MRC Biostatistics Unit

UNIVERSITY OF CAMBRIDGE

Inference procedures in target trial emulation with survival outcomes

Comparison of confidence intervals based on the sandwich variance estimator and two forms of bootstrap

Juliette Limozin, Shaun Seaman, Li Su [1]

[1]MRC Biostatistics Unit, University of Cambridge, UK

European Causal Inference Meeting 2023 18 April 2023

Background

- Target trial emulation:
  - A popular approach for causal inference using longitudinal observational data
  - Applications to survival analysis

Background

- Target trial emulation:
  - A popular approach for causal inference using longitudinal observational data
  - Applications to survival analysis
- Estimation and inference:
  - Interested in estimating and constructing confidence intervals of the per-protocol effect of sustained treatment compared to no treatment on survival outcome
  - Presence of time-varying confounding and dependent loss-to-follow-up

Background

- Target trial emulation:
  - A popular approach for causal inference using longitudinal observational data
  - Applications to survival analysis
- Estimation and inference:
  - Interested in estimating and constructing confidence intervals of the per-protocol effect of sustained treatment compared to no treatment on survival outcome
  - Presence of time-varying confounding and dependent loss-to-follow-up
- $\Rightarrow$ Confidence interval construction is not straightforward

Sequentially emulated trials

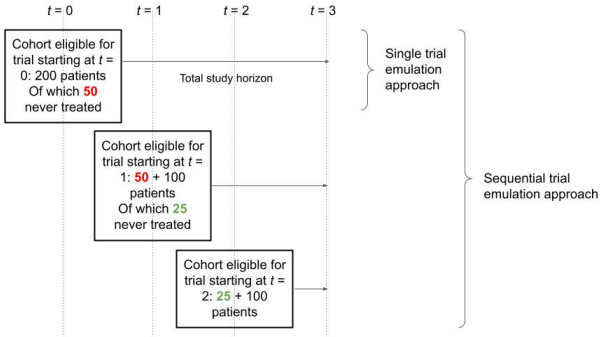
- Instead of having one emulated trial starting at time t_0, we can emulate a sequence of trials starting at times $t_0, t_1, ...,$

Sequentially emulated trials

- Instead of having one emulated trial starting at time t_0, we can emulate a sequence of trials starting at times $t_0, t_1, ...,$
- Each individual is enrolled in all trials they're eligible to (e.g. no prior treatment history)

Sequentially emulated trials

- Instead of having one emulated trial starting at time t_0, we can emulate a sequence of trials starting at times $t_0, t_1, ...,$
- Each individual is enrolled in all trials they're eligible to (e.g. no prior treatment history)



Estimation and Inference

Causal estimand of interest: the marginal risk difference (averaged over target population)

$$MRD(t_k) = \Pr(Y_k^{\bar{A}_k=\bar{1}} = 1) - \Pr(Y_k^{\bar{A}_k=\bar{0}} = 1)$$

- Need to censor patients when they no longer comply with treatment protocol for per-protocol analysis

Estimation and Inference

Causal estimand of interest: the marginal risk difference (averaged over target population)

$$MRD(t_k) = \Pr(Y_k^{\bar{A}_k=\bar{1}} = 1) - \Pr(Y_k^{\bar{A}_k=\bar{0}} = 1)$$

- Need to censor patients when they no longer comply with treatment protocol for per-protocol analysis
- Marginal Structural Models (MSMs) to estimate marginal probabilities $\Pr(Y_k^{\bar{A}_k=\bar{1}} = a), a = \bar{0}, \bar{1},$
- Inverse Probability Weighting (IPW) to account for the time-varying confounders that affect treatment switching and dependent censoring

Estimation and Inference cont.

Confidence interval of the marginal risk difference:

- Using a robust sandwich variance estimator as an estimate of the MSM coefficients' variance-covariance matrix (accounts for correlation induced from individuals being in multiple trials)

Estimation and Inference cont.

Confidence interval of the marginal risk difference:

- Using a robust sandwich variance estimator as an estimate of the MSM coefficients' variance-covariance matrix (accounts for correlation induced from individuals being in multiple trials)
  - Draw sample $\beta_1, ..., \beta_M \sim MVN(\hat{\beta}, \hat{V})$ from a multivariate Normal distribution centered at the point estimate $\hat{\beta}$ of MSM parameters and with robust sandwich estimator $\hat{V}$ as variance-covariance matrix

Estimation and Inference cont.

Confidence interval of the marginal risk difference:

- Using a robust sandwich variance estimator as an estimate of the MSM coefficients' variance-covariance matrix (accounts for correlation induced from individuals being in multiple trials)
  - Draw sample $\beta_1, ..., \beta_M \sim MVN(\hat{\beta}, \hat{V})$ from a multivariate Normal distribution centered at the point estimate $\hat{\beta}$ of MSM parameters and with robust sandwich estimator $\hat{V}$ as variance-covariance matrix
  - Calculate sample marginal risk difference for each β_m, use percentile method to obtain confidence interval

Estimation and Inference cont.

Confidence interval of the marginal risk difference:

- Using a robust sandwich variance estimator as an estimate of the MSM coefficients' variance-covariance matrix (accounts for correlation induced from individuals being in multiple trials)
  - Draw sample $\beta_1, ..., \beta_M \sim MVN(\hat{\beta}, \hat{V})$ from a multivariate Normal distribution centered at the point estimate $\hat{\beta}$ of MSM parameters and with robust sandwich estimator $\hat{V}$ as variance-covariance matrix
  - Calculate sample marginal risk difference for each β_m, use percentile method to obtain confidence interval
- Nonparametric bootstrapping of marginal risk difference estimate, then construct pivot percentile confidence interval

$$L = 2\widehat{MRD} - \widehat{MRD}^*_{(0.975)}, U = 2\widehat{MRD} - \widehat{MRD}^*_{(0.275)}$$

Research motivation

Several considerations:

- Problem with robust sandwich estimator: doesn't account for uncertainty in weight estimation
- Nonparametric bootstrapping can account for this uncertainty but is computationally heavy to construct CIs from, also may get ill-conditioned covariance matrix in certain data resamples which can affect estimation

Research motivation

Several considerations:

- Problem with robust sandwich estimator: doesn't account for uncertainty in weight estimation
- Nonparametric bootstrapping can account for this uncertainty but is computationally heavy to construct CIs from, also may get ill-conditioned covariance matrix in certain data resamples which can affect estimation
- No literature comparing the two methods in the sequential trial emulation context using IPW-MSM
- The Linearised Estimating Function (LEF) bootstrap: a computationally more efficient alternative to nonparametric bootstrapping, but hasn't been applied in the trial emulation context

Research motivation

Linearised estimating function (LEF) bootstrap:

- Instead of re-estimating parameters of a regression model in each bootstrap sample, calculate bootstrap parameter estimates in each sample *b* using a linear approximation of the estimating function $0 = \hat{U}^{(b)}(\hat{\theta}) \approx \hat{U}^{(b)}(\theta) + \frac{\partial \hat{U}^{(b)}(\theta)}{\partial \theta}(\hat{\theta} - \theta)$:

$$\begin{aligned}\hat{\theta}^{(b)_{LEF}} &= \hat{\theta} - \left(\frac{\partial \widehat{U}(\theta)}{\partial \theta}|_{\theta=\hat{\theta}}\right)^{-1} \hat{U}^{(b)}(\hat{\theta}) \\ &\approx \hat{\theta} + \text{vcov}(\hat{\theta})\hat{U}^{(b)}(\hat{\theta})\end{aligned}$$

Research motivation

Linearised estimating function (LEF) bootstrap:

- Instead of re-estimating parameters of a regression model in each bootstrap sample, calculate bootstrap parameter estimates in each sample *b* using a linear approximation of the estimating function $0 = \hat{U}^{(b)}(\hat{\theta}) \approx \hat{U}^{(b)}(\theta) + \frac{\partial \hat{U}^{(b)}(\theta)}{\partial \theta}(\hat{\theta} - \theta)$:

$$\begin{aligned}\hat{\theta}^{(b)_{LEF}} &= \hat{\theta} - \left(\frac{\partial \widehat{U}(\theta)}{\partial \theta}|_{\theta=\hat{\theta}}\right)^{-1} \hat{U}^{(b)}(\hat{\theta}) \\ &\approx \hat{\theta} + \text{vcov}(\hat{\theta})\hat{U}^{(b)}(\hat{\theta})\end{aligned}$$

- can be used for bootstrapping parameters of MSM and/or propensity score model for IPW estimation

Research motivation

Linearised estimating function (LEF) bootstrap:

- Instead of re-estimating parameters of a regression model in each bootstrap sample, calculate bootstrap parameter estimates in each sample *b* using a linear approximation of the estimating function $0 = \hat{U}^{(b)}(\hat{\theta}) \approx \hat{U}^{(b)}(\theta) + \frac{\partial \hat{U}^{(b)}(\theta)}{\partial \theta}(\hat{\theta} - \theta)$:

$$\hat{\theta}^{(b)_{LEF}} = \hat{\theta} - \left(\frac{\partial \widehat{U}(\theta)}{\partial \theta}|_{\theta=\hat{\theta}} \right)^{-1} \hat{U}^{(b)}(\hat{\theta})$$
$$\approx \hat{\theta} + \text{vcov}(\hat{\theta})\hat{U}^{(b)}(\hat{\theta})$$

- can be used for bootstrapping parameters of MSM and/or propensity score model for IPW estimation
- No re-fitting of the treatment process model/MSM in each bootstrap sample $\Rightarrow$ computationally more efficient than nonparametric bootstrap

Research motivation

Aim:

- compare these statistical methods for constructing confidence intervals of the marginal risk difference in the sequentially emulated trial setting for causal inference on survival outcomes, through a simulation study

Simulation framework

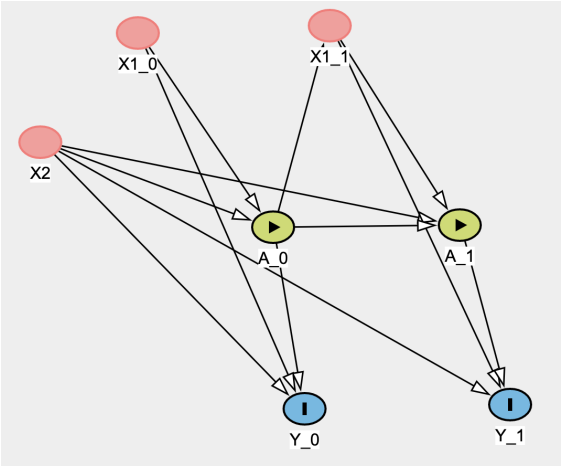
- Notation:
  - $k = 0, 1, ..., 4$ follow-up visit $\Rightarrow$ 5 sequentially emulated trials
  - Y_k outcome indicator: $Y_k = 1$ if event occurred within $[k, k+1)$
  - A_k binary treatment
  - X_{1_k} time-dependent confounder that is dependent on treatment history
  - X_2 time-invariant confounder

Simulation framework

- Notation:
  - $k = 0, 1, ..., 4$ follow-up visit $\Rightarrow$ 5 sequentially emulated trials
  - Y_k outcome indicator: $Y_k = 1$ if event occurred within $[k, k+1)$
  - A_k binary treatment
  - X_{1_k} time-dependent confounder that is dependent on treatment history
  - X_2 time-invariant confounder
- Simulation study:
  - 1000 Monte Carlo simulations per data scenario (81 scenarios total)

Outcome prevalence	Patient sample size	Confounding strength	Treatment prevalence
2-3.5% (low event rate), 4.5-6.5% (medium), 9-12% (high)	200, 1000, 5000	0.1 (low confounding), 0.5 (medium), 0.9 (high)	25-30% (low treatment prevalence), $\approx$ 48-52% (medium), 70-75% (high)

Simulation framework



DAGitty v3.0

Simulation framework

For each simulated dataset:

- Estimated the marginal risk difference of patients in trial 0
  - MSM of marginal risk with weighted pooled logistic model:

$$\begin{aligned}\text{logit}[\Pr(Y_{m,k} = 1 | \bar{A}_{m,k}, X_{1_{m,0}}, X_{2_m}, \bar{Y}_{m,k-1} = 0)] \\ = \text{trial-specific intercept} + \text{visit main effect} \\ + \text{treatment main effect} + \text{main effects of } X_{1_k}, X_2 \\ + \text{interactions between visits and treatment/confounders}\end{aligned}$$

Simulation framework

For each simulated dataset:

- Estimated the marginal risk difference of patients in trial 0
  - MSM of marginal risk with weighted pooled logistic model:

$$\begin{aligned}\text{logit}[\Pr(Y_{m,k} = 1 | \bar{A}_{m,k}, X_{1_{m,0}}, X_{2_m}, \bar{Y}_{m,k-1} = 0)] \\ = \text{trial-specific intercept} + \text{visit main effect} \\ + \text{treatment main effect} + \text{main effects of } X_{1_k}, X_2 \\ + \text{interactions between visits and treatment/confounders}\end{aligned}$$

  - Treatment confounding adjustment with IPW (IPTW): treatment process modelled by

$$\text{logit}[\Pr(A_k = 1 | A_{k-1} = a, X_{1_k}, X_2)] = \gamma_{0,a} + \gamma_{1,a} X_{1_k} + \gamma_{2,a} X_2 \quad (1)$$

Simulation framework

For each simulated dataset:

- Estimated the marginal risk difference of patients in trial 0
  - MSM of marginal risk with weighted pooled logistic model:

$$\begin{aligned}\text{logit}[\Pr(Y_{m,k} = 1 | \bar{A}_{m,k}, X_{1_{m,0}}, X_{2_m}, \bar{Y}_{m,k-1} = 0)] \\ = \text{trial-specific intercept} + \text{visit main effect} \\ + \text{treatment main effect} + \text{main effects of } X_{1_k}, X_2 \\ + \text{interactions between visits and treatment/confounders}\end{aligned}$$

  - Treatment confounding adjustment with IPW (IPTW): treatment process modelled by

$$\text{logit}[\Pr(A_k = 1 | A_{k-1} = a, X_{1_k}, X_2)] = \gamma_{0,a} + \gamma_{1,a} X_{1_k} + \gamma_{2,a} X_2 \quad (1)$$

  - 'True' MRD: two KM curve on large number of simulated trial 0 patients, setting them as either always treated or never treated

Simulation framework

In each data iteration:

- Constructed four types of confidence intervals of marginal risk difference for patients in trial 0:
  - Using the robust sandwich variance estimator

Simulation framework

In each data iteration:

- Constructed four types of confidence intervals of marginal risk difference for patients in trial 0:
  - Using the robust sandwich variance estimator
  - Nonparametric bootstrapping for pivot confidence interval

Simulation framework

In each data iteration:

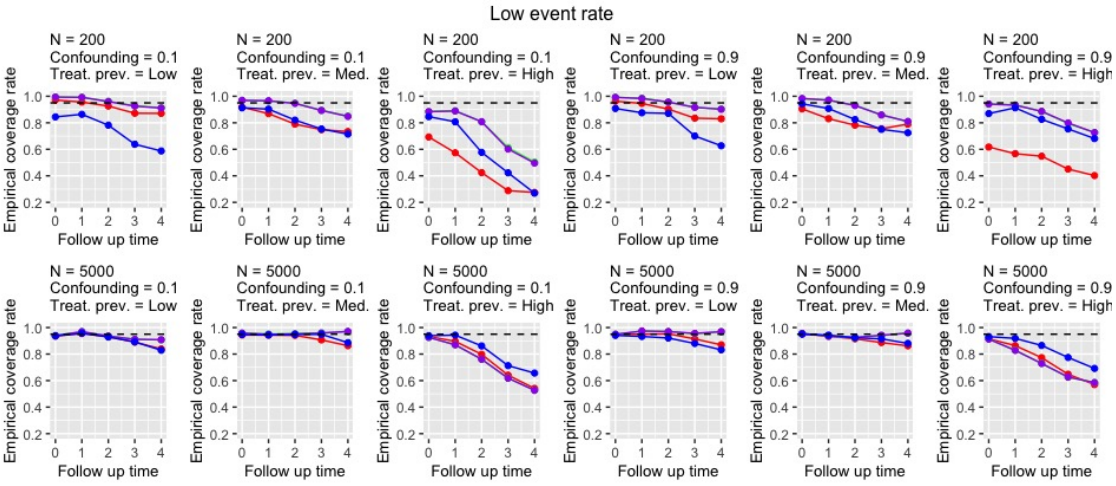
- Constructed four types of confidence intervals of marginal risk difference for patients in trial 0:
  - Using the robust sandwich variance estimator
  - Nonparametric bootstrapping for pivot confidence interval
  - LEF bootstrapping on MSM parameter estimates only (IPW reestimated in each bootstrap sample) for pivot confidence interval

Simulation framework

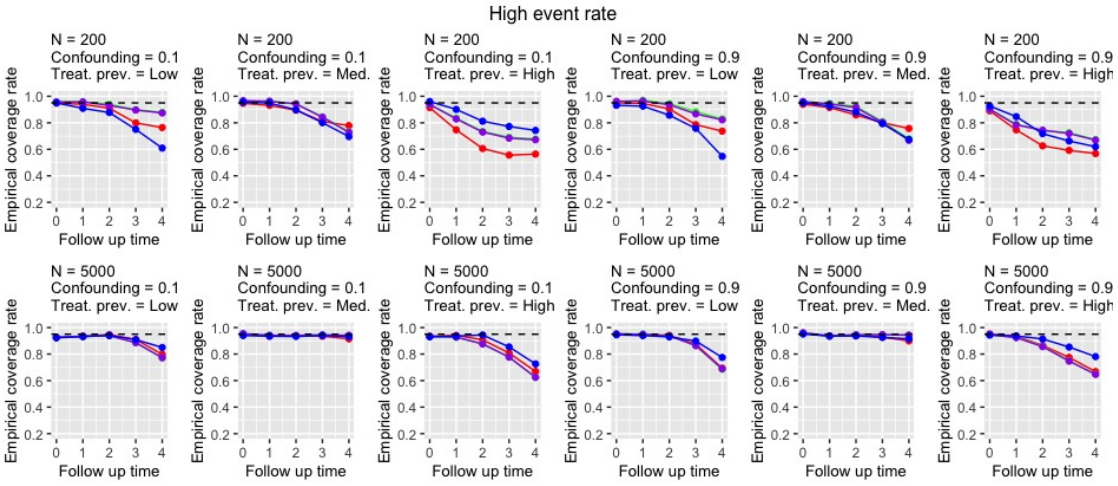
In each data iteration:

- Constructed four types of confidence intervals of marginal risk difference for patients in trial 0:
  - Using the robust sandwich variance estimator
  - Nonparametric bootstrapping for pivot confidence interval
  - LEF bootstrapping on MSM parameter estimates only (IPW reestimated in each bootstrap sample) for pivot confidence interval
  - LEF bootstrapping on IPW **and** MSM parameter estimates for pivot confidence interval

Results


Low event rate
N = 200
Confounding = 0.1
Treat. prev. = Low
N = 200
Confounding = 0.1
Treat. prev. = Med.
N = 200
Confounding = 0.1
Treat. prev. = High
N = 200
Confounding = 0.9
Treat. prev. = Low
N = 200
Confounding = 0.9
Treat. prev. = Med.
N = 200
Confounding = 0.9
Treat. prev. = High
N = 5000
Confounding = 0.1
Treat. prev. = Low
N = 5000
Confounding = 0.1
Treat. prev. = Med.
N = 5000
Confounding = 0.1
Treat. prev. = High
N = 5000
Confounding = 0.9
Treat. prev. = Low
N = 5000
Confounding = 0.9
Treat. prev. = Med.
N = 5000
Confounding = 0.9
Treat. prev. = High
Empirical coverage rate
Follow up time
CI type
Pivot Bootstrap
Pivot LEF both
Pivot LEF outcome
Sandwich

Results


High event rate
N = 200
Confounding = 0.1
Treat. prev. = Low
N = 200
Confounding = 0.1
Treat. prev. = Med.
N = 200
Confounding = 0.1
Treat. prev. = High
N = 200
Confounding = 0.9
Treat. prev. = Low
N = 200
Confounding = 0.9
Treat. prev. = Med.
N = 200
Confounding = 0.9
Treat. prev. = High
N = 5000
Confounding = 0.1
Treat. prev. = Low
N = 5000
Confounding = 0.1
Treat. prev. = Med.
N = 5000
Confounding = 0.1
Treat. prev. = High
N = 5000
Confounding = 0.9
Treat. prev. = Low
N = 5000
Confounding = 0.9
Treat. prev. = Med.
N = 5000
Confounding = 0.9
Treat. prev. = High
Empirical coverage rate
Follow up time
CI type
Pivot Bootstrap
Pivot LEF both
Pivot LEF outcome
Sandwich

Results

- Key takeaways:
  - Lack of data at late follow-up visits introduced large bias and low coverage.

Results

- Key takeaways:
  - Lack of data at late follow-up visits introduced large bias and low coverage.
  - Favour LEF bootstrapping methods in terms of CI coverage for data with small patient sample size, low event rate, low or high treatment prevalence

Results

- Key takeaways:
  - Lack of data at late follow-up visits introduced large bias and low coverage.
  - Favour LEF bootstrapping methods in terms of CI coverage for data with small patient sample size, low event rate, low or high treatment prevalence
  - Extreme weights or outcome results in certain bootstrap samples for small sample and low event rates $\Rightarrow$ nonparametric bootstrap suffering in coverage

Results

- Key takeaways:
  - Lack of data at late follow-up visits introduced large bias and low coverage.
  - Favour LEF bootstrapping methods in terms of CI coverage for data with small patient sample size, low event rate, low or high treatment prevalence
  - Extreme weights or outcome results in certain bootstrap samples for small sample and low event rates $\Rightarrow$ nonparametric bootstrap suffering in coverage
  - Increased event rate & sample size $\Rightarrow$ similar CI coverage for all methods
    - Favour robust sandwich estimator method with large sample size, fastest to compute

Results

- Key takeaways:
  - Lack of data at late follow-up visits introduced large bias and low coverage.
  - Favour LEF bootstrapping methods in terms of CI coverage for data with small patient sample size, low event rate, low or high treatment prevalence
  - Extreme weights or outcome results in certain bootstrap samples for small sample and low event rates $\Rightarrow$ nonparametric bootstrap suffering in coverage
  - Increased event rate & sample size $\Rightarrow$ similar CI coverage for all methods
    - Favour robust sandwich estimator method with large sample size, fastest to compute
- Other observations to note:
  - Higher risk of sandwich variance-based CIs failing to compute due to positive non-definiteness of variance-covariance matrix
  - Computation time: nonparametric bootstrap 18 times slower, LEF bootstrap 4 times slower than sandwich

TrialEmulation R package

On GitHub: github.com/Causal-LDA/TrialEmulation

- Currently being developed by Li Su, Isaac Gravestock (Roche), Shaun Seaman, Roonak Rezvani (U of Surrey)
- Utility:
  - Expands observational data selected for trial emulation into sequential trials, assigning patients to trials they're eligible for
  - Estimates IPW and fits weighted MSM to obtain intention-to-treat and per-protocol effect
  - Provides inference of the marginal risk difference estimate

  $\rightarrow$ Automates the analysis of sequentially emulated trials

Acknowledgements

Thank you to:

- Li Su, MRC Biostatistics Unit, University of Cambridge (supervisor)
- Shaun Seaman, MRC Biostatistics Unit, University of Cambridge (advisor)
- Isaac Gravestock, Roche Pharmaceuticals
- Roonak Rezvani, University of Surrey

References I

Peter C. Austin.

Bootstrap vs asymptotic variance estimation when using propensity score weighting with continuous and binary outcomes.
Statistics in Medicine, 7 2022.
ISSN 0277-6715.
doi: 10.1002/sim.9519.
URL https://onlinelibrary.wiley.com/doi/10.1002/sim.9519.

Goodarz Danaei, Luis A GarcÃa Rodríguez, Oscar FernÃ¡ndez Cantero, Roger Logan, and Miguel A Hernán.

Observational data for comparative effectiveness research: An emulation of randomised trials of statins and primary prevention of coronary heart disease.
Statistical Methods in Medical Research, 22(1):70–96, 2 2013.
ISSN 0962-2802.
doi: 10.1177/0962280211403603.
URL http://journals.sagepub.com/doi/10.1177/0962280211403603.

R.M. Daniel, S.N. Cousens, B.L. De Stavola, M. G. Kenward, and J. A. C. Sterne.

Methods for dealing with time-dependent confounding.
Statistics in Medicine, 32(9):1584–1618, 4 2013.
ISSN 02776715.
doi: 10.1002/sim.5686.
URL https://onlinelibrary.wiley.com/doi/10.1002/sim.5686.

References II

M A Hernán, B Brumback, and J M Robins.
Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men.
Epidemiology (Cambridge, Mass.), 11(5):561–70, 9 2000.
ISSN 1044-3983.
doi: 10.1097/00001648-200009000-00012.

Miguel A. Hernán and James M. Robins.
Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available.
American Journal of Epidemiology, 183(8):758–764, 4 2016.
ISSN 0002-9262.
doi: 10.1093/aje/kwv254.

Miguel A. Hernán, Alvaro Alonso, Roger Logan, Francine Grodstein, Karin B. Michels, Walter C. Willett, JoAnn E. Manson, and James M. Robins.
Observational Studies Analyzed Like Randomized Experiments.
Epidemiology, 19(6):766–779, 11 2008.
ISSN 1044-3983.
doi: 10.1097/EDE.0b013e3181875e61.
URL https://journals.lww.com/00001648-200811000-00002.

References III

Feifang Hu and John D. Kalbfleisch.
The estimating function bootstrap.
Canadian Journal of Statistics, 28(3):449–481, 9 2000.
ISSN 03195724.
doi: 10.2307/3315958.
URL http://doi.wiley.com/10.2307/3315958.

Ruth H. Keogh, Jon Michael Gran, Shaun R. Seaman, Gwyneth Davies, and Stijn Vansteelandt.
Causal inference in survival analysis using longitudinal observational data: Sequential trials and marginal structural models.
10 2021.
URL http://arxiv.org/abs/2110.03117.

Eleanor J Murray, Ellen C Caniglia, and Lucia C Petito.
Causal survival analysis: A guide to estimating intention-to-treat and per-protocol effects from randomized clinical trials with non-adherence.
Research Methods in Medicine & Health Sciences, 2(1):39–49, 1 2021.
ISSN 2632-0843.
doi: 10.1177/2632084320961043.
URL http://journals.sagepub.com/doi/10.1177/2632084320961043.

References IV

James M. Robins, Miguel Ãngel Hernán, and Babette Brumback.
Marginal Structural Models and Causal Inference in Epidemiology.
Epidemiology, 11(5):550–560, 9 2000.
ISSN 1044-3983.
doi: 10.1097/00001648-200009000-00011.
URL http://journals.lww.com/00001648-200009000-00011.

LEF on MSM parameters CI algorithm

1. Fit IPW-MSM to the whole data and generate one point estimate $\hat{\beta}$ of MSM parameters and IP weights w_i
2. Draw B bootstrap samples , recalculate patients' IP weights $w_i^{(b)}$ in each bootstrap sample b
3. Calculate the bth bootstrap parameter estimate $\hat{\beta}^{(b)}$ from each bootstrap sample $b = 1, ..., B$ based on new weights $w_i^{(b)}$:

$$\hat{\beta}^{(b)} = \hat{\beta} - \left[\sum_i w_i X_i X_i^T p_i(\hat{\beta})(1 - p_i(\hat{\beta}))\right]^{-1} \times \sum_i w_i^{(b)}(Y_i - p_i(\hat{\beta}))$$

   where $p_i(\beta) = \exp(X_i^T\beta)/[1 + \exp(X_i^T\beta)]$ is the probability of event for patient i and X_i is the vector of variables for patient i in the outcome model.
4. From each bootstrap sample $b = 1, ..., B$, estimate MRD by setting $\hat{\beta}^{(b)}$ as the MSM parameters.
5. $L = 2\widehat{MRD} - \widehat{MRD}^*_{(0.975)}$, $U = 2\widehat{MRD} - \widehat{MRD}^*_{(0.275)}$

Data generating mechanism

Data simulation variables	n_s: number of patients n_v: number of follow-up visits; $k = 0, ..., n_v$ α_t: treatment prevalence in logit scale α_c: confounding strength of time-dependent variable $X2$ on treatment assignment and outcome α_y: outcome prevalence
Time-dependent confounders	$X_{1_k} = Z2_k - 0.3A_{k-1}, Z2_0, ..., Z2_{n_v} \sim^{iid} N(0, 1)$ a disease activity marker, which is the target of the treatment
Time-independent confounders	Baseline continuous covariate: $X_2 \sim N(0, 1)$
Treatment	$\text{logit}\Pr(A_k = 1 \mid A_{k-1} = a, X_{1_k}, X_2)$ $= \alpha_t + 0.05a + \alpha_c X_{1_k} + 0.2X_2$
Outcome	$\text{logit}\Pr(Y_k = 1 \mid X_{1_k}, X_2, A_k = a)$ $= \alpha_y - 0.5A_k + \alpha_c X_{1_k} + X_2$
Eligibility	$E_k = 1$ if they have not received treatment before visit k and have not had recent occurrence before visit k, 0 otherwise

Marginal risk calculation

Marginal risk of a patient eligible to trial 0 at visit k given treatment history $\bar{a}$, assuming consistency, positivity, sequentially ignorable treatment assignment hold:

$$\Pr(Y_k^{\bar{A}_k=\bar{a}} = 1) = \sum_{x_{1_0}, x_2} \left[\sum_{i=0}^{k} \Pr(Y_i = 1 | \bar{A}_i = \bar{a}, X_{1_0} = x_{1_0}, X_2 = x_2) \times \prod_{j=0}^{i-1} \Pr(Y_j = 0 | \bar{A}_j = a, X_{1_0}, X_2) \right]$$

where $\sum_{x_{1_0}, x_2}$ is the standardisation by X_{1_0}, X_2 which can be obtained from averaging the marginal risk by visit k over the N_0 patients eligible to trial 0