Estimation of ATE (average treatment effect) in Causal Survival Analysis : Practical recommendations

Charlotte VOINOT[1,2], Julie JOSSE[1], Bernard Sebastien[2]
[1] INRIA, Montpellier, France [2] R&D Sanofi, Gentilly, France

Introduction

- Causal survival inferences: mixed between causal inference vs survival analysis. It assesses the **causal effect** of a **treatment** on an outcome which is a **time until an event** occurs in the presence of **censoring**
- HR mainly used in survival analysis but not a causal measure and assumes proportional hazard.

Objectives:

- Comprehensive overview of the different estimators & Implementations
- Practical recommandation for users

Causal estimand

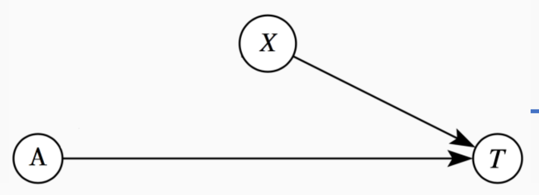
In grey, the observed data $(X_i, A_i, \Delta_i, \widetilde{T}_i)$ with $\widetilde{T}_i = \min(T_i, C_i)$

Covariates		Treatment	Censoring	Status	Outcomes		
X_1	X_2	A	C	Δ	T(0)	T(1)	$\widetilde{T}$
1	24	1	?	1	?	200	200
2	52	0	?	1	100	?	100
1	33	1	200	0	?	?	200

Difference in RMST: Average treatment effect in survival analysis

$$\theta_{RMST}(\tau) = E[\min(T(1),\tau) - \min(T(0),\tau)] = \int_0^\tau S_1(t) - S_0(t)\, dt$$

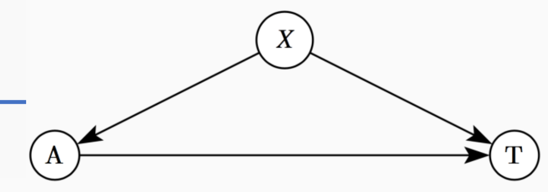
Identifiability assumptions & Estimators



RCT

S.T.U.V.A. $T = AT(1) + (1-A)T(0)$

Observational study



RCT & Independent censoring

- **Random treatment assignment** $A \perp\!\!\!\perp (T(0), T(1), C, X)$
- **Independent censoring** $C \perp\!\!\!\perp T(0), T(1), X, A$

Non-adjusted Kaplan-Meier estimator [1]

$$\hat{S}_{KM}(t|a) = \prod_{j=1, t_j \le t} 1 - \frac{\sum_i I\{\widetilde{T}_i = t_j, \Delta_i = 1, A_i = a\}}{\sum_k I\{\widetilde{T}_k \ge t_j, A_i = a\}}$$

RCT & Dependent censoring

- **Random treatment assignment** $A \perp\!\!\!\perp (T(0), T(1), C, X)$
- **Conditionally independent censoring** $C \perp\!\!\!\perp T(0), T(1) | X, A$
- **Positivity for censoring** $0 < P(C > t \mid X = x, A = a) < 1$

IPCW Kaplan-Meier estimator [2]

$$\hat{S}_{IPCW}(t|a) = \prod_{j=1, t_j \le t} 1 - \frac{\sum_i \widehat{w_i}(t_j, X_i).I\{\widetilde{T}_i = t_j, \Delta_i = 1, A_i = a\}}{\sum_k \widehat{w_k}(t_j, X_k).I\{\widetilde{T}_k \ge t_j, A_i = a\}}$$

with $\widehat{w}_i(t, X_i) = \frac{\Delta^{\tau}{}_i}{\widehat{S_c}(t|X_i, A_i)}$

- $\widehat{S_c}(t|X_i, A_i)$ is the survival function of remain uncensored given the covariate X in the treatment arm A=a.
- $\Delta^\tau = I\{T \wedge \tau < C | A = 1\}$ is the status of the individual truncated at τ.

Every uncensored observation is weighted by the inverse of the probability of remain uncensored given the covariates.

Observational & Independent censoring

- **Unconfoundedness** $A \perp\!\!\!\perp (T(0), T(1)) | X$
- **Positivity for treatment** $1 > P(A = a \mid X = x) > 0$
- **Independent censoring** $C \perp\!\!\!\perp T(0), T(1), X, A$

IPTW Kaplan-Meier estimator [3]

$$\hat{S}_{IPCW}(t|a) = \prod_{j=1, t_j \le t} 1 - \frac{\sum_i \widehat{w_i}(t_j, X_i).I\{\widetilde{T}_i = t_j, \Delta_i = 1, A_i = a\}}{\sum_k \widehat{w_k}(t_j, X_k).I\{\widetilde{T}_k \ge t_j, A_i = a\}}$$

with $\widehat{w}_i(t, X_i) = \frac{A_i}{\hat{e}(X_i)} + \frac{1-A_i}{1-\hat{e}(X_i)}$

Observational & Dependent censoring

- **Unconfoundedness** $A \perp\!\!\!\perp (T(0), T(1)) | X$
- **Positivity for treatment** $1 > P(A = a \mid X = x) > 0$
- **Conditionally independent censoring** $C \perp\!\!\!\perp T(0), T(1) | X, A$
- **Positivity for censoring** $0 < P(C > t \mid X = x, A = a) < 1$

IPTW-IPCW Kaplan-Meier estimator [4]

$\hat{S}_{IPTW-IPCW}$ corresponds to a weighted KM with $\widehat{w}_i(t, X_i) = \frac{\Delta^{\tau}{}_i}{\widehat{S_c}(t|X_i, A_i)}\left(\frac{A_i}{\hat{e}(X_i)} + \frac{1-A_i}{1-\hat{e}(X_i)}\right)$

G-formula plug-in estimator [5]

$$\hat{\theta}_{g-formula}(\tau) = \frac{1}{n}\sum_{1}^{n}[\hat{F}(X_i, A = 1) - \hat{F}(X_i, A = 0)]$$

with $\hat{F}(x, a) = E[T \wedge \tau | X = x, A = a]$

AIPTW-AIPCW estimator [6][7]

$\hat{\theta}_{AIPTW-AIPCW}$ is an **augmented** estimator of $\hat{\theta}_{IPTW-IPCW}$:

- 3 **nuisance models** to compute
- Consistent if at least one of nuisance parameter is consistent (**Double robust**)
- Parametric **convergence** rate

Methodology

Simulations of various scenario:

- Type of study: RCT or Observational study.
- Type of censoring: Independent or Dependent censoring.
- DGP: Parametric (well estimated by cox) and Non-parametric (well estimated by forest).

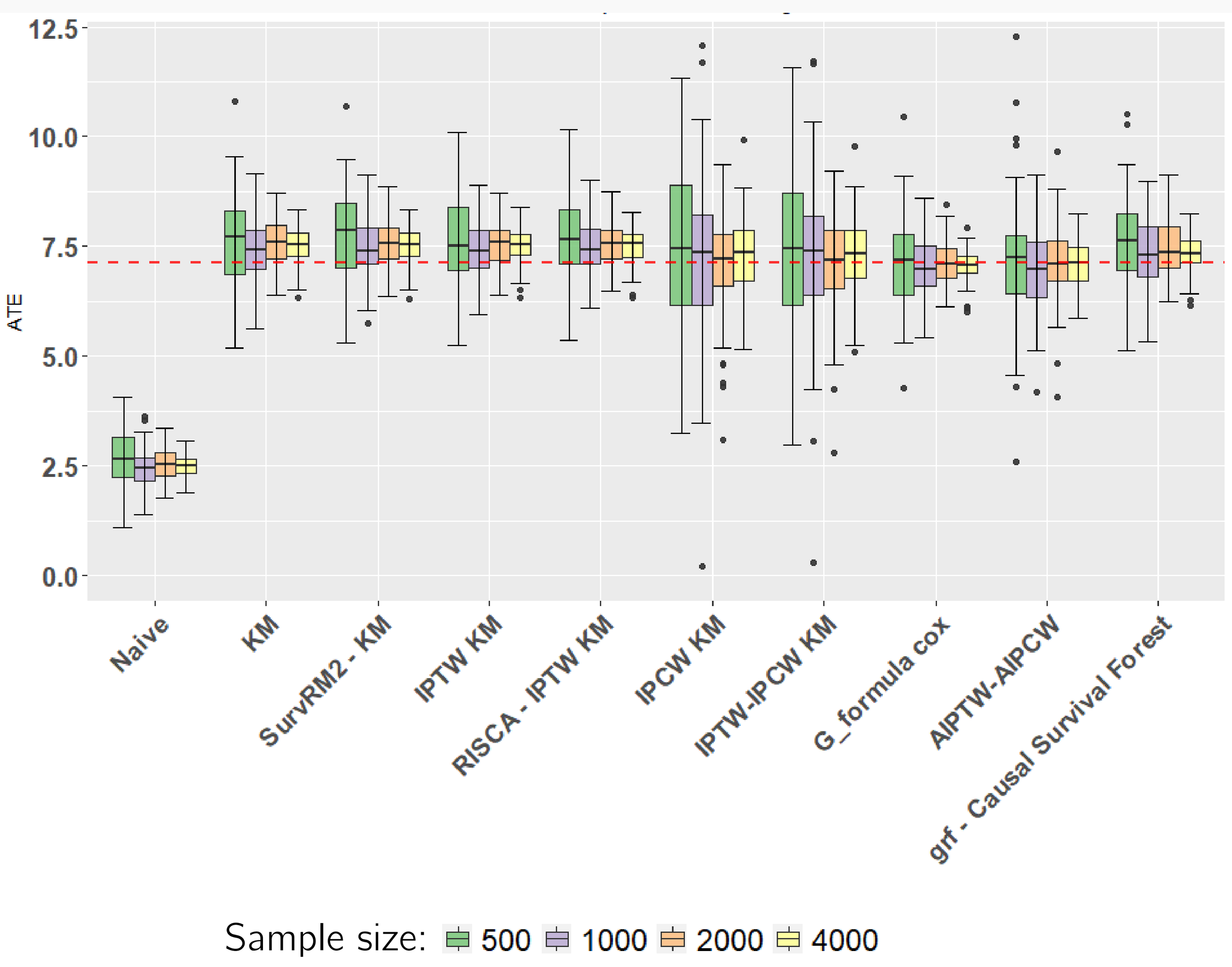
Evaluate and compare efficiency of:

- Self-implemented RMST estimators & Naive estimator $(E(\min(\widetilde{T}(1),\tau) - \min(\widetilde{T}(0),\tau))$.
- RMST Estimators from R-Packages:
  - IPTW Kaplan-Meier from RISCA.
  - Unadjusted Kaplan-Meier from SurvRM2.
  - Causal survival forest from grf.

Simulations results

Boxplot of RMST results from different estimators under well specified nuisance models (nuisance for propensity = « glm », conditional survival & conditional censoring =« cox », number of simulations = 150 and τ=25)

RCT & Dependent censoring: Parametric simulation



- **Biased:**

Naive,
Kaplan Meier (KM),
Rmst2() from SurvRM2,
IPTW Kaplan Meier,
ipw.survival() from RISCA,
Causal survival forest.

- **Unbiased:**

IPCW Kaplan Meier,
IPTW-IPCW Kaplan Meier,
🏆 G-formula,
AIPTW-AIPCW.

Boxplot of RMST results from different estimators under well specified nuisance models (nuisance for propensity = « probability forest », conditional survival & conditional censoring =« survival forest », number of simulations = 150, n.folds = 5, and τ=50)

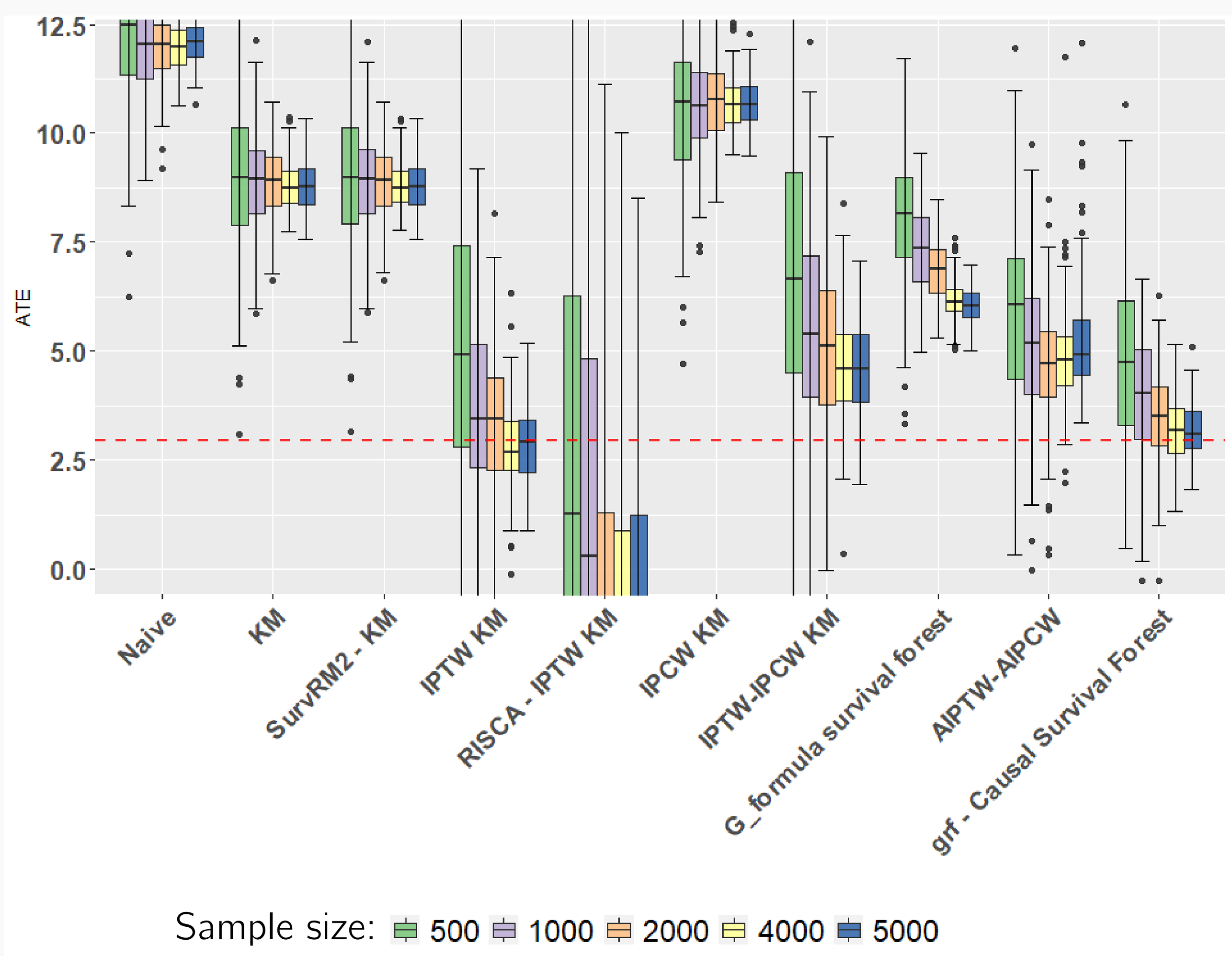
Observational & Dependent censoring: Non-parametric simulation

- **Biased:**

Naive,
Kaplan Meier (KM),
Rmst2() from SurvRM2,
ipw.survival() from RISCA,
IPCW Kaplan Meier,
IPTW-IPCW Kaplan Meier,
G-formula,
AIPTW-AIPCW.

- **Unbiased:**

IPTW Kaplan Meier
🏆 Causal survival forest.



Conclusions & Perspectives

- Few packages available (RISCA, SurvRM2, grf).
- **G-formula** has the **lowest variance** when conditional survival model is **well specified** in parametric simulation. In complex simulation, it converges slowly with survival forest (need a lot of observations).
- **Causal survival forest** is **accurate** for **complex** setting and large sample size and have nice theoretical properties.

REFERENCES

[1] E. L. Kaplan and Paul Meier. "Nonparametric Estimation from Incomplete Observations". In: Journal of the American Statistical Association (1958).
[2] James M. Robins and Dianne M. Finkelstein. "Correcting for Noncompliance and Dependent Censoring in an AIDS Clinical Trial with Inverse Probability of Censoring Weighted (IPCW) Log-Rank Tests". In: Biometrics (2000).
[3] Jun Xie and Chaofeng Liu. "Adjusted Kaplan–Meier estimator and log-rank test with inverse probability of treatment weighting for survival data". In: Statistics in Medicine (2005).
[4] Schaubel, Douglas E., and Guanghui Wei. "Double Inverse-Weighted Estimation of Cumulative Treatment Effects under Nonproportional Hazards and Dependent Censoring." Biometrics 67 (2011).
[5] Robins, James. "A New Approach to Causal Inference in Mortality Studies with a Sustained Exposure Period—Application to Control of the Healthy Worker Survivor Effect." Mathematical Modelling 7 (1986).
[6] Rubin, Daniel, and Mark J. van der Laan. "A Doubly Robust Censoring Unbiased Transformation", The International Journal of Biostatistics (2007).
[7] James M. Robins, Andrea Rotnitzky, and Lue Ping Zhao. "Estimation of Regression Coefficients When Some Regressors Are Not Always Observed." Journal of the American Statistical Association (1994).

Research funding from the Association Nationale Recherche Technologie. CV and BS are employees of Sanofi and may hold stock/stock options. JJ: Nothing to disclose.