A Novel Approach for Minimizing Immortal Time Bias in Observational Studies

Victor A Kiri MPH, PhD[1], Paolo Messina MSc[2], Tyler Knight MS[3], Kathryn P Anastassopoulos, MS[3]
Market Access Consulting, Labcorp Clinical Development Ltd., [1]Maidenhead, UK, [2]Milan, Italy, [3]Princeton, New Jersey, USA

Abstract

- **Background:** Immortal time arises when the determination of a patient's treatment status involves a waiting period. A bias is introduced when this period is unaccounted for in the assessment of treatment effectiveness and/or safety.
- **Objective:** To describe a novel and innovative approach that uses propensity scores (PS) to minimize immortal time bias.
- **Method:** For our novel approach, we classified treatment status at cohort entry (ie, time=t0) based on whether or not the patient initiated treatment at any time during the study period, with actual treatment initiation time as t1 (ie, t1≥t0). We then simulated data using Weibull distributions for multiple scenarios of waiting period (1-24 months), proportion of treated patients (10-50%), confounding level (no to high), median time to event (4.5 to 18 months), treatment effect (hazard ratio [HR] 0.2 to 5), and two independent covariates. We derived two sets of PS on each patient involved in the final analysis dataset (ie, PS0 at t0 and PS1 at t1). First, PS0 at t0 was derived to construct blocks of comparable patients. For each treated patient, the closest untreated patient on PS0 among those who survived up to t1 was selected by 1:1 greedy matching. Second, with t1 as the start of follow-up for the treated and untreated patients in that block, we derived PS1 and used inverse probability treatment weighting (IPTW) to address confounding. The percentages of residual bias from our approach were compared to those from 6 other approaches: 1) unadjusted (Crude), 2) median waiting time (MWT), 3) IPTW, 4) MWT with IPTW (MWT +IPTW), 5) landmark with IPTW (Landmark+IPTW), and 6) time-dependent Cox (TD Cox) with covariate adjustment (TD Cox+covariates).
- **Results:** Our novel approach yielded the lowest median percentage of residual bias across all approaches (4.9% vs. 51.0% for Crude, 30.5% for MWT, 24.2% for IPTW, 13.8% for MWT+IPTW, 17.2% for Landmark+IPTW and 15.8% for TD Cox+covariates).
- **Conclusion:** Our results suggest that the PS is an effective design tool for minimizing immortal time bias.

Background

- Immortal time refers to a period of follow-up during which, by design, death or the study outcome cannot occur.[1]
- In observational studies, immortal time typically arises when the determination of an individual's treatment status involves a delay or waiting period during which follow-up time is accrued. For example:
  - Waiting for a prescription to be dispensed after discharge from hospital when the discharge date represents the start of follow-up.
  - This waiting period is considered immortal because individuals who end up in the treated or exposed group have to survive (be alive and event free) until the treatment definition is fulfilled.
  - If they have an event before taking up treatment they are in the untreated or unexposed group.
  - The immortal time bias is introduced when this period of "immortality" is either misclassified as regards treatment status or ignored in the analysis.
- This bias has been widely described and the pitfalls of inappropriate solutions, including the misuse of the ''events per person-time'' statistics within the ''treatment-switching methodology'' have similarly been reported.[2-3]
- An effective solution should also facilitate comparability between the two treatment groups for the assessment of treatment effect.
- A study design without waiting time will obviously be best and where such is not feasible, then the best alternative may be one that involves the same follow-up definition for both the treated (ie, exposed) and the untreated (ie, unexposed) patients.

Objective

- To describe a novel and innovative approach that uses PS to minimize immortal time bias.

Method

The Novel Approach

- According to sample theories, the scalar propensity score is sufficient to remove bias due to the observed covariates (ie, characteristics and clinical history) between the treated and untreated patients in observational studies.[4]
- Since the most eligible comparators to patients who initiated treatment at cohort entry (ie, time=t0) are the untreated patients with similar propensity scores at cohort entry (ie, PS1), we assumed by extension, that the same is applicable to propensity scores derived at subsequent times of treatment initiation beyond the time of cohort entry (ie, time=t1; t1>t0).[5-8]
- In other words, it is reasonable to restrict our search for the most suitable comparators to a patient who initiated treatment at time t1≥t0, to the group of untreated patients previously identified as comparable to that patient at cohort entry (ie, t0) on the basis of the subsequently derived propensity scores (ie, PS2).
- We propose as the most suitable comparators for each treated patient, those untreated patients who have remained comparable at treatment initiation, from among those previously identified as comparable to that treated patient at cohort entry and have not yet experienced the outcome of interest or lost to follow-up.
  - Thus, the start of follow-up is the same for both the treated and untreated.
  - In other words, every untreated patient with a propensity score within the specified caliper of PS1 for the treated patient will be retained for the second stage.
- This design is similar to the blocking approach adopted elsewhere for the treated patients, which is based on randomly selected start of follow-up times for the untreated patients within six months of the corresponding treatment initiation dates.[9]
- Our novel approach involved greedy matching of treated and untreated patients at 1:1 ratio with a caliper of 0.25.
- Using simulated data, we compared the results from our novel approach with those from the following 6 other known approaches:
  - **Crude:** Without adjusting for immortal time bias or confounding (start of follow-up for untreated=t0)
  - **MWT:** Adjusting for immortal time bias, but without adjusting for confounding (start of follow-up for untreated=t0 + MWT)[10]
  - **IPTW:** Adjusting for confounding using IPTW only (start of follow-up for untreated=t0)
  - **MWT+IPTW:** Adjusting for both immortal time bias and confounding (start of follow-up for untreated=t0 + MWT)
  - **Landmark+IPTW:** Adjusting for both immortal time bias and confounding (start of follow-up=landmark time for all)[11]
  - **TD Cox+covariates:** Adjusting for both immortal time bias and confounding[12-13]
- The estimated HRs from the Cox proportional hazard models were compared to the expected HRs as defined in the simulations and the percent bias was derived for each approach and scenario with each involving 2,000 simulations.
- The results are presented for each approach as the median percent bias, including those from a sensitivity analysis of the impact of the exclusions of patients in the Landmark+IPTW approach.

Simulation

- Time to treatment initiation (TTI) was simulated using a Weibull distribution with different scale and shape parameters to address several scenarios of median waiting time (ie, 1, 3, 6, 12, and 24 months) and percentage of treated patients (ie, 10%, 20%, 30%, 40%, and 50%).
- We defined ti as time of treatment initiation.
- For each scenario, we generated two covariates, age and X at cohort entry, with X as a time-dependent confounder that was updated at treatment initiation. Thus, X was correlated with both TTI and time to event (TTE) within varying confounding settings (ie, none, marginal, moderate, and high).
- TTE was simulated using a Weibull distribution, but with different scale and shape parameters to generate several scenarios of median TTE (ie, 4.5, 9, 13.5, and 18 months) and fixed treatment effects in five different HR settings (ie, 0.2, 0.5, 1, 2, and 5).
- The sample size for each scenario was 10,000.

Abbreviations: MWT=median waiting time; IPTW=inverse probability treatment weighting; PS=propensity scores; TD=time-dependent; cov=covariates; TTI=time to treatment initiation; TTE=time to event

Results

- Our novel approach yielded the lowest median percent (%) residual bias (5th – 95th percentiles) across all approaches, and it was also the most consistent with the lowest standard deviation (SD) (Figure 1):
  - **Novel:** 4.9 (0.4-18.3), SD=6.9
  - **Crude:** 51.0 (10.4-175.8), SD=56.8
  - **MWT:** 30.5 (6.1-104.2), SD=33.9
  - **IPTW:** 24.2 (3.1-94.2), SD=33.7
  - **MWT+IPTW:** 13.8 (2.1-57.9), SD=21.1
  - **Landmark+IPTW:** 17.2 (1.2-72.0), SD=21.3
  - **TD Cox+covariates:** 15.8 (1.1-85.6), SD=31.7
- Additionally, our novel approach was consistently superior with respect to:
  - MWT to treatment initiation (Figure 2)
  - Percentage of treated patients (Figure 3)
  - Confounding level (Figure 4)
  - Median time to event (Figure 5)
  - Treatment effect (Figure 6)

Figure 1. Median (5th - 95th percentiles) residual bias (%) by approach

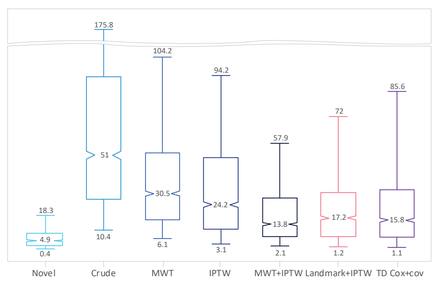


Figure 2. Median residual bias (%) by median waiting time to treatment initiation

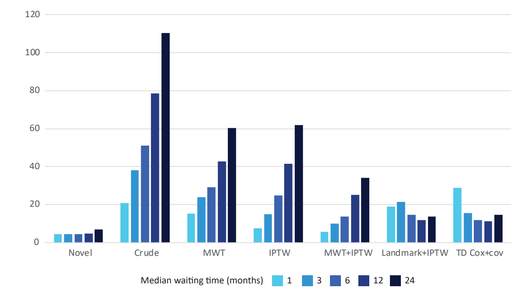


Figure 3. Median residual bias (%) by percentage of treated patients

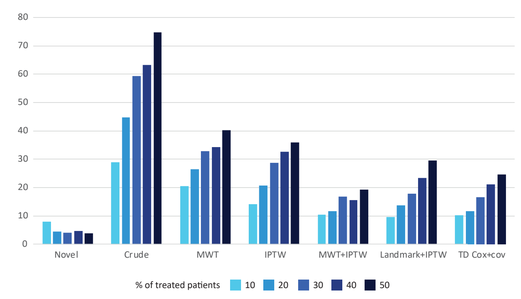


Figure 4. Median residual bias (%) by confounding level

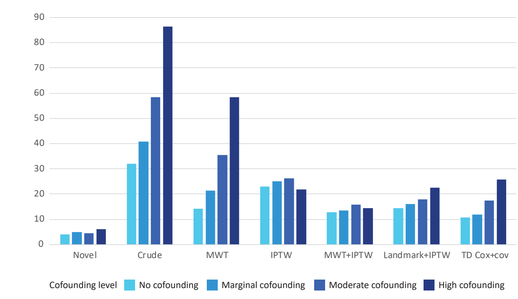


Figure 5. Median residual bias (%) by median time to event

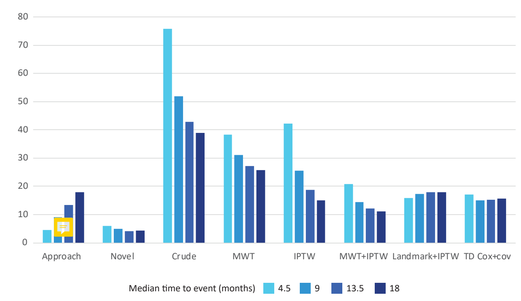
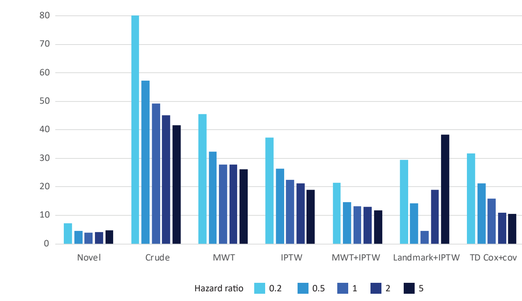
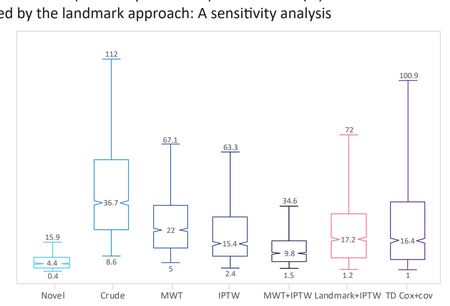


Figure 6. Median residual bias (%) by treatment effect (hazard ratio)



- The results were confirmed across all of the approaches retained by the Landmark approach in our sensitivity analysis (Figure 7).

Figure 7. Median (5th - 95th percentiles) residual bias (%) for all scenarios that were retained by the landmark approach: A sensitivity analysis



Conclusions

- Our novel approach is based on the availability of information for identifying untreated patients that are comparable to the treated patients at cohort entry.
- It also assumes updated information can be obtained at the time of treatment initiation, whereby the latest available sets for both the treated and untreated patients are to be used.
- We acknowledge that information derived at treatment initiation are likely to be more up-to-date for the treated patients than for their corresponding untreated counterparts.
- Although we have used PS-matching by the Greedy approach at 1:1 ratio and 0.25 caliper as the simplest option for identifying comparable untreated patients at cohort entry, other options can be used depending on the features of the data (ie, sample size and proportions of treated/untreated patients).
- The evidence from the simulations, involving different scenarios consistently suggests the novel approach as possibly the most effective among the current alternatives.
- The results from the approach suggest that the PS can serve as an effective study design tool for minimizing immortal time bias.

References

[1]Suissa S. Immortal time bias in pharmacoepidemiology. Am J Epidemiol. 2008;167(4):492-9.[2]Kiri VA, Mackenzie G. Re: "Immortal time bias in pharmacoepidemiology". Am J Epidemiol. 2009;170(5):667-8. [3]Kraemer HC. Events per person-time (incidence rate): a misleading statistic? Stat Med. 2009;28(6):1028–39. [4]Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983;70:41–55.[5]D'Agostino RB Jr. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. Stat Med. 1998;17(19):2265-81. [6]Stürmer T, Joshi M, Glynn RJ, et al. A review of the application of propensity score methods yielded increasing use, advantages in specific settings, but not substantially different estimates compared with conventional multivariable methods. J Clin Epidemiol. 2006;59:437–47. [7]Austin PC. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. Multivariate Behav Res. 2011;46(3):399-424. [8]Austin PC. Advances in propensity score analysis. Stat Methods Med Res. 2020;29(3):641-43.[9]Seeger JD, Walker AM, Williams PL, et al. A propensity score-matched cohort study of the effect of statins, mainly fluvastatin, on the occurrence of acute myocardial infarction. Am J Cardiol. 2003;92(12):1447-51.[10]Feudjo-Tepie MA, Kiri VA, Qizilbash N, et al. Waiting Time Bias in Drug Safety and Effectiveness Cohort Analyses Using Clinical Practice Database. Pharmacoepidemiol Drug Saf. 2003;12(Suppl 1):S141. [11]Agarwal P, Moshier E, Meng Ru M, et al. Immortal Time Bias in Observational Studies of Time-to-Event Outcomes: Assessing Effects of Postmastectomy Radiation Therapy Using the National Cancer Database. Cancer Control. 2018;25:1-7. [12]Cho IS, Chae YR, Kim JH, et al. Statistical methods for elimination of guarantee-time bias in cohort studies: a simulation study. BMC Med Res Method. 2017;17(1):126. [13]Mi X, Hammill BG, Curtis LH, et al. Use of the landmark method to address immortal person-time bias in comparative effectiveness research: a simulation study. Stat Med. 2016;35(26):4824–36.

Presented at ISPOR Europe 2022

labcorp Drug Development