Correcting Informative Censoring in Treatment Effect Estimation with Large Scale, Real-World Health Data

Hsin Yi Chen, BS[1], Ruochong Fan, MA[2], Linying Zhang, PhD[2], George Hripcsak, MD, MS[1]

[1]Department of Biomedical Informatics, Columbia University Irving Medical Center, New York, NY, USA

[2]Institute for Informatics, Data Science and Biostatistics, Washington University in St. Louis, St. Louis, MO, USA

Background

- **Censoring** may occur in observational studies when patients are lost to follow-up or do not experience the event of interest**.** If the timing of censoring is related to the risk of the outcome, it can introduce bias: this is **informative censoring** (Figure 1)
- Failure to adjust for informative censoring can bias effect estimates, as the censored population may systematically differ from the uncensored population in ways related to the outcome risk [1, 2]
- **Inverse probability of censoring weighting (IPCW)** addresses this issue by adjusting for censored subjects with additional weights such that it creates a pseudo-population that would have been observed if censoring had not occurred [3, 4]
- IPCW is analogous to inverse probability of treatment weighing (IPTW) for confounding adjustment, which reduces confounding bias by balancing patient characteristics across treatment groups
- While IPCW has been well-studied in isolation, relatively few studies have explored the combined use of IPTW and IPCW to simultaneously reduce both confounding and censoring biases in treatment effect estimation

Hypothesis: Combining IPTW (confounding adjustment) and IPCW (censoring adjustment) weights will *more effectively* reduce bias in treatment effect estimation in high-dimensional, real-world studies

Methods

- We compared the effects of two classes of anti-hypertensive drugs (ACE inhibitors and thiazide diuretics) on one clinical outcome, acute myocardial infarction (AMI), using a retrospective, observational, comparative new-user cohort design
- We had a total cohort size of 41,996 and a total of 49,104 covariates
- We plotted treatment-stratified Kaplan-Meier survival curves and estimated the comparative treatment effect with Cox proportional hazards models using 4 different weighing methodologies: (1) Unadjusted (no confounding or censoring adjustment), (2) IPTW alone (confounding adjustment), (3) IPCW alone (censoring adjustment) and the (4) combined IPTW+IPCW weights (both confounding and censoring adjustment). Treatment effect estimates were then compared to a published large-scale meta-analysis [5]
- Full workflow diagram of the models we fit can be seen in Figure 2

Results

- Adjusting for censoring alone (IPCW) led to improvements from unadjusted estimates, but the combined weights (IPCW and IPTW) resulted in estimates that were closest to the hazard ratio of the published meta-analysis (Figure 3). However, the improvement from using the combined weights (IPTW+IPCW) compared to IPTW alone was not significant
- Results from plotting the survival curves across the two treatment groups was consistent with the hazard ratio results (Figure 4)

Conclusions

- Adjusting for both censoring and confounding may provide a more unbiased estimate for hazard ratios, but addressing censoring alone is not enough when confounding is present
- While censoring adjustment is theoretically important, but may not be as pervasive of a problem as we thought when it comes to treatment effect estimation, since IPTW inadvertently balances on censoring variables
- Perhaps, where there are factors that influence health outcomes but not treatment decision-making (ex: social determinants of health), censoring adjustment will play a more critical role

References

1. S. W. Lagakos, "General Right Censoring and Its Impact on the Analysis of Survival Data," Biometrics, vol. 35, no. 1, pp. 139–156, 1979.
2. K.-M. Leung, R. M. Elashoff, and A. A. Afifi, "Censoring Issues in Survival Analysis," Annual Review of Public Health, vol. 18, no. 1, pp. 83–104, 1997
3. J. M. Robins and D. M. Finkelstein, "Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests," Biometrics, vol. 56, no. 3, pp. 779–788, Sep. 2000.
4. A. L. Buchanan, M. G. Hudgens, S. R. Cole, B. Lau, and A. A. Adimora, "Worth the Weight: Using Inverse Probability Weighted Cox Models in AIDS Research," AIDS Res Hum Retroviruses, vol. 30, no. 12, pp. 1170–1177, Dec. 2014.
5. Suchard MA, Schuemie MJ, Krumholz HM, You SC, Chen R, Pratt N, et al. Comprehensive comparative effectiveness and safety of first-line antihypertensive drug classes: a systematic, multinational, large-scale analysis. The Lancet. 2019 Nov;394(10211):1816–26.

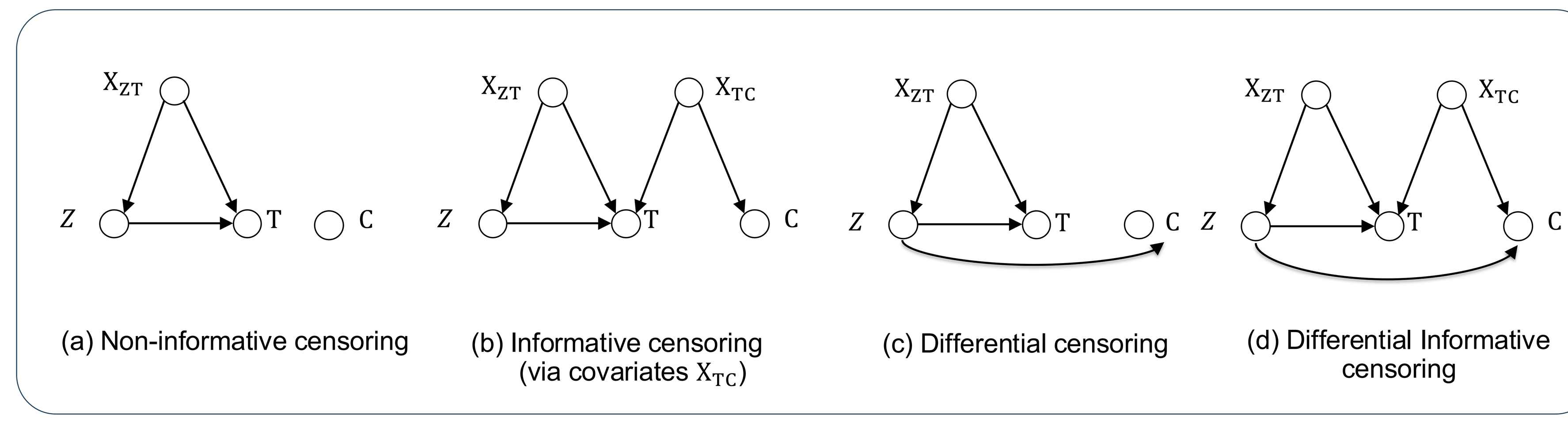


Figure 1. Causal diagrams for survival analyses where Z is the treatment, T is the outcome time, C is the censoring time, and X is a set of covariates which can affect treatment, outcome, and/or censoring. (a) Non-informative censoring in the presence of confounding (X_{ZT} denotes confounders). (b) Informative censoring (via covariates X_{TC}) in the presence of confounding. (c) Differential censoring (informative censoring via treatment) in the presence of confounding. (d) Differential Informative censoring, where there is informative censoring through covariates and treatment, in the presence of confounding.

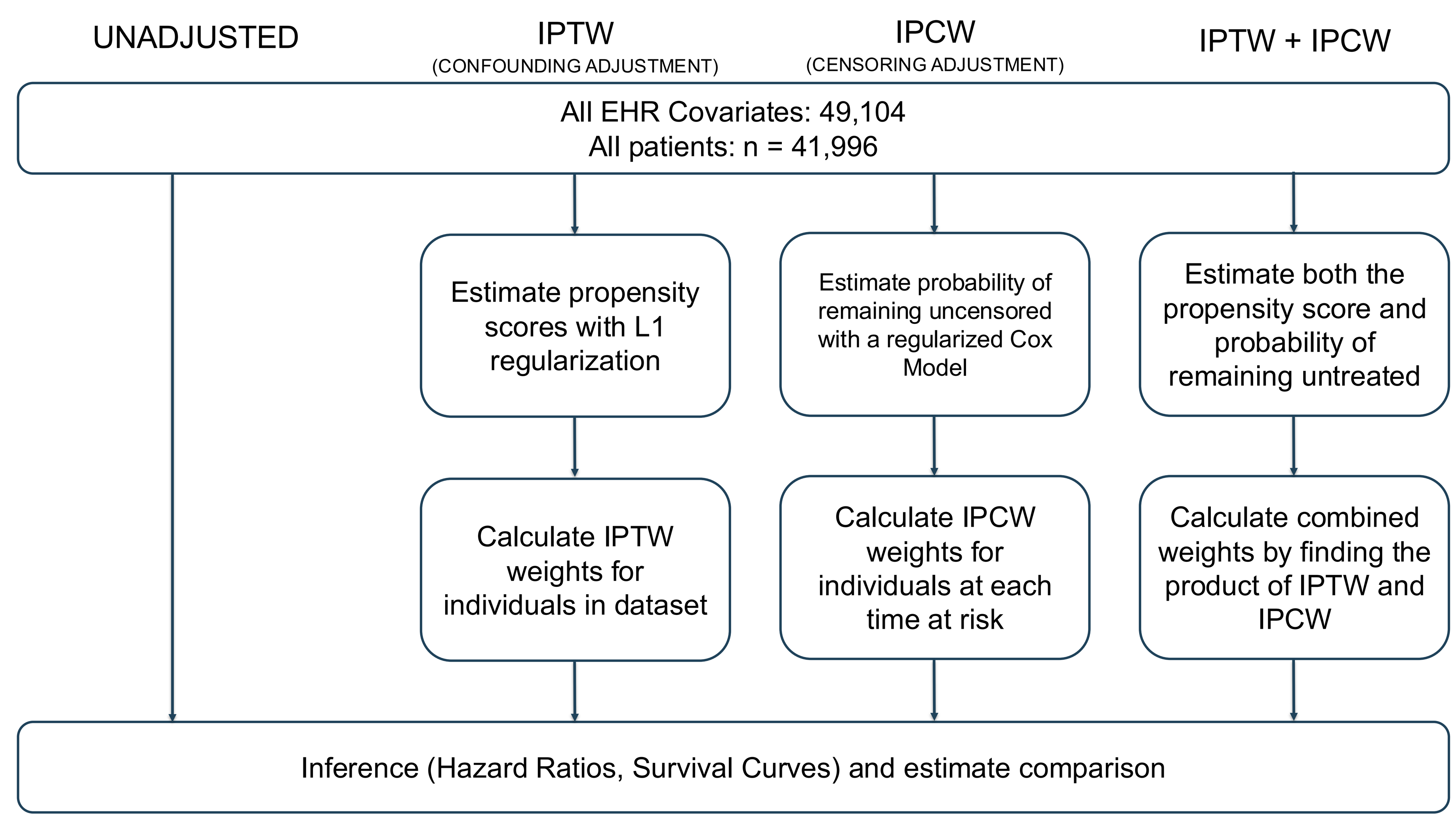


Figure 2. Workflow diagram of our empirical study. We compared four different methods, (1) unadjusted, (2) confounding adjustment with IPTW weights, (3) censoring adjustment with IPCW weights, and (4) combined censoring and confounding adjustment, using the product of IPTW and IPCW weights

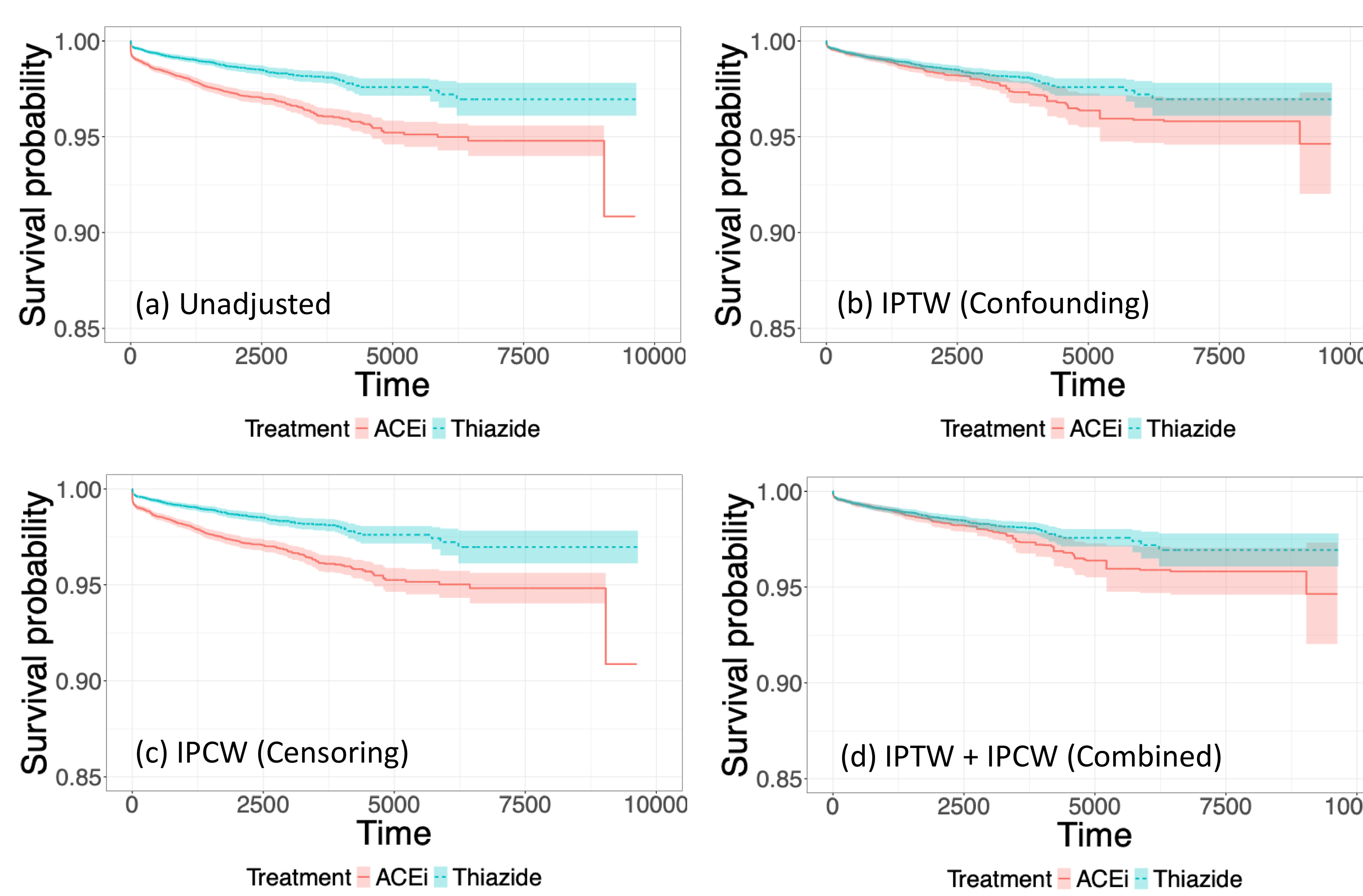
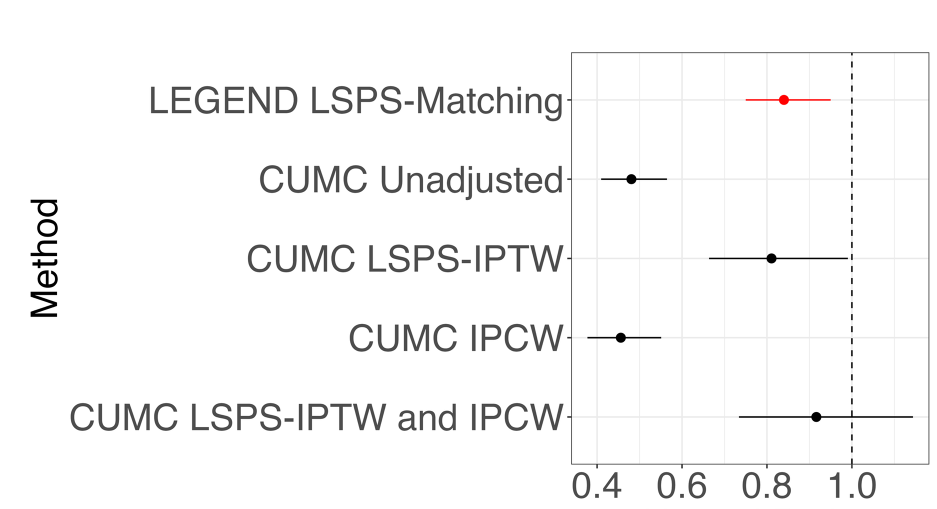


Figure 3. Unadjusted survival curves (a) did not use any weights, while survival curves are weighted by (b) IPTW, (c) IPCW, and (d) combining IPTW and IPCW. The two treatment groups were thiazide diuretics (dotted line) and ACE inhibitors (solid line). Shaded areas denote the 95% confidence interval.



Model	Hazard Ratio [95% CI]
"Gold standard": LEGEND-HTN study (Suchard et al. 2019)	0.84 [0.75, 0.95]
Unadjusted	0.48 [0.41, 0.56]
IPTW (Confounding adjustment only)	0.81 [0.66, 0.99]
IPCW (Censoring adjustment only)	0.45 [0.38, 0.55]
IPCW and IPTW (Confounding and censoring adjustment)	0.92 [0.73, 1.14]

Figure 4. Results of our study comparing four different weighing methodologies to estimate treatment effects. The gold standard model under comparison is the large-scale meta-analysis comparing the same drugs (Thiazide diuretics vs. ACE inhibitors) using data from 20 million patients across 9 databases and 4 countries. Among our models, the IPTW model performed the best, followed by the combined IPTW and IPCW method.

Preprint is now available on medRxiv ☞

Contact: hc3292@cumc.columbia.edu

Columbia University Department of Biomedical Informatics