# Tradeoffs of Target Trial Emulation Approaches for Comparing Repeated Treatment Schedules

Meghan Gerety
Biostatistics PhD Candidate
University of Pennsylvania

ENAR
March 17th, 2026

GGEB
Graduate Group in Epidemiology & Biostatistics

# Motivation

- No randomized controlled trial has compared **annual vs biennial screening mammography** head-to-head, in part because of the long follow-up needed to observe clinically meaningful endpoints
- Currently, different organizations give *differing recommendations*:
  - United States Preventive Services Task Force: biennial screening for women over 40
  - American Cancer Society, Society of Breast Imaging: annual screening for women over 40

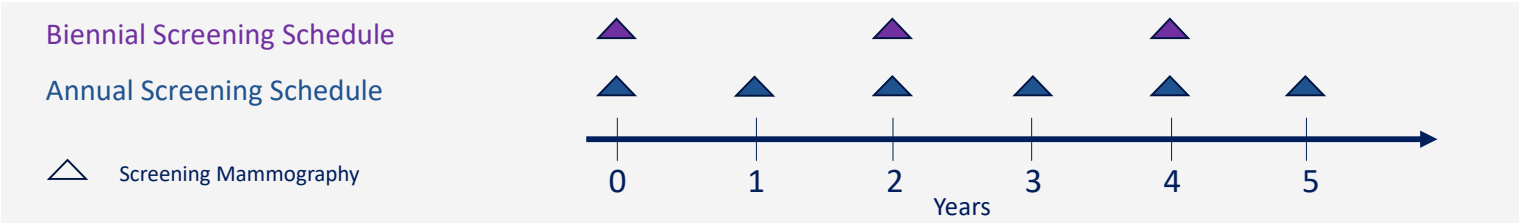


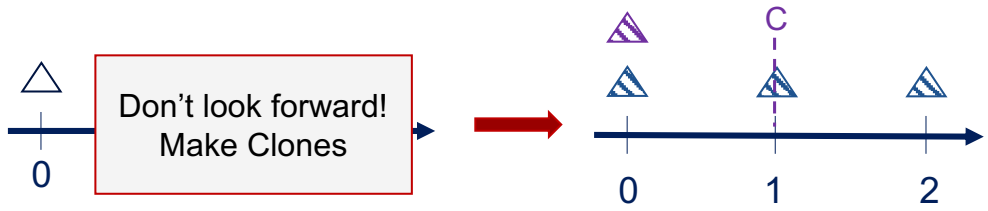
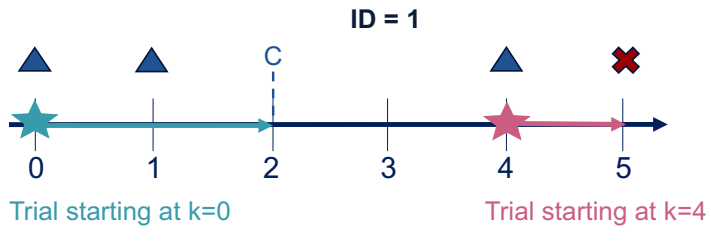

Figure 1: Visualizing annual vs biennial mammography screening schedules.

US Preventive Services Task Force. JAMA. 2024. doi:10.1001/jama.2024.5534
ACS Breast Cancer Screening Guidelines. Accessed August 14, 2024.
Monticciolo DL, et al. Journal of the American College of Radiology. 2021. doi:10.1016/j.jacr.2021.04.021

# Overview of Target Trial Emulation

- **Target trial emulation (TTE)** applies the principles of randomized trials to observational studies to avoid suboptimal study design
  - Two-stage approach

Table 1: Obstacles to emulating a trial protocol in observational data.

<table>
<tr><th>Obstacle</th><th>Common Solution</th></tr>
<tr><td>Individuals are consistent with both treatments arms at time zero</td><td rowspan="2">Clone-Censor-Weight (CCW) Approach<br> 
</td></tr>
<tr><td>Treatment switching and non-adherence occur</td></tr>
<tr><td>Individuals are eligible for the trial at multiple times</td><td>Sequential Trials Approach<br> 

<table>
<tr><th>ID</th><th>Arm</th><th>Time</th><th>Event</th></tr>
<tr><td>1</td><td>A</td><td>2</td><td>0</td></tr>
<tr><td>1</td><td>A</td><td>1</td><td>1</td></tr>
</table></td></tr>
</table>

# Specifying a Treatment Schedule Target Trial

- TTE is **understudied for comparing different schedules of the same repeated treatment**

- We specify **four target trial protocols** for estimating **per-protocol marginal risk differences** between annual and biennial treatment schedules

- Eligibility criteria include being screened at a baseline visit, being at least 40 years of age, having an average risk of breast cancer, and
  - Trial 1. (No additional criteria).
  - Trial 2. Were not screened in the last 2 years.
  - Trial 3. Were not screened in the last 3 years.
  - Trial 4. Were screening naive at the baseline visit.

*Assumptions:* no interference, positivity, consistency, conditional exchangeability of treatment statuses at each visit, uninformative censoring conditional on observed data.

# Data Structure and Notation

- Assume for each individual $i$, $i = 1, \ldots, N$, we observe $k = 0, \ldots, K_i$ data are collected at *regular, annual visits*
- Let $A_{k,i} \in \{0, 1\}$ be a binary treatment indicator and $\boldsymbol{L}_{k,i}$ be a vector of measured covariate values.
  - Overbars indicate history: $\bar{A}_{k,i} = \{A_{0,i}, A_{1,i}, \ldots, A_{k,i}\}$,
  - Particular range of history: $\bar{A}_{(k-t,k),i} = \{A_{k-t,i} \ldots, A_{k,i}\}$
  - Underlines indicate future: $\underline{A}_{k,i} = \{A_{k,i}, A_{k+1,i}, \ldots A_{K_i,i}\}$

# Data Structure and Notation

- Assume for each individual $i$, $i = 1, \ldots, N$, we observe $k = 0, \ldots, K_i$ data are collected at *regular, annual visits*
- Let $A_{k,i} \in \{0, 1\}$ be a binary treatment indicator and $\boldsymbol{L}_{k,i}$ be a vector of measured covariate values.
  - Overbars indicate history: $\bar{A}_{k,i} = \{A_{0,i}, A_{1,i}, \ldots, A_{k,i}\}$,
  - Particular range of history: $\bar{A}_{(k-t,k),i} = \{A_{k-t,i} \ldots, A_{k,i}\}$
  - Underlines indicate future: $\underline{A}_{k,i} = \{A_{k,i}, A_{k+1,i}, \ldots A_{K_i,i}\}$
- In contrast to $A_k$, define $Z_{k,t}$ for $t \geq 1$ such that

$$Z_{k,t} = \begin{cases} 1 & \text{if } \bar{A}_{k,t} = \{A_k, A_{k+1}, \ldots, A_{t-1}, A_t\} = \{1, 1, \ldots, 1, 1\} \\ 0 & \text{if } \bar{A}_{k,t} = \{A_k, A_{k+1}, \ldots, A_{t-1}, A_t\} = \{1, 0, \ldots, I(t-1 \text{ is even}), I(t \text{ is even})\} \end{cases},$$

- Denote adherence for the duration of the study $Z = Z_{0,K_i}$

# Data Structure and Notation

- Assume for each individual $i$, $i = 1, \ldots, N$, we observe $k = 0, \ldots, K_i$ data are collected at *regular, annual visits*
- Let $A_{k,i} \in \{0, 1\}$ be a binary treatment indicator and $\boldsymbol{L}_{k,i}$ be a vector of measured covariate values.
  - Overbars indicate history: $\bar{A}_{k,i} = \{A_{0,i}, A_{1,i}, \ldots, A_{k,i}\}$,
  - Particular range of history: $\bar{A}_{(k-t,k),i} = \{A_{k-t,i} \ldots, A_{k,i}\}$
  - Underlines indicate future: $\underline{A}_{k,i} = \{A_{k,i}, A_{k+1,i}, \ldots A_{K_i,i}\}$
- In contrast to $A_k$, define $Z_{k,t}$ for $t \geq 1$ such that

$$Z_{k,t} = \begin{cases} 1 & \text{if } \bar{A}_{k,t} = \{A_k, A_{k+1}, \ldots, A_{t-1}, A_t\} = \{1, 1, \ldots, 1, 1\} \\ 0 & \text{if } \bar{A}_{k,t} = \{A_k, A_{k+1}, \ldots, A_{t-1}, A_t\} = \{1, 0, \ldots, I(t-1 \text{ is even}), I(t \text{ is even})\} \end{cases},$$

- Denote adherence for the duration of the study $Z = Z_{0,K_i}$
- Let $\lfloor t \rfloor$ indicate the time of the most recent study visit $k$ at or prior to continuous time $t$, $\lfloor t \rfloor \in \{0, 1, \ldots K_i\}$ for individual $i$

# Emulating a Treatment Schedule TTE

- Common obstacles, but also...
  - Non-adherence at treatment *and* schedule level
  - Schedule indicator (common, CCW) functional form?

# Emulating a Treatment Schedule TTE

- Common obstacles, but also...
  - Non-adherence at treatment *and* schedule level
  - Schedule indicator (common, CCW) functional form?

- We consider two emulation approaches, similar to Keogh et al.:
  1. **CCW/Sequential Trials**: common to TTE literature
  2. **MSM-IPTW**: a conventional marginal structural model with inverse-probability-of-treatment weighting

- Different approaches will emulate certain elements differently
  - *Eligibility criteria:* multiple time-zeros?
  - *Treatment allocation:* how do we think about treatment?
  - *Follow-up:* censoring at non-adherence?

# Emulating a Treatment Schedule TTE

- Common obstacles, but also...
  - Non-adherence at treatment *and* schedule level
  - Schedule indicator (common, CCW) functional form?

- We consider two emulation approaches, similar to Keogh et al.:
  1. **CCW/Sequential Trials**: common to TTE literature
  2. **MSM-IPTW**: a conventional marginal structural model with inverse-probability-of-treatment weighting

- Different approaches will emulate certain elements differently
  - *Eligibility criteria:* multiple time-zeros?
  - *Treatment allocation:* how do we think about treatment?
  - *Follow-up:* censoring at non-adherence?

- Both approaches use **empirical standardization**. Specifically,

$$\hat{P}(T^{\underline{a}_0=\boldsymbol{r}} > \tau) = \frac{1}{n_0}\sum_{i\in C_0}\hat{P}(T^{\underline{a}_0=\boldsymbol{r}} > \tau|\bar{\boldsymbol{L}}_{0,i}) = \frac{1}{n_0}\sum_{i\in C_0}\exp\left(-\hat{H}_{T^{\underline{a}_0=\boldsymbol{r}}}(\tau|\bar{\boldsymbol{L}}_{0,i})\right). \quad (1)$$

  where $C_0$ denotes the set of $n_0$ eligible individuals at $k = 0$.

Keogh RH, et al. Stat Med. 2023. doi:10.1002/sim.9718

# Simulation Study Objectives

1. Evaluate the performance of different TTE approaches for comparing fixed schedules of the same repeated treatment

2. Clarify the pitfalls of common TTE analysis tools in this setting

# Data Generating Mechanism

We generate discrete visit information for a 10-year study period, with more granular information on event times (e.g., due to a registry) for $n = 1000$ according to Figure 2.

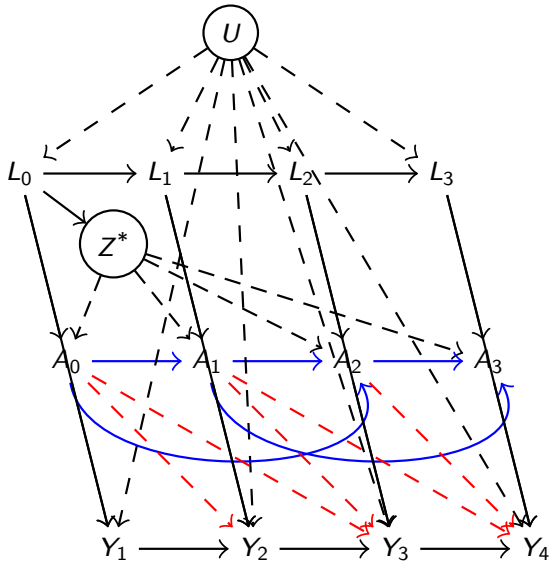



Figure 2: Directed acyclic graph (DAG) depicting relationships between measured covariate $L$, treatment $A$, discrete time outcome $Y$, unmeasured frailty $U$ and unmeasured, underlying screening pattern tendency $Z^*$. Red, dashed lines represent carryover effects under hazard model (2).

**Hazard Model 1.** Without Carryover Effect

$$h\left(t \mid A_{\lfloor t \rfloor}, L_{\lfloor t \rfloor}, U\right) = \beta_0 + \beta_A A_{\lfloor t \rfloor} + \beta_L L_{\lfloor t \rfloor} + \beta_U U \tag{2}$$

**Hazard Model 2.** Carryover Effect

$$h\left(t \mid A_{\lfloor t \rfloor}, \bar{A}_{(\lfloor t \rfloor - 2, \lfloor t \rfloor - 1)}, L_{\lfloor t \rfloor}, U\right) = \beta_0 + \beta_A A_{\lfloor t \rfloor} + \beta_{A1} A_{\lfloor t \rfloor - 1} + \beta_{A2} A_{\lfloor t \rfloor - 2} + \beta_L L_{\lfloor t \rfloor} + \beta_U U \tag{3}$$

# Analytical Approaches

## MSM-IPTW

- Stabilized IPTW include $L_0$ and are consistent with Figure 2:

$$SW_i(t) = \prod_{k=1}^{\lfloor t \rfloor} \frac{P\left(A_k = a_{k,i} \mid L_{0,i}, \bar{A}_{(k-2,k-1),i}, T_i \geq k\right)}{P\left(A_k = a_{k,i} \mid L_{k,i}, \bar{A}_{(k-2,k-1),i}, T_i \geq k\right)}. \tag{4}$$

- Fitted MSM must also include $L_0$:

$$h\left(t \mid A_0 = 1, \bar{A}_{\lfloor t \rfloor - 1}, \bar{L}_{\lfloor t \rfloor}\right) = \beta_0(t) + I(t \geq 1) \sum_{p=1}^{\lfloor t \rfloor} \beta_{Ap}(t) a_{\lfloor t \rfloor - p} + \beta_L(t) l_0. \tag{5}$$

# Analytical Approaches

## CCW/Sequential Trials

- Stabilized IPCW depend on treatment history at trial baseline:

$$SW^{z}_{i,k}(t-k) = \prod_{l=k+1}^{k+\lfloor t \rfloor} \frac{P\left(Z_l = z \mid L_{k,i}, \bar{Z}_{(k,l-1),i} = z, \bar{A}_{(k-2,k-1),i}, T_i \geq k\right)}{P\left(Z_l = z | L_{l,i}, \bar{Z}_{(k,l-1),i} = z, \bar{A}_{(k-2,k-1),i}, T_i \geq k\right)}, \quad z \in \{0,1\}. \tag{6}$$

- Naïvely-fitted MSMs for Trials 1 and 4, respectively, are:

$$h_k(t|A_k = 1, \bar{L}_k, Z) = \beta_0(t) + \beta_Z(t)z + \beta_L(t)l_k, \quad k = 0, \ldots, 9. \tag{7}$$

$$h\left(t \mid A_0 = 1, \bar{A}_{-1} = 0, \bar{L}_0, Z\right) = \beta_0(t) + \beta_Z(t)z + \beta_L(t)l_0. \tag{8}$$

# Adjusting for Pre-Trial Treatment History

Sequential trials *changes the distribution of treatment histories at baseline* when pooling the trials together, differs from $k = 0$

- **With a Carryover Effect**, $\beta_Z(t)$ will be biased under Trials 1 and 2
- **Without a Carryover Effect**, we **still have to account** for the change, think about the target population

  *Sequential Trials, Trial 1:*

$$\hat{P}(T^{\underline{a}_0=\boldsymbol{r}} > \tau | A_0 = 1) = \frac{1}{n_0} \sum_{i \in C_0} \hat{P}(T^{\underline{a}_k=\boldsymbol{r}} > \tau | L_k = l_i, A_k = 1)$$

  *No Sequential Trials:*

$$\hat{P}(T^{\underline{a}_0=\boldsymbol{r}} > \tau | A_0 = 1, \bar{A}_{-1} = 0) = \frac{1}{n_0} \sum_{i \in C_0} \hat{P}(T^{\underline{a}_0=\boldsymbol{r}} > \tau | L_{0,i}, A_0 = 1, \bar{A}_{-1} = 0)$$

An adjusted model for Trial 1 is:

$$h_k\left(t \mid A_k = 1, \bar{A}_{(k-2,k-1)}, \bar{L}_k, Z\right) = \beta_0(t) + \beta_Z(t)z + \beta_{A1}(t)a_{k-1} + \beta_{A2}(t)a_{k-2} + \beta_L(t)l_k. \quad (9)$$

# Results

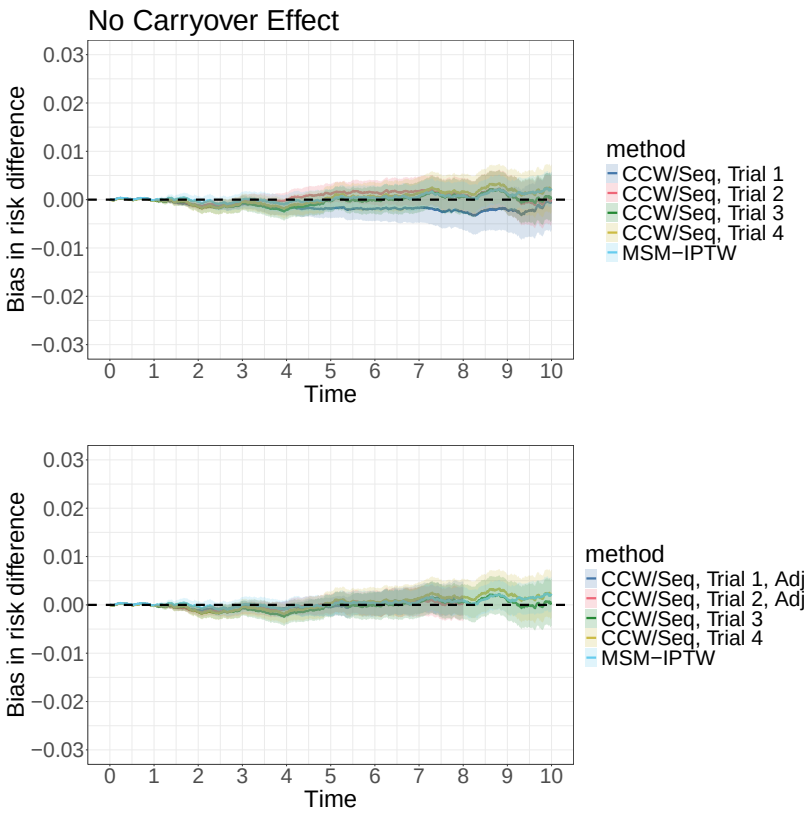



Figure 3: Bias of different emulation approaches under Hazard Model 1.

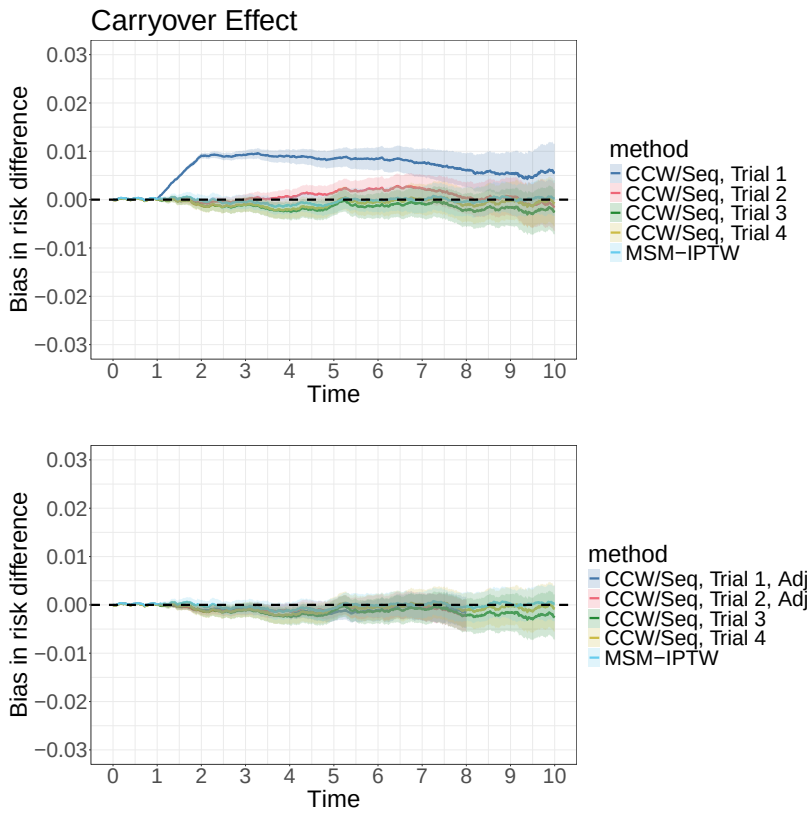



Figure 4: Bias of different emulation approaches under Hazard Model 2.

Bias for the adjusted models is truncated at time 8, as effects are inestimable afterwards.
Results are averaged over 1000 simulated datasets.

# Results

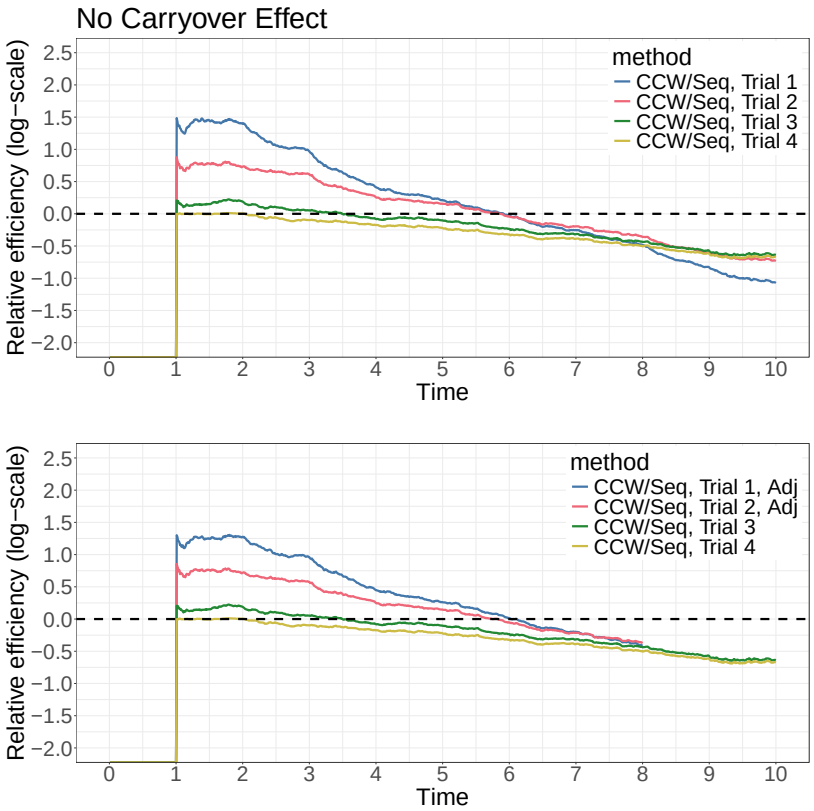



Figure 5: Relative efficiency compared with MSM-IPTW under Hazard Model 1.

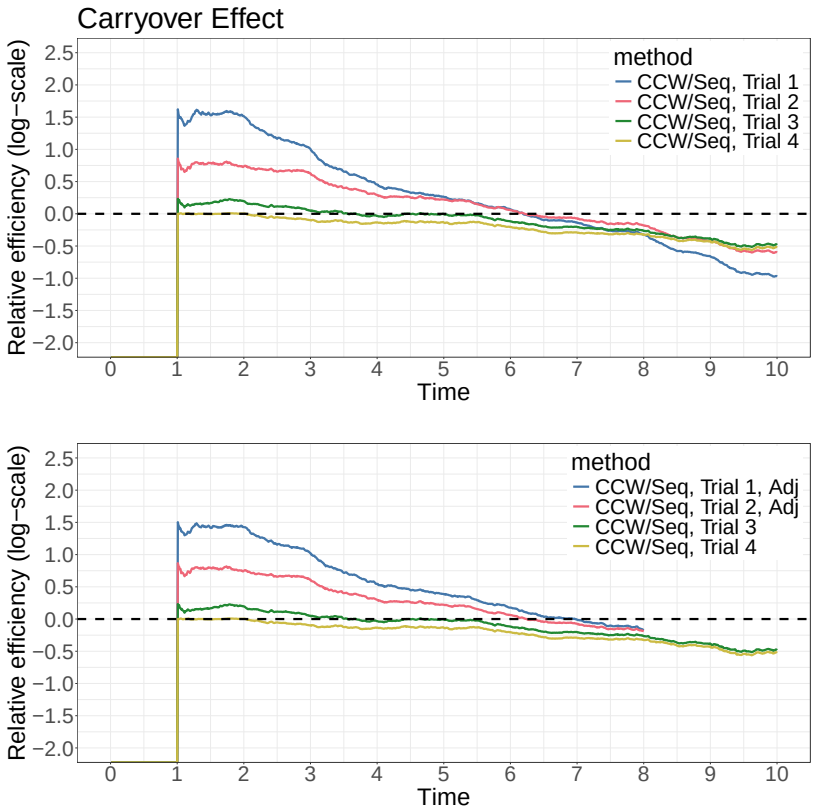



Figure 6: Relative efficiency compared with MSM-IPTW under Hazard Model 2.

Efficiency for the adjusted models is truncated at time 8, as effects are inestimable afterwards.
Results are averaged over 1000 simulated datasets.

# Conclusions

## CCW

1. Relative efficiency favors MSM-IPTW late in study due to censoring at non-adherence
2. Requires correct functional form for *treatment schedule indicator*

## Sequential Trials

1. Estimates can differ between studies with one time-zero and studies that allow multiple time-zeros
2. Including pre-trial treatment history in outcome models permits standardization, mitigates differences and biases
3. Short-term efficiency gains compared to MSM-IPTW

# Discussion

- TTE is not a statistical method! Different approaches exist and may be preferred in different settings
- Assumptions structured differently for the two analytical approaches
- **Limitations**
  - Simulation assumes data are collected at regular, annual study visits
  - Assume covariate data are collected regardless of treatment status
- **Future directions**
  - Incorporate irregular and/or informative visit processes
  - Incremental causal effects (Kennedy, 2019; Ying et al., 2025)
  - Real data examples

Danaei et al. Stat Methods Med Res. 2013. doi:10.1177/0962280211403603
Kennedy EH. JASA. 2019. doi: 10.1080/01621459.2017.1422737.
Ying A, et al. arXiv. 2025. 10.48550/arXiv.2409.13097

# Thank you! Questions?

meghan.gerety@pennmedicine.upenn.edu

Advised by Nicholas J. Seewald, PhD

# The Causal Estimand

- Counterfactual event time under regime $\boldsymbol{z} = (z_0, \ldots, z_{K_i})$: $T^{\underline{a}_0 = \boldsymbol{z}}$.
- We focus on a per-protocol **marginal risk difference**:

$$P(T^{\underline{a}_0=\{1,0,1,0\ldots\}} \leq \tau) - P(T^{\underline{a}_0=\mathbf{1}} \leq \tau) = P(T^{\underline{a}_0=\mathbf{1}} > \tau) - P(T^{\underline{a}_0=\{1,0,1,0\ldots\}} > \tau) \quad (10)$$

$$= P(T^{Z=1} > \tau) - P(T^{Z=0} > \tau), \quad (11)$$

- **Empirical standardization** targets a **specific population** defined by $\bar{\boldsymbol{L}}_0$. Let $C_0$ denote the set of $n_0$ eligibile individuals at $k = 0$.

$$\hat{P}(T^{\underline{a}_0=\boldsymbol{r}} > \tau) = \frac{1}{n_0} \sum_{i \in C_0} \hat{P}(T^{\underline{a}_0=\boldsymbol{r}} > \tau | \bar{\boldsymbol{L}}_{0,i}) = \frac{1}{n_0} \sum_{i \in C_0} \exp\left(-\hat{H}_{T^{\underline{a}_0=\boldsymbol{r}}}(\tau | \bar{\boldsymbol{L}}_{0,i})\right). \quad (12)$$

*Assumptions:* no interference, positivity, consistency, conditional exchangeability of treatment statuses at each visit, uninformative censoring conditional on observed data.

# Positivity

## MSM-IPTW

Stabilized IPTW:

$$SW_i(t) = \prod_{k=1}^{\lfloor t \rfloor} \frac{P\left(A_k = a_{k,i} \mid L_{0,i}, \bar{A}_{(k-2,k-1),i}, T_i \geq k\right)}{P\left(A_k = a_{k,i} \mid L_{k,i}, \bar{A}_{(k-2,k-1),i}, T_i \geq k\right)}.$$

Corresponding positivity assumption:

$$P\left(A_{k,i} = a \mid L_{k,i}, \bar{A}_{(k-1,k-2),i}, T_i \geq k\right) > 0 \quad \text{for } a \in \{0,1\}$$

## CCW/Sequential Trials

Stabilized IPCW:

$$SW^z_{i,k}(t-k) = \prod_{l=k+1}^{k+\lfloor t \rfloor} \frac{P\left(Z_l = z \mid L_{k,i}, \bar{Z}_{(k,l-1),i} = z, \bar{A}_{(k-2,k-1),i}, T_i \geq k\right)}{P\left(Z_l = z | L_{l,i}, \bar{Z}_{(k,l-1),i} = z, \bar{A}_{(k-2,k-1),i}, T_i \geq k\right)}, \quad z \in \{0,1\}.$$

Corresponding positivity assumption:

$$P\left(Z_l = z | L_{l,i}, \bar{Z}_{(k,l-1),i} = z, \bar{A}_{(k-2,k-1),i}, T_i \geq k\right) > 0 \quad \text{for } z \in \{0,1\}$$

These assumptions must be in slightly different analytic samples.

# Data Generating Mechanism

Generate time-fixed $U$ and $Z^*$ and data for the baseline visit $j = 0$:

$$\begin{aligned} U &\sim N(0,1) \\ L_0 \mid U &\sim N(U,1) \\ Z^* \mid L_0 &\sim Bern\left(expit(0 + 0.5L_0)\right) \\ A_0 \mid L_0, Z^* &\sim Bern\left(expit(0 + 0.3L_0 + 0.4Z^*)\right), \end{aligned}$$

Then, for the remaining visits $j = 1, 2, \ldots, 9$,

$$L_j \mid L_{j-1}, A_{j-1}, j, U \sim N\left(0 + 0.8L_{j-1} - A_{j-1} + 0.1j + U, 1\right),$$

$$A_j | A_{j-1}, A_{j-2}, L_j, Z^* \sim Bern\Big(expit\big(0 - 2A_{k-1} + 1.5A_{j-2} + 2.75A_{j-1}Z^* - A_{j-2}Z^* + 0.3L_j + 0.4Z^*\big)\Big).$$

The hazard is then generated from a conditional additive hazards model, dependent on simulation setting.

# Data Generating Mechanism

Table 2: Hazard Model Parameters by Simulation Setting

| | Primary Simulation | Decaying Hazards | Low Hazards Simulation |
|---|---|---|---|
| Parameters in Both Models | $\beta_L = 0.005625$<br>$\beta_U = 0.005625$ | $\beta_L = 0.005625$<br>$\beta_U = 0.005625$ | $\beta_L = 0.001875$<br>$\beta_U = 0.001875$ |
| Model Without Carryover Effect | $\beta_0 = 0.12$<br>$\beta_A = -0.03$ | $\beta_0 = 0.12$<br>$\beta_A = -0.0375$ | $\beta_0 = 0.04$<br>$\beta_A = -0.01$ |
| Model With Carryover Effect | $\beta_0 = 0.15$<br>$\beta_A = -0.03$<br>$\beta_{A1} = -0.03$<br>$\beta_{A2} = -0.03$ | $\beta_0 = 0.015$<br>$\beta_A = -0.045$<br>$\beta_{A1} = \beta_A(1 - \frac{11}{36})$<br>$\beta_{A2} = \beta_{A1}(1 - \frac{11}{36})$ | $\beta_0 = 0.05$<br>$\beta_A = -0.01$<br>$\beta_{A1} = -0.01$<br>$\beta_{A2} = -0.01$ |

# Data Generating Mechanism

The generated population of data, indexed by $j = 0, \ldots, 9$, is re-aligned such that an individual's first treatment is at $k = 0$; can be at different calendar times for different $i$.

| | id | A | L | L.baseline | time | time.stop | event |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 0 | -0.6553571 | -0.6553571 | 0 | 1 | 0 |
| 2 | 2 | 0 | 0.1494684 | -0.6553571 | 1 | 2 | 0 |
| 3 | 2 | 0 | 2.2511630 | -0.6553571 | 2 | 3 | 0 |
| 4 | 2 | 1 | 0.5501216 | -0.6553571 | 3 | 4 | 0 |
| 5 | 2 | 0 | 0.2920004 | -0.6553571 | 4 | 5 | 0 |
| 6 | 2 | 1 | -1.0730618 | -0.6553571 | 5 | 6 | 0 |
| 7 | 2 | 0 | -2.9149301 | -0.6553571 | 6 | 7 | 0 |
| 8 | 2 | 1 | -1.4522084 | -0.6553571 | 7 | 8 | 0 |
| 9 | 2 | 0 | -2.6129599 | -0.6553571 | 8 | 9 | 0 |
| 10 | 2 | 1 | -1.1481563 | -0.6553571 | 9 | 10 | 0 |

Set Study Baseline at $j$=3

| | id | A | L | L.baseline | time | time.stop | event |
|---|---|---|---|---|---|---|---|
| 1 | 2 | 1 | 0.5501216 | 0.5501216 | 0 | 1 | 0 |
| 2 | 2 | 0 | 0.2920004 | 0.5501216 | 1 | 2 | 0 |
| 3 | 2 | 1 | -1.0730618 | 0.5501216 | 2 | 3 | 0 |
| 4 | 2 | 0 | -2.9149301 | 0.5501216 | 3 | 4 | 0 |
| 5 | 2 | 1 | -1.4522084 | 0.5501216 | 4 | 5 | 0 |
| 6 | 2 | 0 | -2.6129599 | 0.5501216 | 5 | 6 | 0 |
| 7 | 2 | 1 | -1.1481563 | 0.5501216 | 6 | 7 | 0 |

Figure 7: Realignment of generated data for analysis.

# Data Generating Mechanism: Screening Patterns

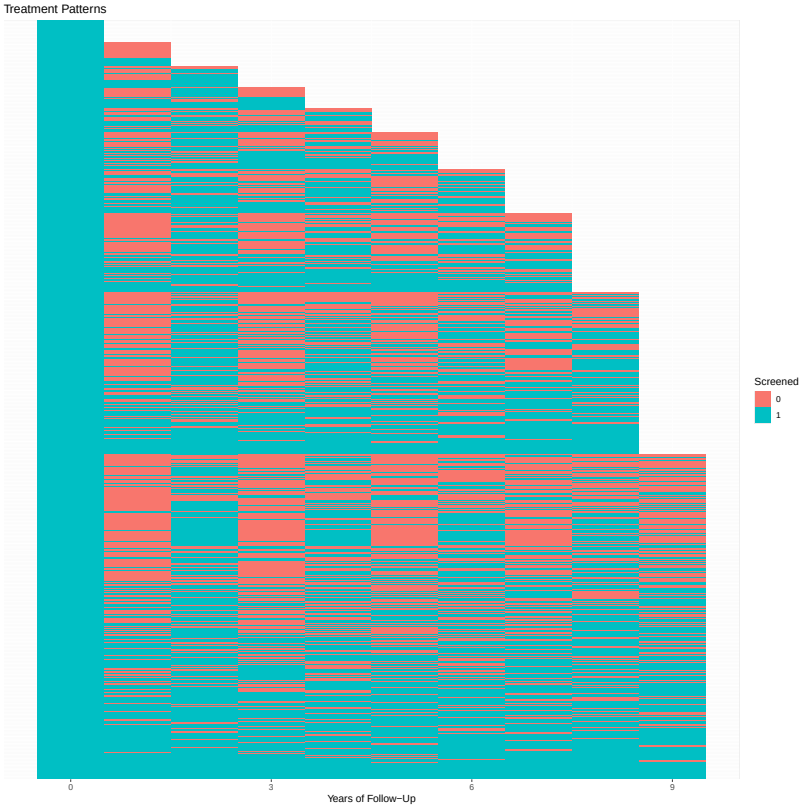



Figure 8: Heatmap of screening patterns for one randomly selected iteration, under Hazard Model 1.

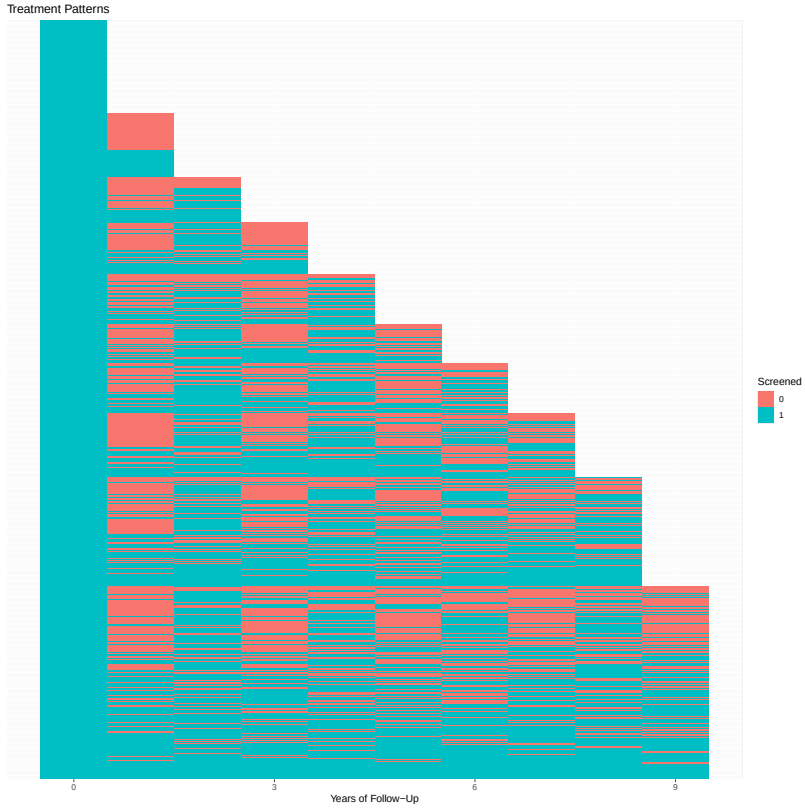



Figure 9: Heatmap of screening patterns for one randomly selected iteration, under Hazard Model 2.

# Data Generating Mechanism: Screening Patterns

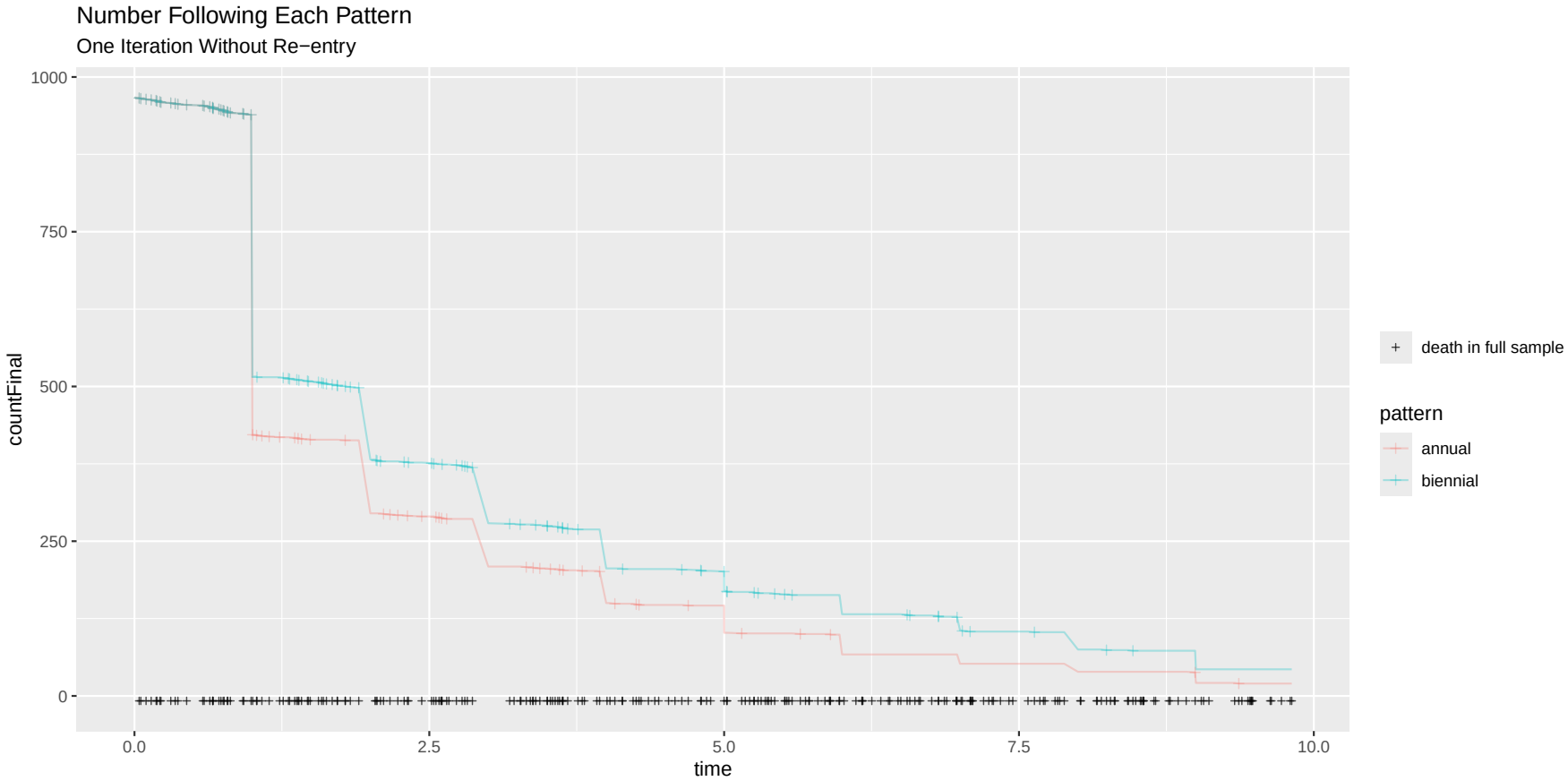



Figure 10: Rugplot of annual and biennial arms for one randomly selected iteration, under Hazard Model 1.

# Data Generating Mechanism: Screening Patterns

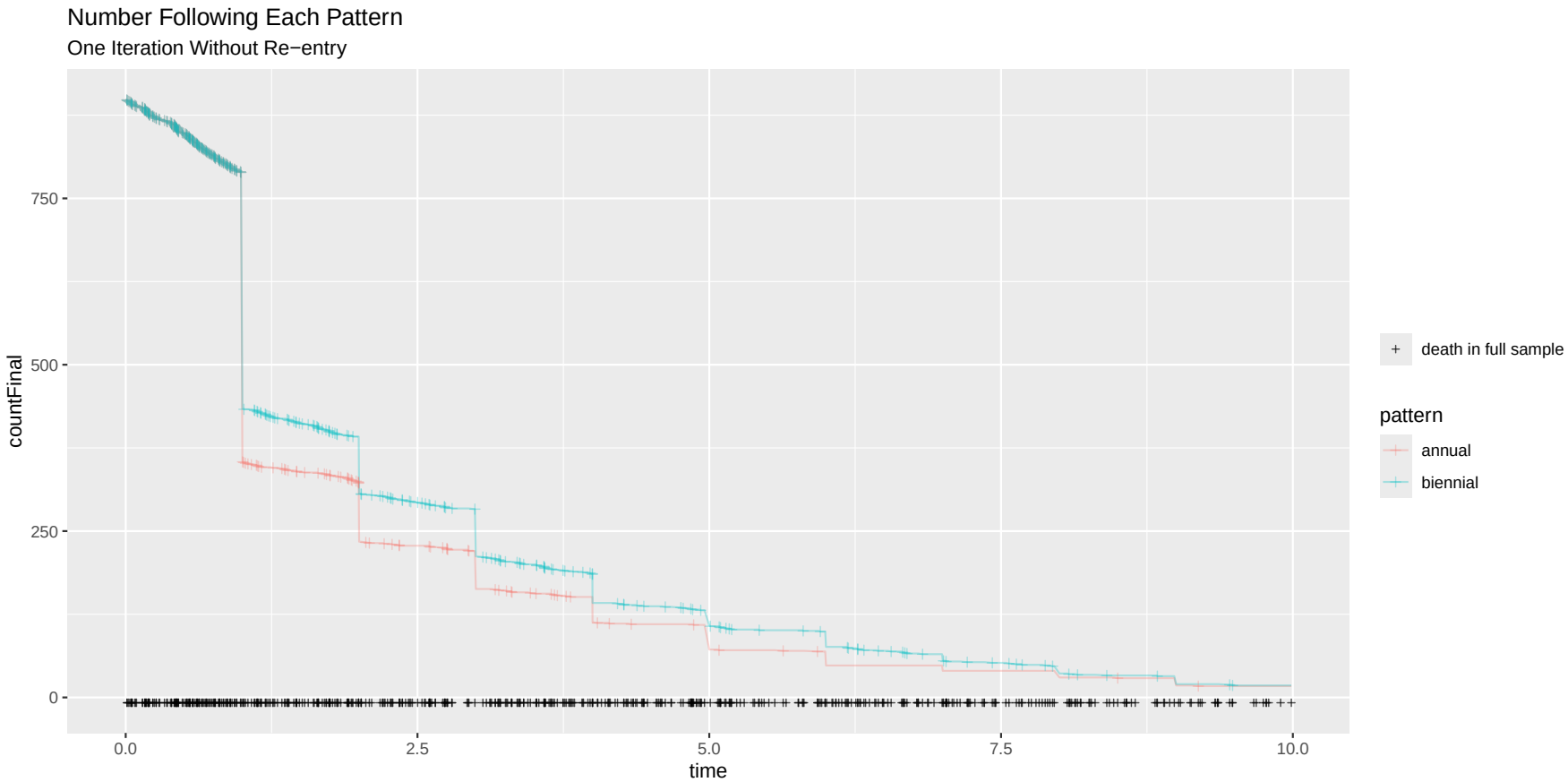



Figure 11: Rugplot of annual and biennial arms for one randomly selected iteration, under Hazard Model 2.

# Additional Results, Original Settings

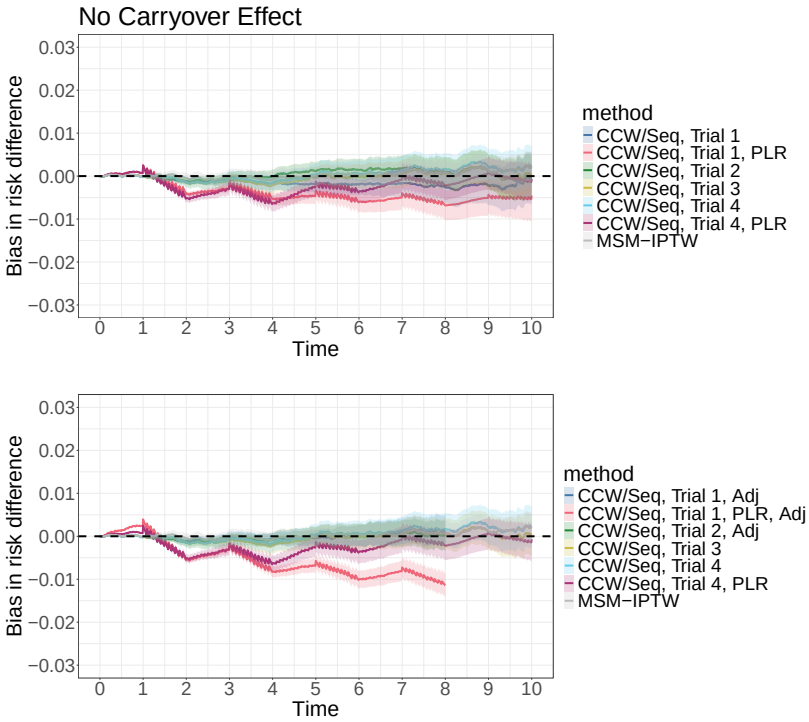



Figure 12: Bias of different emulation approaches, including pooled linear models, under Hazard Model 1.

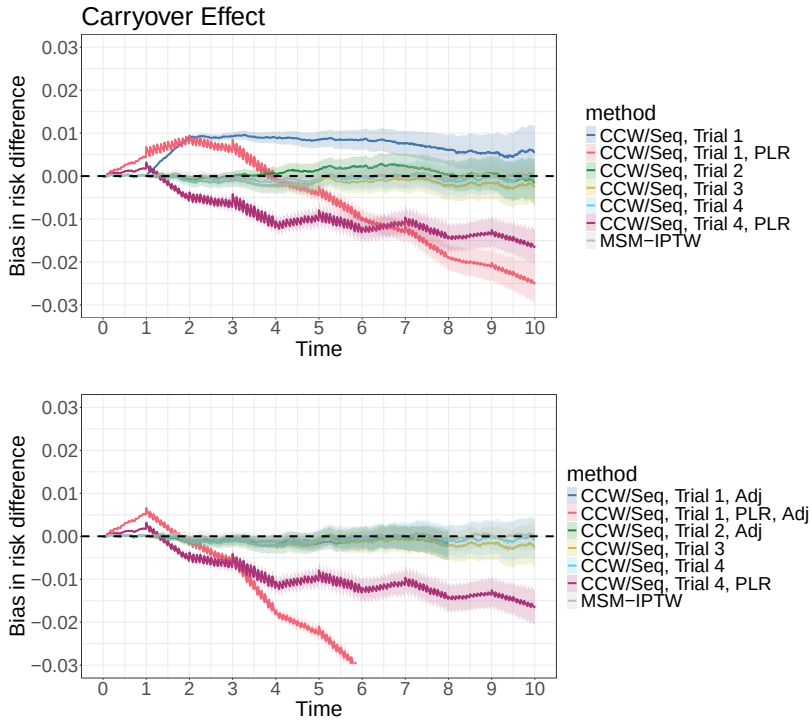



Figure 13: Bias of different emulation approaches, including pooled linear models, under Hazard Model 2.

# Additional Results, Original Settings

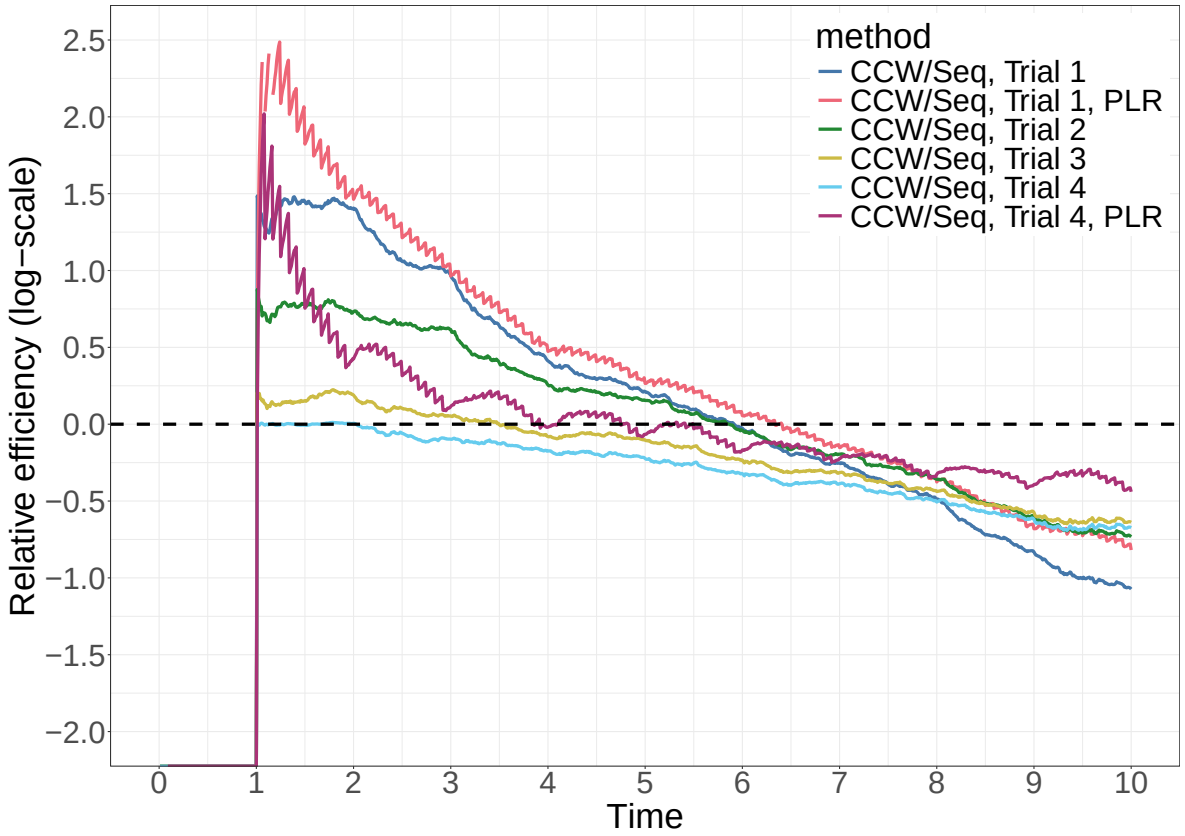



Figure 14: Relative efficiency of analytical approaches, including pooled linear models, compared with MSM-IPTW under Hazard Model 1.

# Additional Results, Decaying Hazards

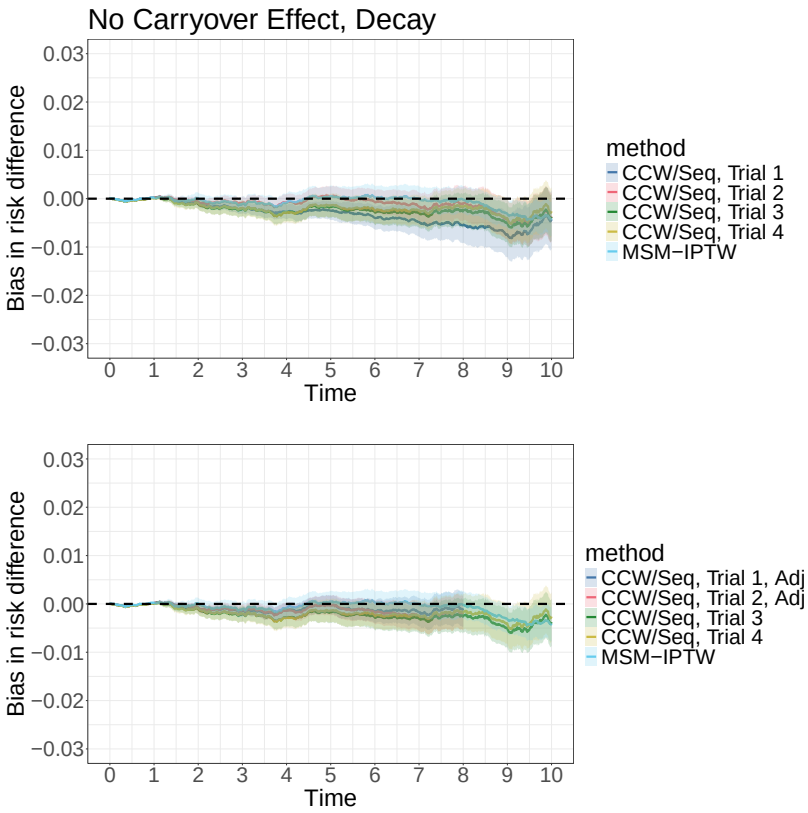



Figure 15: Bias of different emulation approaches under Hazard Model 3, where the treatment effect decays over time.

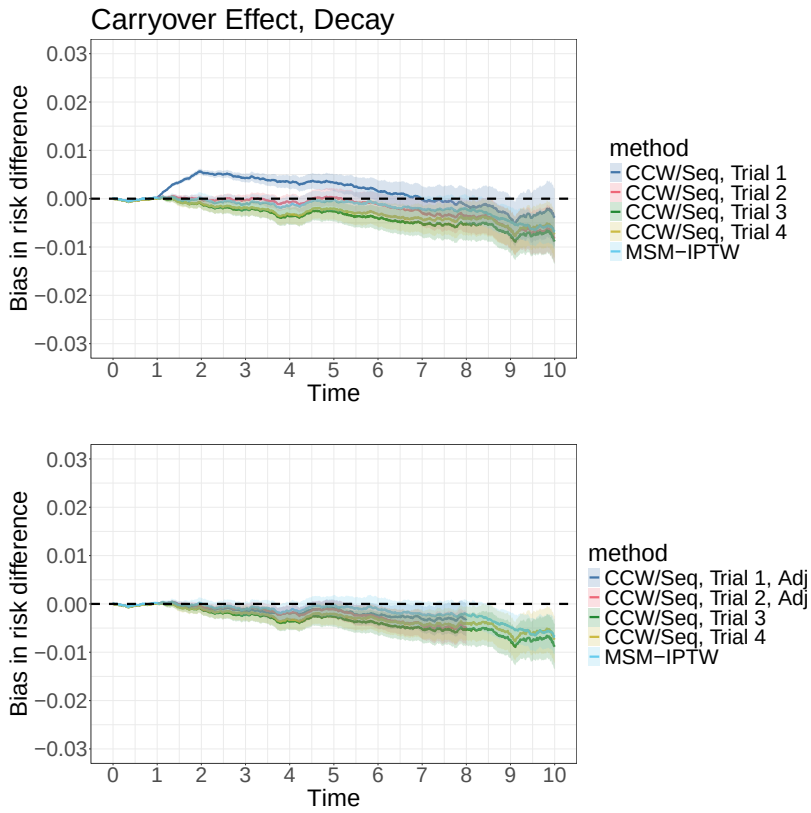



Figure 16: Bias of different emulation approaches under Hazard Model 4, where the treatment effect decays over time.

# Simulation: Varied Treatment Histories at Baseline

- **Objective**: verify that the comparability of MSM-IPTW and CCW (when eligibility is not re-assessed) is not specific to the population at $k = 0$ being treatment-naive.
- Let some individuals be treatment-naive at time-zero, while letting others be treated three years ago
- To avoid differential immortal time bias: generate three pre-baseline visits for all individuals, require no event in this period for eligibility

- Estimates under both analytical approaches will be marginal to the population at $k = 0$, regardless of the distribution of treatment history in that baseline population
- When eligibility after previous treatment is permitted and a sequence of trials is emulated, the distribution of treatment histories pooled over all trial baselines will still look different

# Simulation: Varied Treatment Histories at Baseline

$$U \sim N(0,1)$$
$$L_0|U \sim N(U,1)$$
$$Z^*|L_0 \sim Bern\Big(expit(\tau_0 + \tau_L L_0)\Big)$$
$$A_0 \sim Bern(0.2). \text{ Set } A_1, A_2 = 0.$$

For $j = 1, 2$ (pre-baseline visits):

- $L_j|L_{j-1}, A_{j-1}, j, U \sim N\Big(\delta_0 + \delta_L L_{j-1} + \delta_A A_{j-1} + \delta_T j + U, 1\Big)$

For $j = 3, 4, ...12$ (study visits):

- $L_j|L_{j-1}, A_{j-1}, j, U \sim N\Big(\delta_0 + \delta_L L_{j-1} + \delta_A A_{j-1} + \delta_T j + U, 1\Big)$
- $A_j|A_{j-1}, A_{j-2}, L_j, Z^* \sim$
  $Bern\Big(expit(\beta_0 + \beta_1 A_{k-1} + \beta_2 A_{j-2} + \beta_3 A_{j-1} Z^* + \beta_4 A_{j-2} Z^* + \beta_L L_j + \beta_Z Z^*)\Big)$

# Additional Results, Varied Treatment Histories at Baseline

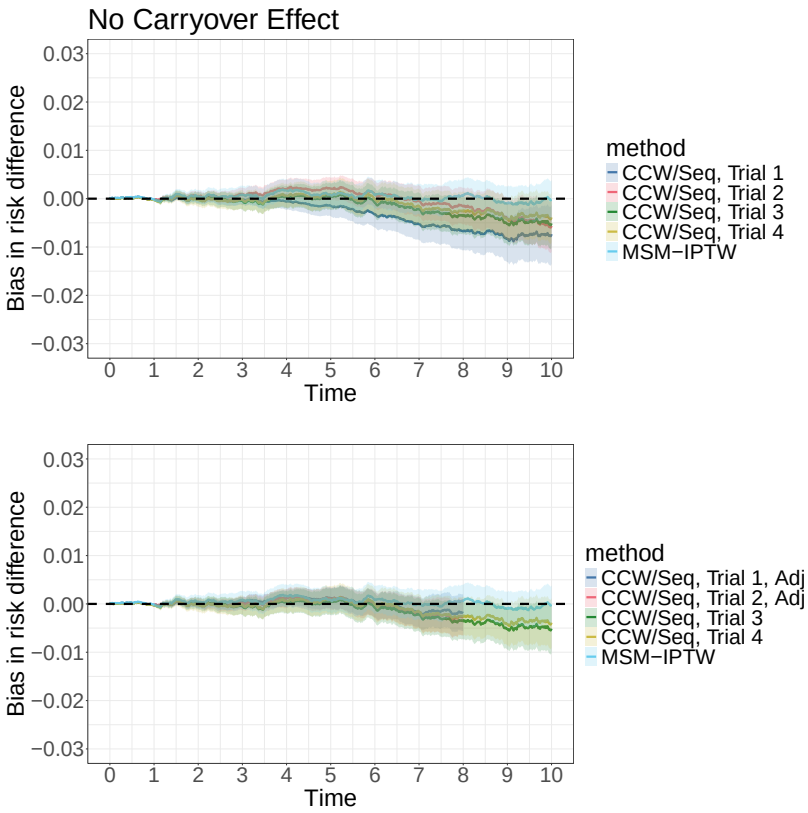



Figure 17: Bias of different emulation approaches under Hazard Model 1 when there are variable treatment histories at baseline.

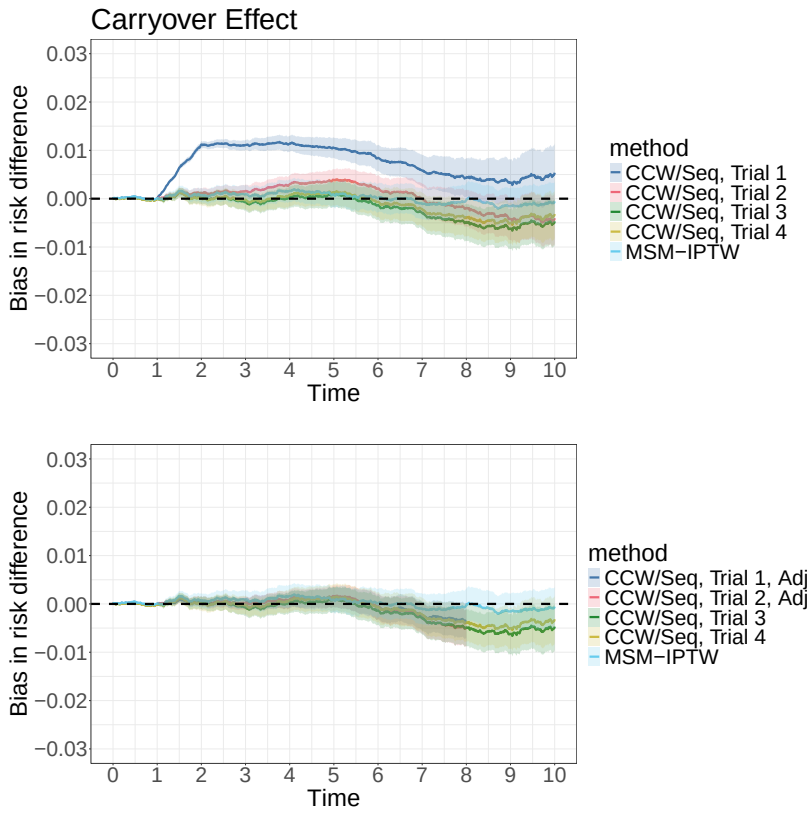



Figure 18: Bias of different emulation approaches under Hazard Model 2 when there are variable treatment histories at baseline.

# Additional Results, Lower Hazards

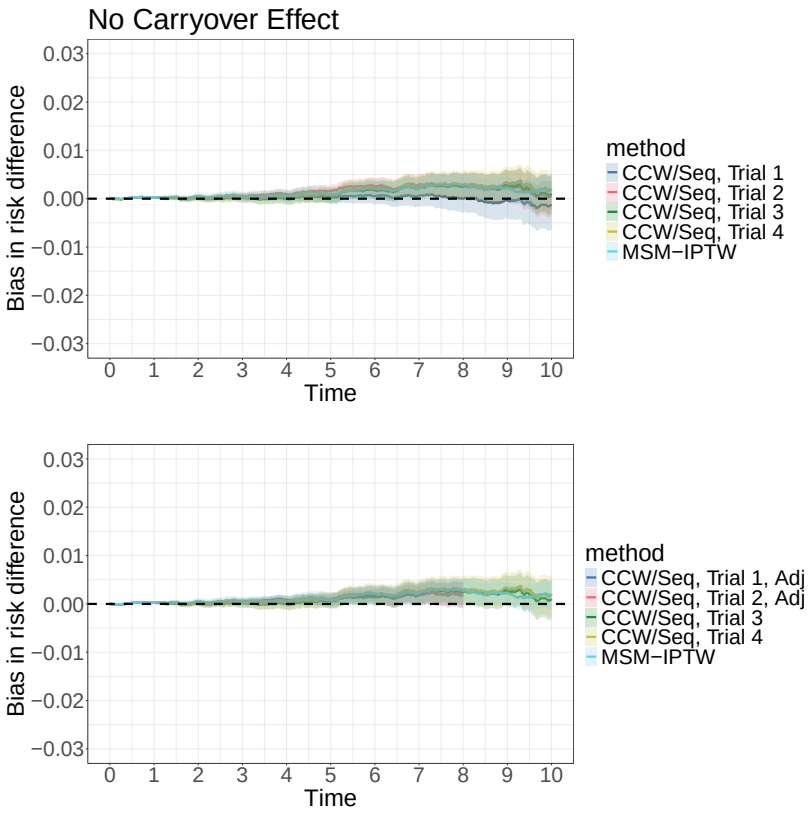



Figure 19: Bias of different emulation approaches under Hazard Model 1 with lower hazards.

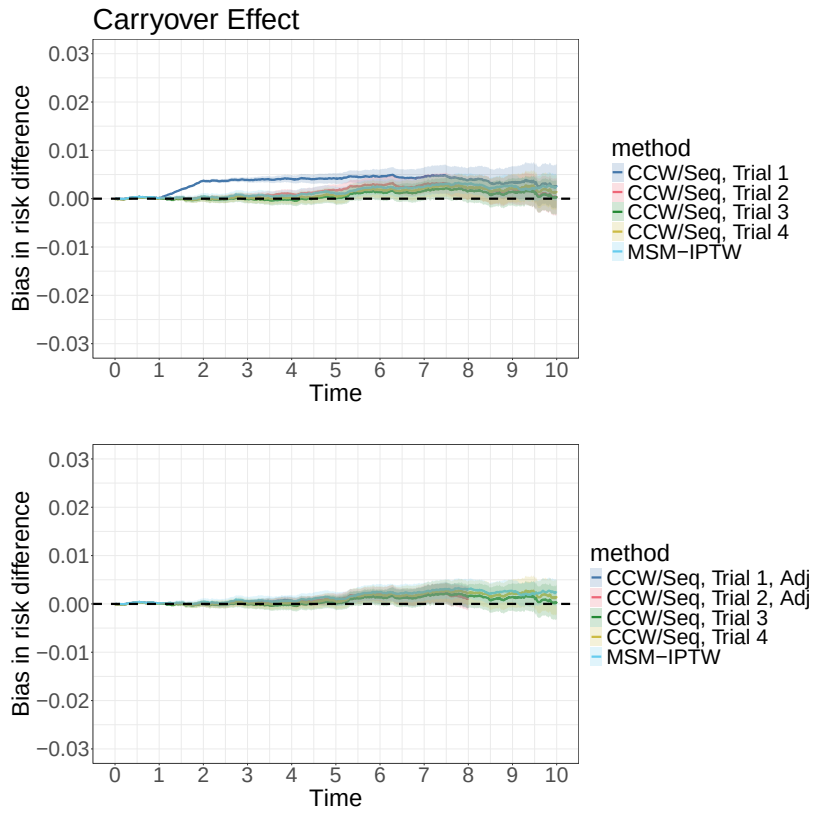



Figure 20: Bias of different emulation approaches under Hazard Model 2 with lower hazards.

# Connecting Model Forms: An Aside

- These models are **parametrized differently** and **make use of the data differently**, but still recover the same estimand
- For eq. (8): $h\left(t \mid A_0 = 1, \bar{A}_{-1} = 0, \bar{L}_0, Z\right) = \beta_0(t) + \beta_Z(t)z + \beta_L(t)L_0$, we estimate $B_V(t) = \int_0^t \beta_V(u)du$ for $V \in \{0, Z, L\}$.
- Define design matrix $\mathbf{X}(t)$ where $\boldsymbol{X}_i(t) = Y^R_{t,i}(1, Z_i, L_{0,i})$ for at-risk process $Y^R_{t,i} = I(T_i \geq t)$.
- Let $dN_i(t) = N_i(t^- + dt) - N_i(t^-)$ where $t^-$ is the time instant immediately before time $t$.
- Let the vector $dN(t)$ have element $i$ equal to $dN_i(t)$. The $ith$ element of $dN(t)$ is 1 if subject $i$ dies at time $t$.
- We estimate the vector $\boldsymbol{B}(t)$ with

$$\hat{\mathbf{B}}(t) = \sum_{T_i \leq t} \left[\mathbf{X}'(T_i)\mathbf{X}(T_i)\right]^{-1} \mathbf{X}'(T_i)dN(T_i) \tag{13}$$

# Connecting Model Forms: An Aside

- For MSM-IPTW, for times $t \in [2, 3)$, the design matrix is given by:

$$\boldsymbol{X}_i(t) = Y^R_{t,i}\left(1, A_{1,i}(t), A_{2,i}(t), L_{0,i}\right)$$

Consider the following steps:

- *Re-parameterize* the MSM-IPTW model based on regime indicators, for example for times $t \in [2, 3)$:

$$\begin{aligned} h\left(t \mid A_0 = 1, \bar{A}_{\lfloor t \rfloor - 1}, \bar{L}_{\lfloor t \rfloor}\right) = \gamma_0(t) I(A_0 = 1, A_1 = 0, A_2 = 1) + \gamma_1(t) I(A_0 = 1, A_1 = 1, A_2 = 0) + \\ \gamma_2(t) I(A_0 = 1, A_1 = 0, A_2 = 0) + \gamma_3(t) I(A_0 = 1, A_1 = 1, A_2 = 1) + \gamma_L(t) L_0. \end{aligned} \tag{14}$$

This results in design matrix:

$$\begin{aligned} \boldsymbol{X}_i(t) = Y^R_{t,i}\big(I(A_{0,i} = 1, A_{1,i} = 0, A_{2,i} = 1), I(A_{0,i} = 1, A_{1,i} = 1, A_{2,i} = 0), \\ I(A_{0,i} = 1, A_{1,i} = 0, A_{2,i} = 0), I(A_{0,i} = 1, A_{1,i} = 1, A_{2,i} = 1), L_{0,i}\big) \end{aligned} \tag{15}$$

- *Subset* the data to those consistent with the annual or biennial regime through time $t$:

$$\begin{aligned} \boldsymbol{X}_i(t) &= Y^R_{t,i}\left(I(A_{0,i} = 1, A_{1,i} = 0, A_{2,i} = 1), I(A_{0,i} = 1, A_{1,i} = 1, A_{2,i} = 1), L_{0,i}\right) \\ &= Y^R_{t,i}\left(I(Z_i = 0), I(Z_i = 1), L_{0,i}\right) \Rightarrow Y^R_{t,i}\left(1, Z_i, L_{0,i}\right) \end{aligned}$$

- The role of baseline covariate distributions and collapsibility