# Multiply robust estimation of causal effects with noncompliance and time-to-event outcomes

Fan Li

Department of Biostatistics
Yale University School of Public Health

https://lifan90.com/

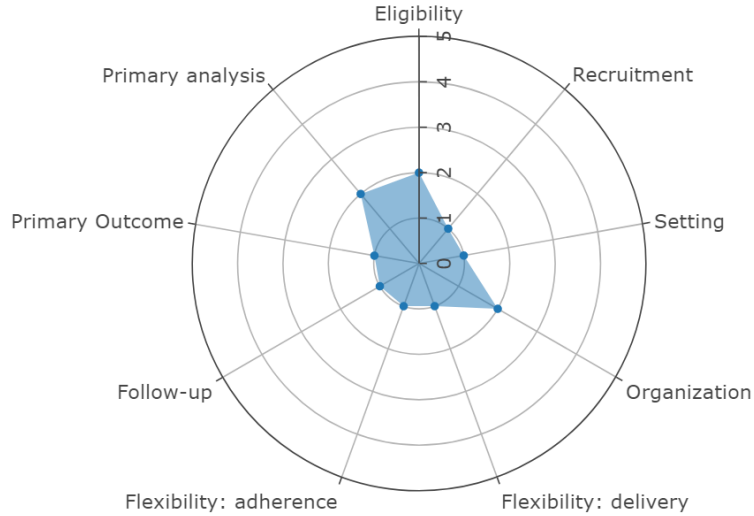
15th Annual University of Pennsylvania Conference on Statistical Issues in Clinical Trials
–"Advances in Time-to-Event Analyses in Clinical Trials"

April 17, 2023

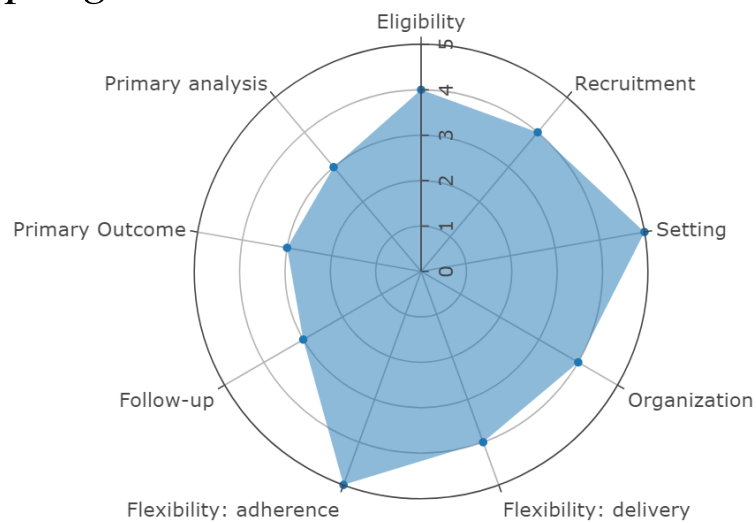
# Background: pragmatic trials

- Pragmatic trials often compare interventions or alternative delivery approaches to evaluate impact on outcomes under routine practice conditions
- Challenges arise as we move toward pragmatism
  - PRECIS-II
  - treatment effect estimands as increasingly important concept[1]
- Many pragmatic trials report null intention-to-treat (ITT) effect
- Provide an example to go beyond ITT estimand in a recent pragmatic trial with survival outcome

(a) *traditional RCT*



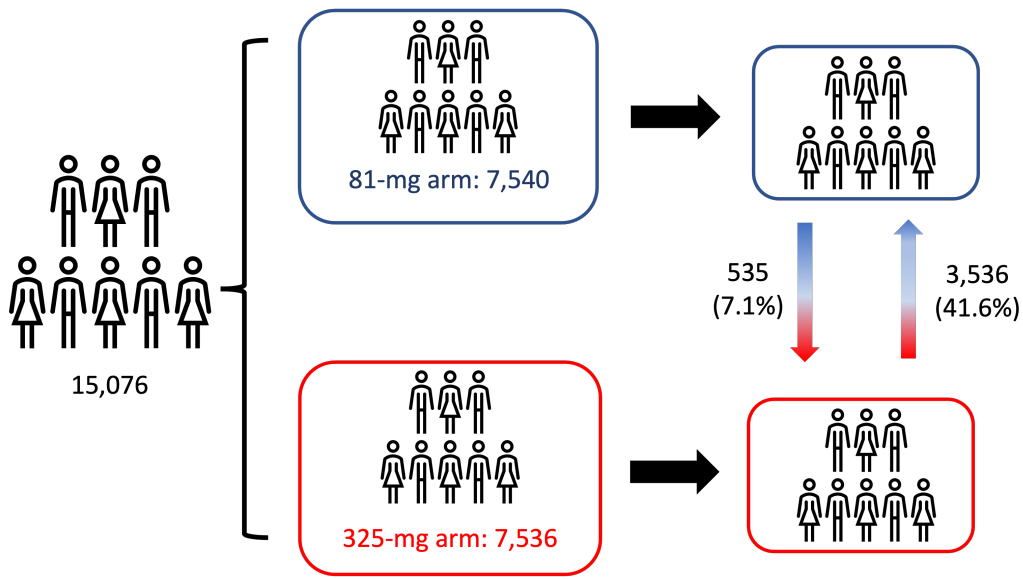

(b) *pragmatic trial*




[1]Kahan, B. C., Cro, S., Li, F., & Harhay, M. O. (2023). Eliminating ambiguous treatment effects using estimands. American Journal of Epidemiology, kwad036.

# An example: ADAPTABLE pragmatic trial

- ADAPTABLE (*Aspirin Dosing: A Patient-Centric Trial Assessing Benefits and Long-Term Effectiveness*) is a pragmatic trial to study the effectivenss of two aspirin dosing strategies for patients with established cardiovascular diseases[2].
- **Treatment assignment:** 81-mg v.s. 325-mg aspirin dosage
- **Outcome:** a composite outcome of death from any cause and hospitalization for stroke or myocardial infraction (time-to-event)
- **Complication:** not all patients take the assigned aspirin dosage.




[2]Jones, W., Mulder, H., and others (2021). Comparative effectiveness of aspirin dosing in cardiovascular disease. *New England Journal of Medicine*, 384(21), 1981-1990.

# Average treatment effect

- **Baseline covariates:** age, gender, race, ethnicity, medical and diseases history, aspirin dosage use prior to the trial, *etc*.
- Up until the maximum follow-up time (38 months after randomization), the outcome occurred in 7.5% and 7.4% of participants in the 81-mg arm and the 325-mg arm, respectively.
  - relatively rare events
- As expected, the primary analysis found a null average treatment effect
- To confirm no overall effect, we applied a covariate-adjusted approach to report the ITT effect
  - augmented estimator for the causal survival curves (Bai et al, 2013; Zhang and Schaubel, 2012)

# ITT effect estimates

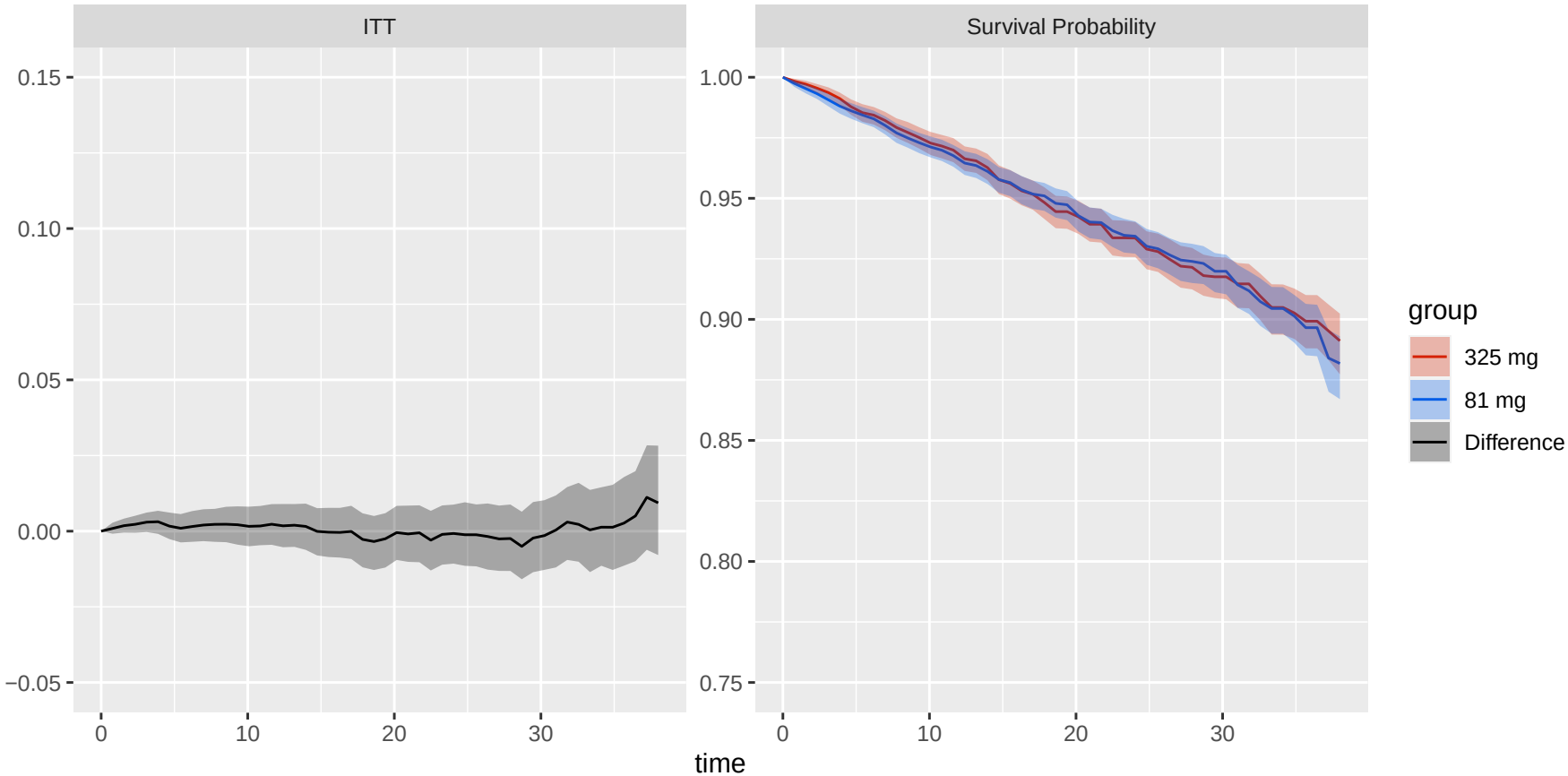



Figure: The ITT effects and corresponding survival probability curve, ADAPTABLE trial, 2016–2020. The ITT effects are obtained based on the doubly robust estimator by Bai et al. (2013)[3].

[3]Bai, X., Tsiatis, A. A., & O'Brien, S. M. (2013). Doubly-robust estimators of treatment-specific survival distributions in observational studies with stratified sampling. *Biometrics*, 69(4), 830-839.

# Motivating questions

Recognizing the potential impact of noncompliance

- Is one specific aspirin dosage more effective than the other among compliers?
  - the treatment efficacy
- Is there treatment effect heterogeneity among patients with different compliance behaviour types that contribute to the null ITT effect?
  - 'direct effect' of treatment assignment (possibly due to other mechanisms that are unmeasured)
- To what extent the study results are sensitive to unverifiable assumptions?

**Goal**: seek model-robust methods and sensitivity strategies to address noncompliance with a survival outcome

# Notation

- Data structure
  - $\boldsymbol{X}$: pre-treatment covariates
  - $Z$: treatment assignment (1 if 325-mg aspirin and 0 if 81-mg aspirin)
  - $S$: actual treatment received
  - $T$: survival time of interest—partially observed due to right censoring. Instead, we only observe $(U, \delta) = \Big( \min(T, C), \mathbb{I}(T \leq C) \Big)$, where $C$ is the censoring time
- Pursue potential outcomes framework and define
  - $S(z)$: the potential value of $S$ when setting the assignment to $z \in \{0, 1\}$.
  - $T(z)$ & $C(z)$: the potential value of $T$ and $C$ when setting the assignment to $z \in \{0, 1\}$.

# Estimands

- Under the principal stratification framework, we partition the study population into 4 principal strata, based on the joint potential values $G = \{S(1), S(0)\}$:
  - $G = \{1, 1\}$: always high-dose (325-mg) takers
  - $G = \{0, 0\}$: always low-dose (81-mg) takers
  - $G = \{1, 0\}$: compliers
  - $G = \{0, 1\}$: defiers
- abbreviate the above four principal strata as $\{a, n, c, d\}$, respectively.
- Define the **Principal Survival Causal Effect (PSCE)**:

$$\Delta_g(u) = \mathcal{S}_{1,g}(u) - \mathcal{S}_{0,g}(u)$$

where $\mathcal{S}_{z,g} = \mathbb{P}(T(z) > u | G = g)$ and $g \in \{a, n, c, d\}$.

  - $\Delta_c(u)$: 'efficacy' of the treatment in the ideal compliance condition.
  - $\Delta_a(u)$ and $\Delta_n(u)$: 'direct effect' due to treatment assignment
  - do not pursue strata-specific HR as estimands

# Assumptions

Five structural assumptions to identify the PSCE.

**Set I: standard assumptions in causal survival analysis** (Chen et al, 2001; Zhang and Schaubel, 2012; Bai et al. 2013)

A1. (SUTVA) For all $i$, we have $S_i(z) = S_i$, $T_i(z) = T_i$, and $C_i(z) = C_i$ if patient $i$ was assigned to treatment $Z_i = z$.

A2. (Conditional randomization) $\{T(1), T(0), S(1), S(0)\} \perp Z|\boldsymbol{X}$.

- in ADAPTABLE, a stronger version, $\{T(1), T(0), S(1), S(0), \boldsymbol{X}\} \perp Z$, is satisfied by study design.
- use a more general assumption to allow considerations of conditional randomized trials and observational studies.

A3. (Covariate-dependent censoring) $T(z) \perp C(z)|\{Z = z, S, \boldsymbol{X}\}$.

# Assumptions - cont'd

**Set II: additional assumptions**

A4. (Monotonicity) $S_i(1) \geq S_i(0)$ for all $i$.

- excludes defiers and can be plausible for ADAPTABLE

A5. (Prinicipal ignorability) For all $u \geq 0$, we have

$$\mathbb{P}\left(T(1) \geq u | G = a, \boldsymbol{X}\right) = \mathbb{P}\left(T(1) \geq u | G = c, \boldsymbol{X}\right),$$
$$\mathbb{P}\left(T(0) \geq u | G = c, \boldsymbol{X}\right) = \mathbb{P}\left(T(0) \geq u | G = n, \boldsymbol{X}\right).$$

- assumes sufficient baseline information to capture characteristics influencing both the noncompliance behavior and the potential outcomes
- extension of Ding and Lu. (2016); Jiang et al. (2022) to causal survival analysis
- a stronger version: $G \perp T(z) | \boldsymbol{X}$ for $z = 0, 1$

Both unverifiable from observed data and sensitivity analysis can help.

# Possible working models

1. **Assignment mechanism**:

   $\mathcal{M}_\pi$: $\pi(\boldsymbol{X}) = \mathbb{P}(Z = 1|\boldsymbol{X})$, the propensity score (Rosenbaum and Rubin, 1983).

   - in ADAPTABLE, $\pi(\boldsymbol{X}) = 0.5$ by randomization, but modeling this process may gain efficiency (Zeng et al, 2021; Li, Buchanan, Cole, 2022)

2. **Noncompliance**:

   $\mathcal{M}_e$: $e_g(\boldsymbol{X}) = \mathbb{P}(G = g|\boldsymbol{X})$ for $g \in \{a, c, n\}$ the principal scores (Ding and Lu, 2016).

   - define $e_g = E[e_g(\boldsymbol{X})]$ as strata proportion
   - by monotonicity, we have $e_c(\boldsymbol{X}) = p_1(\boldsymbol{X}) - p_0(\boldsymbol{X})$, $e_a(\boldsymbol{X}) = p_0(\boldsymbol{X})$, and $e_n(\boldsymbol{X}) = 1 - p_1(\boldsymbol{X})$, where $p_z(\boldsymbol{X}) = \mathbb{P}(S = 1|Z = z, \boldsymbol{X})$ is the observed probability for receiving the 325-mg dosage.

- Can fit logistic regressions to obtain $\widehat{\pi}(\boldsymbol{X})$ and $\widehat{p}_z(\boldsymbol{X})$.

# Possible working models - cont'd

3. **Censoring**:

   $\mathcal{M}_C$: $\mathcal{S}^C_{zs}(u|\boldsymbol{X}) = \mathbb{P}(C \geq u|Z = z, S = s, \boldsymbol{X})$

   - survival function of the censoring time within observed cell $(Z = z, S = s)$.

4. **Survival outcome (of interest)**:

   $\mathcal{M}_T$: $\mathcal{S}_{zs}(u|\boldsymbol{X}) = \mathbb{P}(T \geq u|Z = z, S = s, \boldsymbol{X})$

   - survival function of the time-to-event outcome of interest within observed cell $(Z = z, S = s)$

- Can fit Cox proportional hazards models to obtain $\widehat{\mathcal{S}}^C_{zs}(u|\boldsymbol{X})$ and $\widehat{\mathcal{S}}_{zs}(u|\boldsymbol{X})$

  - working models and no attempt to interpret HR as causal parameter

# Moment estimators

- We do not need all models to point identify PSCEs
- We characterized 3 moment-type estimators of $\Delta_g(u)$, each depending on part of, but not all of, the 4 working models $\mathcal{M}_\pi$, $\mathcal{M}_e$, $\mathcal{M}_C$, and $\mathcal{M}_T$.

| | $\mathcal{M}_\pi$ (IPTW) | $\mathcal{M}_e$ (PSW) | $\mathcal{M}_C$ (IPCW) | $\mathcal{M}_T$ (OM) |
|---|---|---|---|---|
| $\widehat{\Delta}_g^{(1)}(u)$ | ✓ | ✓ | ✓ | |
| $\widehat{\Delta}_g^{(2)}(u)$ | | ✓ | | ✓ |
| $\widehat{\Delta}_g^{(3)}(u)$ | ✓ | | | ✓ |

- $\widehat{\Delta}_g^{(1)}(u)$, $\widehat{\Delta}_g^{(2)}(u)$, and $\widehat{\Delta}_g^{(3)}(u)$ consistent to $\Delta_g(u)$ if $\mathcal{M}_{\pi+e+C}$, $\mathcal{M}_{e+T}$, and $\mathcal{M}_{\pi+T}$ is correctly specified, respectively
  - $\mathcal{M}_{a+b}$ is the intersection model, hence all 3 estimators are only singly robust

# Combing multiple working models with complete data

- Ideal to leverage all working models to improve robustness and efficiency
- Suppose the true failure time can be observed (no censoring) with the complete data $\mathcal{A} = \{\boldsymbol{X}, Z, S, T\}$.
- Adapted from Jiang et al. (2022)[4], a multiply robust estimator of $S_{1,c}(u)$ can be obtained by solving the following estimating equation,

$$\mathbb{P}_n \left[ \psi_{1,c}^{mr}(\mathcal{A}) \right] = 0,$$

where

$$\begin{aligned}\psi_{1,c}^{mr}(\mathcal{A}) = & \frac{e_c(\boldsymbol{X})}{\widetilde{p}_1 - \widetilde{p}_0} \frac{S}{p_1(\boldsymbol{X})} \frac{Z}{\pi(\boldsymbol{X})} \mathbb{I}(T \geq u) - S_{1,c}(u) \\ & + \frac{e_c(\boldsymbol{X})}{\widetilde{p}_1 - \widetilde{p}_0} \left(1 - \frac{Z}{\pi(\boldsymbol{X})}\right) S_{11}(u|\boldsymbol{X}) \\ & + \frac{S_{11}(u|\mathbf{X})}{\widetilde{p}_1 - \widetilde{p}_0} \left[ \frac{1-Z}{1-\pi(\boldsymbol{X})} (S - p_0(\boldsymbol{X})) + \frac{p_0(\boldsymbol{X})}{p_1(\boldsymbol{X})} \frac{Z}{\pi(\boldsymbol{X})} (p_1(\boldsymbol{X}) - S) \right],\end{aligned}$$

- The first row: a weighting-based estimating equation of $S_{1,c}(u)$; the last two rows: augmented zero-mean terms based on outcome modeling.

[4] Jiang, Z., Yang, S., & Ding, P. (2022). Multiply robust estimation of causal effects under principal ignorability. *JRSSB*, 84(4), 1423-1445.

# Accommodate right-censoring

- Observed data subject to censoring: $O = \{X, Z, S, U, \delta\}$
- Key idea: identify a set of unbiased estimating functions that depend only on the observed data $O$:

$$\psi_{1,c}(O) = \frac{\delta \psi_{1,c}^{mr}(\mathcal{A})}{S_{ZS}^{C}(U|X)} + \int h(r, Z, S, X) dM_{ZS}^{C}(r|X),$$

where $h(t, Z, S, X)$ can be arbitrary function and $dM_{ZS}^{C}(t|X)$ is the censoring process martingale (within in cell defined by $(Z, S)$).

  - The resulting estimator is an *augmented inverse probability of weighted complete-case* (AIPWCC) estimator (Tsiatis, 2006)

- The optimal choice of $h(t, Z, S, X)$ to maximize efficiency is

$$h(t, Z, S, X) = \frac{\mathbb{E}[\psi_{1,c}^{mr}(\mathcal{A})|T \geq t, Z, S, X]}{S_{ZS}^{C}(t|X)}$$

# The proposed estimator

- After some algebra, the proposed estimator has the following explicit form

$$\begin{aligned}\widehat{\mathcal{S}}^{mr}_{1,c}(u) = \mathbb{P}_n \Bigg\{ & \frac{\widehat{e}_c(\boldsymbol{X})}{\widehat{p}_1 - \widehat{p}_0} \frac{S}{\widehat{p}_1(\boldsymbol{X})} \frac{Z}{\widehat{\pi}(\boldsymbol{X})} \left[ \frac{\mathbb{I}(U \geq u)}{\widehat{\mathcal{S}}^C_{11}(u|\boldsymbol{X})} + \widehat{\mathcal{S}}_{11}(u|\boldsymbol{X}) \int_0^u \frac{d\widehat{M}^C_{11}(r|\boldsymbol{X})}{\widehat{\mathcal{S}}_{11}(r|\boldsymbol{X}) \widehat{\mathcal{S}}^C_{11}(r|\boldsymbol{X})} \right] \\ & + \frac{\widehat{\mathcal{S}}_{11}(u|\boldsymbol{X})}{\widehat{p}_1 - \widehat{p}_0} \left[ \frac{1 - Z}{1 - \widehat{\pi}(\boldsymbol{X})} (S - \widehat{p}_0(\boldsymbol{X})) + \frac{\widehat{p}_0(\boldsymbol{X})}{\widehat{p}_1(\boldsymbol{X})} \frac{Z}{\widehat{\pi}(\boldsymbol{X})} (\widehat{p}_1(\boldsymbol{X}) - S) \right] \\ & + \frac{\widehat{e}_c(\boldsymbol{X})}{\widehat{p}_1 - \widehat{p}_0} \left( 1 - \frac{Z}{\widehat{\pi}(\boldsymbol{X})} \right) \widehat{\mathcal{S}}_{11}(u|\boldsymbol{X}) \Bigg\},\end{aligned}$$

where $\mathbb{P}_n[V] = \frac{1}{n} \sum_{i=1}^n V_i$ is the empirical average operator

- Similar ideas to estimate counterfactual survival functions in other strata

- Bootstrap variance

# Multiple robustness

- **Result**. (Multiple robustness)

  *Suppose that Assumptions 1–5 hold. For all* $z \in \{0, 1\}$ *and* $g \in \{c, n, a\}$, $\widehat{\mathcal{S}}^{mr}_{z,g}(u)$ *is consistent to* $\mathcal{S}_{z,g}(u)$ *under the union model* $\mathcal{M}_{\pi+e+C} \cup \mathcal{M}_{\pi+T} \cup \mathcal{M}_{e+T}$. *As a consequence,* $\widehat{\Delta}^{mr}_{g}(u)$ *is also consistent to* $\Delta_g(u)$ *under* $\mathcal{M}_{\pi+e+C} \cup \mathcal{M}_{\pi+T} \cup \mathcal{M}_{e+T}$ *for all* $g \in \{c, n, a\}$.

- **remark 1**: the result is general and can be applied to observational study (where modeling assignment $\mathcal{M}_\pi$ is necessary for removing confounding bias)

- **remark 2**: in ADAPTABLE, because $\mathcal{M}_\pi$ is known under randomization, $\widehat{\Delta}^{mr}_{g}(u)$ becomes a doubly robust estimator, in that consistency holds under the union model $\mathcal{M}_{e+C} \cup \mathcal{M}_T$
  - correct weighting or correct outcome modeling, but not necessarily both

- simulations confirm the robustness property

# Application to ADAPTABLE

- Working models (all models adjusting for all baseline characteristics)
  - $\mathcal{M}_\pi$: logistic regression
  - $\mathcal{M}_e$: logistic regression
  - $\mathcal{M}_C$: Cox regression
  - $\mathcal{M}_T$: Cox regression
- Proportion of each principal strata estimated from principal scores (Figure)
- For any $V \in \boldsymbol{X}$, we calculate its estimated mean and variance within each strata
  - $\widehat{\mathbb{E}}[V|G=g] = \mathbb{P}_n\left[\frac{\widehat{e}_g(\boldsymbol{X})}{\widehat{e}_g}V\right]$
  - $\widehat{\text{Var}}[V|G=g] := s_g^2$ given by the principal score weighted variance
  - useful in describing and comparing demographic and clinical characteristics among different compliance strata.

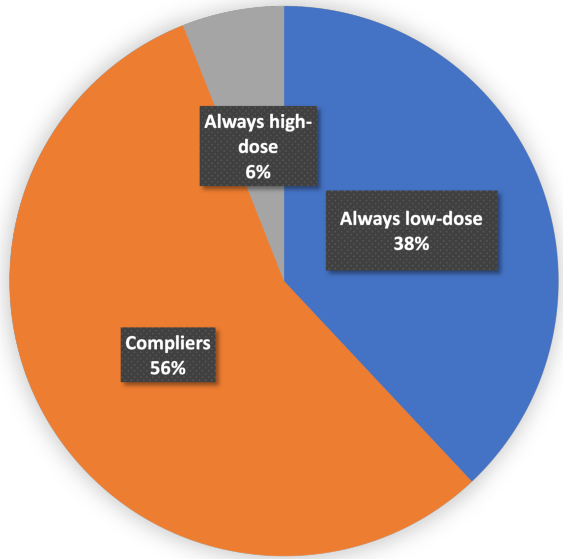



**Figure**:The estimated proportions of each strata (i.e., $\widehat{e}_g$)

# Strata characteristics

- **Always low-dose:** older, higher prevalence of cardiovascular diseases and worse medical history, more 81-mg aspirin users
- **Compliers:** less non-internet users, more white patients
- **Always high-dose:** youngest, less medical conditions, more 325-mg aspirin users

Table 3: Mean and standard deviation of baseline characteristics among always low-dose takers, compliers, and always high-dose takers, ADAPTABLE trial, 2016–2020.

| Variable | Always low-dose takers | Compliers | Always high-dose takers | Max ASD* |
|---|---|---|---|---|
| Age (years) | 67.70 (9.34) | 66.44 (9.36) | 66.02 (9.56) | 0.18 |
| Female sex | 0.32 (0.47) | 0.29 (0.45) | 0.28 (0.45) | 0.09 |
| White race | 0.78 (0.41) | 0.89 (0.31) | 0.80 (0.40) | 0.29 |
| Hispanic ethnicity | 0.04 (0.20) | 0.03 (0.16) | 0.04 (0.19) | 0.08 |
| Non-internet users | 0.21 (0.41) | 0.09 (0.28) | 0.24 (0.42) | 0.41 |
| Current smokers | 0.10 (0.31) | 0.08 (0.28) | 0.14 (0.35) | 0.18 |
| P2Y12 inhibitor | 0.25 (0.43) | 0.22 (0.41) | 0.22 (0.42) | 0.07 |
| **Medical History** | | | | |
| Myocardial infarction | 0.39 (0.49) | 0.35 (0.48) | 0.38 (0.49) | 0.09 |
| Atrial fibrillation | 0.09 (0.28) | 0.08 (0.27) | 0.10 (0.30) | 0.09 |
| Percutaneous coronary intervention | 0.46 (0.50) | 0.40 (0.49) | 0.39 (0.49) | 0.15 |
| Coronary artery disease | 0.96 (0.19) | 0.95 (0.21) | 0.94 (0.24) | 0.12 |
| Coronary-artery bypass grafting | 0.26 (0.44) | 0.25 (0.43) | 0.25 (0.43) | 0.02 |
| Hypertension | 0.89 (0.32) | 0.85 (0.35) | 0.90 (0.30) | 0.13 |
| Hyperlipidemia | 0.91 (0.28) | 0.90 (0.30) | 0.88 (0.32) | 0.10 |
| Peripheral artery disease | 0.29 (0.45) | 0.22 (0.41) | 0.25 (0.44) | 0.16 |
| Congestive heart failure | 0.27 (0.44) | 0.22 (0.41) | 0.26 (0.44) | 0.12 |
| History of bleeding | 0.11 (0.31) | 0.07 (0.26) | 0.10 (0.30) | 0.12 |
| **Aspirin use before trial** | | | | |
| Prior dose: 81 mg | 0.90 (0.30) | 0.87 (0.34) | 0.45 (0.50) | 1.11 |
| Prior dose: 325 mg | 0.08 (0.27) | 0.11 (0.32) | 0.47 (0.50) | 0.97 |

*Max ASD is the maximum pairwise absolute standardized difference across the three principal strata.

# Balance check

- Why balance check? To empirically check evidence for principal score model adequacy

- Balance metrics: The *weighted standardized mean differences* (SMDs) of a given covariate $V$ across the four observed $(Z, S)$-strata, $(Z = 1, S = 1)$, $(Z = 0, S = 1)$, $(Z = 1, S = 0)$, and $(Z = 0, S = 0)$:

$$\text{SMD}_c = \frac{1}{s_c}\left|\mathbb{P}_n\left[\frac{ZS\mathcal{W}_{1,c}(\boldsymbol{X})V}{\mathbb{P}_n[ZS]} - \frac{(1-Z)(1-S)\mathcal{W}_{0,c}(\boldsymbol{X})V}{\mathbb{P}_n[(1-Z)(1-S)]}\right]\right|,$$

$$\text{SMD}_n = \frac{1}{s_n}\left|\mathbb{P}_n\left[\frac{Z(1-S)\mathcal{W}_{1,n}(\boldsymbol{X})V}{\mathbb{P}_n[Z(1-S)]} - \frac{(1-Z)(1-S)\mathcal{W}_{0,n}(\boldsymbol{X})V}{\mathbb{P}_n[(1-Z)(1-S)]}\right]\right|,$$

$$\text{SMD}_a = \frac{1}{s_a}\left|\mathbb{P}_n\left[\frac{ZS\mathcal{W}_{1,a}(\boldsymbol{X})V}{\mathbb{P}_n[ZS]} - \frac{(1-Z)S\mathcal{W}_{0,a}(\boldsymbol{X})V}{\mathbb{P}_n[(1-Z)S]}\right]\right|,$$

  where $\mathcal{W}_{z,g}(\boldsymbol{X})$ for $z \in \{0, 1\}$ and $g \in \{c, a, n\}$ are specified weights.

- When $\mathcal{W}_{z,g}(\boldsymbol{X}) = 1$, the SMDs measure the systematic difference of $X$ across different $(Z, S)$-strata, and therefore reflects heterogeneity of $X$ due to patients' noncompliance behavior
- When $\mathcal{W}_{z,g}(\boldsymbol{X})$ is set to the (true) principal score weight, SMDs should be 0

# Balance check - cont'd

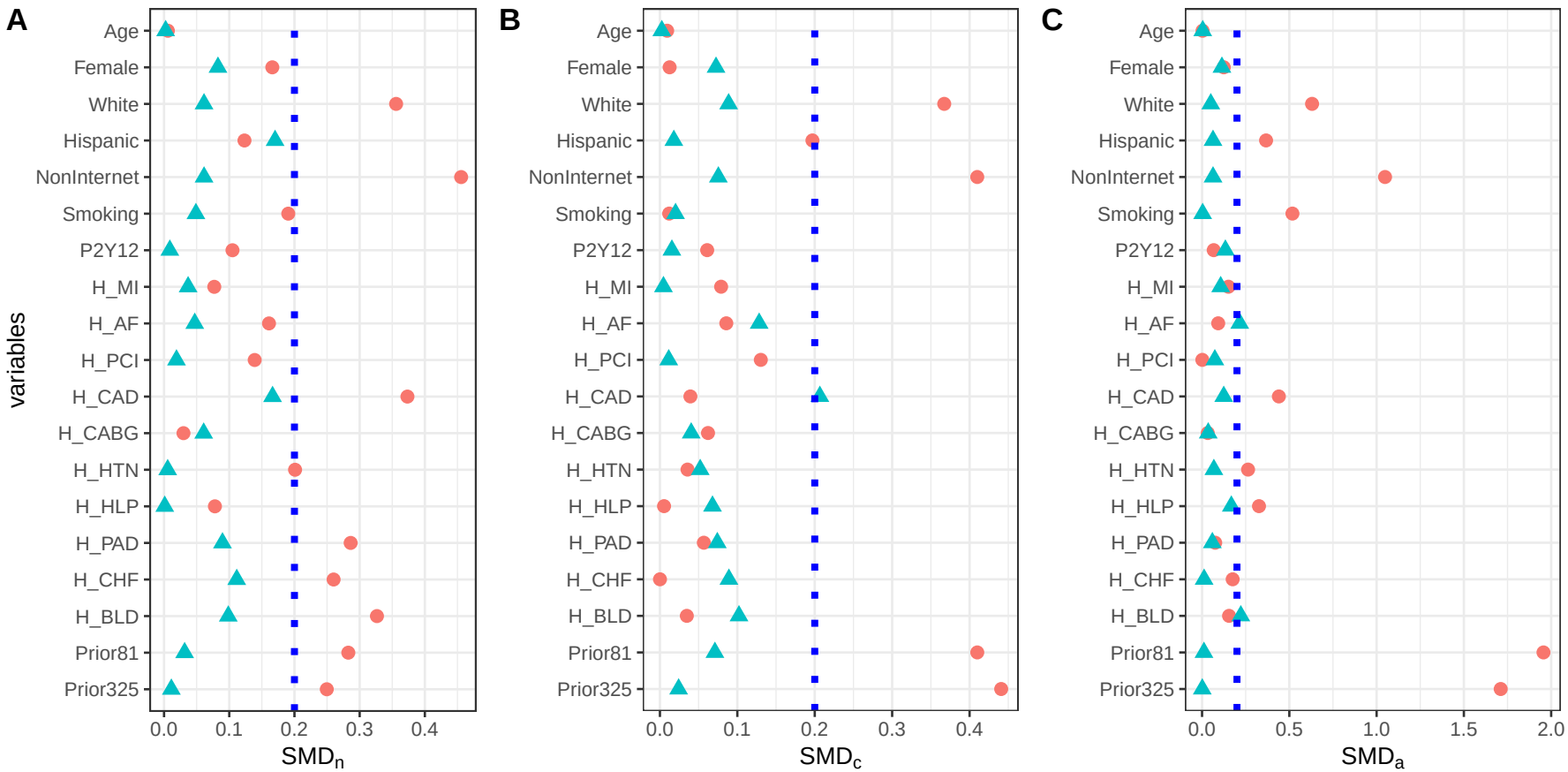



Figure: Balance check for baseline characteristics. The red '●' symbol indicates the unweighted SMDs and the blue '▲' symbol indicates the weighted SMDs by the principal score weighting.

# Principal causal effect estimates

- Implement the proposed estimator of the PSCEs among the always low-dose, compliers, and always high-dose

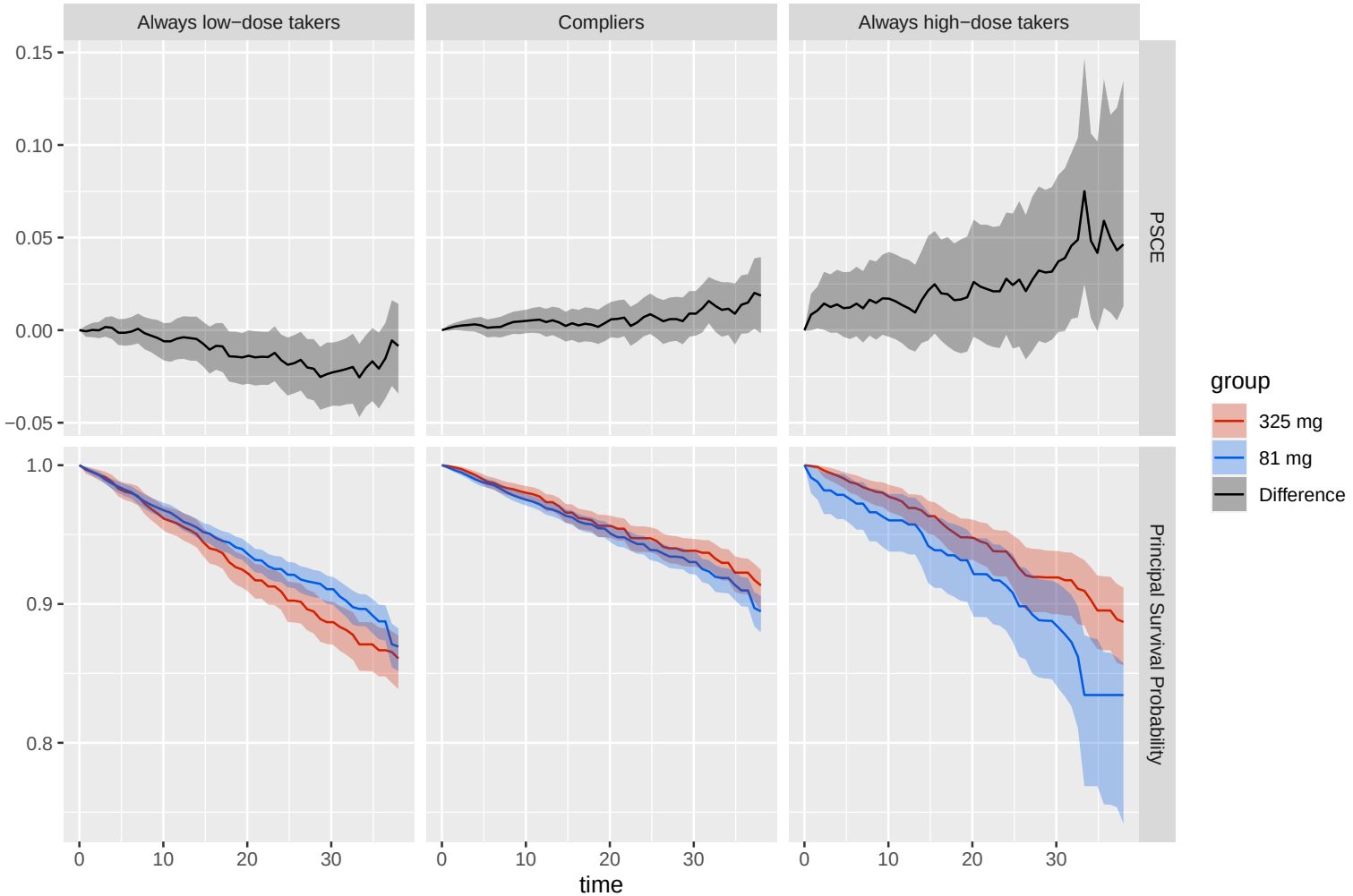



Figure: The principal survival causal effects (PSCEs) and the principal survival probability curves

# Sensitivity analysis for principal ignorability (PI)

- PI holds if there is no residual confounding between $G$ and $T$ contional on $\boldsymbol{X}$
  - PI violated if there exists unmeasured confounding ($U$) between $G$ and $T$
  - One potential $U$: digestive disease, which may increase the risk of death. Moreover, people with digestive diseases may prefer not to take high aspirin dosage due to its risk on bleeding.
- Consider sensitivity functions to measure departure from PI:

$$\varepsilon_1(t, \boldsymbol{X}) = \frac{\mathbb{P}(T(1) \geq t | G = c, \boldsymbol{X})}{\mathbb{P}(T(1) \geq t | G = a, \boldsymbol{X})} = \exp\left\{\xi_1 \times \left(\frac{t}{t_{\max}}\right)\right\}$$

$$\varepsilon_0(t, \boldsymbol{X}) = \frac{\mathbb{P}(T(0) \geq t | G = c, \boldsymbol{X})}{\mathbb{P}(T(0) \geq t | G = n, \boldsymbol{X})} = \exp\left\{\xi_0 \times \left(\frac{t}{t_{\max}}\right)\right\},$$

- We use $\{\xi_1, \xi_0\}$ to control the pattern of the two sensitivity functions

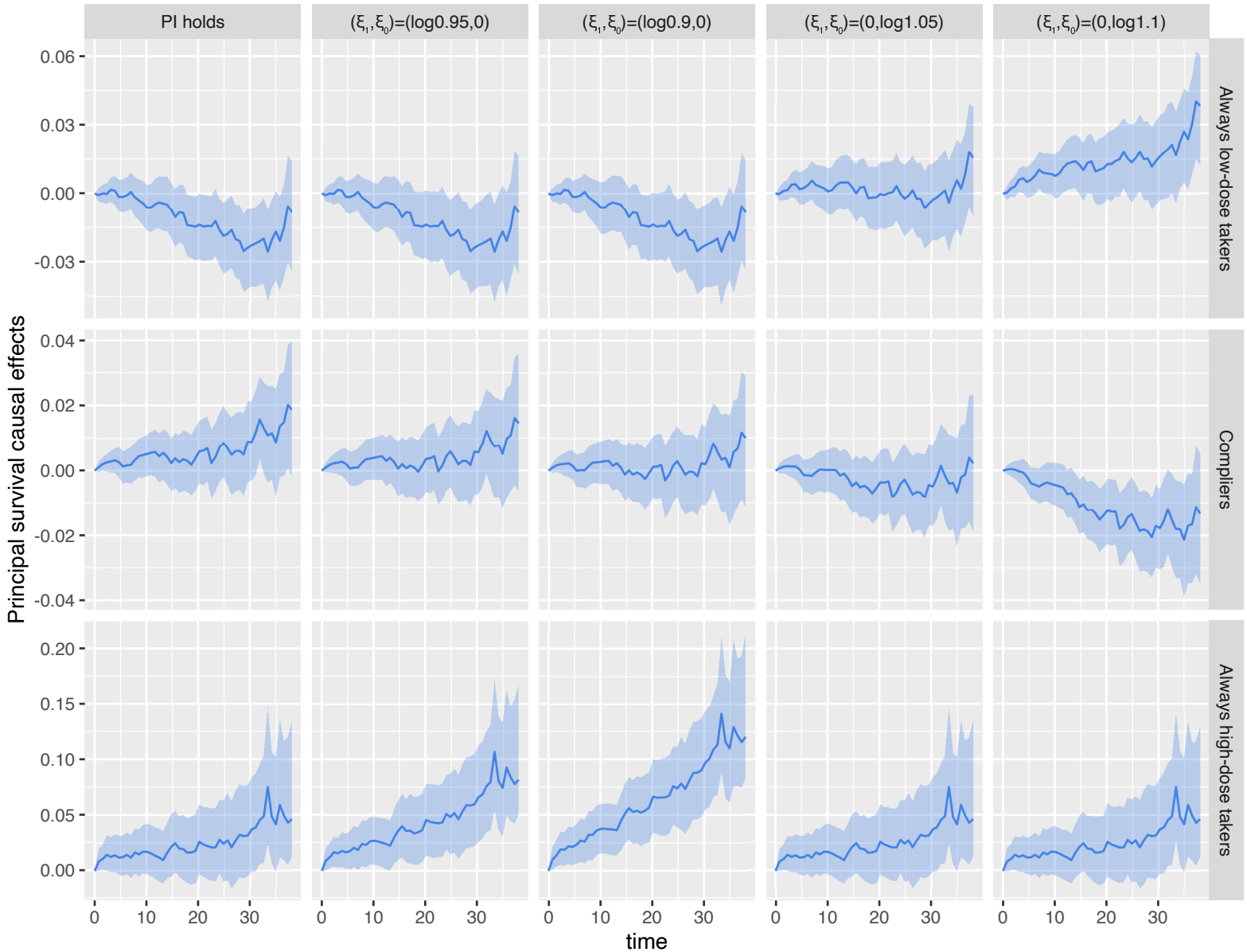
| Scenario | $\xi_1$ | $\xi_0$ | Interpretation |
|---|---|---|---|
| Benchmark | 0 | 0 | PI holds |
| I | − | + | always low-dose takers < compliers < always high-dose takers |
| II | + | − | always high-dose < compliers < always low-dose takers |
| III | − | − | compliers < {always high-dose, always low-dose takers} |
| IV | + | + | {always high-dose, always low-dose takers} < compliers |

Note: 'A<B' means that A is less healthier than B due to unobserved confounders (i.e., $U$).

# Choice of sensitivity parameter

- For fixed $\{\xi_1, \xi_0\}$, we develop a bias-corrected estimator for PSCE, $\widehat{\Delta}_g^{bc}(u)$, which is consistent under $\mathcal{M}_{\pi+e+C} \cup \mathcal{M}_{\pi+T}$.
  - Under randomization, doubly robust ($\mathcal{M}_{e+C} \cup \mathcal{M}_T$)
- A plausible scenario: always low-dose takers often less healthier than compilers; always high-dose taker often the healthiest.
- We choose $\xi_1 \in [\log(0.9), 0]$ and $\xi_0 \in [0, \log(1.1)]$, corresponding to Scenario I in previous table
  - since $\xi_1$ and $\xi_0$ defined based on conditional causal survival function, there is a bound on these values
  - a simple choice but can extend to $\xi_1(\boldsymbol{X})$ and $\xi_0(\boldsymbol{X})$

# Sensitivity analysis: results



# Discussion

- Carried out sensitivity analysis for monotonicity (A4) – results robust to this assumption in ADAPTABLE (back-up slides)
- **Some conclusions under PI**:
  - (1) always high-dose takers appear to benefit
  - (2) compliers slightly benefit
  - (3) always low-dose takers appear not to benefit
- (2) and (3) may change if PI does not hold (depending on how rich the collected baseline covariates are)
  - mechanisms of these effects require further studies
- **Limitations**:
  - composite outcomes
  - complete case analysis without addressing treatment discontinuation

# Acknowledgements


- Collaborators
  - Chao Cheng (Yale)
  - Yueqi Guo (Duke)
  - Bo Liu (Duke)
  - Lisa Wruck (Duke/DCRI)
  - Fan Li (Duke)
- The authors gratefully acknowledge the Patient-Centered Outcomes Research Institute (PCORI) contract **ME-2019C1-16146**. The contents of this presentation do not necessarily represent the view of PCORI. We thank the clinical input and motivating questions from investigators of ADAPTABLE, and in particular, Sunil Kripalani, Schuyler Jones, Hillary Mulder.

# References


Bai, X., Tsiatis, A. A., & O'Brien, S. M. (2013). Doubly-robust estimators of treatment-specific survival distributions in observational studies with stratified sampling. Biometrics, 69(4), 830-839.

Baker, S. G. (1998). Analysis of survival data from a randomized trial with all-or-none compliance: estimating the cost-effectiveness of a cancer screening program. Journal of the American Statistical Association, 93(443), 929-934.

Chen, P. Y., & Tsiatis, A. A. (2001). Causal inference on the difference of the restricted mean lifetime between two groups. Biometrics, 57(4), 1030-1038.

Cuzick, J., Sasieni, P., Myles, J., & Tyrer, J. (2007). Estimating the effect of treatment in a proportional hazards model in the presence of non-compliance and contamination. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 69(4), 565-588.

Ding, P., & Lu, J. (2017). Principal stratification analysis using principal scores. Journal of the Royal Statistical Society. Series B (Statistical Methodology), 757-777.

Frangakis, C. E., & Rubin, D. B. (1999). Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. Biometrika, 86(2), 365-379.

Liu, B., Wruck, L., & Li, F. (2023). Principal stratification with time-to-event Outcomes. arXiv preprint arXiv:2301.07672.

Jones, W. S., Mulder, H., Wruck, L. M., Pencina, M. J., Kripalani, S., Munoz, D., ... & Hernandez, A. F. (2021). Comparative effectiveness of aspirin dosing in cardiovascular disease. New England Journal of Medicine, 384(21), 1981-1990.

# References - cont'd


Jiang, Z., Yang, S., & Ding, P. (2022). Multiply robust estimation of causal effects under principal ignorability. Journal of the Royal Statistical Society Series B: Statistical Methodology, 84(4), 1423-1445.

Loeys, T., & Goetghebeur, E. (2003). A causal proportional hazards estimator for the effect of treatment actually received in a randomized trial with all-or-nothing compliance. Biometrics, 59(1), 100-105.

Nie, H., Cheng, J., & Small, D. S. (2011). Inference for the effect of treatment on survival probability in randomized trials with noncompliance and administrative censoring. Biometrics, 67(4), 1397-1405.

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70(1), 41-55.

Tsiatis, A. A. (2006). Semiparametric theory and missing data.

Wei, B., Peng, L., Zhang, M. J., & Fine, J. P. (2021). Estimation of causal quantile effects with a binary instrumental variable and censored data. Journal of the Royal Statistical Society. Series B, Statistical methodology, 83(3), 559.

Yu, W., Chen, K., Sobel, M. E., & Ying, Z. (2015). Semiparametric transformation models for causal inference in time to event studies with all-or-nothing compliance. Journal of the Royal Statistical Society. Series B, Statistical Methodology, 77(2), 397.

Zeng, S., Li, F., Wang, R., & Li, F. (2021). Propensity score weighting for covariate adjustment in randomized clinical trials. Statistics in medicine, 40(4), 842-858.

Zhang, M., & Schaubel, D. E. (2012). Double-robust semiparametric estimator for differences in restricted mean lifetimes in observational studies. Biometrics, 68(4), 999-1009.

Li, F., Buchanan, A. L., & Cole, S. R. (2022). Generalizing trial evidence to target populations in non-nested designs: Applications to aids clinical trials. Journal of the Royal Statistical Society. Series C, Applied statistics, 71(3), 669.

# Simulation study (ignorable assignment)

- Scenario 1–4: the union model $\mathcal{M}_{\pi+e+C} \cup \mathcal{M}_{\pi+T} \cup \mathcal{M}_{e+T}$ is correctly specified.
- Scenario 5: all models are misspecified.

Table 1: Simulation results of $\widehat{\mathcal{S}}^{mr}_{0,c}(u)$, where $\mathcal{M}_e, \mathcal{M}_\pi, \mathcal{M}_T$ and $\mathcal{M}_C$ indicate whether the working model for principal score, propensity score, time-to-event and time-to-censoring is correctly specified (denoted by a 'T' label) or misspecified (denoted by a 'F' label).

| | $u$ | $\mathcal{S}_{0,c}(u)$ | $\widehat{\mathcal{S}}^{mr}_{0,c}(u)$ | Bias | Monte Carlo SE | Coverage Rate |
|---|---|---|---|---|---|---|
| Scenario 1 | 1 | 0.695 | 0.695 | 0.000 | 0.050 | 0.946 |
| $\mathcal{M}_e$ T | 2 | 0.517 | 0.518 | 0.000 | 0.048 | 0.960 |
| $\mathcal{M}_\pi$ T | 3 | 0.397 | 0.395 | -0.002 | 0.046 | 0.960 |
| $\mathcal{M}_T$ T | 4 | 0.309 | 0.306 | -0.003 | 0.039 | 0.968 |
| $\mathcal{M}_C$ T | 5 | 0.245 | 0.242 | -0.003 | 0.036 | 0.964 |
| Scenario 2 | 1 | 0.695 | 0.693 | -0.002 | 0.048 | 0.960 |
| $\mathcal{M}_e$ T | 2 | 0.517 | 0.517 | 0.000 | 0.048 | 0.972 |
| $\mathcal{M}_\pi$ T | 3 | 0.397 | 0.394 | -0.002 | 0.046 | 0.970 |
| $\mathcal{M}_T$ F | 4 | 0.309 | 0.307 | -0.003 | 0.041 | 0.960 |
| $\mathcal{M}_C$ T | 5 | 0.245 | 0.243 | -0.002 | 0.037 | 0.970 |
| Scenario 3 | 1 | 0.695 | 0.693 | -0.002 | 0.172 | 0.936 |
| $\mathcal{M}_e$ F | 2 | 0.517 | 0.517 | -0.001 | 0.133 | 0.946 |
| $\mathcal{M}_\pi$ T | 3 | 0.397 | 0.394 | -0.003 | 0.108 | 0.956 |
| $\mathcal{M}_T$ T | 4 | 0.309 | 0.306 | -0.004 | 0.084 | 0.966 |
| $\mathcal{M}_C$ F | 5 | 0.245 | 0.241 | -0.003 | 0.068 | 0.960 |
| Scenario 4 | 1 | 0.695 | 0.692 | -0.003 | 0.042 | 0.962 |
| $\mathcal{M}_e$ T | 2 | 0.517 | 0.517 | 0.000 | 0.044 | 0.954 |
| $\mathcal{M}_\pi$ F | 3 | 0.397 | 0.395 | -0.002 | 0.043 | 0.962 |
| $\mathcal{M}_T$ T | 4 | 0.309 | 0.308 | -0.002 | 0.038 | 0.968 |
| $\mathcal{M}_C$ F | 5 | 0.245 | 0.244 | -0.001 | 0.035 | 0.976 |
| Scenario 5 | 1 | 0.695 | 0.761 | 0.066 | 0.024 | 0.246 |
| $\mathcal{M}_e$ F | 2 | 0.517 | 0.600 | 0.083 | 0.029 | 0.228 |
| $\mathcal{M}_\pi$ F | 3 | 0.397 | 0.479 | 0.082 | 0.029 | 0.234 |
| $\mathcal{M}_T$ F | 4 | 0.309 | 0.387 | 0.078 | 0.029 | 0.264 |
| $\mathcal{M}_C$ F | 5 | 0.245 | 0.318 | 0.073 | 0.029 | 0.328 |

# Back-up: Basline table

- Balance check for baseline characteristics in each treatment arm

Table 2: Mean and standard deviation of baseline characteristics stratified by the treatment assignment group, ADAPTABLE trial, 2016–2020.

| Variable | 81-mg Group ($N = 5239$) | 325-mg Group ($N = 4791$) | ASD* |
|---|---|---|---|
| Age (years) | 67.00 (9.45) | 66.77 (9.32) | 0.02 |
| Female sex | 0.29 (0.46) | 0.31 (0.46) | 0.04 |
| White race | 0.84 (0.37) | 0.85 (0.36) | 0.02 |
| Hispanic ethnicity | 0.03 (0.18) | 0.03 (0.17) | 0.01 |
| Non-internet users | 0.15 (0.35) | 0.14 (0.35) | 0.02 |
| Current smokers | 0.09 (0.29) | 0.10 (0.30) | 0.01 |
| P2Y12 inhibitor | 0.23 (0.42) | 0.23 (0.42) | 0.00 |
| **Medical History** | | | |
| Myocardial infarction | 0.37 (0.48) | 0.36 (0.48) | 0.01 |
| Atrial fibrillation | 0.08 (0.27) | 0.09 (0.28) | 0.03 |
| Percutaneous coronary intervention | 0.42 (0.49) | 0.42 (0.49) | 0.01 |
| Coronary artery disease | 0.95 (0.21) | 0.96 (0.20) | 0.04 |
| Coronary-artery bypass grafting | 0.25 (0.44) | 0.25 (0.43) | 0.02 |
| Hypertension | 0.87 (0.34) | 0.87 (0.33) | 0.01 |
| Hyperlipidemia | 0.90 (0.30) | 0.91 (0.29) | 0.01 |
| Peripheral artery disease | 0.24 (0.43) | 0.25 (0.44) | 0.03 |
| Congestive heart failure | 0.23 (0.42) | 0.25 (0.43) | 0.04 |
| History of bleeding | 0.08 (0.28) | 0.09 (0.29) | 0.03 |
| **Aspirin use before trial** | | | |
| Prior dose: 81 mg | 0.85 (0.36) | 0.86 (0.35) | 0.02 |
| Prior dose: 325 mg | 0.12 (0.33) | 0.12 (0.33) | 0.01 |

*ASD: the absolute standardized difference.

# Back-up: Sensitivity for monotonicity

- If monotonicity does not hold, defiers exist and we have four strata $G \in \{c, n, a, d\}$.
- The following sensitivity parameter captures the deviation from monotonicity:

$$\zeta = \frac{\mathbb{P}(G = d|\boldsymbol{X})}{\mathbb{P}(G = c|\boldsymbol{X})},$$

  —ratio between the probability of defiers and compliers given $\boldsymbol{X}$.

- $\zeta$ takes values from 0 to $\infty$, and monotonicity holds with $\zeta = 0$.
  - if we further assume that the treatment assignment has a positive effect on the treatment receipt (i.e., $\mathbb{E}[S(1) - S(0)] \geq 0$), then $\zeta$ is bounded by

$$0 \leq \zeta \leq 1 - \frac{p_1 - p_0}{\min(p_1, 1 - p_0)},$$

    where $p_z = \mathbb{E}[p_z(\boldsymbol{X})]$.
  - in ADAPTABLE, the estimated bound is $\zeta \in [0, 0.103]$.
- For a fixed value $\zeta$, we developed a modified estimator, which is consistent under $\mathcal{M}_{\pi+e+C} \cup \mathcal{M}_{\pi+T} \cup \mathcal{M}_{e+T}$ for any $g \in \{c, n, a, d\}$.

# Back-up: Sensitivity for monotonicity