Application of Time-to-Event Methods in the Assessment of Safety in Clinical Trials

by KL Moore and MJ van der Laan

Final Presentation

William Jen Hoe Koh

June 4, 2013

Overview

- Introduction to survival analysis
- Motivations
  1. Scientific objectives
  2. Statistical objectives
- Targeted Maximum Likelihood Estimation (TMLE)
- Simulation Results
- Discussion

Survival Analysis

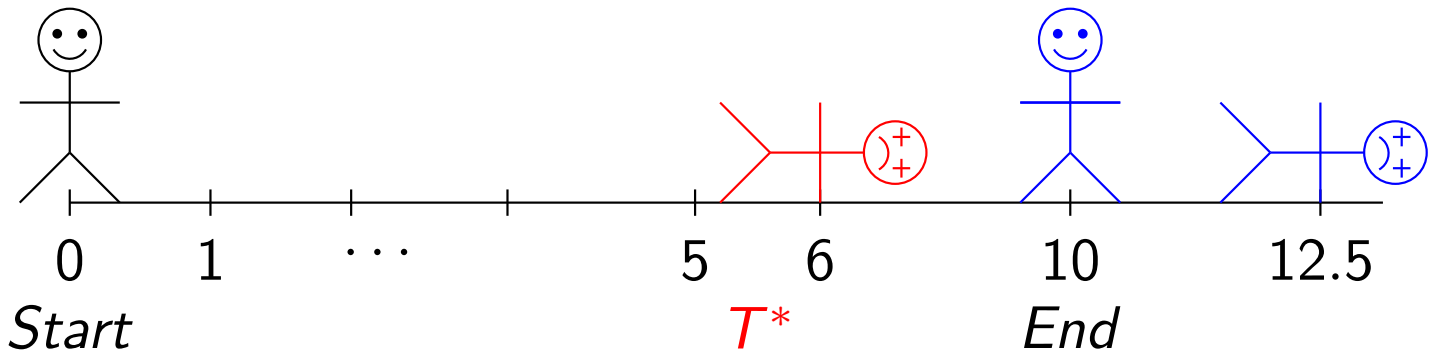
$$\lambda(t) = \lim_{\Delta t \to 0} \frac{\Pr[t \leq T < t + \Delta t | T \geq t]}{\Delta t}$$

$$S(t) = Pr(T > t) = 1 - F(t)$$

Event of Interest: Infection/AE at clinic visit

Right censored data: Non-ignorable missing data.

Discrete Failure Time



$$\lambda(t_j) = \Pr(T = t_j | T > t_{j-1})$$

T^*: "True" failure time (Unobserved due to discrete follow-up)

$T = t_j$ if $T^* \in [t_{j-1}, t_j)$ with t_j for $j = 1, \cdots, 10$.

$\widetilde{T} = min(T, C)$: where C is our censoring time.

$\Delta = I(T \leq C)$: Indicator of subject **not** being censored

Scientific Motivation: RCT

Scientific question: $\Pr(T > t_k|A = 1) - \Pr(T > t_k|A = 0)$

Simplest analysis: Kaplan Meier Survival curves

What if we have a set of potential baseline covariates W that predicts outcome?

Adjusted analysis: Cox-PH or Logistic regression (More assumptions)

**A is the treatment indicator, W is our baseline covariate.

Statistical Motivation: Marginal vs Conditional

Scientific question: $\Pr(T > t_k|A = 1) - \Pr(T > t_k|A = 0)$

Marginal: $P(T > t|A = a) = S_0(t|A = a)$

Probability of survival past time t for treatment a for the entire population.

Conditional: $P(T > t|A = a, W) = S_0(t|A = a, W)$

Probability of survival past time t for treatment a while holding W fixed.

Potential problem: Adjusting for a covariate in survival analysis do not provide a marginal interpretation even though there is gain in efficiency (Hernández et al., 2006)

Paper's Objectives: Marginal vs Conditional

Scientific question: $\Pr(T > t_k | A = 1) - \Pr(T > t_k | A = 0)$

1. Use of baseline covariate adjustment to estimate marginal treatment specific survival at a fixed time point.

2. Exploit baseline covariates to gain efficiency.

3. Provide a consistent estimator under certain censoring mechanisms.
   - Random censoring (missing completely at random)
   - Baseline covariates that predicts censoring outcomes (missing at random)

Proposed estimator: Targeted MLE

$$\Psi_a^{TMLE}(p_0)(t_k) = \overbrace{\Pr(T_a > t_k)}^{Marginal} = \mathbb{E}_W[\overbrace{S_0(t_k|A=a,\mathbf{W})}^{Conditional}]$$

$$\Psi_{AD}^{TMLE}(p_0)(t_k) = \Psi_1^{TMLE}(p_0)(t_k) - \Psi_0^{TMLE}(p_0)(t_k)$$

where p_0 is the density distribution of $O = (W, A, \widetilde{T}, \Delta)$

- $\Psi_a^{TMLE}(p_0)(t_k)$: Marginal parameter
- Marginal is obtained by averaging the conditional survival over observed baseline covariates W
- Uses the influence curve (IC) which has mean 0 at the true parameter value. Can be used as an estimating equation.

Causal framework

You take the red pill - you stay in Wonderland and I show you how deep the rabbit hole goes.

You take the blue pill - the story ends, you wake up in your bed and believe whatever you want to believe.

Morpheus, The Matrix

- T^r: potential survival outcome of taking the red pill
- T^b: potential survival outcome of taking the blue pill

Causal framework for TMLE

In reality (movie), only 1 outcome (story) plays out. The counterfactual definition is useful for applying it to the entire population. We treat the counterfactual component as missing.

The average treatment effect of taking the red pill instead of the blue is

$$\begin{aligned}
&\widehat{\Psi}_r^{TMLE}(p_0)(t_k) - \widehat{\Psi}_b^{TMLE}(p_0)(t_k) \\
&= \overbrace{\widehat{S}(t_k|A=r)}^{Marginal} - \overbrace{\widehat{S}(t_k|A=b)}^{Marginal} \\
&= \mathbb{E}_W[\widehat{S}(t_k|A=r,\mathbf{W}) - \widehat{S}(t_k|A=b,\mathbf{W})]
\end{aligned}$$

Influence curve: Notation heavy

For an individual on treatment arm $A = a$

$$IC_a(p_0)(t_k) = \sum_{t \leq t_k} \overbrace{[I(\widetilde{T} = t, \Delta = 1) - I(\widetilde{T} \geq t)\lambda(t|A = a, W)]}^{\text{Residual from fitting the hazard model}} \times$$

$$\underbrace{h_a(t, A, W)}_{\text{Covariate to remove bias}} + \underbrace{S(t_k|A = a, W) - \Psi_a(p_0)(t_k)}_{\text{Expectation is 0}}$$

$$h_a(t, A, W) = -\frac{I(A = a)I(t \leq t_k)}{g^0(A = a|W)\bar{G}^0(t_-|A, W)} \frac{S(t_k|A, W)}{S(t|A, W)}$$

$g^0(A = a|W)$: Treatment mechanism

$\bar{G}^0(t_-|A, W) = \Pr(C \geq t|A, W)$: Censoring mechanism

Key Results for TMLE: Doubly robust

Q_0

Q_{10}: Distribution of the baseline covariates W.

$\lambda(t|A, W)$: Conditional distribution of hazard given treatment and baseline covariates.

g_0

$g_1(A|W)$: Treatment mechanism

$\bar{G}(.|A, W)$: Conditional distribution of censoring given treatment and baseline covariates. (Censoring mechanism)

Either Q_0 or g_0 is correct, then TMLE is consistent

Inference

Asymptotic Distribution

$$\sqrt{n}\left(\widehat{\Psi^*}(t_k) - \Psi(p_0)(t_k)\right) \rightarrow N(0, \sigma^2)$$

Asymptotic variance

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} \widehat{IC}(\hat{g}, \hat{G}, \hat{\lambda})^2$$

Wald CI

$$\hat{\psi}^*(t_k) \pm 1.96\frac{\hat{\sigma}}{\sqrt{n}}$$

Plug-in estimators!!

1. Perform a parametric estimation using `glm` fit to estimate $\lambda(t|A, W)$

$$\text{logit}[\hat{\lambda}^0(t|A, W)] = \sum_{i=1}^{K} \alpha_i I(t = i) + \beta_A A + \boldsymbol{\beta}_W W$$

2. Refit the above model using $\hat{\lambda}^0(t|A, W)$ as an offset and include the "clever" covariate $\hat{h}^0(t, A, W) = \{\hat{h}_0, \hat{h}_1\}$

$$\text{logit}[\hat{\lambda}^1(t|A, W)] = \text{logit}[\hat{\lambda}^0(t|A, W)] + \epsilon^T \hat{h}^0(t, A, W) \quad (1)$$

where $\epsilon = \{\epsilon_1, \epsilon_2\}$

TMLE as Plug-in estimators continued

3. Use current estimate of $\widehat{\lambda}^1(t|A,W)$ to update $\hat{h}^1(t,A,W)$

$$\widehat{S}(t_k|A,W) = \prod_{t \leq t_k} [1 - \widehat{\lambda}^1(t|A,W)]$$

4. Iterate 2 & 3 until $\widehat{\epsilon} \to 0$
5. Plug-in final estimate of $\widehat{S}_i^*(t_k|A=i,W)$ for $i = 0, 1$

$$\widehat{\Psi}_a(p_0)(t_k) = \frac{1}{n} \sum_{i=1}^{n} \widehat{S}^*(t_k|A=a,W_i)$$

Simulation: Target parameter $\widehat{\Psi}_1(t_k) - \widehat{\Psi}_0(t_k)$

Scenario	Censoring	% Censored*	β_w
1	No censoring	0	Weak
2	MCAR	34.2	Weak
3	MAR	35.5	Weak
4	No censoring	0	Strong
5	MCAR	37.5	Strong
6	MAR	36.9	Strong

MCAR: i.e a subject has some probability of being censored and is independent of the covariate.
MAR: i.e. a subject has some probability of being censored depending on the covariate W.

Reference analysis: Kaplan Meier survival estimate at each time point

*Censoring proportion based on observed data in the paper is 33%

Simulation results 1

Addressing objective 1 & 3

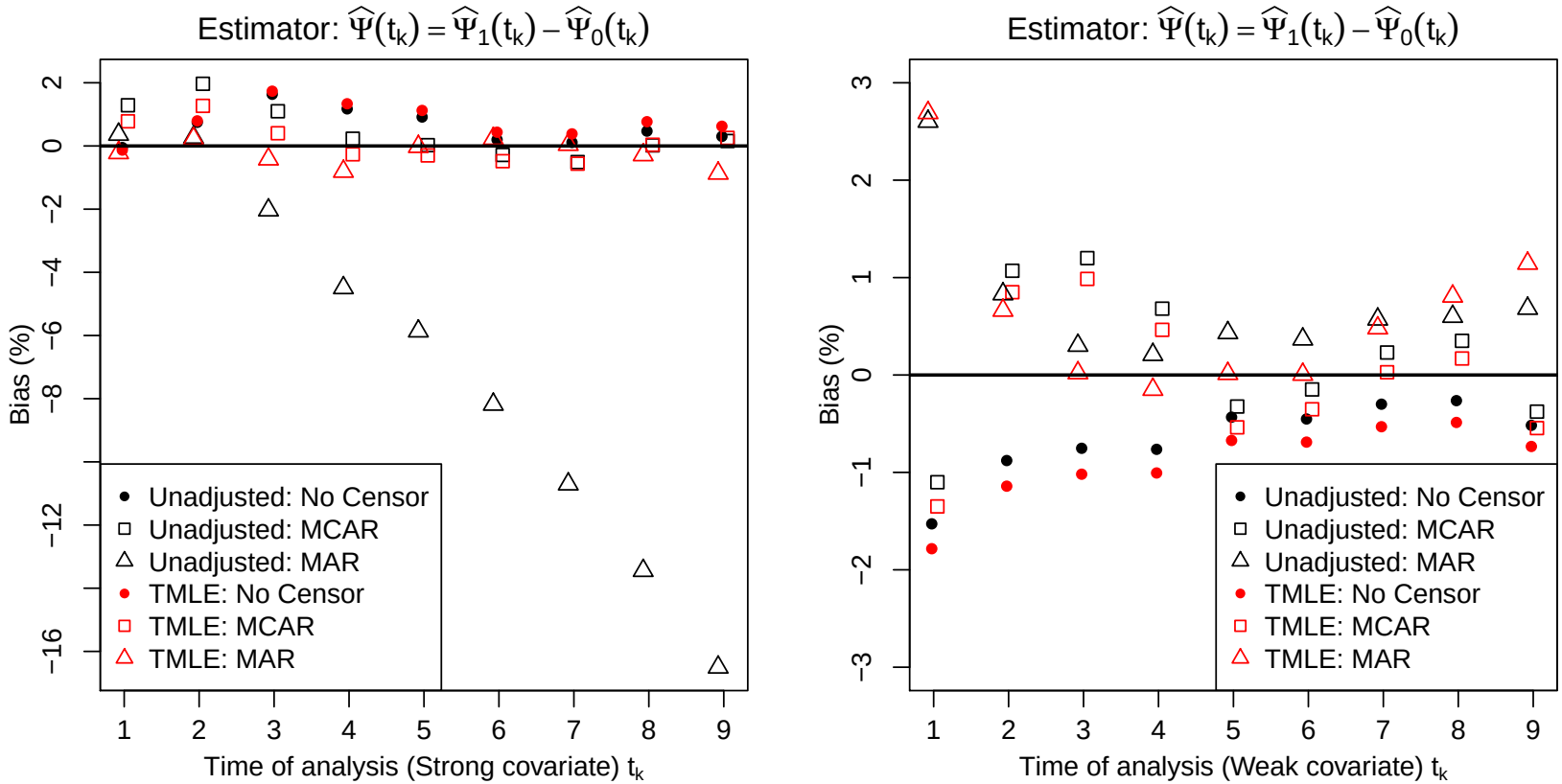


Figure: Left: Bias in unadjusted estimator with respect to the truth when censoring is strongly associated with covariate. Minimal bias when we have weak covariate and informative censoring.

Simulation results 2

Addressing statistical objective 2 & 3

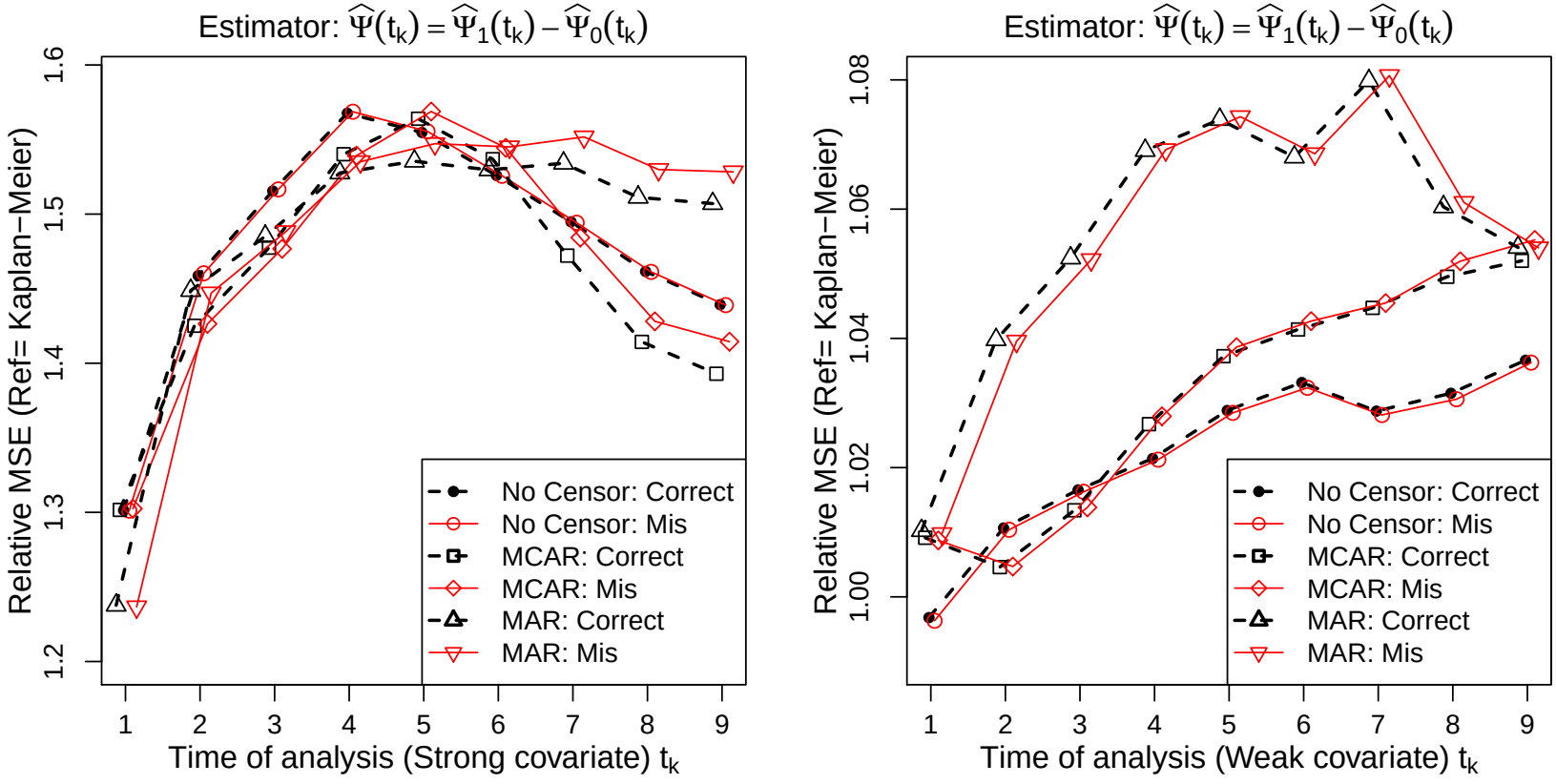


Figure: Comparison of MSE: Under model misspecification of hazard, we still obtain gain in efficiency.

Take home messages

Pros

- "Correctly" estimate the marginal parameter with covariate adjustment and efficiency gain over unadjusted estimator
- (Doubly) Robust to model misspecification
- Unifies concepts in causal inference and survival

Disadvantages

- Complicated
- Potential problem in observational studies.

Potential extensions

- Collaborative TMLE: Construct a sequence of TMLE estimators
- How does this method perform when we use group sequential designs?

Thank you Laina

Thank you all specially for the feedback and support throughout the quarter.

In particular, I would like to thank

- Patrick, Jon
- Ken for his 570 slides on causal inference
- Peter for suggesting the paper

Hernández, A. V., Eijkemans, M., and Steyerberg, E. W. (2006). Randomized controlled trials with time-to-event outcomes: how much does prespecified covariate adjustment increase power? *Annals of epidemiology*, 16(1):41.

Moore, K. and van der Laan, M. (2009a). Increasing power in randomized trials with right censored outcomes through covariate adjustment. *Journal of biopharmaceutical statistics*, 19(6):1099–1131.

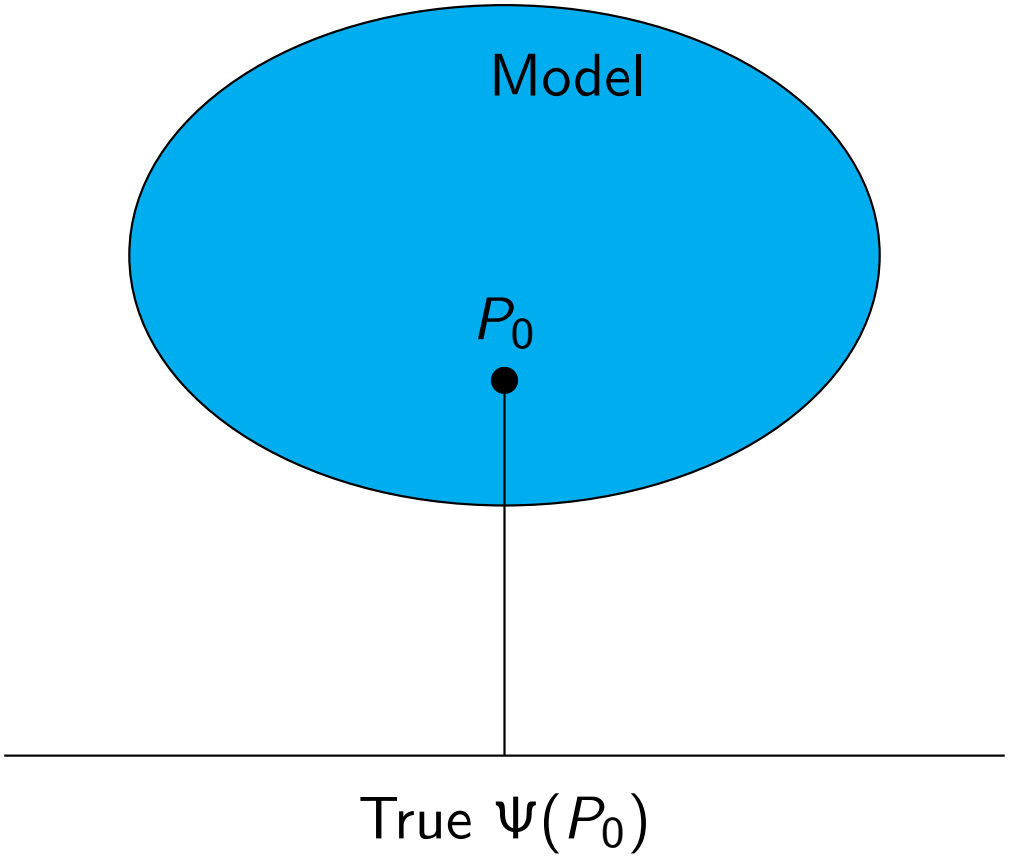
Moore, K. L. and van der Laan, M. J. (2009b). Application of time-to-event methods in the assessment of safety in clinical trials (chapter 20). In Peace, K. E., editor, *Design and Analysis of Clinical Trials with Time-to-Event Endpoints*, pages 455–482. Chapman and Hall/CRC; Biostatistics Series.

Stitelman, O. and van der Laan, M. (2010). Collaborative targeted maximum likelihood for time to event data. *The international journal of biostatistics*, 6(1):Article–21.

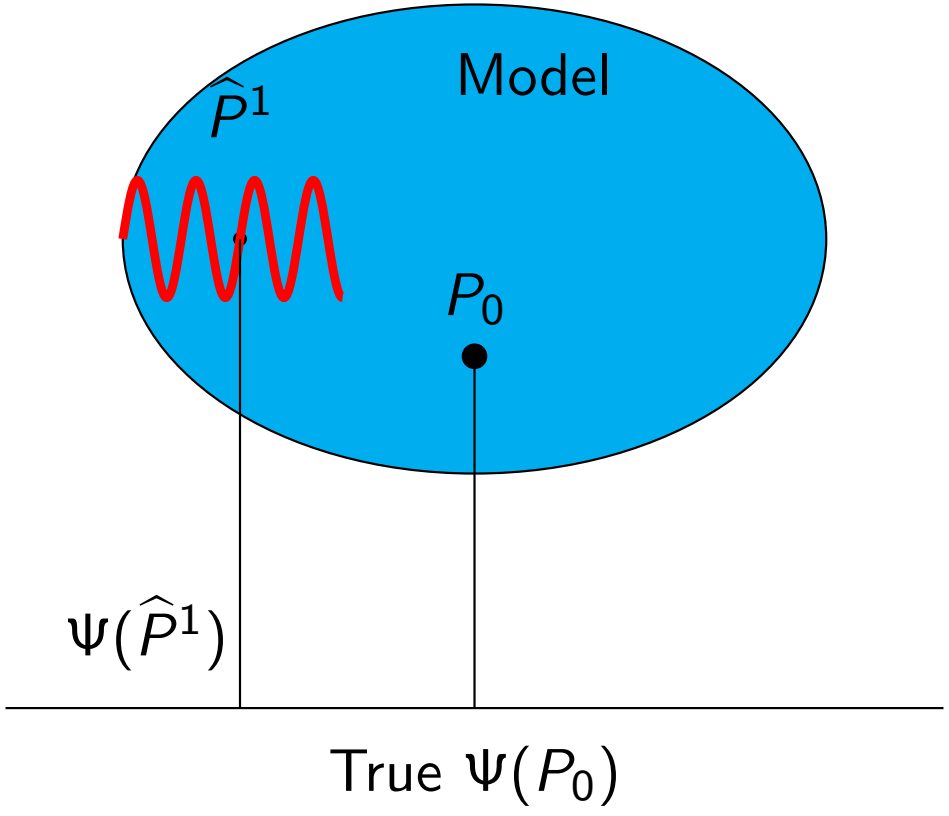
Van der Laan, M. J. and Robins, J. M. (2003). *Unified methods for censored longitudinal data and causality*. Springer Verlag.

Zhang, M. and Gilbert, P. B. (2010). Increasing the efficiency of prevention trials by incorporating baseline covariates. *Statistical communications in infectious diseases*, 2(1).

Intuition behind TMLE

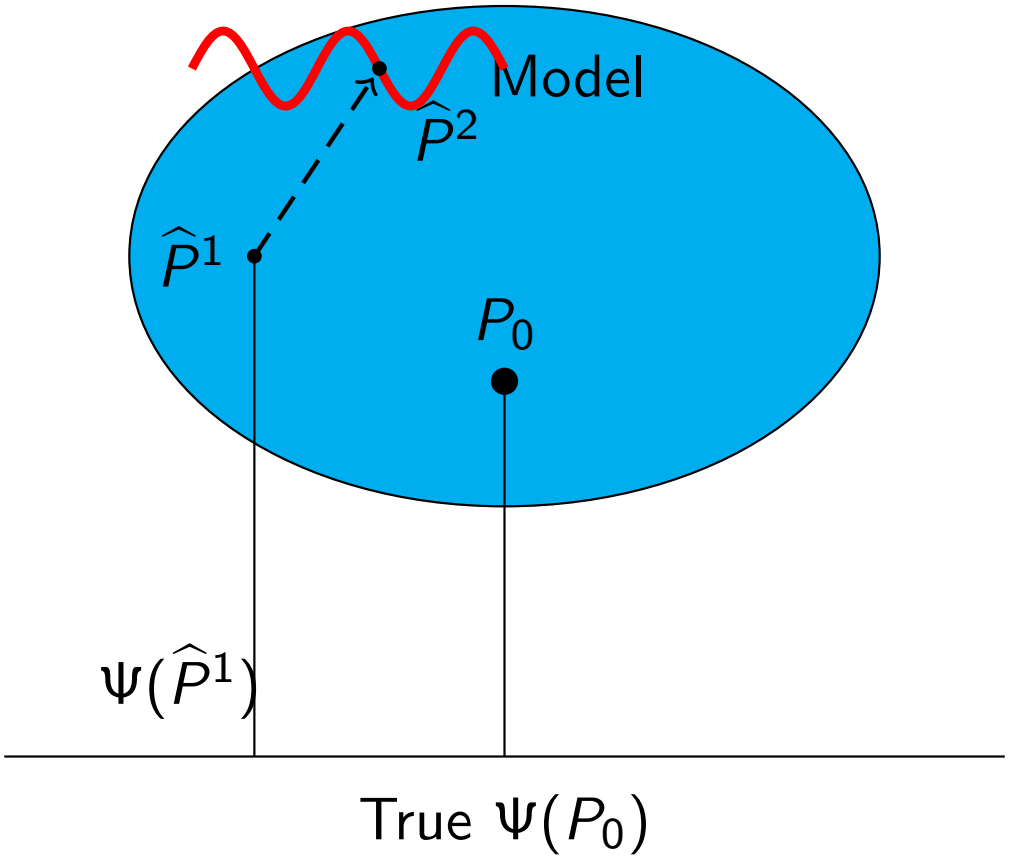


Intuition behind TMLE



$$\text{logit}[\hat{\lambda}^0(t|A,W)] = \alpha(t) + m(A,W)$$

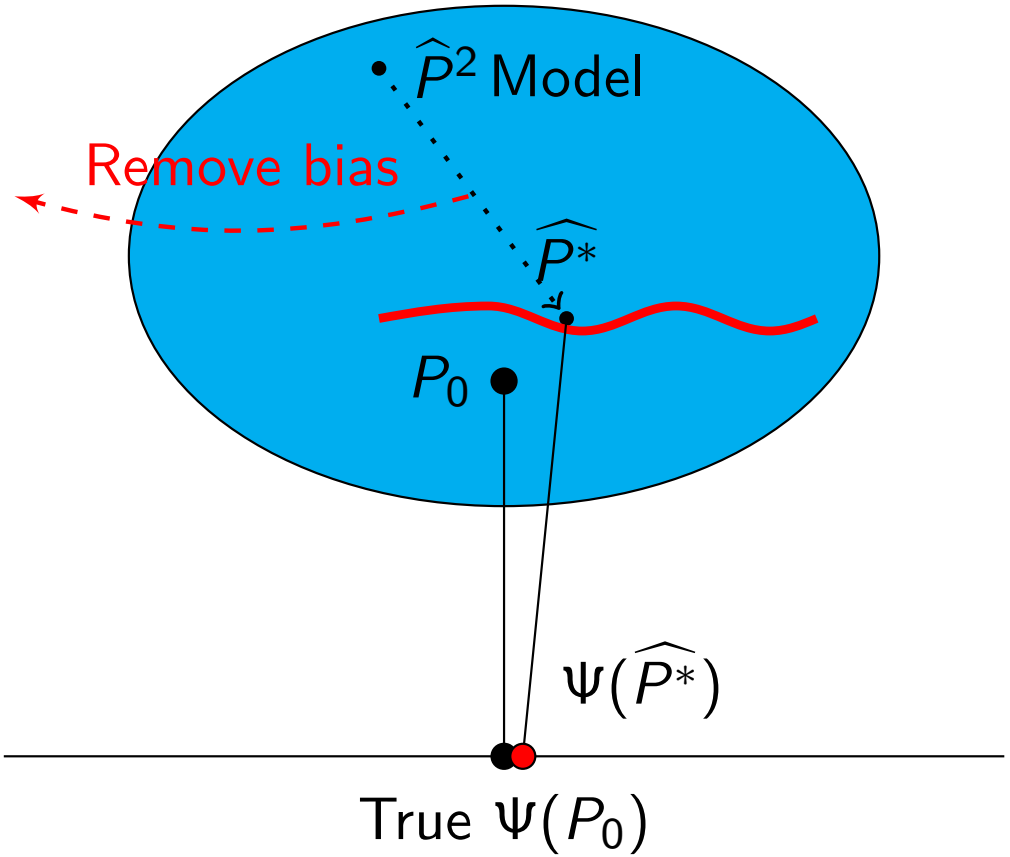
Intuition behind TMLE



$$\text{logit}[\hat{\lambda}^1(t|A,W)] = \text{logit}[\hat{\lambda}^0(t|A,W)] + \epsilon^T \hat{h}^0(t,A,W) \qquad (2)$$

Update $\widehat{\lambda}^1(t|A,W)$, $\hat{h}^1(t,A,W)$.

Intuition behind TMLE



Repeat previous step and iterate ϵ until $\widehat{\epsilon}^* \longrightarrow 0$

Plug-in final estimate of $\widehat{S}_i^*(t_k|A=i,W)$ for $i=0,1$