# Efficient and robust transfer learning of optimal individualized treatment regimes with right-censored survival data

Pan Zhao

April 18, 2023

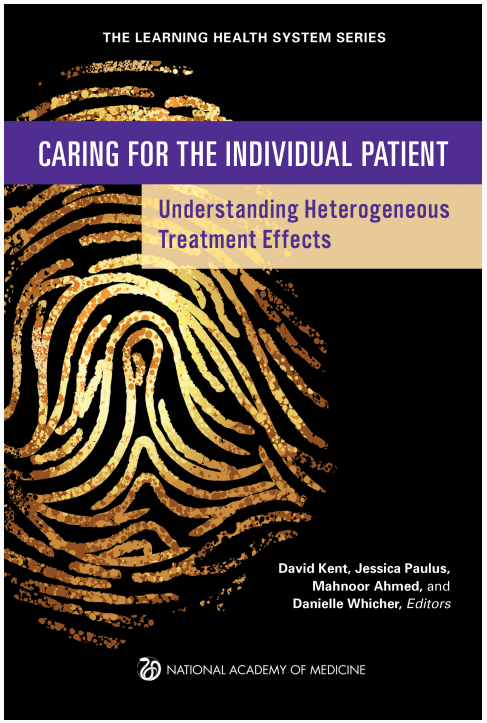
joint work with Julie Josse (Inria) and Shu Yang (NCSU)

https://arxiv.org/abs/2301.05491

# Precision medicine

Caring for the Individual Patient:
Understanding Heterogeneous Treatment Effects [a]

- one-size-fits-all approaches are inadequate
- treatments should be tailored to individuals based on heterogeneity of clinical characteristics and their personal preferences



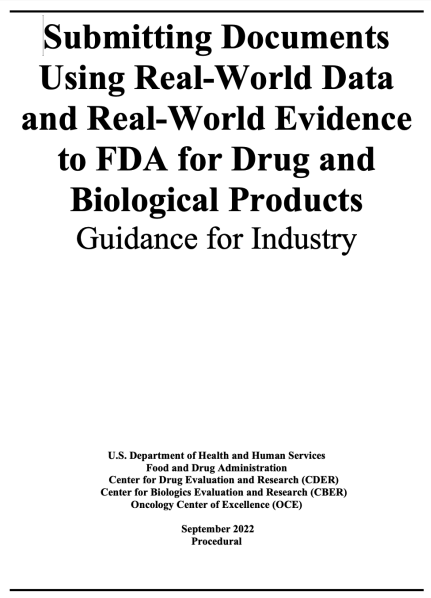

[a] National Academy of Medicine, 2018

An individualized treatment regime (ITR) is a (deterministic) decision rule that assigns personalized treatments based on patients' individual characteristics

# Randomized Control Trial (RCT) & Real-World Data (RWD)

Cures enhances our ability to modernize clinical trial designs, including use of real-world evidence: [a]

Electronic health record (EHR) data, medical claims data, product or disease registry data, etc. [b]


**Submitting Documents Using Real-World Data and Real-World Evidence to FDA for Drug and Biological Products**
Guidance for Industry

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)
Oncology Center of Excellence (OCE)

September 2022
Procedural

[a] 21st Century Cures Act, 2016
[b] FDA Guidance

Complementary features of different data sources:

**Randomized control trial** :

- gold standard (internal validity)
- restrictive inclusion/exclusion criteria, etc.

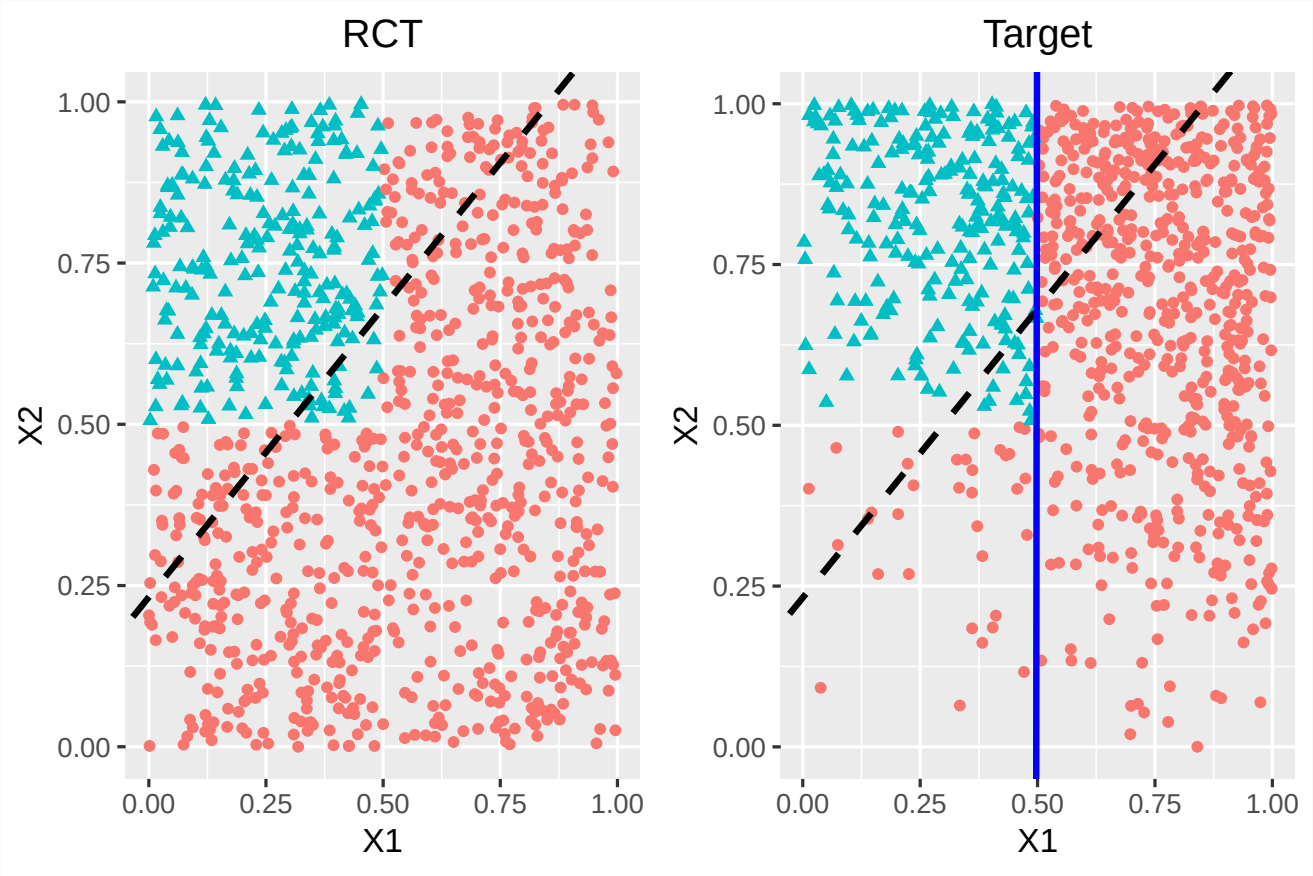
**Observational study (RWD)** :

- representative of **target pop.** (external validity)
- confounding bias

# Transfer learning

**Cannot directly apply the RCT-optimal ITR on another population**

Covariate shift between the RCT and target population




Dash line: RCT-optimal ITR. Blue line: target optimal ITR.
Red points: positive treatment effects

# Summary of contributions

- Heterogeneity within and across populations
- Data integration: e.g. combining randomized trials and observational data (target)

  Goal: learn the optimal ITR that generalizes well to the target population (**transfer learning**)

- For interpretability and transparency, use parametric (linear) ITRs (true optimal ITR not necessarily parametric)
- Clinical outcome: **right-censored survival data**
- Construction of efficient and robust estimators using semi-parametric efficiency theory

# Related works

Large literature on optimal ITR using trial or observational data:

- Q-learning (Robins 2004)
- A-learning (Murphy 2003)
- Direct value search (Zhang et al. 2012)
- Classification perspective (Zhao et al. 2012)
- Tree or list-based ITRs (Laber & Zhao 2015)
- Survival data (Goldberg & Kosorok 2012)

Transfer learning / generalizability ...

- Distributionally robust optimization (Mo et al. 2021)
- Sensitivity analysis (Sahoo et al. 2022)
- Combine trials and observational studies (Colnet et al. 2023)

Limitations: only use a single data; few development for survival data

# Setup: vanilla causal survival analysis

Data:

- covariates $X \in \mathcal{X} \subseteq \mathbb{R}^p$, treatment $A \in \mathcal{A} = \{0, 1\}$
- $T(a)$ potential outcome of survival time under treatment $a$
- censoring time $C$, observed time $U = \min\{T, C\}$, event indicator $\Delta = I\{T \leq C\}$

Consider some deterministic transformation function $y(\cdot)$

- survival probability at time $t$: $y(T) = I\{T \geq t\}$
- restricted mean survival time (RMST): $y(T) = \min(T, L)$ with some pre-specified maximal time horizon $L$

# Setup: vanilla causal survival analysis

Data:

- covariates $X \in \mathcal{X} \subseteq \mathbb{R}^p$, treatment $A \in \mathcal{A} = \{0, 1\}$
- $T(a)$ potential outcome of survival time under treatment $a$
- censoring time $C$, observed time $U = \min\{T, C\}$, event indicator $\Delta = I\{T \leq C\}$

Consider some deterministic transformation function $y(\cdot)$

- survival probability at time $t$: $y(T) = I\{T \geq t\}$
- restricted mean survival time (RMST): $y(T) = \min(T, L)$ with some pre-specified maximal time horizon $L$

Under classical causal assumptions: consistency, positivity, unconfoundedness, conditionally independent censoring,

$\Rightarrow$ Identification of $E[y(T(a))]$ by the outcome regression (OR), inverse probability weighting (IPW), and also doubly robust (DR) formulas

# Setup: ITR and value function

ITR $d(x) : \mathcal{X} \to \mathcal{A}$ a mapping from the covariate space $\mathcal{X}$ to the treatment space $\mathcal{A}$

Define the potential outcome $T(d)$ under any ITR $d \in \mathcal{D}$:

$$T(d) = d(X)T(1) + (1 - d(X))T(0),$$

$\Rightarrow$ The value function of $d$ is $V(d) = E[y(T(d))]$

$\Rightarrow$ The optimal ITR is $d^{\text{opt}} = \arg\max_{d \in \mathcal{D}} V(d)$

# Setup: ITR and value function

ITR $d(x) : \mathcal{X} \to \mathcal{A}$ a mapping from the covariate space $\mathcal{X}$ to the treatment space $\mathcal{A}$

Define the potential outcome $T(d)$ under any ITR $d \in \mathcal{D}$:

$$T(d) = d(X)T(1) + (1 - d(X))T(0),$$

$\Rightarrow$ The value function of $d$ is $V(d) = E[y(T(d))]$

$\Rightarrow$ The optimal ITR is $d^{\text{opt}} = \arg\max_{d \in \mathcal{D}} V(d)$

Consider a class of parametric ITRs $d_\eta$, indexed by $\eta$. Let $V(\eta) = V(d_\eta)$. We focus on **linear ITRs**:

$$\mathcal{D}_\eta = \{d_\eta : d_\eta(X) = I\{\eta^T \tilde{X} \geq 0\}, |\eta_{p+1}| = 1\}, \text{ where } \tilde{X} = (1, X^T)^T$$

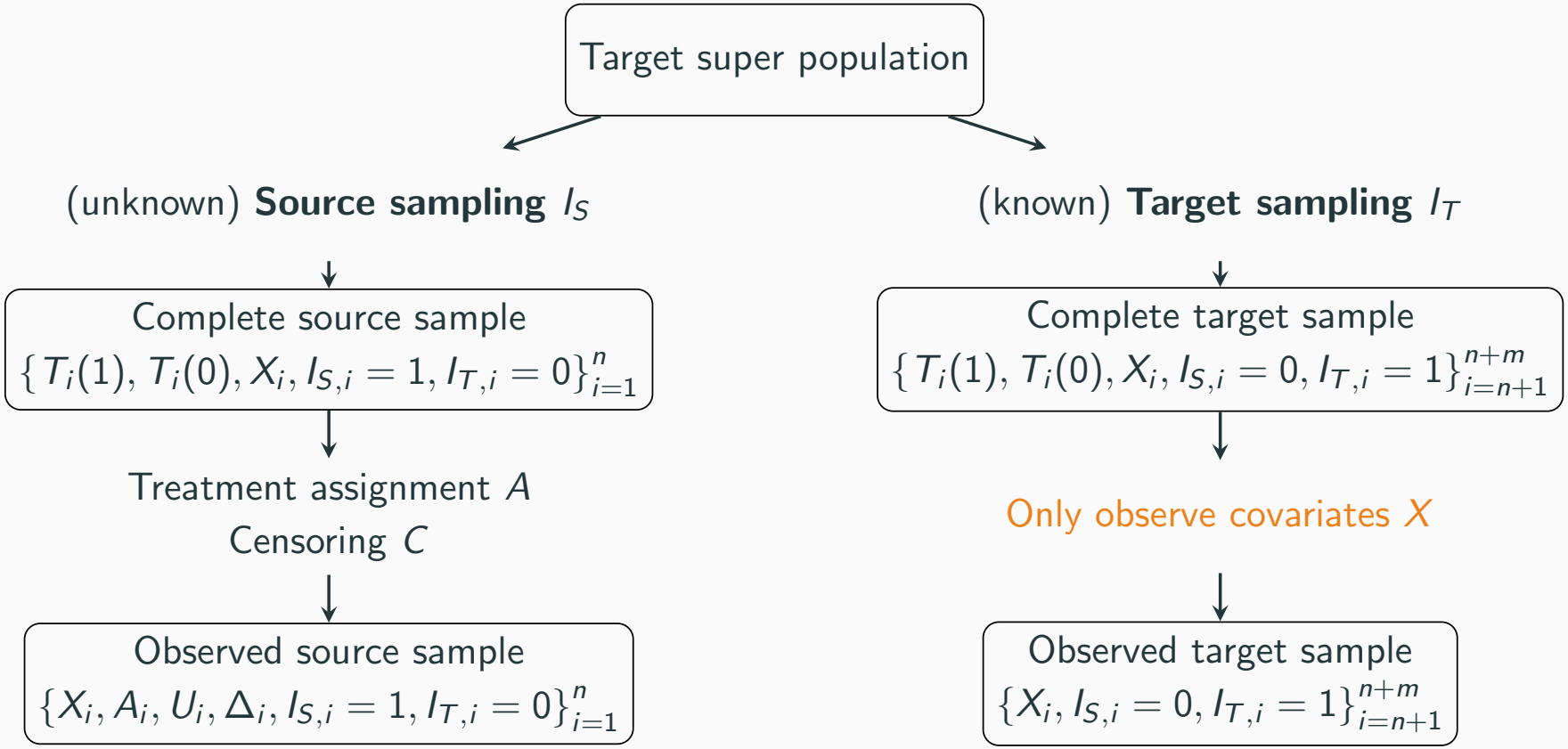
Population parameter $\eta^*$ indexing the optimal ITR:

$$\eta^* = \arg\max V(\eta),$$

and the optimal value function is $V(\eta^*)$

# Setup: illustration of data structure

2 data sources: RCT (source) and observational data (target)




observed time $U = \min\{T, C\}$, event indicator $\Delta = I\{T \leq C\}$

# Identification assumptions for generalization

Assume standard causal assumptions hold for the RCT (source), and

- Survival mean exchangeability
  $E[y(T(a)) \mid X, I_S = 1] = E[y(T(a)) \mid X]$ for every $a \in \mathcal{A}$,
- Positivity of Source Inclusion $0 < Pr(I_S = 1 \mid X) < 1$ almost surely,
- Known target design The target sample design weight
  $e(x) = 1/Pr(I_T = 1 \mid X = x)$ is known by design.

In our framework, we have the key identity that for any $g(X)$

$$E\left[\frac{I_S}{\pi_S(X)} g(X)\right] = E[I_T\, e(X) g(X)] = E[g(X)],$$

where $\pi_S(X) = Pr(I_S = 1 \mid X)$ is the sampling score.

$\Rightarrow$ this motivates the calibration weighting (CW) approach.

# Identification results for generalization

The value function $V(d)$ can be identified by the IPW formulas: [1]

$$E\left[ \underbrace{\frac{I_S}{\pi_S(X)}}_{\text{IPSW/CW}} \underbrace{\frac{I\{A = d(X)\}}{\pi_d(X)} \frac{\Delta y(U)}{S_C(U \mid A, X)}}_{\text{naive IPW}} \right],$$

and the OR formulas:

$$E[\underbrace{I_T\, e(X)}_{\text{target}} \underbrace{E[y(T) \mid A = d(X), X, I_S = 1]}_{\text{Outcome Regression in source}}], \qquad \text{ORt}$$

$$E\left[ \underbrace{\frac{I_S}{\pi_S(X)}}_{\text{CW}} E[y(T) \mid A = d(X), X, I_S = 1] \right], \qquad \text{CW-OR}$$

[1] $\pi_d(x) = Pr(A = d(x) \mid X = x, I_S = 1)$, $S_C(t \mid a, x) = Pr(C > t \mid A = a, X = x)$

# Calibration weighting

**Assign weights $q$ to subjects in the source to empirically match the target population**

Convex optimization problem:

$$\min \sum_{i=1}^{n} q_i \log q_i$$

subject to the balancing constraint:

$$\underbrace{\sum_{i=1}^{n} q_i g(X_i)}_{\text{source}} = \underbrace{\sum_{i=n+1}^{n+m} e(X_i) g(X_i)}_{\text{target}}$$

and $q_i > 0, \sum_{i=1}^{n} q_i = 1$.

- other objective functions can be used
- specify $g(x)$: mean and higher order moments; sieve approximation
- favorable robustness

# An efficient and robust estimation procedure

We also derive the efficient influence function (EIF) of $V(d)$:

- construction of efficient estimator: **Augmented Calibration Weighting** (ACW) not displayed for simplicity
- double robustness: consistent under $\mathcal{M}_1 \bigcup \mathcal{M}_2$
- cross-fitting procedure for flexible ML methods

4 nuisance parameters:

$\mathcal{M}_1$ survival outcome model: $\mu(a, x) = E[y(T) \mid A = a, X = x]$

$\mathcal{M}_2$
- treatment assignment: $\pi_A(x) = Pr(A = 1 \mid X = x)$
- source sampling: $\pi_S(x) = Pr(I_S = 1 \mid X = x)$
- censoring: $S_C(t \mid a, x) = Pr(C > t \mid A = a, X = x)$

Doubly robustness, semiparametric efficiency hold for general ITR $d$ without parametric restriction

# Asymptotic properties I

Assume $S(t;\eta) = Pr(T(d_\eta) > t)$ is twice continuously differentiable in a neighborhood of $\eta^*$, and the margin condition $Pr(0 < |\eta^T \tilde{X}| < \delta) = O(\delta)$, [2] we have

- $\hat{S}(t;\eta) \to S(t;\eta)$ for any $\eta$ and $0 < t \leq L$
- $\sqrt{N}\left\{\hat{S}(t;\eta) - S(t;\eta)\right\}$ converges weakly to a mean zero Gaussian process for any $\eta$
- $N^{1/3} \|\hat{\eta} - \eta^*\|_2 = O_p(1)$
- $\sqrt{N}\left\{\hat{S}(t;\hat{\eta}) - S(t;\eta^*)\right\} \to \mathcal{N}(0, \sigma^2_{t,1})$

Under certain conditions, same results hold for the cross-fitted estimator.

Readily extended to a broad class of functionals of survival distributions (Yang et al. 2021)

[2] There exists some constant $\delta_0 > 0$ such that the big-$O$ term is uniform in $0 < \delta < \delta_0$

# Asymptotic properties II

In comparison with $\hat{V}_{DR}$ the standard doubly robust estimator using only source sample:

1) When **covariate distributions of the source and target populations are the same**, both $\sqrt{N}\{\hat{V}_{DR}(\eta) - V(\eta)\}$ and $\sqrt{N}\{\hat{V}_{CF}(\eta) - V(\eta)\}$ are asymptotically normal with mean zero and same variance.

2) If $d^{\text{opt}} \in \mathcal{D}_\eta$, i.e., $d^{\text{opt}} = d_{\eta^*}$, then despite covariate shift

- both the maximizers of $\hat{V}_{DR}(\eta)$ and $\hat{V}_{CF}(\eta)$ converge to $\eta^*$,
- however, $\hat{V}_{DR}(\eta)$ is a biased estimator of $V(\eta)$.

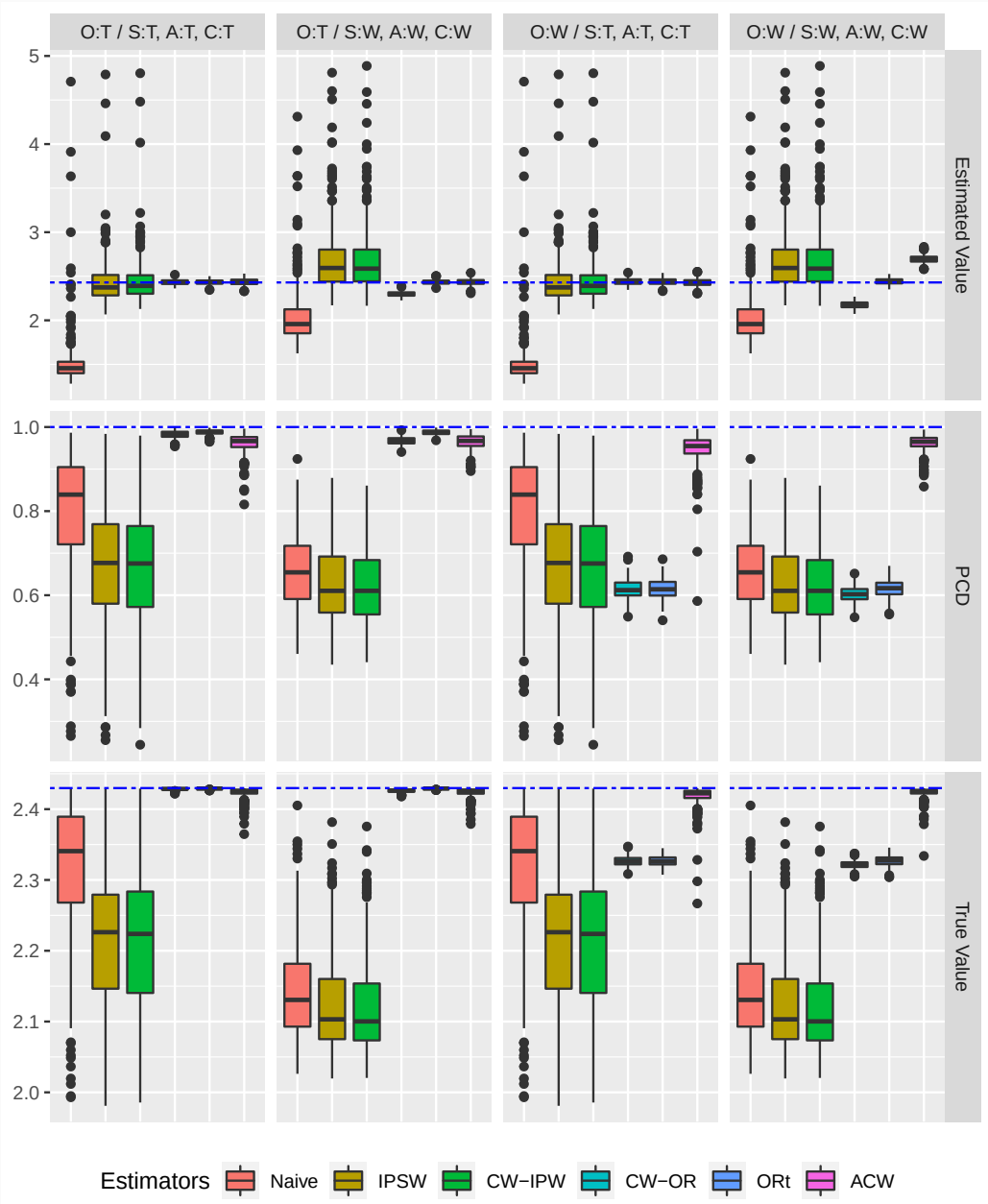
# Simulations: (semi)parametric models

O: outcome
S: source sampling
A: treatment assignment
C: censoring

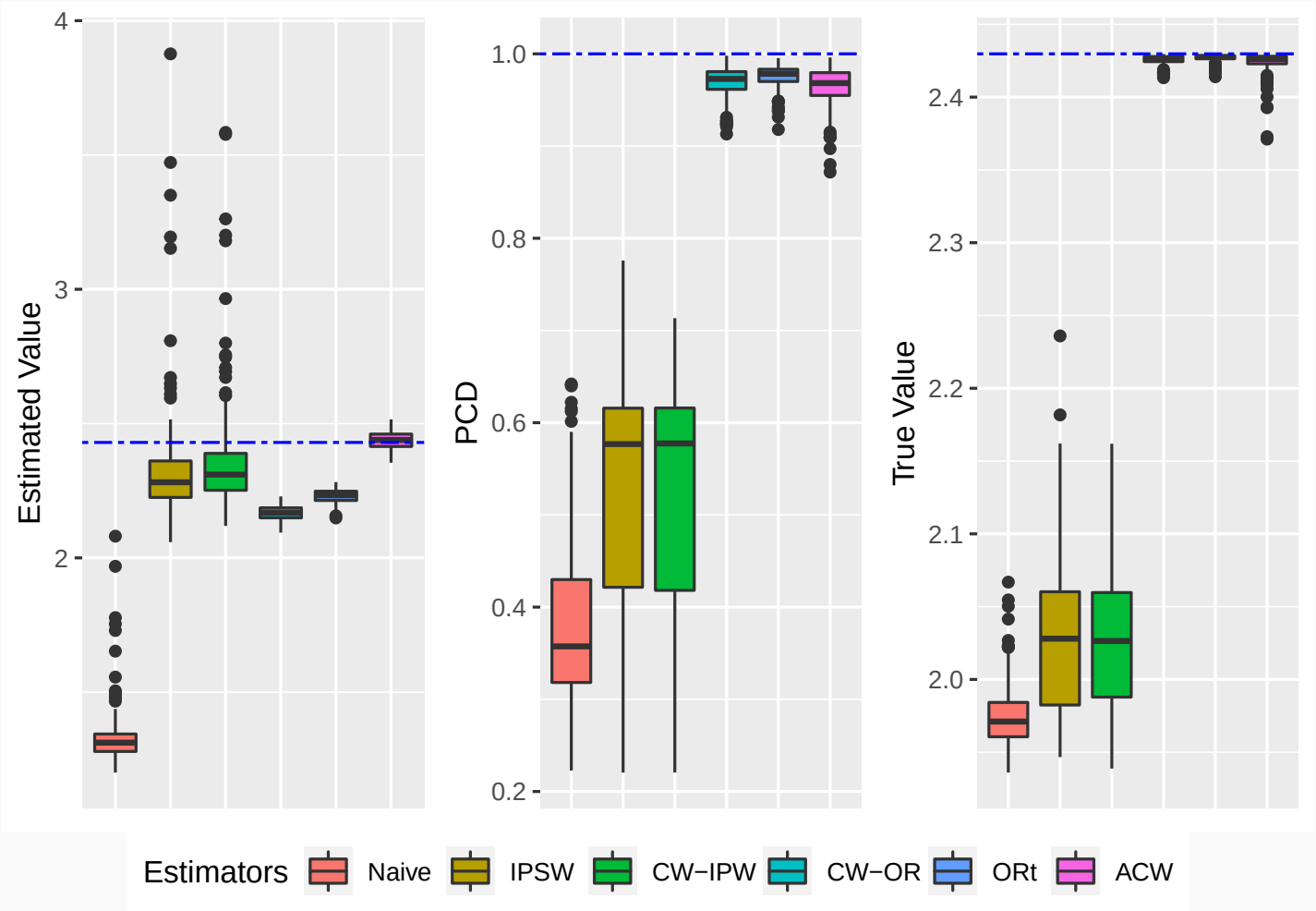
T/W: correctly/mis-specified

**Different methods**:

- Naive: source only
- weighting: IPSW, CW-IPW
- outcome regression: CW-OR, ORt
- **ACW**: our proposal



# Simulations: machine learning methods




Random forest (`grf`) for nuisance parameters estimation

Sample size: $n \simeq 3000, m \simeq 8000$

# Real data application

Sodium bicarbonate therapy for patients with severe metabolic acidaemia in the intensive care unit

- BICAR-ICU: multi-center, open-label, randomized controlled, phase 3 trial
- OS: prospective, multiple-center observational study

| | SEPSIS | AKIN | SOFA | SEX | AGE |
|---|---|---|---|---|---|
| BICAR-ICU ($n = 387$) | 236 (60.98%) | 181 (46.77%) | 10.12 (3.72) | 237 (61.24%) | 63.95 (14.41) |
| OS ($m = 193$) | 99(51.30%) | 75 (38.86%) | 9.10 (4.54) | 122 (63.21%) | 62.73 (17.49) |

Summary of baseline characteristics of the BICAR-ICU trial sample and the OS sample. Mean (standard deviation) for continuous and number (proportion) for the binary covariate.

# Real data application

We consider the class of linear ITRs that depend on five variables

$$\mathcal{D} = \{I\{\eta_1 + \eta_2 \text{SEPSIS} + \eta_3 \text{AKIN} + \eta_4 \text{SOFA} + \eta_5 \text{SEX} + \eta_6 \text{AGE} > 0\} : \eta_1, \ldots, \eta_6 \in \mathbb{R}, |\eta_6| = 1\},$$

with the aim to maximize the RMST within 28 days in ICU stay.

- estimated optimal ITR $\hat{\eta}_{\text{ACW}} = (22.9, -36.1, 87.4, -9.8, 33.7, 1.0)^T$
- estimated value function $\hat{V}(\hat{\eta}_{\text{ACW}}) = 19.52$ days, with 95% confidence interval $[17.74, 21.30]$
- RCT-optimal ITR $\hat{V}(\hat{\eta}_{\text{DR.RCT}}) = 15.37$