. But this generalization of double-subscript R can be helpful in other settings, as we described in Web Appendix A:

> For exmample, in the HeartSteps II MRT (Liao et al., 2020), the proximal outcome was whether there is an activity bout in the 30 minutes following a decision point. There was a decision point every 5 minutes, so we denote by $Y_{t,\Delta=6}$ the proximal outcome using our notation. Here $R_{t,t+1}$ is whether an activity bout occurred between t and $t+1$, $R_{t,t+2}$ is whether an activity bout occurred between t and $t+2$, and so on. Here (**??**) holds with $\Delta = 6$.

4. ***Drivers of efficiency***

Figure 2 gives me the sense that Δ drives efficiency, but I suspect that the efficiency gains are a complex function of base rates, treatment probability, and reference distribution. Is there any way you can characterize the efficiency gain using the asymptotic normality results? For example, if $\Delta = 10$ but the probability of treatment is $p = 1/100$ then the approach may not help much. But if $p = 1/2$ then it seems to be very important. So characterizing the efficiency gain in terms of these trade-offs (even under some assumptions) would help the reader understand the gains more explicitly. The empirical studies do show efficiency gains but do not summarize the key drivers of this gain for me.

Reply. [TQ: Ask Yihan to run some simulations first (mainly with different base rates). Then see if we can analytically derive the relative efficiency (perhaps under some assumptions).]

- Revise the DGM (e.g., incorporating another tuning parameter) so that the success probability under no treatment can be tuned, too.
- Try to derive the relative efficiency theoretically, even under some assumptions (e.g., Markovian assumption that R_{t+1} only depends on A_t or A_t and X_t), so that we can provide some insight into what drives the relative efficiency.
- Consider develop an estimator for various reference distributions (e.g., some 0's and then follow the MRT), and see how the relative efficiency depends on the length of the 0's. This requires some additional methodology development.

Minor comment

- *page 5: "a[n] MRT"*

  Reply. Thanks for your comment. We corrected the typo.

- *page 19: rates are approximate due to non-linear transforms, no? So should these statements be 'approximately'?*

  Reply. Thanks for your comment. You are correct and we added "approximately" to the text.

References

Samuel L Battalio, David E Conroy, Walter Dempsey, Peng Liao, Marianne Menictas, Susan Murphy, Inbal Nahum-Shani, Tianchen Qian, Santosh Kumar, and Bonnie Spring. Sense2stop: A micro-randomized trial using wearable sensors to optimize a just-in-time-adaptive stress management intervention for smoking relapse prevention. Contemporary Clinical Trials, page 106534, 2021.

Lauren Bell, Claire Garnett, Tianchen Qian, Olga Perski, Henry WW Potts, and Elizabeth Williamson. Notifications to improve engagement with an alcohol reduction app: protocol for a micro-randomized trial. JMIR research protocols, 9(8):e18690, 2020.

Audrey Boruvka, Daniel Almirall, Katie Witkiewitz, and Susan A Murphy. Assessing time-varying causal effect moderation in mobile health. Journal of the American Statistical Association, 113(523):1112–1121, 2018.

F Richard Guo, Thomas S Richardson, and James M Robins. Discussion of estimating time-varying causal excursion effects in mobile health with binary outcomes. Biometrika, 108 (3):541–550, 2021.

Predrag Klasnja, Eric B Hekler, Saul Shiffman, Audrey Boruvka, Daniel Almirall, Ambuj Tewari, and Susan A Murphy. Microrandomized trials: An experimental design for developing just-in-time adaptive interventions. Health Psychology, 34(S):1220, 2015.

Predrag Klasnja, Dori E Rosenberg, Jing Zhou, Jane Anau, Anirban Gupta, and David E Arterburn. A quality-improvement optimization pilot of barifit, a mobile health intervention to promote physical activity after bariatric surgery. Translational Behavioral Medicine, 11(2):530–539, 2021.

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(1):1–22, 2020.

Inbal Nahum-Shani, Mashfiqui Rabbi, Jamie Yap, Meredith L Philyaw-Kotov, Predrag Klasnja, Erin E Bonar, Rebecca M Cunningham, Susan A Murphy, and Maureen A Walton. Translating strategies for promoting engagement in mobile health: A proof-of-concept microrandomized trial. Health Psychology, 40(12):974, 2021.

Romain Neugebauer, Mark J van der Laan, Marshall M Joffe, and Ira B Tager. Causal inference in longitudinal studies with history-restricted marginal structural models. Electronic journal of statistics, 1:119, 2007.

Doina Precup. Eligibility traces for off-policy policy evaluation. Computer Science Department Faculty Publication Series, page 80, 2000.

Tianchen Qian, Hyesun Yoo, Predrag Klasnja, Daniel Almirall, and Susan A Murphy. Estimating time-varying causal excursion effects in mobile health with binary outcomes. Biometrika, 108(3):507–527, 2021.

Tianchen Qian, Ashley E Walton, Linda M Collins, Predrag Klasnja, Stephanie T Lanza, Inbal Nahum-Shani, Mashfiqui Rabbi, Michael A Russell, Maureen A Walton, and Hyesun Yoo. The microrandomized trial for developing digital interventions: Experimental design and data analysis considerations. Psychological Methods, 2022.

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

James M Robins. Optimal structural nested models for optimal sequential decisions. In Proceedings of the second seattle Symposium in Biostatistics, pages 189–326. Springer, 2004.

James M Robins, Miguel Angel Hernan, and Babette Brumback. Marginal structural models and causal inference in epidemiology. Epidemiology, 11(5):550–560, 2000.