Let's talk Re@l: Immortal Time Bias - Zeroing in on time-zero

Deepak Parashar *(University of Warwick)* and **Jixian Wang** *(Bristol-Myers Squibb)*
[*With input from* Peter Almgren, Alessandro Guasconi, Elizabeth Merrall, Barbara Torlinska, Qing Wang, Josephine Wolfram, Rima Izem]
Note: Opinions are those of the authors not necessarily all members of the SIG, nor the companies they represent.

While statisticians seldom come across the concept of eternity or immortality in their exploration of data-driven scientific quest, there are anomalies where this concept makes a subtle presence in medical studies. That is enough to send them into a state of caution as to how best to comprehend such an incongruity.

In epidemiological and clinical survival studies, a typical outcome of interest is an event (such as death) and one can compute the patients' survival. In order to measure time duration, one needs a start time (**time-zero**) and a finish time (time when a certain event or censoring occurs). In clinical trials, time-zero is often defined as the treatment start time (designed to be immediately after randomization), while in many observational studies the choice may be less obvious. There can be a period during which the event of interest (e.g. death) cannot occur, such as a delay between diagnosis and treatment initiation. If we analyse patients who initiate treatment and then calculate time to event including this period of delay, we introduce a period of time during which the subject is **'immortal'**. This can lead to biased results if the 'immortal time' is either excluded from the study or misclassified into the exposure group. To mitigate such a bias, researchers aim to design studies such that patients are assigned to exposure groups based on their baseline data at the time of eligibility determination rather than later data. The issue can be further compounded when data from observational studies are used to contribute external control arm data in clinical trials, which requires comparability, including alignment of time-zero, between the data sources.

The Challenge

At the Real-World Data (RWD) SIG, we delved into the challenge of time-zero and considered how best to mitigate the potential misalignment of time-zero. Figure 1 below illustrates the complexity of comparing survival or time-to-event endpoints between a clinical trial and RWD source.

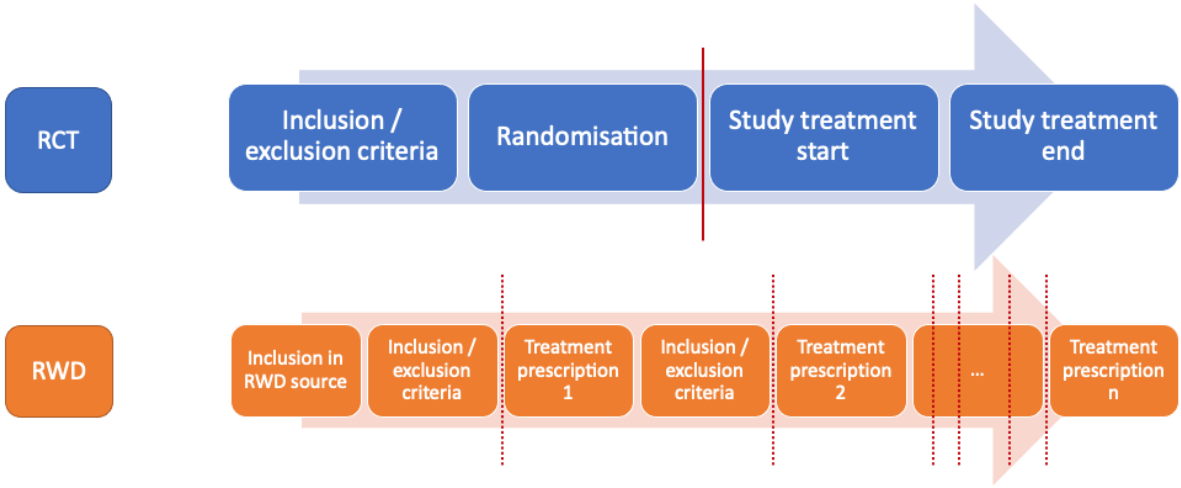


Figure 1: Illustration of well-defined time-zero in a Randomised Controlled Trial and time-zero ambiguity in RWD.

The challenge here is that baseline of therapy initiation may not match for subjects in a clinical trial and those in a RWD source. While in a clinical trial, patients are assigned treatment or control arm at randomisation, defined as time-zero, the corresponding definition may be ambiguous in a RWD source since patients could be eligible multiple times and could be following a sequence of treatment prescriptions. Does one then consider the first eligible time, or randomly select an eligible time, or consider the last line of therapy, or something else? Although here we illustrate the issue in the context of a comparison to trial data that ambiguity also arises for between-cohort comparisons within RWD, with the issue further exacerbated when comparing to an untreated cohort. We also note that similar situations can sometime be encountered within an RCT too, e.g., there is a delay between randomisation and receiving treatment. Or, for the case of single-arm trials, where time-zero would be designated as Day 1 following the inclusion criteria being met, and there is a delay in start of treatment.

Recommendation

Drawing upon the RWD SIG members' experience of the issue of time-zero across multiple therapeutic areas, such as Dermatitis *(estimation of disease trajectories over time)*, Multiple Sclerosis *(comparing rate of malignancies between exposed / unexposed groups in a registry case-control study; with long latency; if the patient switches treatment they contribute twice to time at risk in both the original and switched treatments)*, Vaccines *(time to Covid-19 related hospitalisation; vaccination cohort matched 1:10 to participants with no evidence of Covid-19 vaccine)*, Endocrinology *(estimation of weight over time; time-zero needed for aligning risk trajectories under different treatments for hyperthyroidism in a real-world study)*, and Oncology *(considering multiple starting times for each patient to mitigate time-zero)*, the authors' recommendation to mitigate time-zero is:

1. Allow for 'multiple' baselines for patients in RWD, with a possible adjustment for within-patient correlation.
2. Adjust for patient baseline differences using causal inference models such as propensity matching, inverse probability weighting etc.

A hypothetical scenario

The example is adapted from Wang et al. (2023), in which data of four hypothetical patients are given (table 1).

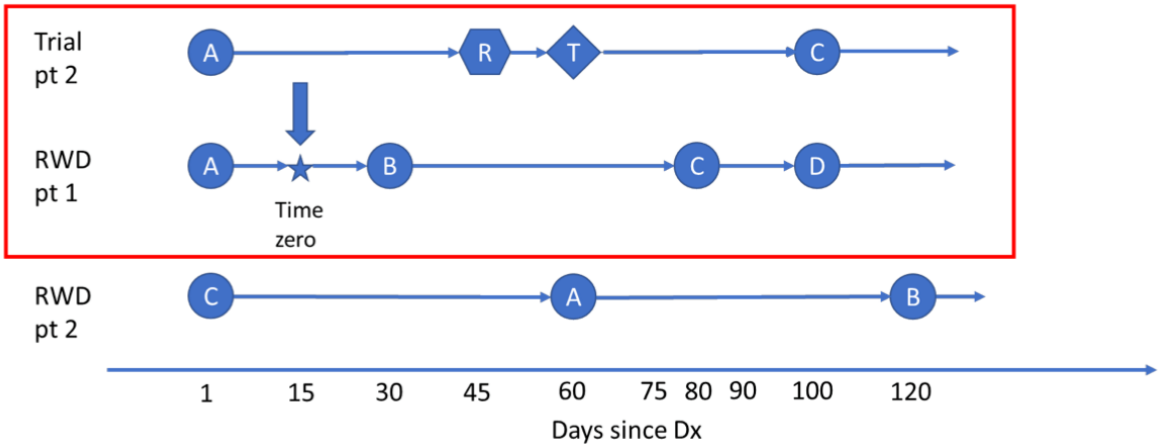
Table 1 A hypothetical dataset for trial and RWD patients

Source	ID	Randomization or treatment start and change/ days since diagnosis			
		1	2	3	4
Trial	1	Control A/1	Control C/30	Randomization/75	Trial treatment/90
Trial	2	Control A/1	Randomization/45	Trial treatment/60	Control C/100
...					
RWD	1	Control A/1	Control B/30	Control C/80	Control D/100
RWD	2	Control C/1	Control A/60	Control B/120	
...					

Suppose we want to compare trial treatment to standard of care. Multiple approaches can be used to determine the matches and corresponding time-zero, as shown in the three panels below. The first approach (Panel 1) is to match by previous treatment A and to align the time of randomization and the time 0 in the RWD. Note that if the gap between time 0 and treatment

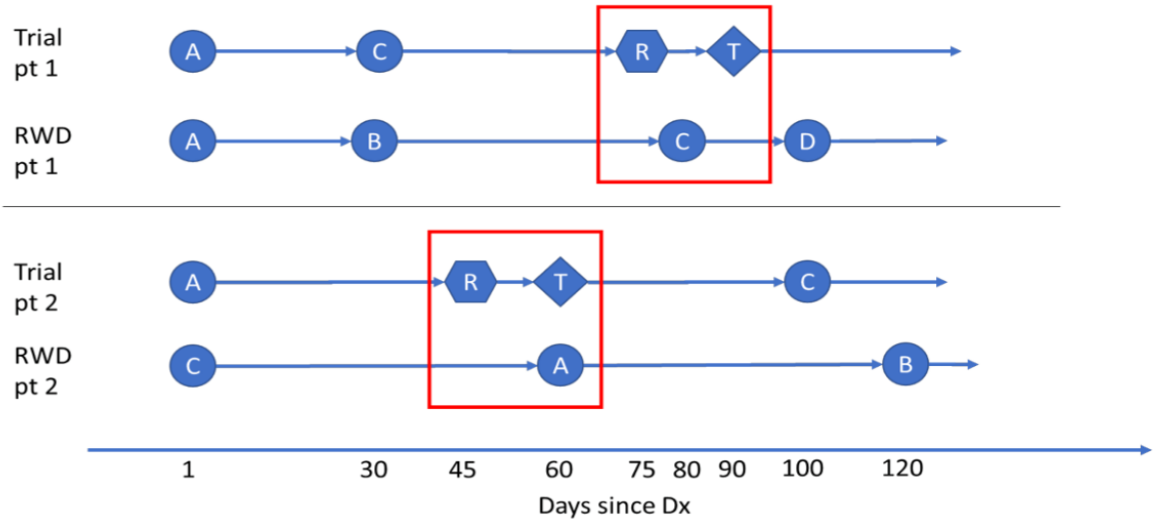
(here 15 days) is not short relative to disease status change, immortal bias could occur in both sources.

Panel 1: Trial patient 2 and RWD patient 1 are matched based on receiving prior treatment of A. Time zero of RWD patient 1 is 15 days before treatment B was initiated.



The next one (Panel 2) is to match by number of previous treatments, hence is more flexible. This is a matching at "macro scale", assuming those with the same number of previous treatments are comparable and (as for prior approach) assign time 0 by matching the time from randomization to trial/new SOC treatment).

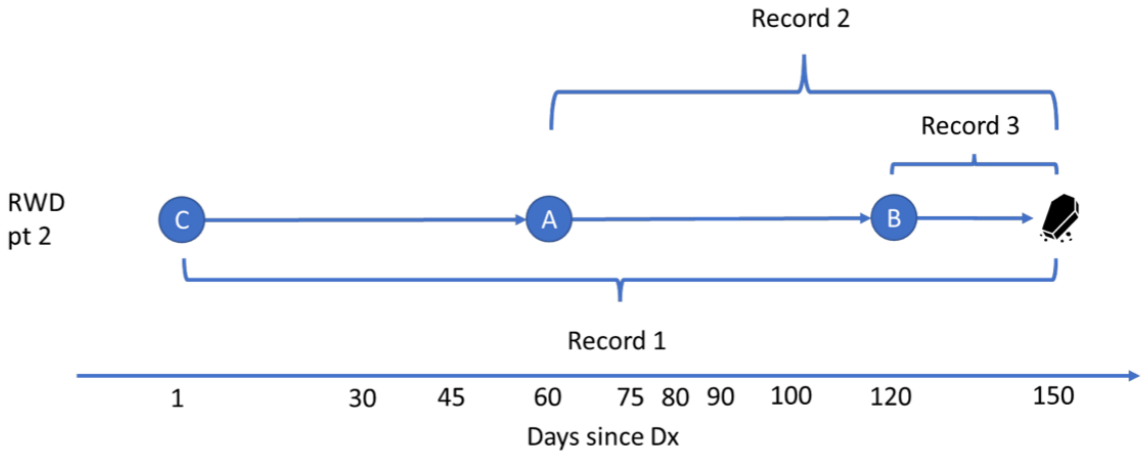
Panel 2: Matching is based on the number of previous treatments.



One way to use all possible choices is to create multiple records one for each choice of time-zero (Panel 3), if the comparator of the trial treatment is non-specific standard of care. Adjustment is needed to deal with correlation between them and comparability with the trial patients. One such method is inverse probability weighting. The propensity score (PS) here is the probability of being in the trial, given baselines and history, hence each record has its own weight. Artificial censoring may also be needed in some situations(Hernán et al.

2016). For example, if the comparator excludes treatment A, then Record 1 should be censored at 60 days.

Panel 3: RWD patient 2 contributes to three correlated records.



This approach (Panel 3) constitutes an example of the afore-mentioned recommendation to mitigate time-zero.

References

1. Suissa, S. (2008). Immortal time bias in pharmaco-epidemiology. Am J Epidemiol. 167, 492–499.
2. Suissa, S., Moodie, E. E., & Dell'Aniello, S. (2017). Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. Pharmacoepidemiology and drug safety, 26(4), 459–468.
3. Hernán, M. A., Sauer, B. C., Hernández-Díaz, S., Platt, R., & Shrier, I. (2016). Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. Journal of clinical epidemiology, 79, 70–75.
4. Hernán, M. A., \& Robins, J. M. (2016). Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. American journal of epidemiology, 183(8), 758–764.
5. Mi, X., Hammill, B. G., Curtis, L. H., Lai, E. C.-C., and Setoguchi, S. (2016) Use of the landmark method to address immortal person-time bias in comparative effectiveness research: a simulation study. Statist. Med., 35: 4824– 4836.
6. Backenroth, D. (2021). How to choose a time zero for patients in external control arms. Pharm. Stat. 20(4):783-792.
7. Wang, J., Zhang, H., & Tiwari, R. (2023). Statistical Challenges for Causal Inference Using Time-to-Event Real-World Data. In *Real-World Evidence in Medical Product Development* (pp. 233-254). Cham: Springer International Publishing.