American Journal of Epidemiology, 2024, 193, 835–845
https://doi.org/10.1093/aje/kwad241
Advance access publication date December 7, 2023
Original Contribution

The application of target trials with longitudinal targeted maximum likelihood estimation to assess the effect of alcohol consumption in adolescence on depressive symptoms in adulthood

Yan Liu [1], Mireille E. Schnitzer [1,2,*], Ronald Herrera [3], Iván Díaz [4], Jennifer O'Loughlin [2,5], Marie-Pierre Sylvestre [2,5]

[1]Faculté de Pharmacie, Université de Montréal, Montréal, QC H3T1J4, Canada
[2]Département de Médecine sociale et préventive, École de Santé publique, Université de Montréal, Montréal, QC H3N1X9, Canada
[3]Bayer AG, Berlin 13353, Germany
[4]Department of Population Health, NYU Langone Health, New York, NY 10016, United States
[5]Centre de Recherche du Centre Hospitalier de l'Université de Montréal, Montréal, QC H2X0A9, Canada
*Corresponding author: Mireille E. Schnitzer, Faculte de Pharmacie, Université de Montréal, Pavillon Jean-Coutu, 2940 Chemin de la Polytechnique, Office 2236, PO Box 6128, Station Centre-ville, Montreal, QC H3C 3J7, Canada (mireille.schnitzer@umontreal.ca)

Abstract

Time-varying confounding is a common challenge for causal inference in observational studies with time-varying treatments, long follow-up periods, and participant dropout. Confounder adjustment using traditional approaches can be limited by data sparsity, weight instability, and computational issues. The Nicotine Dependence in Teens Study is a prospective cohort study, and we used data from 21 data collection cycles carried out from 1999 to 2008 among 1294 students recruited from 10 high schools in Montreal, Quebec, Canada, including follow-up into adulthood. Our aim in this study was to estimate associations of timing of alcohol initiation and cumulative duration of alcohol use with depression symptoms in adulthood. Based on the target trials framework, we defined intention-to-treat and as-treated parameters in a marginal structural model with sex as a potential effect-modifier. We then used the observational data to emulate the trials. For estimation, we used pooled longitudinal target maximum likelihood estimation, a plug-in estimator with double-robust and local efficiency properties. We describe strategies for dealing with high-dimensional potential drinking patterns and practical positivity violations due to a long follow-up time, including modifying the effect of interest by removing sparsely observed drinking patterns from the loss function and applying longitudinal modified treatment policies to represent the effect of discouraging drinking.

Introduction

Estimation of the effect of time-varying exposures in observational studies becomes methodologically challenging in the presence of time-dependent confounding, requiring statistical methods beyond the standard approaches.[1,2] Robins[3] proposed marginal structural models (MSMs) which model the potential outcome under an assigned treatment history (or "pattern"). Hernán and Robins[4] proposed the target trials framework to define causal effects, in particular MSM parameters, by means of a mapping of the observational analysis onto an analysis of a hypothetical randomized controlled trial. The parameters of an MSM can be estimated with inverse probability of treatment weighting,[2] G-computation,[5,6] augmented inverse probability of treatment weighting estimators,[7,8] and more recently longitudinal targeted maximum likelihood estimation (LTMLE).[9–11] LTMLE has the advantage of double-robustness, meaning that the estimator is consistent if either the models for treatments (and censoring) or the models for outcomes are correctly specified. LTMLE can also readily incorporate machine learning in the process of generating the initial estimates while providing valid statistical inference,[11] thus reducing the chance of incurring model misspecification bias.

Though LTMLE has been successfully applied in different contexts,[12–16] there exist data sparsity and high-dimensionality challenges.[17,18] One solution to these challenges lies in defining hypothetical longitudinal interventions that shift an individual's propensity score, making that person more or less likely to be exposed, corresponding to exposure encouragement or discouragement. The intervention can also be applied conditional on the observed exposure—for example, only discouraging exposure for persons who actually were exposed.[19,20]

In this paper, we demonstrate the application of LTMLE in a complex substantive application. We consider an observational cohort study from Montréal, Quebec, Canada, called Nicotine Dependence in Teens (NDIT).[21,22] Associations between alcohol use in adolescence and later risk-taking behaviors have been established in longitudinal studies.[23,24] Regarding assessment of the longitudinal effect of alcohol use on depressive symptoms in early adulthood, to the best of our knowledge, no study has adjusted for time-varying confounding using causal inference

Received: November 16, 2022. Accepted: December 4, 2023

methods.[25-28] We aim to study the effect of time of alcohol-drinking initiation and cumulative duration of drinking in adolescence on depression in young adulthood. We take into account time-varying confounders that can also be caused by drinking in adolescence, including depressive symptoms,[25,29] smoking,[30,31] stress,[32,33] and participation in team sports.[34,35] We define 2 target trials that recruit adolescents who have not yet initiated regular drinking. Using working MSMs, we correspondingly define the "intention-to-treat" (ITT) and "as-treated" (AT) effects, respectively, and investigate effect modification by sex. The working MSM represents a projection of a true causal relationship between exposures and the outcome onto a low-dimensional linear model.[36] We then estimate the parameters of the 2 MSMs using G-computation and LTMLE. We describe high-dimensionality and sparsity challenges encountered when estimating the AT effect and explore ways to address them.

Methods

NDIT data

The NDIT Study is a prospective longitudinal investigation of 1294 grade 7 students recruited from 10 Montréal-area high schools in 1999-2000.[22] Self-report questionnaires were administered from grade 7 to grade 11 at each of the 10 schools every 3 months, for a total of 20 study cycles from 1999 to 2005 (i.e., during the 5 years of high school). Mail or in-person questionnaires were administered in 2007/2008 (cycle 21) when participants were aged 20.4 years, on average. The data collected included repeated measures of a wide range of sociodemographic, substance use, psychosocial, lifestyle, and physical and mental health variables. Figure S1 in Appendix S1 presents the structure of the follow-ups in the NDIT Study.

Parents and legal guardians provided informed consent, and all students provided assent and then consent in adulthood. The study was approved by the Ethics Research Committee of the Centre de Recherche du Centre Hospitalier de l'Université de Montréal.

Exposure

Participants were asked, "During the past 3 months, how often did you drink alcohol (beer, wine, hard liquor)?". We considered a participant exposed to regular alcohol use if the participant answered "once or a couple of times a week" or "usually every day" (alternatives were "never," "a bit to try," or "once or a couple of times a month"). Therefore, "alcohol use" in this paper refers to "at least weekly use." In defining the population of interest, we excluded all participants reporting regular alcohol use at time 0. We then defined exposure initiation as the time when participants first became exposed. We correspondingly denoted the binary exposure over time as $A_t, t \in (0, \cdots, 19)$, with $A_t = 1$ representing exposed and $A_t = 0$ representing unexposed.

Censoring

We denoted the censoring indicators as $C_t, t \in (1, \cdots, 20)$. A participant was censored by time t, denoted $C_t = 1$, when they were lost to follow-up or when they skipped more than 1 entire year of follow-up; otherwise, $C_t = 0$.

Covariates

Baseline covariates. As baseline variables, we included sociodemographic characteristics, including sex, school indicator, mother's education, whether the participant lived in a single-parent home, whether the participant spoke French at home, and country of birth, which were assessed in the first data collection cycle. In addition, we also included as baseline covariates self-esteem, impulsivity, and novelty-seeking, measured in the 12th study cycle (April 2002 for grade 9 students; average age = 15 years), since they were considered personal traits and unlikely to vary considerably over time. We denoted the baseline variables as $\boldsymbol{W}$.

Time-varying covariates. The time-varying covariates $\boldsymbol{L}_t, t \in (1, \cdots, 20)$ were measured between exposures A_t and $A_{(t+1)}$ and included current depressive symptoms, participation in team sports, family-related stress (on a validated 4-point scale, with higher values indicating more stress), other types of stress (validated 4-point scale), worry about weight, and ever having smoked.

Detailed information on all covariates is given in Appendix S2.

Outcomes

The outcome Y, depression symptoms, was measured using the Major Depression Inventory (MDI) in 2007/2008. This scale measures depression symptoms over the past 2 weeks with a score range of 0-50, where greater values indicate more severe symptoms.[37,38] A detailed list of items included in the MDI is presented in Appendix S3.

Data structure

Given the above, the following represents the observed data structure:

$$O = \{\boldsymbol{W}, \boldsymbol{L}_1, A_1, \boldsymbol{L}_2, C_2, A_2 \cdots, \boldsymbol{L}_{19}, C_{19}, A_{19}, \boldsymbol{L}_{20}, C_{20}, Y\}.$$

Definition of target parameter

Target trial

We define 2 target trials, with corresponding ITT and AT parameters of interest. Both trials recruit participants who had not initiated regular alcohol drinking at the beginning of grade 7. The first target trial randomizes drinking initiation to one of the first 19 follow-up time points. The second target trial randomizes drinking (yes/no) at each of the 19 time points during follow-up. The depression score outcome is measured at the follow-up time in adulthood. The parameters of interest are the coefficients of a linear regression conditional on sex, drinking assignment, and their interaction.

Parameter of interest

In the observational study, we characterize counterfactuals under the different types of hypothetical interventions. In the ITT trial, the intervention is initiation time, where the analysis ignores changes in subsequent alcohol use. This leads to 20 possible patterns, denoted $\bar{\boldsymbol{a}}$, that can be represented by vectors of length 19 of zeros followed by ones. For example, a vector of 19 zeros represents never initiating alcohol use; the vector $(0, 1, \ldots, 1)$ represents initiation at the second time point. In contrast, the AT study randomizes drinking at each time, so that a pattern $\bar{\boldsymbol{a}}$, any vector of length 19 of zeros and ones, is assigned to each participant. Therefore, the AT study has 2^{19} potential treatment patterns. We define a specific treatment pattern $\bar{\boldsymbol{a}}^d$, for $\bar{\boldsymbol{a}}^d \in \mathcal{D}$, where $\mathcal{D}$ is the set of all possible patterns in either the ITT trial or the AT trial.

Define $Y(\bar{\boldsymbol{a}}^d)$ as the counterfactual outcome that would have been observed under some treatment pattern $\bar{\boldsymbol{a}}^d = (a_1^d, \cdots, a_{19}^d)$. Because of a lack of data support for specific drinking initiation

times or patterns over the 19 time points, it was deemed infeasible to contrast mean counterfactual outcomes under specific trajectories. Thus, the parameters of interest are defined through working MSMs to summarize how the mean counterfactual outcome varies as a function of different interventions, and the baseline covariate sex.

The working MSM is

$$\mathbb{E}[Y(\overline{\boldsymbol{a}}^d)|\mathbf{sex}] = m(\mathbf{sex}, \overline{\boldsymbol{a}}^d; \boldsymbol{\beta}^d) = \beta_0^d + \beta_1^d \mathbf{sex} + \beta_2^d \mathbf{cum}(\overline{\boldsymbol{a}}^d) + \beta_3^d \{\mathbf{sex} \times \mathbf{cum}(\overline{\boldsymbol{a}}^d)\}, \quad (1)$$

where $\mathbf{cum}(\overline{\boldsymbol{a}}^d)$ counts the number of exposed time points in the pattern and $\mathbb{E}[Y(\cdot)|\mathbf{sex}]$ represents the mean counterfactual outcome under some intervention, in a sex subgroup such that $\mathbf{sex} = 1$ denotes female. The true parameter values $\boldsymbol{\beta}^d$ minimize the expectation of a squared error loss function, summing over all patterns in either the ITT or AT space (see Appendix S4), corresponding to the parameters estimated in the hypothetical target trials. Equation (1) thus contrasts the counterfactual mean depression scores given different alcohol initiation times or cumulative usage. Thus, we model the average counterfactual outcome under any drinking initiation time (ITT) or one of the 2^{19} drinking patterns (AT) in order to contrast how the expected outcome differs under one additional time point of drinking.

These parameters are estimable under the assumptions defined in Appendix S5. In particular, the positivity assumption requires that at every time point, all individuals must have a positive probability of initiating drinking (ITT) or continuing to follow any drinking pattern (AT). Even if theoretically satisfied, if these probabilities are estimated to be close to 0, this amounts to practical positivity violations (or sparsity) and the estimation relies on extrapolation or smoothing across covariate strata.[39]

Estimation Methods

In the application of the causal inference methods to obtain point estimates of the parameters of interest, we assumed independence between study participants. However, the variance estimation adjusted for clustering by school.[12] Our handling of baseline and time-varying covariate missingness involved multiple imputation by chained equations (see Appendix S2). We imputed 10 databases and then averaged the point estimates and computed the SEs using Rubin's rules.

Sequential G-computation

We use the notation $\overline{\boldsymbol{L}}_t$ to denote the history of time-dependent covariates up to time t, and likewise $\overline{\boldsymbol{A}}_t$ represents the history of the exposure $A_1, \cdots, A_t$. Define $\overline{Q}_t(\overline{\boldsymbol{a}}^d)$ as the mean counterfactual outcome at time $t \in (21, \cdots, 1)$ had past exposures been set to $\overline{\boldsymbol{a}}_t^d$, given the covariate history. Note that the coefficients in equation (1) correspond to a regression of $\overline{Q}_1(\overline{\boldsymbol{a}}^d)$ on sex and the respective summary of exposure according to the MSM form under the loss function in Appendix S4. To apply G-computation, we first rescaled the continuous outcome Y to (0, 1) and defined $\overline{Q}_{21}(\overline{\boldsymbol{a}}^d) = Y$. Then sequentially, for each time t, we fitted logistic regressions conditioning on exposure and covariate history using uncensored participants (Appendix S6). Finally, for each pattern $\overline{\boldsymbol{a}}^d$ belonging to the ITT or AT regimen space, we obtained estimates of $\overline{Q}_1(\overline{\boldsymbol{a}}^d)$ for all participants by predicting from the model fit under the imputed exposure pattern. We then stacked the vectors $\overline{Q}_1(\overline{\boldsymbol{a}}^d)$ for each pattern $\overline{\boldsymbol{a}}^d$ and regressed this vector on baseline covariates and summaries of regular drinking exposure using linear regression according to the MSM. SEs were then estimated by clustered bootstrap with 150 resamples.[5,12] The detailed algorithm of sequential G-computation is given in Table S1 in Appendix S7.

Longitudinal targeted maximum likelihood estimation

LTMLE requires estimates of conditional probabilities of treatment and censoring to update the initial estimates of each $\overline{Q}_t(\overline{\boldsymbol{a}}^d)$ with the objective of satisfying the efficient influence function estimating equation.[40] This results in double-robustness and asymptotic local efficiency.[41] First we used logistic regression models to estimate the censoring and treatment probabilities stratified by time conditional on the baseline and recent time-varying covariates, and lagged exposure for uncensored participants (Appendix S6). We defined w_t^d as the cumulative weight, which is the cumulative product of the inverse of treatment and censoring probabilities from time 1 to time t under the treatment pattern $\overline{\boldsymbol{a}}^d$. We used stabilized weights (Appendix S8), which result in a weaker positivity assumption.[39] Because we observed practical positivity violations in the AT analysis, we also used post-hoc weight truncation. SEs were estimated based on the influence function.

LTMLE allows for the integration of machine learning to increase the chances of consistent estimation under regularity conditions.[10] Super Learner (SL) is a method that uses V-fold cross-validation to find an optimal convex combination of the predictions of a library of candidate algorithms defined by the user.[42] We therefore used SL to estimate the $\overline{Q}_t(\overline{\boldsymbol{a}}^d)$'s and the exposure and censoring probabilities. Each SL library contained the mean (SL.mean), multivariate adaptive regression splines (SL.earth), generalized additive models (SL.gam), generalized linear models (GLMs) (SL.glm), and the least absolute shrinkage and selection operator (LASSO) (SL.glmnet). We used the default hyperparameters but customized these functions (aside from the mean) by adding terms for interaction between treatment and sex in the SL wrappers. We present the pooled LTMLE algorithm for estimating the parameters of an MSM[11] in Table 1. The subscript n is used to denote an estimate of a quantity.

Since we hand-coded the pooled LTMLE algorithm for the clustered setting,[12] we verified its correctness using 2 simulations with 2 time points and clustered observations and estimated ITT and AT parameters, described in Appendix S9. We simulated 500 data sets where we generated 5000 participants in 50 clusters, with random intercepts in the outcome model. We verified the unbiasedness of the LTMLE and also compared SE estimators assuming independence and clustering (respectively) using the influence function–based sandwich estimator and clustered bootstrap, respectively, showing that the clustered versions are needed under random effects. The full data-generation is shown and results are given in Tables S2 and S3, respectively, in Appendix S9.

Challenges and strategies due to high-dimensionality and sparsity in the AT analysis

The main challenges in the AT analysis involved computational issues introduced by the very large number of potential treatment patterns. Recall that we have $|\mathcal{D}| = 2^{19} = 524\,288$ potential treatment patterns, which would thereby produce several very large stacked vectors and matrices when we perform the pooled TMLE procedure. The vectors are of length $|\mathcal{D}| \times n = 2^{19} \times 1231 = 645\,398\,528$ for each time t in the update step (steps 3.2-3.4 of the pooled LTMLE algorithm in Table 1). However, objects of this size cannot be stored in the R memory or be easily manipulated in R.

Table 1. Pooled longitudinal targeted maximum likelihood estimation algorithm.[a]

Step	Pooled LTMLE algorithm for both ITT and AT
1	Estimate every component of w_t^d for $t = 1, \cdots, 20$ to obtain the estimated weights $w_{t,n}^d$ for each treatment pattern $\overline{a}^d$ that belongs to the ITT or AT regimen space $\mathcal{D}$.
2	Define $\overline{Q}_{21,n}^{d*} = Y$, where Y is rescaled to $(0, 1)$.
3	Then, iteratively for $t = 20, \cdots, 1$:
3.1	Initial estimate of $\overline{Q}_{t,n}^d$: With uncensored participants, regress $\overline{Q}_{t+1,n}^{d*}$ on the treatment and covariate history. When $t = 20$, predict outcomes by setting $\overline{\mathbf{A}}_{19} = \overline{\mathbf{a}}_{19}^d$; otherwise predict while setting $\overline{\mathbf{A}}_t = \overline{\mathbf{a}}_t^d$ for each treatment pattern and each subject. Define $\overline{Q}_{t,n}^d$ as the stacked vector of predictions with length $n \times \mid \mathcal{D} \mid$.
3.2	Construct a covariate matrix for each subject and each treatment pattern $\overline{\mathbf{a}}^d$, $\boldsymbol{h}_t(\overline{\mathbf{a}}^d, \mathbf{sex}) = 1(\overline{\mathbf{A}}_t = \overline{\mathbf{a}}_t^d, C_t = 0) \times [\partial m(\boldsymbol{\beta}, \mathbf{sex}, \overline{\mathbf{a}}^d)/\partial \boldsymbol{\beta}]$, where $\partial m(\boldsymbol{\beta}, \mathbf{sex}, \overline{\mathbf{a}}^d)/\partial \boldsymbol{\beta}$ in our example equals $1, \mathbf{sex}, \mathbf{cum}(\overline{\mathbf{a}}_t^d), \mathbf{sex} \times \mathbf{cum}(\overline{\mathbf{a}}_t^d)$. For a treatment pattern $\overline{\mathbf{a}}^d$, the dimension of $\boldsymbol{h}_t(\overline{\mathbf{a}}^d, \mathbf{sex})$ is the same as the dimension of $\boldsymbol{\beta}$. Thus, for all possible patterns, the $\boldsymbol{h}_t(\overline{\mathbf{a}}^d, \mathbf{sex})$ is of dimension $(n \times \mid \mathcal{D} \mid) \times 4$.
3.3	Update $\overline{Q}_{t,n}^d$ to $\overline{Q}_{t,n}^{d*}$ by fitting an intercept-free weighted pooled logistic regression of $\overline{Q}_{t+1,n}^{d*}$ on the covariate matrix produced from the previous step with offset $\mathrm{logit}(\overline{Q}_{t,n}^d)$ and $w_{t,n}^d$ as weights for each $\overline{\mathbf{a}}^d$, $\mathrm{logit}(\overline{Q}_{t,n}^{d*}) = \mathrm{logit}(\overline{Q}_{t,n}^d) + \boldsymbol{\epsilon}\boldsymbol{h}_t(\overline{\mathbf{a}}^d, \mathbf{sex})$.
3.4	Generate $\overline{Q}_{t,n}^{d*}$ by making predictions for every subject under each pattern $\overline{\mathbf{a}}^d$ using the logistic model fitted in step 3.3. $\overline{Q}_{t,n}^{d*}$ is the stacked vector of updated predictions and has length $n \times \mid \mathcal{D} \mid$.
4	Rescale $\overline{Q}_{1,n}^{d*}$ (length $n \times \mid \mathcal{D} \mid$) to the original scaling of Y.
5	The coefficients are estimated by fitting a pooled linear regression of $\overline{Q}_{1,n}^{d*}$ on stacked covariates and all treatment patterns $\mathcal{D}$, in the ITT or AT space, corresponding to the MSM in equation (1).

Abbreviations: AT, as-treated; ITT, intention-to-treat; LTMLE, longitudinal targeted maximum likelihood estimation; MSM, marginal structural model.
[a]Variables: $w_{t,n}^d$, the estimate of cumulative weight up to time t under treatment pattern $\overline{\mathbf{a}}^d$; $\overline{Q}_{t,n}^d$, the estimate of the mean counterfactual outcome at time t had past exposures been set to $\overline{\mathbf{a}}^d$; $\overline{Q}_{t,n}^{d*}$, the updated estimate of the mean counterfactual outcome; $\mathcal{D}$, the ITT or AT regimen space; $\boldsymbol{\epsilon}$, a vector parameter.

The parameters of interest are defined in terms of the minimization of an expected squared error loss function calculated over all 2^{19} patterns (see Appendix S4). However, in order to tackle this issue, we propose a pragmatic strategy that redefines the parameters of interest by minimizing this expected loss over the patterns that are most supported by the data. Let $\mathcal{D}_t$ be the set of patterns that were supported by data, that is, observed to be followed by at least 1 individual up to time t, for each time $t = 1, \cdots, 19$. Table S4 in Appendix S10 gives the cardinality (size) of $\mathcal{D}_t$ at each time point. To test sensitivity, we performed the analysis involving all supported patterns up to times 19, 18, 17, and 16, such that $|\mathcal{D}_t| = 227$, 494, 936, and 1688 patterns were included, respectively. On our local computer, we could not realistically go further than $t = 16$ for the LTMLE analysis with GLMs. We focused on $\mathcal{D}_{17}$ (936 patterns), since this was the largest number of patterns that could be incorporated in the LTMLE analysis with SL.

Longitudinal modified treatment policies

Inspired by the hypothetical interventions based on the natural value of treatment first discussed by Robins et al.[43] and then formalized by other researchers,[44,45] Díaz et al.[20] proposed longitudinal modified treatment policies (LMTPs). An LMTP involves a hypothetical intervention at each time point which can be expressed as a deterministic or random function of the observed treatment and the unit's covariate history. In this paper, we apply an incremental propensity score intervention based on the risk ratio scale[46,47] that shifts the propensity scores to discourage alcohol use. Specifically, the intervention assigns a new exposure the likelihood of which can be determined hypothetically by the user-defined risk ratio value. Under this intervention, effects can be identified and estimated under weak conditions on the propensity score[46]; consequentially, sparsity does not destabilize the analysis as much. We used the "lmtp" R package[20,48] (also available from the Comprehensive R Archive Network (https://cran.r-project.org/)), which implements LTMLE for LMTP while simultaneously adjusting for censoring. We chose 5 risk ratios for alcohol use [0.1, 0.3, 0.5, 0.7, 0.9] and then estimated the mean MDI score according to sex and compared the mean MDI scores under risk ratios [0.1, 0.5, 0.9] with the mean MDI score without any intervention. We applied the same SL algorithms as in the section "Longitudinal targeted maximum likelihood estimation." Further details are given in Appendix S11.

Results

The data set included 1294 participants. We excluded participants who only completed the first follow-up cycle, skipped the first year of the study, or reported alcohol consumption at baseline, leaving 1231 participants in the analysis (Figure S2 in Appendix S10). Baseline and time-varying characteristics of the 1231 participants are presented in Table 2. There was missingness in the baseline covariates. Table 3 shows the cumulative numbers and percentages of censored persons, initiators, and actual exposed participants at each time point (study cycle). Note that at $t = 1$, there was no censoring due to the exclusion criteria.

ITT analysis

The range of the cumulative stabilized weights was 0.22-72.38, with a mean of 1.02, so no truncation was applied. In Figure 1 and Table S5 in Appendix S10, female sex was associated with more severe depressive symptoms in all estimates. LTMLE-GLM suggested a stronger sex association than G-computation (4.5 [95% CI, 3.4-5.7] vs 3.5 [95% CI, 2.2-4.9]). In addition, LTMLE using a GLM and SL indicated that each unit of earlier alcohol initiation increased the expected depression symptoms score in adulthood among males by approximately 0.05 (GLM: 0.05 [95% CI, 0.0-0.1]; SL: 0.03 [95% CI, 0.0-0.1]) and decreased depression symptoms score in females by approximately 0.14 (GLM: −0.14 [95% CI, −0.2 to −0.1]; SL: −0.13 [95% CI, −0.2 to −0.1]). G-computation estimated no statistically significant association between alcohol initiation and depression; comparatively, LTMLE had a similar point estimate but a lower SE.

AT analysis

We performed the AT analysis with 227, 494, 936, and 1688 patterns, respectively, where the models in the pooled LTMLE

Table 2. Baseline characteristics and time-varying covariates of 1231 participants in the analytical sample at time $t = 1$, Nicotine Dependence in Teens Study, 1999-2008.

Variable	Classification	No. of participants		%		Median (IQR)	No. with missing data
		Category 1	Category 2	Category 1	Category 2		
			Baseline				
Sex	Female vs male	644	587	52.3	47.7		
Single-parent family	Yes vs no	148	1083	12.0	88.0		
French-speaking home	Yes vs no	368	863	29.9	70.1		
Country of birth	In Canada vs outside Canada	1132	99	92.0	8.0		
Mother's education	Less than university vs some university	523	419	55.5	44.5		289
Self-esteem[a]	Numerical					2.7 (2.2-2.9)	276
Impulsivity[a]	Numerical					2.1 (1.6-2.9)	326
Novelty-seeking[a]	Numerical					2.9 (2.3-3.4)	324
			L_1				
Weight worry	Yes vs no	427	736	36.7	63.3		68
Sports participation	Yes vs no	750	437	63.2	36.8		44
Ever smoking	Yes vs no	365	862	29.7	70.3		4
Current depressive symptoms	Numerical					2.0 (1.7-2.5)	55
Family stress	Numerical					1.2 (1.0-1.4)	59
Other stress	Numerical					1.4 (1.2-1.6)	52

Abbreviation: IQR, interquartile range.
[a]These 3 covariates, considered time-invariant, were measured in study cycle 12.

procedure were fitted using GLMs. To evaluate the sensitivity to weight truncation, we set fixed bounds at 1000, 5000, and 10 000 on the cumulative stabilized weights. These values were determined a posteriori based on the observed weight distribution. Table 4 shows the median values and interquartile ranges of the cumulative stabilized weights and the percentages of truncated cumulative weights. The bounding affected 7%-12% of participants.

Figure 2 and Table S6 in Appendix S10 show the estimated coefficients using LTMLE with GLM at 3 levels of truncation. The estimated counterfactual mean of MDI score in females was around 5 points higher than in males and was stable over different numbers of patterns. A single added time period of drinking in adolescence was also associated with increased depression levels in adulthood, but the estimate waned with greater numbers of patterns and less restrictive bounds. For instance, when bounding

Table 3. Numbers of censoring and exposure events occurring at 20 follow-up time points, Nicotine Dependence in Teens Study, 1999-2008.

Data collection cycle	Censoring			Alcohol use			
	No. of censored participants	Cumulative no. of censored participants	Cumulative % of censored participants	Cumulative no. with drinking initiation	Cumulative % with drinking initiation	No. of exposed participants	% of participants exposed
0				0	0	0	0
1	0	0	0.0	30	2.4	30	2.4
2	9	9	0.7	56	4.5	40	3.2
3	20	29	2.4	68	5.5	43	3.5
4	85	114	9.3	101	8.2	60	4.9
5	7	121	9.8	151	12.3	100	8.1
6	5	126	10.2	188	15.3	111	9.0
7	10	136	11.0	210	17.1	103	8.4
8	81	217	17.6	233	18.9	116	9.4
9	10	227	18.4	257	20.9	113	9.2
10	2	229	18.6	276	22.4	113	9.2
11	7	236	19.2	284	23.1	104	8.4
12	87	323	26.2	289	23.5	119	9.7
13	8	331	26.9	317	25.8	136	11.0
14	2	333	27.1	340	27.6	124	10.1
15	3	336	27.3	358	29.1	134	10.9
16	39	375	30.5	377	30.6	170	13.8
17	2	377	30.6	394	32.0	171	13.9
18	5	382	31.0	411	33.4	167	13.6
19	7	389	31.6	425	34.5	184	14.9
20	169	558	45.3	425	34.5		

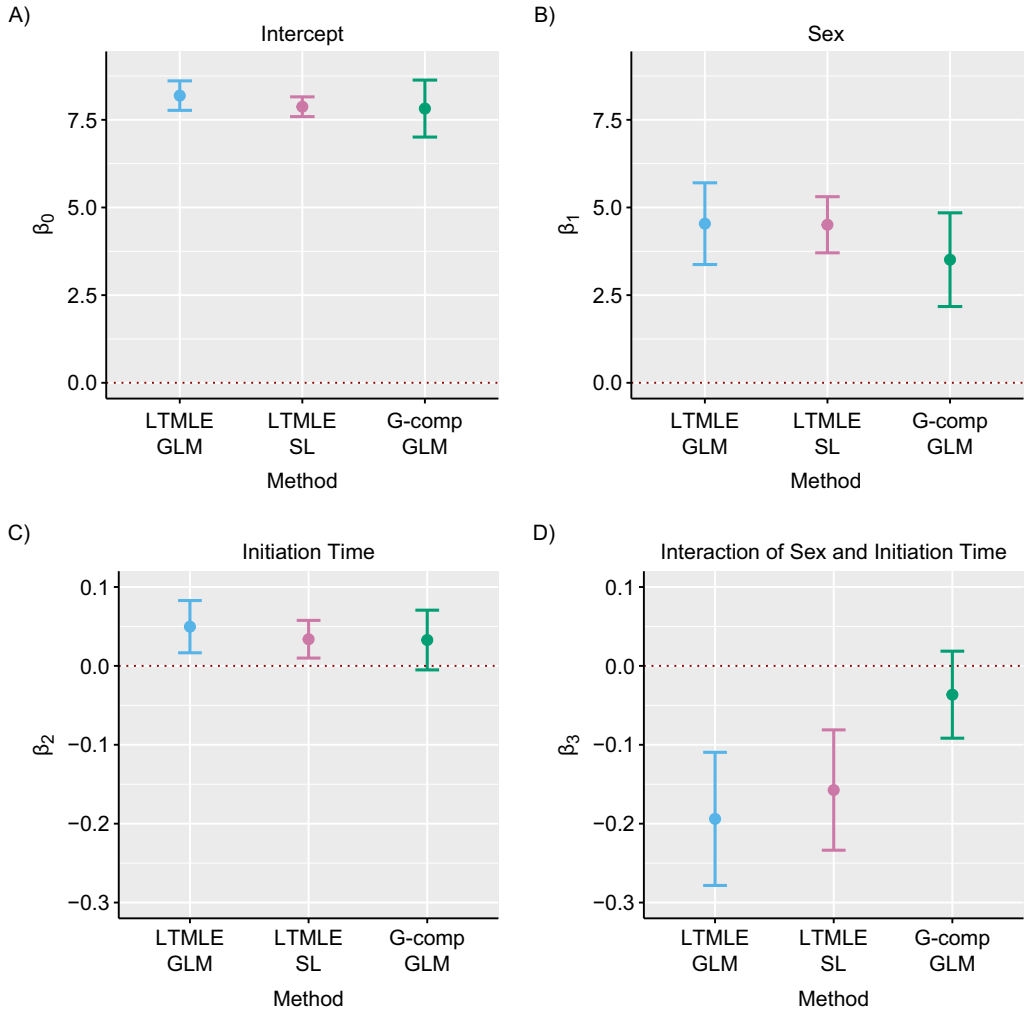

Figure 1. Intercept and coefficient estimates from longitudinal targeted maximum likelihood estimation (LTMLE) fitted with generalized linear models (GLMs; blue) and Super Learner (SL; pink), respectively, and G-computation (G-comp; green) in the intention-to-treat analysis, Nicotine Dependence in Teens Study, 1999-2008. A) Intercept; B) coefficient of sex; C) coefficient of initiation time; D) coefficient of the interaction between sex and initiation time. The y-axis represents the Major Depression Inventory (MDI) score. The intercept represents the average MDI score for males in the absence of drinking initiation; the coefficient of sex represents the average increment of MDI score for females compared with males in the absence of drinking initiation; the coefficient of initiation time represents the average increment of MDI score for males for a 1-time-point earlier initiation time; and the coefficient of the interaction between sex and initiation time represents the average increment of MDI score for females for a 1-time-point earlier initiation time. The reference sex was male. Bars show 95% confidence intervals.

by 10 000, the cumulative exposure coefficient estimate gradually decreased from 1.0 (227 patterns) to 0.60 (1688 patterns). The point estimates of the interaction term were stable with different numbers of treatment patterns but increased with less restrictive weights (from 0.13 under 227 patterns to 0.39 under 1688 patterns); the CI widths also increased with less. SEs decreased as more patterns were included, a result of more data support for the MSM parameters under the loss function.

To better understand the impact of the weights, we compared the LTMLE results with GLM and SL, respectively, with the parametric sequential G-computation estimator (Figure 3 and Table S7 in Appendix S10) using 936 patterns. TMLE with SL produced

Table 4. Median values of the cumulative stabilized weights and percentage of truncated cumulative weights with generalized linear models for 227, 494, 936 and 1688 patterns, Nicotine Dependence in Teens Study, 1999-2008.

No. of patterns	Median CSW (IQR)	% of truncated cumulative stabilized weights		
		1000 bound	5000 bound	10 000 bound
227	1.45 (0.91-18.50)	0.11	0.08	0.07
494	1.70 (0.93-23.64)	0.11	0.08	0.07
936	1.87 (0.94-27.21)	0.11	0.08	0.07
1688	2.09 (0.95-32.94)	0.12	0.08	0.07

Abbreviations: CSW, cumulative stabilized weight; IQR, interquartile range.

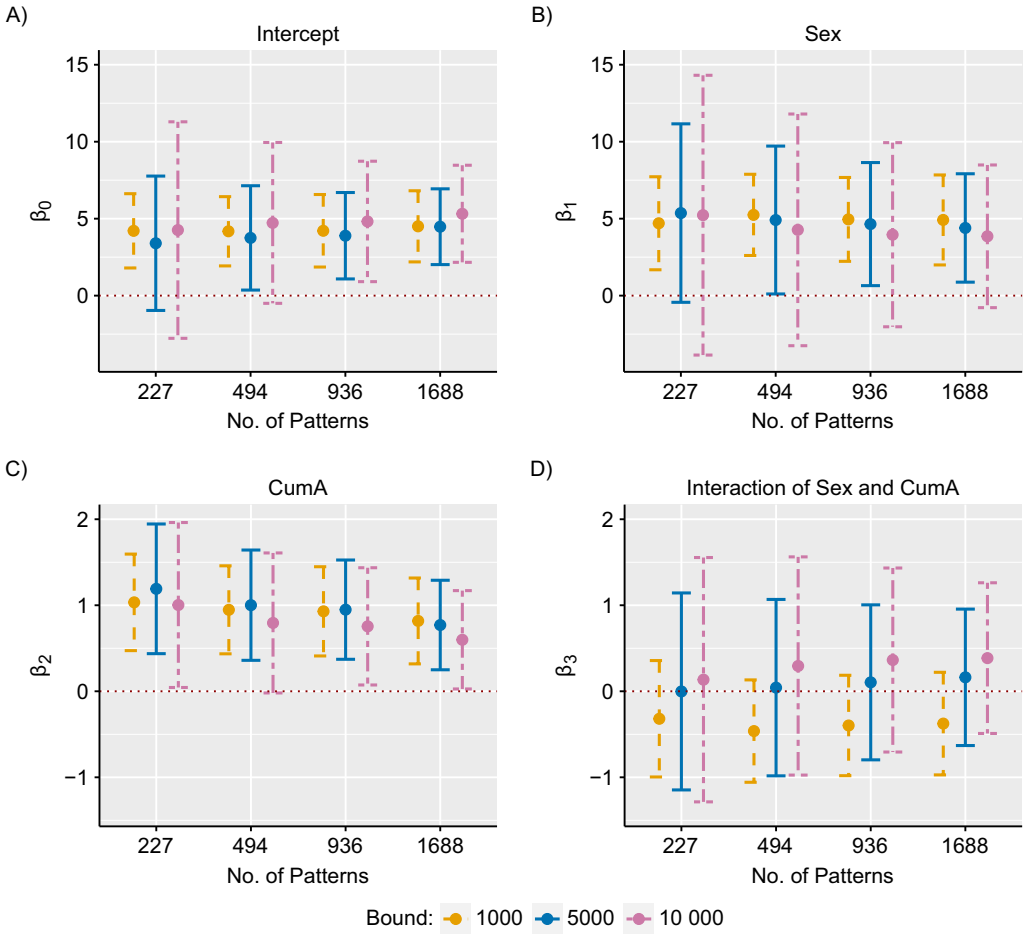

Figure 2. Intercept and coefficient estimates from longitudinal targeted maximum likelihood estimation with a generalized linear model in the as-treated analysis including 227, 494, 936, and 1688 patterns under bounds of 1000 (– – –; gold), 5000 (—; blue), and 10 000 (– - –; pink) on the cumulative stabilized weights, Nicotine Dependence in Teens Study, 1999-2008. A) Intercept; B) coefficient of sex; C) coefficient of cumulative alcohol use (CumA); D) coefficient of the interaction between sex and CumA. The y-axis represents the Major Depression Inventory (MDI) score. The intercept represents the average MDI score for males in the absence of drinking; the coefficient of sex represents the average increment of MDI score for females compared with males in the absence of drinking; the coefficient of CumA represents the average increment of MDI score for males for 1 additional time period of drinking; and the coefficient of the interaction between sex and CumA represents the average increment of MDI score for females for 1 additional time period of drinking. The reference sex was male. Bars show 95% confidence intervals.

smaller effects of sex than TMLE with GLM under all 3 bounds. G-computation produced much narrower CIs that contained the null value for both the main effect of cumulative duration (0.03; 95% CI, −0.10 to 0.17) and the interaction term (−0.14; 95% CI, −0.34 to 0.07). The weights had an important impact on the point estimate for the coefficient of cumulative exposure, with LTMLE suggesting an association when using GLM but not with SL. Only LTMLE with SL under the 10 000 bound indicated that sex modified the effect of cumulative exposure (1.72; 95% CI, 0.44-3.00).

LMTP estimates

Using LTMLE, we estimated mean MDI scores and 95% CIs in males and females separately under 5 hypothetical LMTP interventions in which the propensity scores were shifted to discourage alcohol use at each time point and under no intervention (ie, no shift for the propensity score) (Table S8 in Appendix S11). For both males and females, comparisons of expected outcomes under each LMTP intervention and no intervention did not suggest an impact (Table S9).

Discussion

Our study demonstrates how to apply target trials and modified treatment policies to define causal effects in a challenging longitudinal problem, using LTMLE for estimation. Our analysis involved detailed information on alcohol initiation and use in adolescents and depression in adulthood, with 21 follow-up time points, censoring, and many baseline and time-dependent confounders. Analytical challenges in the AT analysis included highly variable weights induced by data sparsity[17] and high-dimensional potential exposure patterns. To tackle these challenges, we used 2 approaches to modify the target parameter: (1) an ad hoc approach to remove patterns with less data support from the loss function and (2) defining longitudinal interventions shifting the propensity scores to discourage drinking.

Our LTMLE with GLM and SL analyses suggested that earlier initiation of alcohol drinking was associated with increased depression in males and reduced symptoms in females. LTMLE with GLM also indicated that cumulative duration of drinking was associated with increased depression similarly in males and females; LTMLE with SL had similar point estimates but wider CIs that included the null. Further, this AT analysis was hampered

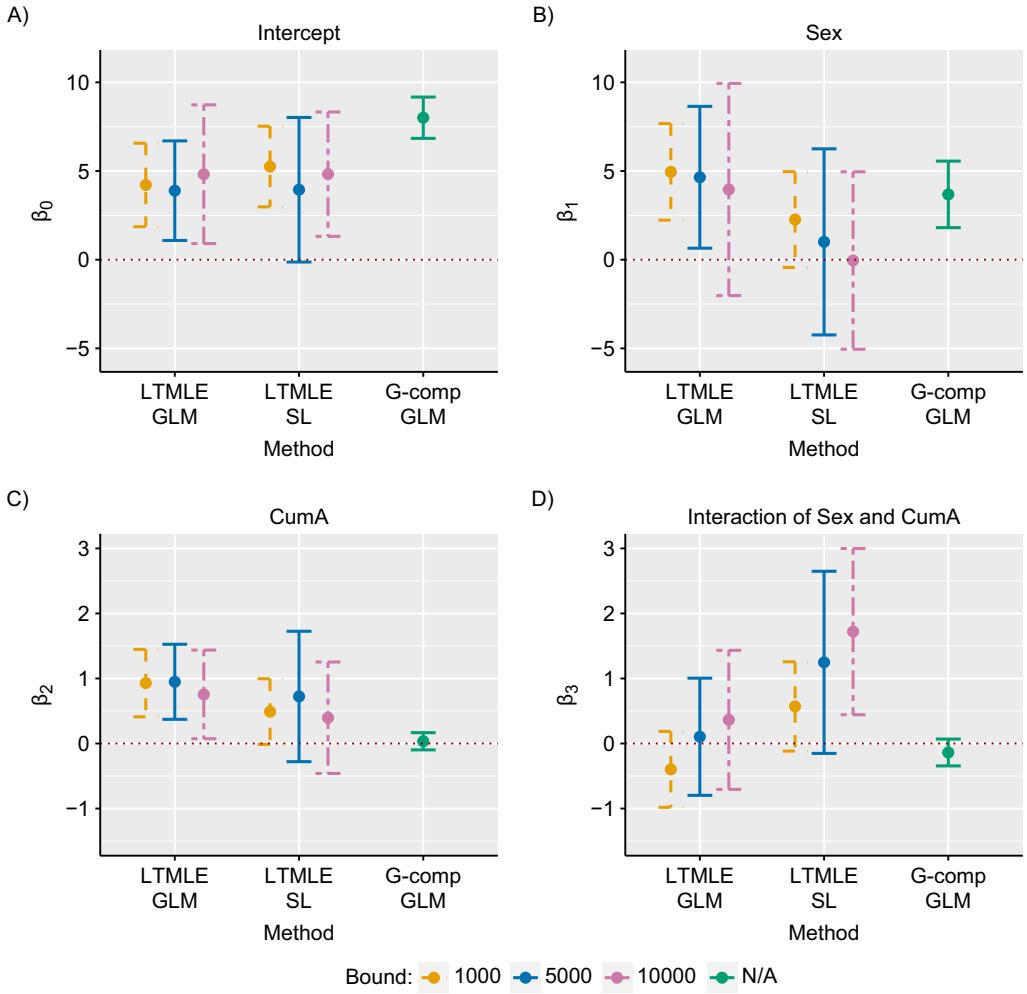

Figure 3. Intercept and coefficient estimates from longitudinal targeted maximum likelihood estimation (LTMLE) with generalized linear models (GLMs) and Super Learner (SL) in the as-treated analysis including 936 patterns under bounds of 1000 (– – –; gold), 5000 (—; blue), and 10 000 (– - –; pink) on the cumulative stabilized weights, Nicotine Dependence in Teens Study, 1999-2008. A) Intercept; B) coefficient of sex; C) coefficient of cumulative alcohol use (CumA); D) coefficient of the interaction between sex and CumA. The y-axis represents the Major Depression Inventory (MDI) score. The intercept represents the average MDI score for males in the absence of drinking; the coefficient of sex represents the average increment of MDI score for females compared with males in the absence of drinking; the coefficient of CumA represents the average increment of MDI score for males for 1 additional time period of drinking; and the coefficient of the interaction between sex and CumA represents the average increment of MDI score for females for 1 additional time period of drinking. The reference sex was male. Bounding of weights for G-computation (G-comp; green) was not applicable (N/A). Bars show 95% confidence intervals.

by sparsity and was sometimes sensitive to the weight bounds, though less sensitive to the number of patterns included in the loss function. For better insight, we employed sequential G-computation, which uses the same estimation procedure as LTMLE without the weighting component, and noted sometimes important differences in the point estimates. Because LTMLE is doubly robust, influential weights suggest misspecified models for the outcome—this is because if the outcome models were correctly specified, the weights would not be informative for the outcome residuals and so the update step would not modify the estimation. While we stratified treatment and censoring models by time point, model-smoothing over time points may provide more weight stability at the risk of increased bias.[18] However, it is not clear to what point the instability of the weights inserted bias into the analysis, and thus we found that it was not possible to draw a strong conclusion from the analysis of the (weakly identified) AT MSM parameters. This is why we defined and estimated a causal parameter under an LMTP intervention discouraging alcohol use to various degrees, which was robust to the previous sparsity issues. From this analysis, we did not conclude that such an intervention would have an effect. Similar LMTP parameters have been proposed elsewhere.[19]

Our truncation levels were determined a posteriori. Indeed, truncation levels are often chosen based on ad hoc criteria in practice. In a recent paper, Gruber et al.[49] suggested a $\sqrt{n}\ln(n)/5$ upper bound in the average treatment effect setting. Ju et al.[50] proposed collaborative TMLE for adaptive propensity score truncation, and one of us (M.E.S.) developed collaborative TMLE in the longitudinal setting[51] that could potentially be applied for adaptive propensity score truncation. Other perspectives and approaches have also been described.[39,52,53]

Limitations of our analysis included the interference assumption's probably being violated to some extent (eg, within school classes), since drinking behavior is transmissible in adolescents, as perceived peer norms have a direct effect on alcohol use.[54] Second, though we adjusted for many relevant confounders, the "no unmeasured confounders" assumption was probably unmet, since we did not have a complete profile on personal

circumstances that would affect the timing of drinking in adolescence and depression in adulthood. Third, due to the limited computational ability of our local computers, we included at most 1688 treatment patterns in the AT analysis which changed the parameter of interest, potentially leading to bias. Finally, our MSMs may have smoothed incorrectly over the effects of drinking exposure if the effects varied by time point. If so, this would have resulted in less interpretable coefficients, though the estimation and inference would remain valid for these parameters.[36]

Our study contributes to a growing body of literature on the application of robust longitudinal causal inference methods. While these methods have many important theoretical properties, data sparsity is a common challenge. We thus encourage epidemiologists and applied statisticians to explore recently proposed parameter definitions and estimation methods that weaken positivity assumptions, leading to more robust results.

Acknowledgments

This work was presented orally at the 50th Annual Meeting of the Statistical Society of Canada, Ottawa, Ontario, Canada, May 28-31, 2023.

Supplementary material

Supplementary material is available at *American Journal of Epidemiology* online.

Funding

For study cycles 1-21, the NDIT Study was supported by the Canadian Cancer Research Society (grants 010271 and 017435). M.E.S. was supported by a Tier 2 Canada Research Chair. M.P.S. was supported by a Chercheur-Boursier J2 award from the Fonds de Recherche du Quebec–Santé.

Conflict of interest

The authors declare no conflicts of interest.

Data availability

Open-access R code for the simulation is available at https://github.com/Yan2020729/Simulations-of-LTMLE-algorithm.

References

1. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11(5):550-560. https://doi.org/10.1097/00001648-200009000-00011
2. Hernán MÁ, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*. 2000;11(5):561-570. https://doi.org/10.1097/00001648-200009000-00012
3. Robins JM. Marginal structural models. In: *1997 Proceedings of the Section on Bayesian Statistical Science*. American Statistical Association; 1997:1-10.
4. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183(8):758-764. https://doi.org/10.1093/aje/kwv254
5. Snowden JM, Rose S, Mortimer KM. Implementation of G-computation on a simulated data set: demonstration of a causal inference technique. *Am J Epidemiol*. 2011;173(7):731-738. https://doi.org/10.1093/aje/kwq472
6. Vansteelandt S, Keiding N. Invited commentary: G-computation—lost in translation? *Am J Epidemiol*. 2011;173(7):739-742. https://doi.org/10.1093/aje/kwq474
7. Scharfstein DO, Rotnitzky A, Robins JM. Adjusting for nonignorable drop-out using semiparametric nonresponse models. *J Am Stat Assoc*. 1999;94(448):1096-1120. https://doi.org/10.1080/01621459.1999.10473862
8. Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models. *Biometrics*. 2005;61(4):962-973. https://doi.org/10.1111/j.1541-0420.2005.00377.x
9. Van der Laan MJ, Rose S. *Targeted Learning in Data Science: Causal Inference for Complex Longitudinal Studies*. Springer-Verlag; 2018.
10. Van der Laan MJ, Gruber S. Targeted minimum loss based estimation of causal effects of multiple time point interventions. *Int J Biostat*. 2012;8(1):9. https://doi.org/10.1515/1557-4679.1370
11. Petersen M, Schwab J, Gruber S, et al. Targeted maximum likelihood estimation for dynamic and static longitudinal marginal structural working models. *J Causal Inference*. 2014;2(2):147-185. https://doi.org/10.1515/jci-2013-0007
12. Schnitzer ME, Van der Laan MJ, Moodie EE, et al. Effect of breastfeeding on gastrointestinal infection in infants: a targeted maximum likelihood approach for clustered longitudinal data. *Ann. Appl Stat*. 2014;8(2):703-725. https://doi.org/10.1214/14-AOAS727
13. Kreif N, Tran L, Grieve R, et al. Estimating the comparative effectiveness of feeding interventions in the pediatric intensive care unit: a demonstration of longitudinal targeted maximum likelihood estimation. *Am J Epidemiol*. 2017;186(12):1370-1379. https://doi.org/10.1093/aje/kwx213
14. Brooks JC, Van der Laan MJ, Singer DE, et al. Targeted minimum loss-based estimation of causal effects in right-censored survival data with time-dependent covariates: warfarin, stroke, and death in atrial fibrillation. *J Causal Inference*. 2013;1(2):235-254. https://doi.org/10.1515/jci-2013-0001
15. Tran L, Yiannoutsos CT, Musick BS, et al. Evaluating the impact of a HIV low-risk express care task-shifting program: a case study of the targeted learning roadmap. *Epidemiol Methods*. 2016;5(1):69-91. https://doi.org/10.1515/em-2016-0004
16. Tran L, Yiannoutsos C, Wools-Kaloustian K, et al. Double robust efficient estimators of longitudinal treatment effects: comparative performance in simulations and a case study. *Int J Biostat*. 2019;15(2):20170054. https://doi.org/10.1515/ijb-2017-0054
17. Schomaker M, Luque-Fernandez MA, Leroy V, et al. Using longitudinal targeted maximum likelihood estimation in complex settings with dynamic interventions. *Stat Med*. 2019;38(24):4888-4911. https://doi.org/10.1002/sim.8340
18. Rudolph JE, Benkeser D, Kennedy EH, et al. Estimation of the average causal effect in longitudinal data with time-varying exposures: the challenge of nonpositivity and the impact of model flexibility. *Am J Epidemiol*. 2022;191(11):1962-1969. https://doi.org/10.1093/aje/kwac136
19. Kennedy EH. Nonparametric causal effects based on incremental propensity score interventions. *J Am Stat Assoc*. 2019;114(526):645-656. https://doi.org/10.1080/01621459.2017.1422737
20. Díaz I, Williams N, Hoffman KL, et al. Nonparametric causal effects based on longitudinal modified treatment policies. *J Am Stat Assoc*. 2023;118(542):846-857. https://doi.org/10.1080/01621459.2021.1955691
21. O'Loughlin J, Karp I, Koulis T, et al. Determinants of first puff and daily cigarette smoking in adolescents. *Am J Epidemiol*. 2009;170(5):585-597. https://doi.org/10.1093/aje/kwp179

22. O'Loughlin J, Dugas EN, Brunet J, et al. Cohort profile: the Nicotine Dependence in Teens (NDIT) Study. *Int J Epidemiol*. 2015;44(5):1537-1546. https://doi.org/10.1093/ije/dyu135
23. Grant BF, Dawson DA. Age at onset of alcohol use and its association with DSM-IV alcohol abuse and dependence: results from the National Longitudinal Alcohol Epidemiologic Survey. *J Subst Abuse*. 1997;9:103-110. https://doi.org/10.1016/S0899-3289(97)90009-2
24. Torres JM, Rudolph KE, Sofrygin O, et al. Longitudinal associations between having an adult child migrant and depressive symptoms among older adults in the Mexican Health and Aging Study. *Int J Epidemiol*. 2018;47(5):1432-1442. https://doi.org/10.1093/ije/dyy112
25. Mason WA, Kosterman R, Haggerty KP, et al. Dimensions of adolescent alcohol involvement as predictors of young-adult major depression. *J Stud Alcohol Drugs*. 2008;69(2):275-285. https://doi.org/10.15288/jsad.2008.69.275
26. Rohde P, Lewinsohn PM, Kahler CW, et al. Natural course of alcohol use disorders from adolescence to young adulthood. *J Am Acad Child Adolesc Psychiatry*. 2001;40(1):83-90. https://doi.org/10.1097/00004583-200101000-00020
27. Marmorstein NR. Longitudinal associations between alcohol problems and depressive symptoms: early adolescence through early adulthood. *Alcohol Clin Exp Res*. 2009;33(1):49-59. https://doi.org/10.1111/j.1530-0277.2008.00810.x
28. Boden J, Blair S, Newton-Howes G. Alcohol use in adolescents and adult psychopathology and social outcomes: findings from a 35-year cohort study. *Aust N Z J Psychiatry*. 2020;54(9):909-918. https://doi.org/10.1177/0004867420924091
29. Edwards AC, Heron J, Dick DM, et al. Adolescent alcohol use is positively associated with later depression in a population-based UK cohort. *J Stud Alcohol Drugs*. 2014;75(5):758-765. https://doi.org/10.15288/jsad.2014.75.758
30. Tercyak KP, Rodriguez D, Audrain-McGovern J. High school seniors' smoking initiation and progression 1 year after graduation. *Am J Public Health*. 2007;97(8):1397-1398. https://doi.org/10.2105/AJPH.2006.094235
31. Jackson KM, Sher KJ, Cooper ML, et al. Adolescent alcohol and tobacco use: onset, persistence and trajectories of use across two samples. *Addiction*. 2002;97(5):517-531. https://doi.org/10.1046/j.1360-0443.2002.00082.x
32. Goldbach JT, Berger Cardoso J, Cervantes RC, et al. The relation between stress and alcohol use among Hispanic adolescents. *Psychol Addict Behav*. 2015;29(4):960-968. https://doi.org/10.1037/adb0000133
33. Meca A, Zamboanga BL, Lui PP, et al. Alcohol initiation among recently immigrated Hispanic adolescents: roles of acculturation and sociocultural stress. *Am J Orthopsychiatry*. 2019;89(5):569-578. https://doi.org/10.1037/ort0000352
34. Lorente FO, Souville M, Griffet J, et al. Participation in sports and alcohol consumption among French adolescents. *Addict Behav*. 2004;29(5):941-946. https://doi.org/10.1016/j.addbeh.2004.02.039
35. Mays D, DePadilla L, Thompson NJ, et al. Sports participation and problem alcohol use: a multi-wave national sample of adolescents. *Am J Prev Med*. 2010;38(5):491-498. https://doi.org/10.1016/j.amepre.2010.01.023
36. Neugebauer R, van der Laan MJ. *Locally Efficient Estimation of Nonparametric Causal Effects on Mean Outcomes in Longitudinal Studies*. (U.C. Berkeley Division of Biostatistics Working Paper Series. University of California, Berkeley; 2003. Working Paper 134).
37. Chaiton M, Contreras G, Brunet J, et al. Heterogeneity of depressive symptom trajectories through adolescence: predicting outcomes in young adulthood. *J Can Acad Child Adolesc Psychiatry*. 2013;22(2):96-105. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3647625/
38. Bech P, Timmerby N, Martiny K, et al. Psychometric evaluation of the Major Depression Inventory (MDI) as depression severity scale using the LEAD (Longitudinal Expert Assessment of all Data) as index of validity. *BMC Psychiatry*. 2015;15(1):1-7. https://doi.org/10.1186/s12888-015-0529-3
39. Petersen ML, Porter KE, Gruber S, et al. Diagnosing and responding to violations in the positivity assumption. *Stat Methods Med Res*. 2012;21(1):31-54. https://doi.org/10.1177/0962280210386207
40. Van der Laan MJ, Rubin D. Targeted maximum likelihood learning. *Int J Biostat*. 2006;2(1):11. https://doi.org/10.2202/1557-4679.1043
41. Van der Laan MJ, Robins JM. *Unified Methods for Censored Longitudinal Data and Causality*. Springer-Verlag; 2003.
42. Van der Laan MJ, Polley EC, Hubbard AE. Super learner. *Stat Appl Genet Mol Biol*. 2007;6(1):25. https://doi.org/10.2202/1544-6115.1309
43. Robins JM, Hernán MA, Siebert U. Effects of multiple interventions. In: Ezzati M, Lopez AD , Rodgers A, et al., eds. *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors*. Vol. 1. World Health Organization; 2004:2191-2230.
44. Richardson TS, Robins JM. *Single World Intervention Graphs (SWIGs): a Unification of the Counterfactual and Graphical Approaches to Causality*. (CSSS Working Paper no. 128). Center for Statistics and the Social Sciences, University of Washington; 2013.
45. Young JG, Hernán MA, Robins JM. Identification, estimation and approximation of risk under interventions that depend on the natural value of treatment using observational data. *Epidemiol Methods*. 2014;3(1):1-19. https://doi.org/10.1515/em-2012-0001
46. Díaz I, Hoffman KL, Hejazi NS. Causal survival analysis under competing risks using longitudinal modified treatment policies. *Lifetime Data Anal*. 2024;30:213-236. https://doi.org/10.1007/s10985-023-09606-7
47. Wen L, Marcus JL, Young JG. Intervention treatment distributions that depend on the observed treatment process and model double robustness in causal survival analysis. *Stat Methods Med Res*. 2023;32(3):509-523. https://doi.org/10.1177/09622802221146311
48. Williams N, Díaz I. lmtp: an R package for estimating the causal effects of modified treatment policies *Obs Stud*. 2023;9(2):103-122. https://doi.org/10.1353/obs.2023.0019
49. Gruber S, Phillips RV, Lee H, et al. Data-adaptive selection of the propensity score truncation level for inverse-probability–weighted and targeted maximum likelihood estimators of marginal point treatment effects. *Am J Epidemiol*. 2022;191(9):1640-1651. https://doi.org/10.1093/aje/kwac087
50. Ju C, Schwab J, van der Laan MJ. On adaptive propensity score truncation in causal inference. *Stat Methods Med Res*. 2019;28(6):1741-1760. https://doi.org/10.1177/0962280218774817
51. Schnitzer ME, Sango J, Ferreira Guerra S, et al. Data-adaptive longitudinal model selection in causal inference with collaborative targeted minimum loss-based estimation. *Biometrics*. 2020;76(1):145-157. https://doi.org/10.1111/biom.13135
52. Wang Y, Petersen ML, Bangsberg D, et al. *Diagnosing Bias in the Inverse Probability of Treatment Weighted Estimator Resulting From Violation of Experimental Treatment Assignment*. (U.C. Berkeley Division of Biostatistics Working Paper Series. University of California, Berkeley; 2006. Working Paper 211).

53. Bembom O, van der Laan MJ. *Data-Adaptive Selection of the Truncation Level for Inverse-Probability-of-Treatment-Weighted Estimators*. (U.C. Berkeley Division of Biostatistics Working Paper Series, Working Paper 230). University of California, Berkeley; 2008.

54. Scheier LM, Botvin GJ. Expectancies as mediators of the effects of social influences and alcohol knowledge on adolescent alcohol use: a prospective analysis. *Psychol Addict Behav*. 1997;11(1):48-64. https://doi.org/10.1037/0893-164X.11.1.48