Quantifying bias from dependent left truncation in survival analyses of real world data

Arjun Sondhi PhD, Olivier Humblet PhD, Akshay Swaminathan
Flatiron Health, New York, NY

Background

- In many EHR-derived databases, patients are only observed if they satisfy a certain entry criteria, such as undergoing a biomarker testing procedure
- **Left truncation (LT)** arises because we do not observe any patients who experienced a disqualifying event (such as death) before satisfying the entry criteria
- When analyzing overall survival (OS), failure to account for LT results in bias, since patients observed in the database had to at least live long enough to qualify for entry

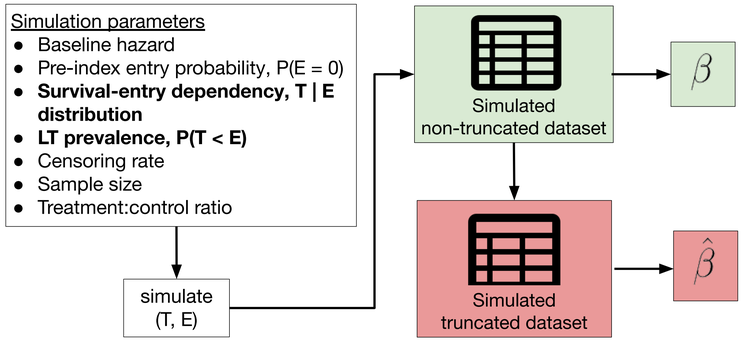
Independent vs dependent left truncation

- Risk set adjustment for survival analyses can adjust for **independent LT,** where the survival time T and entry time E variables are independent
- This condition is testable over the observed space, but is often not satisfied (e.g. patients biomarker-tested later in their disease course may die sooner)
- Under **dependent LT**, standard methods are not guaranteed to unbiasedly estimate the marginal survival distribution
- Magnitude and direction of this bias are generally unexplored

Contributions

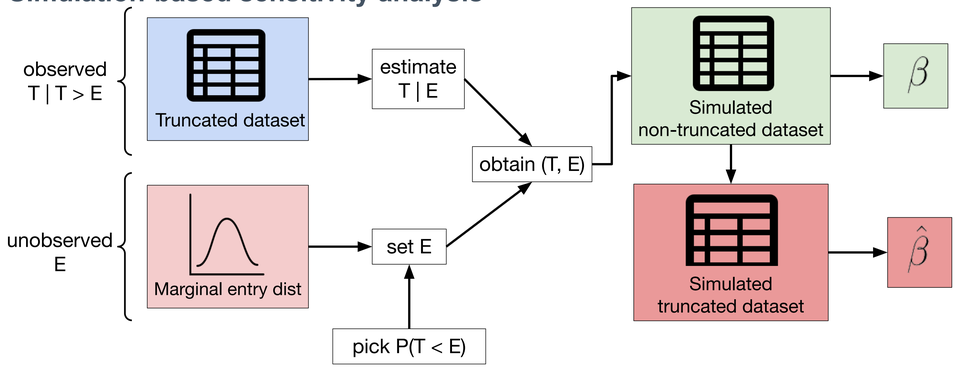
- We implement simulation studies examining analyses involving survival data subject to dependent left truncation, and quantify the resulting inferential bias
- We characterize general trends in the bias, with the goal of determining when such analyses may still provide useful results
- We outline a procedure for conducting sensitivity analyses given an arbitrary dataset subject to dependent left truncation, to determine how bias in results would change when varying levels of left truncation

Simulation study of dependent left truncation



- Considered 3 designs corresponding to survival analyses performed with real world (RW) data subject to LT: (1) comparing two RW cohorts, (2) comparing a RW cohort to a non-truncated trial cohort, (3) estimating survival in a single RW cohort

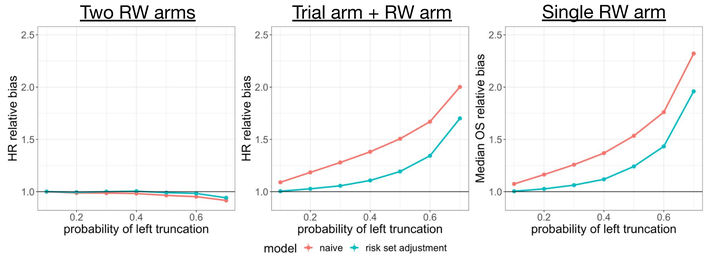
Simulation-based sensitivity analysis



- Given an observed truncated dataset, we can estimate the conditional T | E distribution
- Then, for a particular truncation probability, we can simulate "complete" datasets, and compare truncated vs non-truncated estimates to assess sensitivity to dependent LT bias

Results

- Estimator bias varies substantially between analytic settings
- Estimation of absolute quantities (e.g. median survival time) more subject to bias than that of relative quantities (e.g. hazard ratios)
- Low bias when comparing two RW arms having similar truncation
- When comparing a truncated RW arm to a non-truncated arm, the estimated hazard ratio is **biased upwards**, providing **conservative inference**
- High bias in estimating median survival in a single RW cohort



Conclusions

- Dependent LT can result in **high bias and invalid coverage of confidence intervals**, particularly for **absolute estimands**
- However, when comparing survival of a non-truncated trial cohort to a truncated RW cohort, the bias in the hazard ratio is guaranteed to be upwards
  - I.e. the trial treatment effect **biased towards the null**, more likely to be declared non-significant
  - In drug development, this can still provide an acceptable signal of treatment efficacy
- Our simulation settings cannot cover every possible data-generating process; we recommend that researchers apply **simulation-based sensitivity analyses** based on their own datasets
  - Most simulation parameters are estimable from truncated dataset
  - Examples and guidance for implementation can be found in manuscript

Presented at the 40th ASA Biopharmaceutical Section Regulatory-Industry Statistics Workshop, Sept 21-24, 2021. For additional information, contact Arjun Sondhi at arjun@flatiron.com.

A pre-print of this manuscript is available on medRxiv:
https://www.medrxiv.org/content/10.1101/2021.08.02.21261492v1