A Comprehensive Framework for the Evaluation of Individual Treatment Rules From Observational Data

François Grolleau MD (francois.grolleau@aphp.fr), François Petit PhD, Raphaël Porcher PhD

Université Paris Cité, Centre of Research in Epidemiology and Statistics, INSERM U1153.

Objectives

Individual Treatment Rules (ITRs) are maps $r\colon \mathcal{X} \to \{0;1\}$ assigning patients a treatment option. Our objective is to develop a framework to evaluate the population-level benefit of implementing either i) new ITRs or ii) already partially implemented ITRs using observational data. For this purpose, we introduce the concept of stochastic implementation of deterministic ITRs.

Setup and Notations

Pretreatment covariates: X, binary treatment: A, outcome: Y, the propensity score: $\pi(X) = \mathbb{E}[A|X]$. We introduce $S = 1$ if physicians followed the ITR to prescribe the treatment, $S = 0$ otherwise and the stochastic implementation function $\rho(X) = \mathbb{E}[S|X]$. A patient with covariates X exists in five counterfactual worlds. Random variables and corresponding functions have superscript

- $^{a=1}$ in the world where s/he receives treatment,
- $^{a=0}$ in the world where s/he receives control,
- $^{s=1}$ in the world where the ITR is implemented,
- $^{s=0}$ in the world where it is not implemented,
- * in the world with stochastic implementation.

We denote the propensity score under no implementation $\psi(X) = \mathbb{E}[A^{s=0}|X]$ and the treatment effect function $\tau(x) = \mathbb{E}[Y^{a=1} - Y^{a=0}|X = x]$.

Assumptions (Figure 1)

Consistency assumptions

$$Y^{s=1} = r(X)Y^{a=1} + \big(1 - r(X)\big)Y^{a=0}$$
$$Y^{s=0} = A^{s=0}Y^{a=1} + \big(1 - A^{s=0}\big)Y^{a=0}$$
$$A^{*} = S^{*}A^{s=1} + \big(1 - S^{*}\big)A^{s=0}$$
$$Y^{*} = A^{*}Y^{a=1} + \big(1 - A^{*}\big)Y^{a=0}$$

Exchangeability assumptions

$$\{Y^{a=1}, Y^{a=0}\} \perp\!\!\!\perp A^{s=0}|X$$
$$\{Y^{a=1}, Y^{a=0}\} \perp\!\!\!\perp A^{*}|X$$
$$A^{s=0} \perp\!\!\!\perp S^{*}|X$$

Overlap assumptions

$$0 < \epsilon_{inf} < \pi(x) < \epsilon_{sup} < 1 \text{ and}$$
$$0 < \epsilon_{inf} < \psi(x) < \epsilon_{sup} < 1, \quad \forall\, x \in \mathcal{X}.$$

Estimands of Interest

Average Rule Effect (ARE)

$$\Delta(r) = \mathbb{E}[Y^{s=1} - Y^{s=0}]$$

Average Implementation Effect (AIE)

$$\Lambda(r, \rho) = \mathbb{E}[Y^{*} - Y^{s=0}]$$

Maximal Implementation Gain (MIG)

$$\Gamma(r, \rho) = \mathbb{E}[Y^{s=1} - Y^{*}]$$

We distinguish between two situations:

The new ITR situation: The ITR was just released, and treatment decision was never based on it in the population from our data. Here, $S \equiv 0$, $A = A^{s=0}$, and $Y = Y^{s=0}$ such that $\pi \equiv \psi$.

The partially implemented ITR situation: The ITR was available, and treatment decision was based on it for *some* patients in the population from our data. Here, $S = S^{*}$, $A = A^{*}$, and $Y = Y^{*}$ such that $\rho \equiv \rho^{*}$ and $\pi \equiv \pi^{*}$.

References

[1] B. Zang, A. Tsiatis et al. *A robust method for estimating optimal treatment regimes*, Biometrics (2012)

[2] F. Grolleau, R. Porcher et al. *Personalization of renal replacement therapy initiation*, Critical Care (2022)

[3] M. Jordan, R. Jacobs. *Hierarchical mixtures of experts and the EM Algorithm*, Neural Computation (1994)

New ITR Situation

Letting $\mathcal{C}_i^r = \mathbb{1}\{r(X_i) = A_i\}$ as in [1], we obtain a double-robust estimator for the ARE:

$$\widehat{\Delta}_{AIPW}(r) = n^{-1}\sum_{i=1}^{n}\Big\{\frac{\mathcal{C}_i^r Y_i}{\hat{\pi}(X_i)\mathcal{C}_i^r + \{1-\hat{\pi}(X_i)\}(1-\mathcal{C}_i^r)} - \frac{\mathcal{C}_i^r - [\hat{\pi}(X_i)\mathcal{C}_i^r + \{1-\hat{\pi}(X_i)\}(1-\mathcal{C}_i^r)]}{\hat{\pi}(X_i)\mathcal{C}_i^r + \{1-\hat{\pi}(X_i)\}(1-\mathcal{C}_i^r)}\hat{\mathbb{E}}[Y|X = X_i, A = r(X_i)] - Y_i\Big\}.$$

This estimator exhibits regularity, unbiasedness and consistency. Its asymptotic variance can be derived from M-estimation theory while bootstrap confidence intervals achieve nominal coverage. We propose to numerically explore the population-level benefit the ITR will have under different stochastic implementation schemes via $\widehat{\Lambda}_{ITE}(r, \rho^{*}) = n^{-1}\sum_{i=1}^{n}\rho^{*}(X_i)\{r(X_i) - \hat{\pi}(X_i)\}\hat{\tau}(X_i)$ and $\widehat{\Gamma}_{ITE}(r, \rho^{*}) = n^{-1}\sum_{i=1}^{n}\{1 - \rho^{*}(X_i)\}\{r(X_i) - \hat{\pi}(X_i)\}\hat{\tau}(X_i)$. Below, we illustrate our approach on the MIMIC-III database for a newly released ITR concerned with the initiation of renal replacement therapy in patients with acute kidney injury [2].

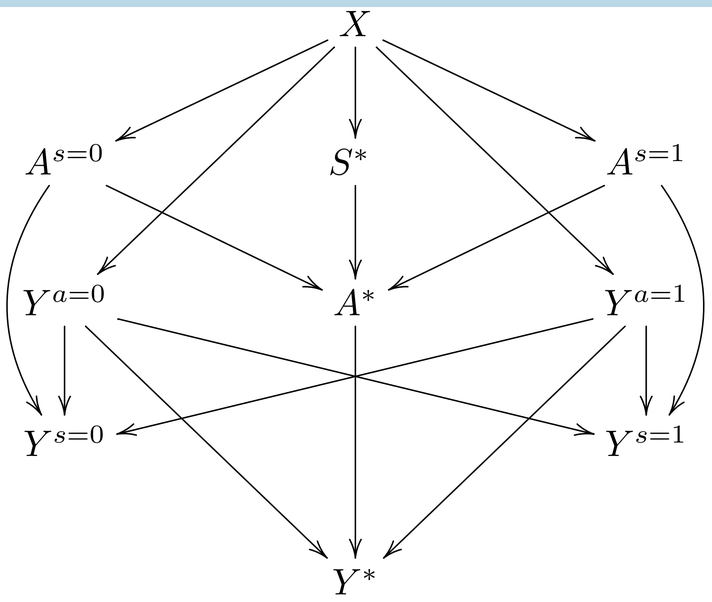

Figure 1: The probabilistic graphical model associated with the data generating mechanism in the world where the ITR is stochastically implemented.

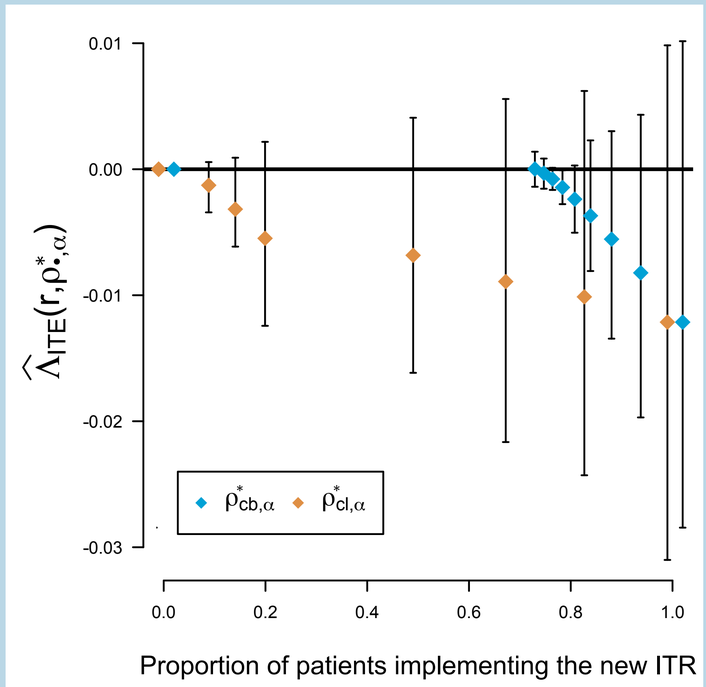

Figure 2: Evaluation of a new ITR.

In **Figure 2**, blue diamonds evaluate the ITR under a cognitive bias implementation scheme (i.e., $\rho^{*}_{cb,\alpha}(x) = \{1 - |r(x) - \pi(x)|\}^{\frac{1}{2}\log\frac{\alpha+1}{1-\alpha}}$) while orange diamonds evaluate the ITR under a confidence level implementation scheme (i.e., $\rho^{*}_{cl,\alpha}(x) = \mathbb{1}[\{\tilde{\tau}(x) - q_{1-\alpha/2}se_{\tilde{\tau}(x)}\}\{\tilde{\tau}(x) + q_{1-\alpha/2}se_{\tilde{\tau}(x)}\} > 0]$).

Partially Implemented ITR Situation

In this situation, S is a latent variable. This means that on top of the fundamental problem of causal inference, we face the challenge of unknown stochastic implementation functions ψ. Our estimation approach is built upon the result below.

Lemma. *In the partially implemented situation, the propensity score has the following form*

$$\pi(x) = \rho(x)r(x) + \{1 - \rho(x)\}\psi(x).$$

The procedure we propose to estimate the ARE and AIE relies on estimating ψ using a mixture of experts (**Figure 3**) fitted via an expectation-maximization algorithm [3]. To investigate the properties of our estimators, we conduct Monte Carlo simulations where we vary the sample size and data generating process. Our estimators appear unbiased and consistent (**Figure 4**) while bootstrap confidence intervals achieve nominal coverage. On the MIMIC-III database, we illustrate our approach for a partially implemented ITR that recommends renal replacement therapy initiation in the most severe patients. (**Figure 5**).

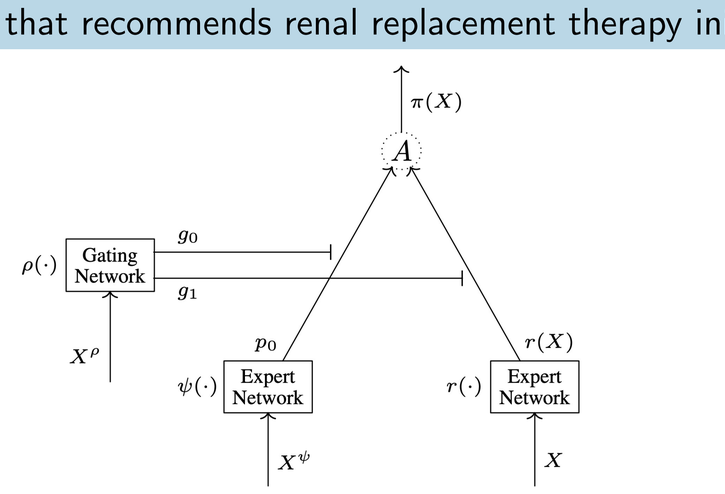

Figure 3: Graphical representation for the mixture of experts used to estimate ψ.

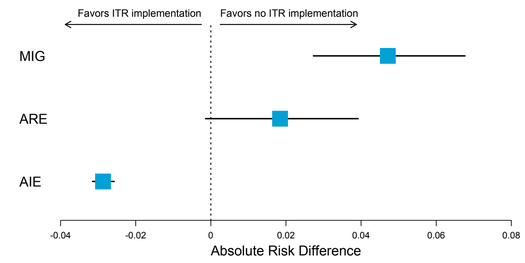

Figure 5: Evaluation of a partially implemented ITR.

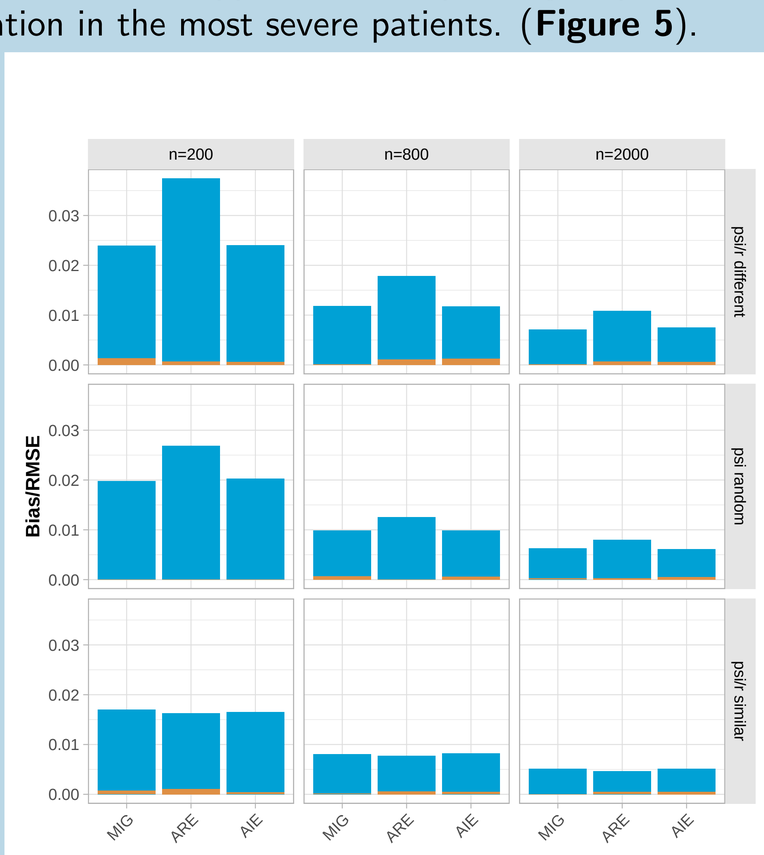

Figure 4: Results of Monte-Carlo simulations.