IJMPR 2025, 9(3), 11-25

International Journal of Modern Pharmaceutical Research
www.ijmpronline.com

ISSN: 2319-5878
IJMPR
Research Article
SJIF Impact Factor: 6.669

THE ENTRY-TIME COMPARABILITY RETENTION (ETCR) METHODOLOGY FOR HANDLING IMMORTAL TIME BIAS IN OBSERVATIONAL STUDIES

Victor Kiri[1,2]* and Paolo Messina[3]

[1]Market Access Consulting & HEOR, RWE & Analytics, Fortrea, Maidenhead, UK.
[2]Faculty of Pharmaceutical Sciences, University of Port Harcourt, Choba, Nigeria.
[3]Market Access Consulting & HEOR, RWE & Analytics, Fortrea, Milan, Italy.

Article Received on: 05/01/2025
Article Revised on: 25/01/2025
Article Accepted on: 15/02/2025

***Corresponding Author**
Victor Kiri
Market Access Consulting & HEOR, RWE & Analytics, Fortrea, Maidenhead, UK.

ABSTRACT

Background: To propose the Entry-time Comparability Retention (ETCR) approach, a simple methodology that uses propensity scores to address the problem of immortal time bias with realistic assumptions. **Methods**: The exposure status of each member of a cohort is classified at cohort entry, based on whether the patient received treatment during the study period (i.e. exposed) or remained on standard care (i.e. unexposed). Two sets of propensity scores are derived on each patient: 1) at cohort entry- to construct blocks of comparable patients by propensity score matching each exposed patient to as many of the unexposed patients as possible, and 2) at treatment initiation- to identify those unexposed patients to be retained and be assigned the same start of follow-up to address the problem of immortal time bias. Using simulated data, the performance of the ETCR methodology with simple blocks of 1:n greedy matching, is compared with two of the leading alternatives- the Landmark design and the time-dependent Cox regression (TDCR) approach. **Results**: When compared with the crude model, the ETCR resulted in the largest reduction in the mean residual bias (i.e. >27% vs 14% and 9% by the Landmark and TDCR respectively). Its estimates were consistently the least sensitive to variations in the different scenarios of our simulations. **Conclusions**: Propensity scores can be utilized as an effective design tool for creating and retaining comparable treatment groups to minimize immortal time bias. The ETCR does not introduce additional assumption, and it is particularly useful for studies involving a "no treatment" strategy.

KEYWORDS: Immortal time, propensity score, confounding, time-dependent Cox regression, Landmark, inverse probability treatment weighting.

1. BACKGROUND

Immortal time refers to a period of follow-up during which, by study design, death or the study outcome cannot occur.[1-16] Immortal time typically arises in observational studies when the determination of a patient's treatment status involves a period before treatment initiation in studies, which are based on the "no treatment" strategy- namely, exposed versus unexposed patients; the latter group including those who remain on standard care for the disease.[2,17,18] A bias is introduced when this period is not adequately accounted for in the assessment of the treatment effect. In the context of patients waiting for a specific intervention, this bias is generally referred to as waiting time bias in the literature.[19-24] In the context of chart reviews, where we assess the safety and effectiveness of a particular drug by comparing the medical records of those who are treated against those on standard care during the study period, the bias arises when we ignore the period of standard care before the start of treatment with the drug.

There are two main settings in observational studies involving the comparison of exposed and unexposed groups, where immortal time may arise: (1) patients in a queue having to wait for an intervention (e.g. heart transplantation), with the enrolment date on the waiting list representing the start of cohort entry and (2) patients in routine clinical practice whose treatments do not start from the date of diagnosis of the condition (i.e. cohort entry) but rather, when their physicians decide to intervene in compliance with the treatment guidelines for that specific disease.[25] In both settings, immortal time bias may arise if the start of cohort entry is considered as the start of follow-up in the evaluation of treatment effect. This is because unlike the unexposed, none of the exposed could have experienced the outcome of interest whilst waiting to receive the treatment, which makes the period an immortal time for such patients. In routine clinical practice, the reason for prescribing a treatment by the physician may be associated with the risk of occurrence of the outcome of interest and so immortal time may pose only one source of bias.[26-27] In this regard, immortal time bias does not apply to the

intention-to-treat (ITT) analysis in randomized controlled trials (RCTs), where exposure status is defined at the start of follow-up even if the exposure is scheduled sometime after the study baseline.[28-29]

The literature on immortal time bias offers several methods for addressing the problem, each with its own set of assumptions. These can be summarily categorized as those that involve (1) the time-dependent Cox models, (2) the Landmark approach, and (3) the matching-on-time methods.[28] Since immortal time bias is only one among many other sources of bias in observational studies, an effective solution should also facilitate comparability between the two groups being compared, with the best approach resulting in the maximum possible reduction of the different sources of bias. Indeed, this point is particularly important in the assessment of drug safety where bias, such as exposure misclassification, can be far more impactful.

The aim of this paper is to propose the Entry-time Comparability Retention (ETCR) method as a simple approach for addressing immortal time bias using the propensity score (PS) methodology.[30] Propensity score is usually employed to reduce the problem of selection bias in observational studies, and in this context, it can be defined as the probability that a person with a given set of relevant observed characteristics will be assigned to the exposed group as against being unexposed. The ETCR approach, which is a design tool, does not introduce any additional assumption to the study design and it can incorporate any of the current methods for addressing the problems of confounding. The approach involves derivation of the PS at two different time points based on the characteristics and clinical histories of these patients: firstly, at cohort entry, to create blocks of comparable exposed and unexposed patients so as to assign the same start of follow-up for members of each block at treatment/exposure initiation and hence, minimize immortal time bias; secondly, at treatment initiation in its traditional role of a balancing tool to retain the comparability between the treatment/exposure groups. To our knowledge, the ETCR methodology is the first to adopt the propensity score specifically as a study design tool for resolving the problem of immortal time bias. Using simulated data, the new methodology is compared with two of the leading approaches in the literature, namely, the Landmark methodology and the most common form of the time-dependent Cox regression methodology, both of which involve assumptions that make them more susceptible to exposure misclassification bias.

2. METHODS

2.1 Notations

We simulated a cohort of patients with cohort entry at time t_0 and treatment initiation at time t_1 for those who received the treatment of interest. We defined two treatment groups for comparisons: the "unexposed" as those who remained on standard care (also referred to as the "untreated") and the "exposed" as those who received the treatment at $t_1 \geq t_0$ (also referred to as the "treated"). The treatment status is denoted by the variable Z: Z=0 for the unexposed and Z=1 for the exposed and ours is not the first to adopt this classification for comparative assessment of the bias.[31] Indeed, studies involving the assessment of treatment effect between exposed/treated versus unexposed/untreated patients in retrospectively collected data are not uncommon; the literature suggests most observational studies on drug safety are based on such comparisons and they include virtually all single-arm trials that involve external controlled arms- a growing trend in drug research on rare diseases for regulatory submissions as all are based on this treatment strategy. In this strategy, the intention-to-treat framework effectively and accurately represents the real-life exposure status as observed- being the hallmark of the real-world evidence generation process, that makes it distinctively different from the controlled, randomized trial setting.

2.2 Landmark (LA) Methodology

In the Landmark approach, a predefined, common time point is fixed for all patients for the classification of treatment/exposure status. This time point is referred to as the landmark and is the start of follow-up.[28,32-38] Immortal time bias is assumed to be resolved by the approach since both the treatment determination point and start of follow-up are the same. However, exposure misclassification may occur since the exposed whose time of initiation is after the landmark time will be considered as "unexposed. For example, if the landmark is set at 6 months from cohort entry, then those who initiated treatment after 6 months will be classified as unexposed. The approach is therefore suitable for settings where the interest is in whether treatment is started within the landmark period. The median waiting time from the simulated data was assigned as the landmark time in our comparisons.

2.3 Time-dependent Cox Regression (TDCR) Methodology

The Time-dependent Cox regression (TDCR) approach is whereby Z is considered as a time-dependent binary variable for treatment initiation. In common with most applications of the approach, we have used the step-function definition of Z such that Z=0 from cohort entry until treatment initiation when it changes to Z=1.[39-41] This, being the most common and simplest form of application of the TDCR approach- is one-directional in the sense that it does not allow for patients who switched to the treated group upon initiating treatment, to switch back to the untreated group when they stop the treatment.

2.4 The Entry-Time Comparability Retention (ETCR) Methodology

The approach assumes that the most eligible unexposed comparators to each exposed patient at treatment initiation (t_1) are those who were comparable at cohort entry (t_0), based on propensity scores. The method,

which involves three steps, relies on the PS to identify the suitable comparators, as illustrated in Figure 1.

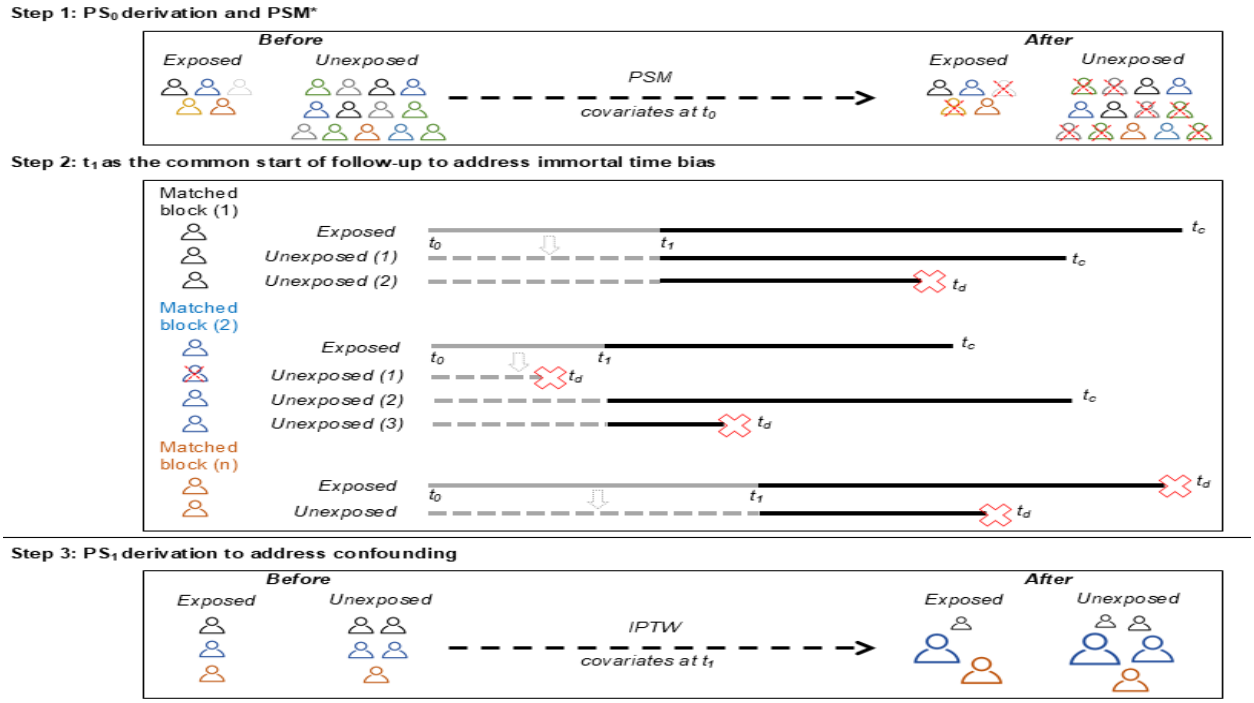

Figure 1: The Entry-Time Comparability Retention Method.

Abbreviations: IPTW= inverse probability treatment weighting; PS_0= propensity score at t_0; PS_1= propensity score at t_1; PSM= propensity score matching; t_0=cohort entry time; t_1= treatment initiation time; t_c= censoring time; t_d= event time.

Step 1: Deriving PS_0

At cohort entry (t_0), derive the first set of propensity scores (PS_0) from the covariates. Create blocks of exposed and comparable unexposed patients (1:n). Select unexposed patients with a PS_0 within a defined caliper of the exposed patient's PS_0.

Step 2: Aligning Follow-up Start

Assign the exposed patient's treatment initiation time as the common start date for follow-up in each block. Retain unexposed patients with a PS_0 within the caliper, provided they haven't experienced the outcome or been lost to follow-up before the treatment initiation time (t_1).

Step 3: Deriving PS_1

At treatment initiation (t_1), derive the second set of propensity scores (PS_1) based on updated covariate values for all patients retained after steps 1 and 2.

We used matching without replacement in our illustrations, though the design also supports matching with replacement. This approach is similar to blocking methods used for treated patients, where follow-up start dates for unexposed patients are randomly selected within six months of the treatment initiation date.[42] In the ETCR method, immortal time bias is addressed by using t_1 as the start of follow-up for both exposed and unexposed patients. Confounding is then managed using the second set of propensity scores (PS_1) for treatment effect assessment.

2.5 Rationale of the ETCR Methodology

The usefulness of any observational study relies heavily on the extent to which the study data accurately captures and represents the experiences of the study patients in real-life clinical practice. Thus, for those involving retrospectively collected data, their usefulness would depend on the extent to which such data accurately represent what has been observed, for the generation of real-world evidence. In other words, the definition of treatment status that is based on what has been observed will not by itself, introduce selection bias, especially if the purpose is to avoid the risk of exposure misclassification, as it is the situation with the counter-matched nested case-control design. The risk of exposure misclassification ought to be of greater concern to us than that of immortal time bias especially when using such data to assess drug safety. According to the literature, we do indeed continue to pay greater attention to this particular risk when conducting observational studies for regulatory submissions, where we are required to ascertain the validity of the assumptions associated with the statistical techniques we utilize.

For the ETCR approach, we consider the patients who are not exposed to the specific drug of interest as those who remained on standard care (SC) and define PS_0 as the propensity of being exposed (i.e. treated) at cohort entry. Thus, the PS_0 is an observation-based statistical "construct" used only for the creation of blocks of

patients who are comparable at the start of the study. In other words, PS_0 serves only as a design tool and it is not involved in the analytical process for the assessment of treatment effect for which the PS_1 is used to compare the effects between the treated patients and those on SC (i.e. unexposed) without immortal time bias. Applied to retrospective data, with x=0 as unexposed and x=1 as exposed, both p(x=0) and p(x=1) are defined over the same observation period for the PS_0, such as to accurately ascertain the status of each patient and hence, minimize the risk of exposure misclassification. In addition, both p(x=1) and p(x=0) are based on the same classification period for the PS_1, and the subsequent treatment assessment process does not rely on the intention-to-treat assumption, which may not accurately reflect the setting in real-world clinical practice. This is particularly useful in the assessment of safety outcomes, where the so-called treatment-switching (i.e. time-dependent Cox) approach, which is based on strong statistical assumptions about the risk profiles of the patients that are often ignored, may not be appropriate. The core assumption of causal inference in the context of treatment switching is sequential exchangeability, which is inherently difficult to satisfy in the context of a non-random process, such as we find in routine clinical practice- involving data on prescription drugs that are based on treatment guidelines (i.e. not any random process).[23]

2.6 Simulation Settings

Two sets of simulations were performed using Weibull distributions to generate (1) time-to-treatment initiation data and (2) survival time data.[43]

Time-to-treatment initiation: We assigned different values for time-to-treatment initiation (t_1) and censoring time (t_c) to generate datasets, involving 5 scenarios of median waiting time (1, 3, 6, 12 and 24 months) and 4 scenarios of percentage of treated subjects (10, 30, 50 and 75%), such that a subject was considered as exposed if $t_1 \leq t_c$ and as unexposed if otherwise.

Survival time: Using different shape parameters (K) for the Weibull distribution (i.e. failure rates of 0.75 for the setting of decreasing rate over time, 1.0 for constant rate over time and 1.5 for increasing rate over time), we simulated the time to event, involving 4 scenarios of median time-to-event (6, 9, 12 and 18 months) and 5 scenarios of treatment effect (hazard ratios of 0.5, 0.75, 1.0, 1.5 and 2.0), such that a subject was considered as a case (i.e. had the event) at time t_d if $t_d \leq t_c$ and as censored if otherwise.

Thus, each subject could be classified as any one of the following: exposed censored (t_1,t_c), exposed case (t_1,t_d), unexposed censored (t_0,t_c) and unexposed case (t_0,t_d)- the terms in parenthesis denoting the start and end of follow-up times.

Two additional covariates (age and x) were simulated for each subject from a Normal distribution at t_0 and at t_1 as follows:

$$Age_{t_1} = Age_{t_0} + (t_1 - t_0)$$

$$x_{t_1} = \max\left(x_{t_0} - c\sqrt{(t_1 - t_0)}; 1\right)$$

Here, x was also a factor in the calculation of the scale parameter for the survival simulation time (st):

$$\lambda_{st} = \sqrt[n]{x_{t_1}}$$

in order to obtain 4 scenarios of confounding - none, marginal, moderate and high confounding, whereby x differed between the treated vs untreated groups by 10, 25 and 40% respectively.

Waiting time bias: This was possible when the waiting time (t_1-t_0) was not accounted for in the definition of start of follow-up.

3. STATISTICAL ANALYSIS

We compared the results from the ETCR methodology with those from the LA and TDCR approaches with failure as endpoint and hazard ratio as the estimate of interest. Both the ETCR and LA methodologies involved the inverse probability treatment weights (IPTW) approach whereas the TDCR model with treatment group as a binary variable, included age and x as additional time-dependent covariates. For the ETCR methodology, PS_0 was calculated using the greedy nearest neighbour matching method at a 1:n ratio with a calliper of 0.25. PS_1 was used for the IPTW where average treatment effect (ATE) weights were calculated at t_1 via a logistic regression model

$$Z\,(1) = Age_{t_1} + x_{t_1} + Age_{t_1} * x_{t_1}$$

The ATE weights for the exposed and unexposed subjects were calculated respectively as

$$\frac{1}{p_j} \; and \; \frac{1}{1 - p_j}$$

with p_j as the propensity score for subject j.

We used the Cox proportional hazards (PH) models to estimate the Hazard ratios (HRs) for both the ETCR and LA methods. The performance of each method was assessed using the residual bias ($\hat{\varepsilon}$) calculated for each scenario as follows

$$\hat{\varepsilon}_i = \log(\widehat{HR}_i) - \log(HR_i)$$

In order to calculate the residual bias, we assumed that the true value of the HR for each scenario was the one determined at the time-to-event simulation stage before applying the time-to-treatment initiation delay (i.e. waiting time bias). As part of this assumption, all simulations included a minimum immortal time (i.e. median time-to-treatment initiation of 1 month). Consequently, the true value of the HR is inherently biased.

To compare the effect of the adjustments proposed by the three methodologies (namely, the ETCR, LA and TDCR), we compared the residual bias of each with the unadjusted (crude HR) from the Cox model. The mean bias is also presented overall and for each scenario. The difference of the mean bias between each of the three compared methods and the crude (non-adjusted) one is presented overall and for the descending, stable and ascending failure rates simulations; 95% Confidence intervals (CI) were estimated using the percentile interval bootstrapping method resampling each result 5000 times.

The analyses were conducted in SAS version 9.4 for Windows (SAS Institute, Inc, Cary, NC).

4. RESULTS

We generated N=4,800 datasets for the different scenarios, each with n=10,000 subjects.

Table 1 shows the mean bias across all scenarios, with the ETCR yielding the lowest indexes in every comparison. In general, the ETCR was found to reduce the bias (vs. the CRUDE model) of more than 27% (-0.276) with a residual mean bias across all scenarios of 0.141, that is even lower (0.100) when the simulated time-to-event failure rate is stable. The ETCR overall mean bias was followed by the LA (0.276), the TDCR (0.325) and the CRUDE (0.417).

Table 1: Overall Results by Approach.

Approach	Failure rate	n	Mean Bias (Diff)	Bootstrap 95% CI
ETCR	*Overall*	*4800*	*0.141 (-0.276)*	*0.097-0.172*
	Stable (k=1)	1600	0.100 (-0.287)	0.095-0.105
	Decreasing (k=0.75)	1600	0.167 (-0.140)	0.161-0.174
	Increasing (k=1.5)	1600	0.155 (-0.402)	0.150-0.160
LA	*Overall*	*4800*	*0.276 (-0.141)*	*0.250-0.311*
	Stable (k=1)	1600	0.270 (-0.117)	0.257-0.283
	Decreasing (k=0.75)	1600	0.302 (-0.005)	0.289-0.315
	Increasing (k=1.5)	1600	0.257 (-0.300)	0.246-0.269
TDCR	*Overall*	*4800*	*0.325 (-0.092)*	*0.294-0.365*
	Stable (k=1)	1600	0.303 (-0.084)	0.289-0.317
	Decreasing (k=0.75)	1600	0.316 (+0.009)	0.302-0.330
	Increasing (k=1.5)	1600	0.356 (-0.201)	0.341-0.370
CRUDE	*Overall*	*4800*	*0.417 (ref.)*	*0.301-0.570*
	Stable (k=1)	1600	0.387 (ref.)	0.377-0.398
	Decreasing (k=0.75)	1600	0.307 (ref.)	0.298-0.315
	Increasing (k=1.5)	1600	0.557 (ref.)	0.539-0.576

The consistency of the performance of the ETCR methodology in comparison with the other approaches across all the scenarios is visibly demonstrated by the scatter plots in Figure 2; the residual bias is more densely distributed towards the lowest values and with a much lower variation when compared to the two other approaches.

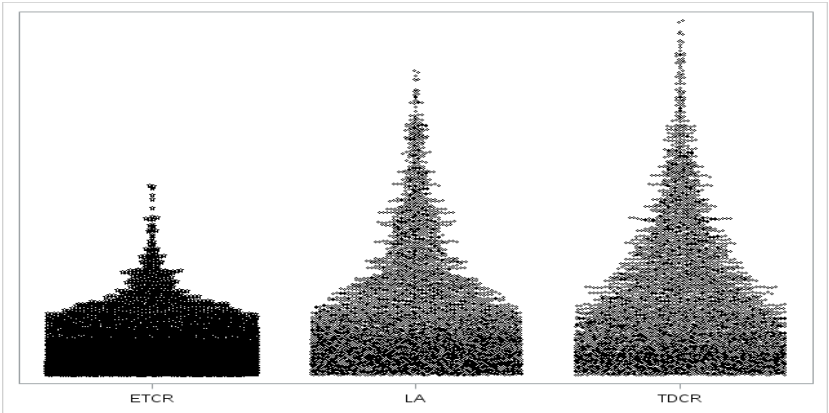

Figure 2: Bias Clouds.

Each point on the plot represents a different simulation, with the vertical axis showing the magnitude of residual bias. The plot helps identify the overall distribution of bias across the tested simulation scenarios using the three different methodologies (ETCR, LA, and TDCR).

The pattern of consistency in its comparative advantage over the others is similarly illustrated in the following two scenarios:

Figure 3 shows that the larger the median waiting time, the larger the residual bias, except for the TDCR approach, where the reverse is the case. However, the impact is most marginal with the ETCR methodology with mean bias ranging from 0.113 to 0.177 when compared with those for the LA (i.e. 0.183 to 0.395) and the TDCR (i.e. 0.198 to 0.496).

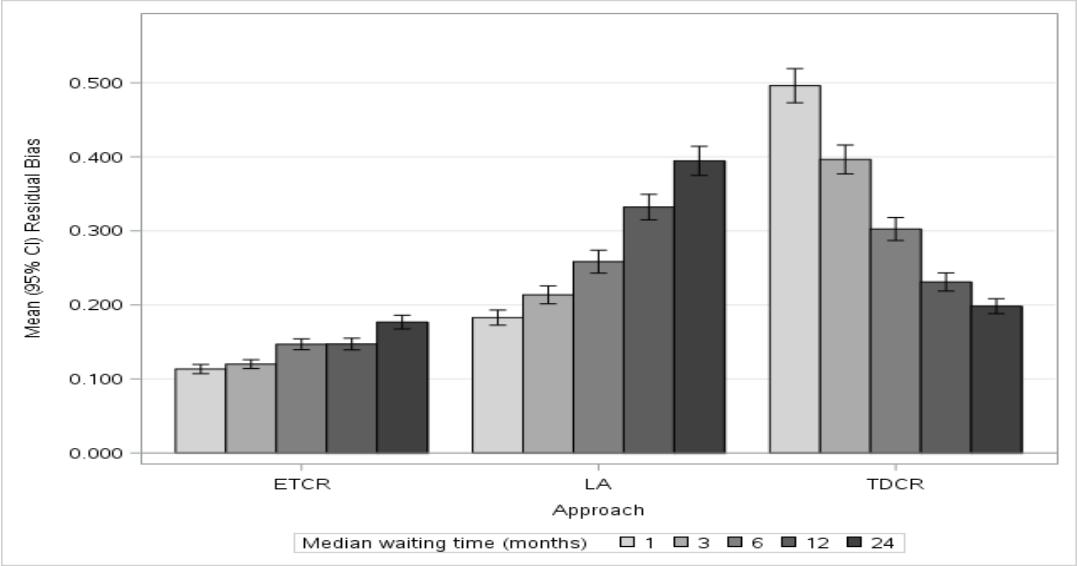

Figure 3: Mean Bias (95% CI) by Median Waiting Time to Treatment Initiation.

This plot represents the average residual bias across different scenarios of median waiting time, showing any changes in bias and trends based on updates to the simulated condition (i.e., median waiting time). The residual bias and corresponding 95% CIs are shown on the vertical axis.

Figure 4 indicates that the methods that use the IPTW (ETCR and LA) are less sensitive to higher levels of confounding- consistently yielding the similar bias levels, the levels increase with increasing confounding levels for the TDCR method.

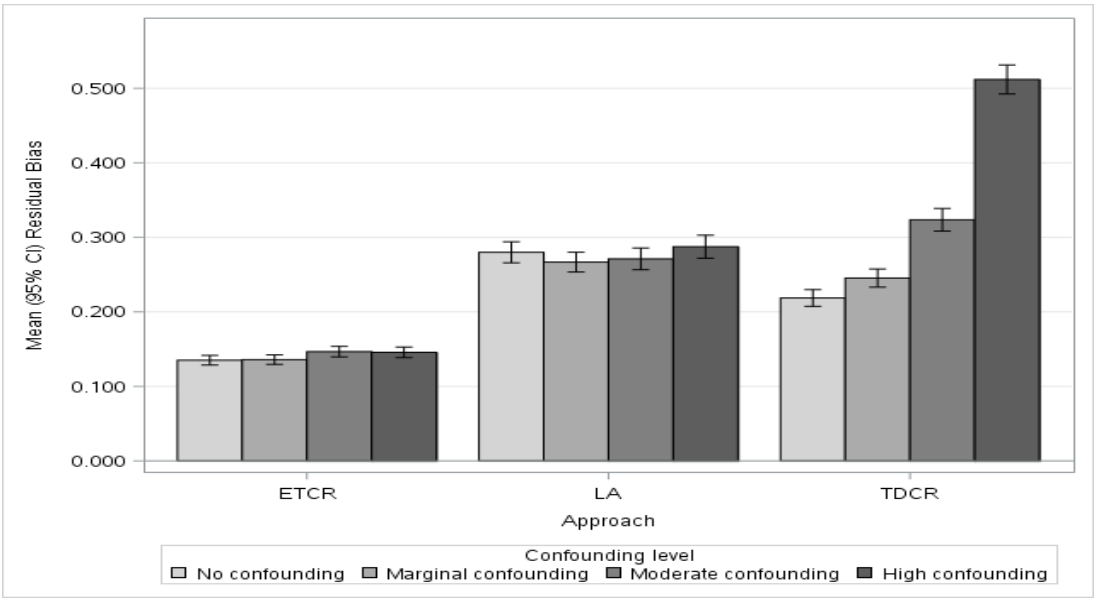

Figure 4: Mean Bias (95% CI) by Confounding Level.

This plot represents the average residual bias across different scenarios of confounding level, showing any changes in bias and trends based on updates to the simulated condition (i.e., confounding level). The residual bias and corresponding 95% CIs are shown on the vertical axis.

According to Figure 5, the higher the percentage of treated subjects, the larger the residual bias. However, the ETCR methodology demonstrates lower sensitivity with a smaller difference between the highest and lowest mean residual bias values of 0.059. In comparison, the LA and TDCR methodologies reported higher differences of 0.275 and 0.384, respectively.

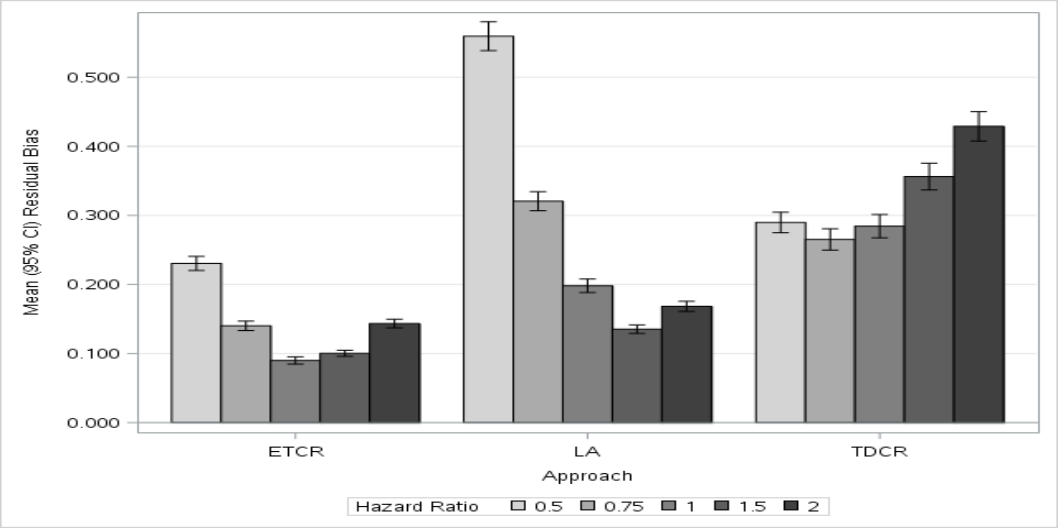

Figure 5: Mean Bias (95% CI) by Treatment Effect (Hazard Ratio).

This plot represents the average residual bias across different scenarios of treatment effect, showing any changes in bias and trends based on updates to the simulated condition (i.e., treatment effect). The residual bias and corresponding 95% CIs are shown on the vertical axis.

Other illustrations of the comparative advantages of the ETCR over the two alternative methods at different (i.e. varying) failure rate, median time-to-event, and treatment effect are provided as Figures 6-8.

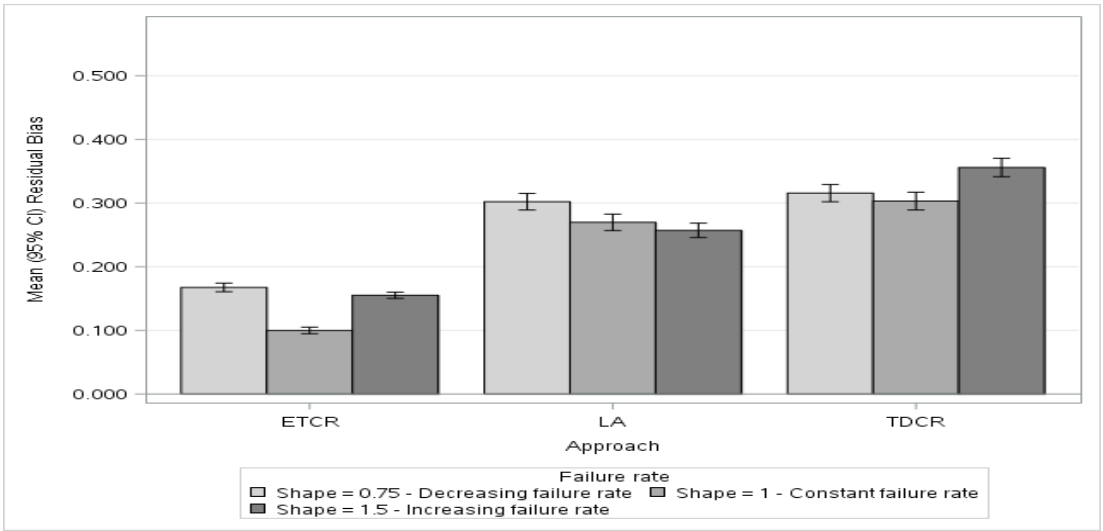

Figure 6: Mean Bias (95% CI) by Weibull Failure Rate.

Abbreviations: CI = confidence interval; ETCR = entry-time comparability retention; LA = landmark; TDCR = time-dependent Cox regression.

This plot represents the consistency of the superiority of the ETCR methodology over the others regardless of whether the underlying event rate is constant or not.

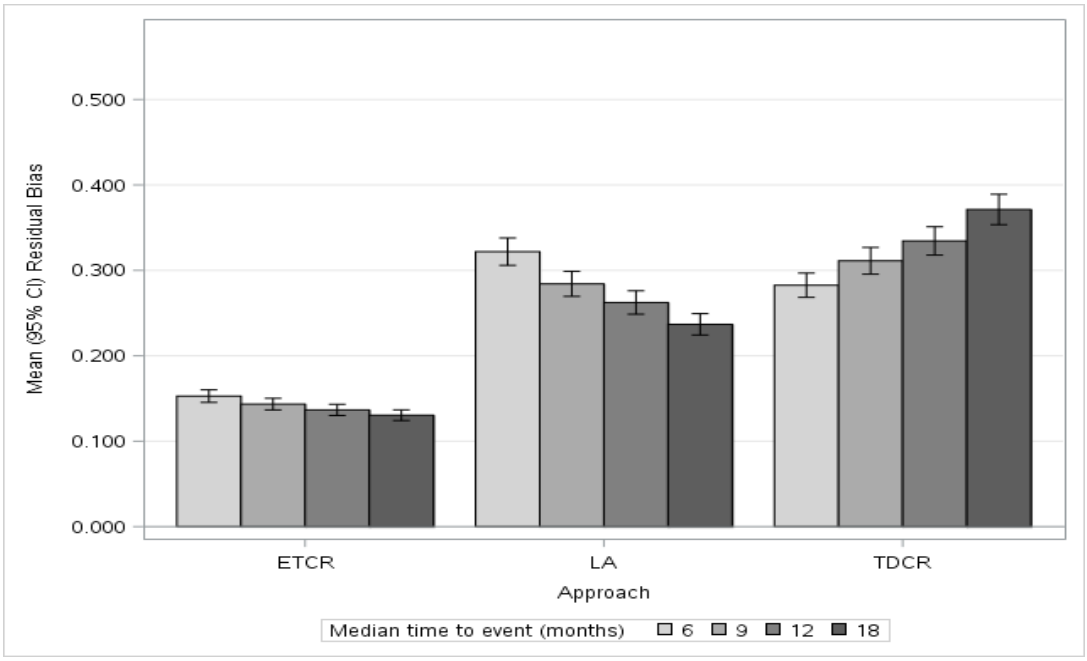

Figure 7: Mean Bias (95% CI) by Median Time to Even.

Abbreviations: CI = confidence interval; ETCR = entry-time comparability retention; LA = landmark; TDCR = time-dependent Cox regression.

This plot indicates a trend of marginally decreasing bias as the median time-to-event increases, except for the TDCR approach for which a reversed pattern is observed. The impact remains most minimal for the ETCR methodology.

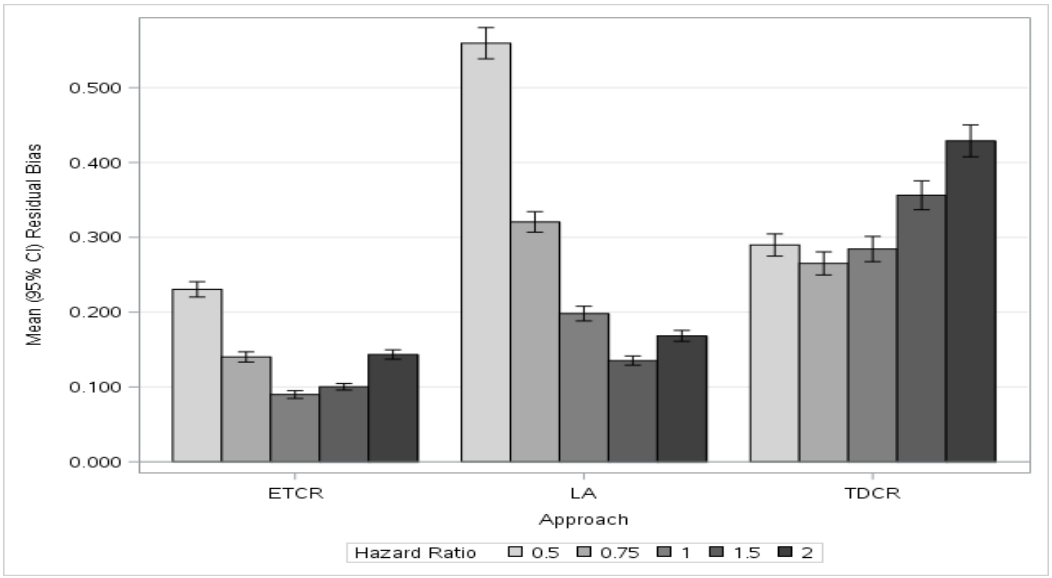

Figure 8: Mean Bias (95% CI) by Treatment Effect (Hazard Ratio).

Abbreviations: CI = confidence interval; ETCR = entry-time comparability retention; LA = landmark; TDCR = time-dependent Co.

This plot indicates the ETCR approach as consistently resulting in the lowest residual bias across the five treatment effect scenarios; each approach revealing a bathtub-like pattern that reflects the combined effect of varying forms of the Weibull distribution (i.e. decreasing, constant and increasing failure rates) and the varying hazard ratios we employed to generate the data.

5. DISCUSSION

Immortal time bias arises when the period prior to treatment initiation by a patient is either mishandled or ignored in the assessment of treatment effect in a setting where some patients are not treated. Waiting time bias, which usually arises from a random process such as a waiting list, is a subset of this bias. Indeed, appropriate

classification of the bias is essential for selecting an appropriate analytical solution to the problem, particularly because each potential solution usually involves its own set of assumptions. For example, the time-dependent Cox regression models, which collectively constitute the most appealing methodology for addressing this bias, are valid only if the condition that the change of treatment status be independent of the patient's risk profile can be realistically satisfied.[2,44-45] Thus, in settings where the change in risk profile is the reason for treatment initiation, as is often the situation in routine clinical practice, it may not be appropriate to assign the accrued persons-time for the immortal period to the untreated (unexposed) group, as has been suggested and subsequently criticized elsewhere.[1-2,44,46-47]

The pitfalls of the misuse of the time-dependent Cox models and the "events per person-time" statistics within this context, have already been reported elsewhere.[2,44-45,48] Indeed, one of the major shortcomings of these models in the handling of immortal time bias is the inherent assumption that the delay in treatment initiation (i.e. intervention) automatically makes it appropriate to consider the treatment effect as time-dependent (i.e. as varying over time). The impact of the delay in treatment initiation on treatment effectiveness or safety cannot be so readily accounted for by such assumption-laden statistical models. In other words, in the real-world evidence generation space, assumptions of time-dependent treatment effects are not so easily ascertainable or demonstrable as valid. Our illustrations are based on the most simplistic, most common version of the time-dependent Cox regression models.

In its February 2023 guidance for industry "*considerations for the design and conduct of externally controlled trials for Drug and Biological Products*", the FDA acknowledges the problem of immortal time bias as particularly challenging for situations in which "no treatment" is the treatment strategy for the external control arm.[49] Indeed, most comparative observational studies involve a similar strategy- namely, of assessing the effect of the treatment of interest between the patients who were treated (received the intervention) and those who were not treated (did not receive the intervention) and we can include most of the retrospective chart reviews we conduct in this category. There is therefore nothing wrong with the "no treatment" strategy itself. In other words, any difficulty we encounter with the analysis of data that involved the "no treatment" strategy is not because one of the groups involved patients who were never treated; we need not be unduly concerned about this reality in our definition of the treatment groups if such is based on the exposure status as observed. The problem of exposure misclassification should not be sacrificed to address that of immortal time.

Understanding the target population of our estimand is crucial for the interpretation of the study findings. The ETCR facilitates comparative analysis of patients with the same start of follow-up, based on the actual immortal time. This differentiates it from the LA estimand, where patient survival until a predetermined landmark time is a requisite, irrespective of treatment status, as that is only determined at the landmark time. The ETCR estimand is also different from that of the TDCR, wherein the target population includes all patients in the cohort and each is classified as either unexposed or exposed during follow-up, based on whether or not the patient has initiated treatment at each particular time of interest. Thus, it is easy to understand why the TDCR approach will not be feasible and the reason the Landmark approach may be problematic in trials with externally controlled arms that involve "no treatment" as the treatment strategy for the external control arm.

Another key aspect to the interpretation of the study findings is the appropriateness of the analytical approach within the context of the intention-to-treat assumption, which is inherent in the analysis, but is more suitable for randomized clinical trials. For example, the Landmark approach, which involves a fixed cut-off time window and invokes an intention-to-treat assumption, may not be appropriate in certain observational studies, particularly those about drug safety, because of the more important problem of exposure misclassification.[50] Since immortal time is only one among many sources of bias in observational studies, it is imperative that any effective solution to the problem also facilitates comparability between the treatment groups.[2] In other words, the best approach should result in the maximum possible reduction of the different sources of bias.

The third key aspect within the intention-to-treat framework is the use of the time-dependent Cox regression models with treatment effect as time-varying, whereby the proportionality assumption is by default, considered as necessarily valid. The literature suggests this assumption may not be valid if treatment effect is time varying.[23] Indeed, where we wish to handle exposure/treatment as a time-dependent/time-varying factor in our observational study, especially in our assessment of the effects of prescription drugs, statistical theory demands that we first demonstrate the underlying risk of the event of interest is constant over time, with treatment having a time-varying multiplicative effect on the risk. This is because failure of the proportionality assumption could have implications on our interpretation of the estimated treatment effect from such models. Indeed, there is evidence that where patients initiate treatment for clinical reason(s), the intention-to-treat (ITT) analysis would generally underestimate the true treatment effect, especially with such models.[50]

In this paper, the ETCR methodology is proposed as a new approach for addressing immortal time bias in observational studies without any additional assumption. It is essentially a design tool that uses two sets of propensity scores, firstly as a tool at cohort entry for

identifying the unexposed patients (i.e. on standard care) that are eligible for assignment of the same start of follow-up time with their comparable exposed patients and secondly, in its traditional role as a balancing tool to facilitate comparability between the treatment groups, based on the updated patient characteristics and clinical history at treatment initiation; in the process, we are able to minimize the problem of immortal time bias and also minimize those biases associated with confounding. The ETCR therefore allows for the use of any of the current analytical methods with the second set of propensity scores, such as matching and weighting. The results from our analysis of the simulated data for different scenarios, in comparison with the Landmark and time-dependent Cox methodologies, consistently indicate the new approach as the most effective in terms of residual bias. The results also confirmed the problem of exposure misclassification associated with the Landmark approach as well as that of over-adjustment associated with the time-dependent Cox method for addressing immortal time bias and the ETCR methodology as the least sensitive to the different simulated scenarios. The new methodology is distinctively different from the prescription time-distribution matching (PTDM) approach, which is based on random sampling of the unexposed subjects, and which has been demonstrated elsewhere as possibly inferior to the time-dependent Cox method.[14] It is also different from the method of matching with propensity scores derived with time-dependent covariates as well as those that utilized the time-dependent Cox models to create time-dependent propensity scores.[51-54]

Our paper follows a common practice of using simulated data to compare a new method for handling immortal time bias with the time-dependent Cox regression (TDCR) approach. In the process, we are mindful of the literature which suggests the assumptions for appropriate applications of the time-dependent Cox models are unlikely to be valid in the setting of prescription therapies, where treatment decisions are usually based on treatment guidelines, which inform or ought to inform the clinician's assessment of the risk.[23,44-45,50] Indeed, evidence from clinical practice suggests there are several settings where longer waiting times may be associated with elevation of the risk of interest.[55-57] In other words, the fundamental requirement of the time-dependent Cox models, which is assumed when used on such data- namely, that change of treatment status should be independent of the risk of the outcome of interest (i.e. the underlying risk is unchanged at the time of treatment initiation) is unlikely to be realistic in the setting of the intention-to-treat framework of routine clinical practice. Similarly, with the time-dependent Cox models, the proportionality assumption which is vital in our assessment and interpretation of treatment effect may not necessary be valid. Our criticism of the use of time-dependent Cox models for immortal time bias is restricted to those that consider the delay in treatment initiation as necessarily making treatment a time-dependent variable. Thus, we have nothing against those that are used for the handling of time-dependent covariates.[58] We can think of no rationale for the assumption of a time-varying effect for the treatment/intervention primarily because of delay in its initiation by the physician. Indeed, the problems associated with the adoption of untested assumptions to address the problem of immortal time bias as described, similarly apply to those approaches that involve the sequential Cox and marginal structural models.[59-64] Others in this category include the version of the "target trial emulation" approach that involved multiple emulated trials over the study follow-up period.[65] The major limitation of the latter approach is the inherent assumption that the risk profiles of the patients involved in the multiple emulated trials- up to their treatment initiation dates, would remain either unchanging as if each such patient is a clone of itself, with the same risk profile at each of these trials or any changes in the risk profiles can be adequately accounted for with a suitable statistical model. In other words, the multiple emulated trials approach is based on the statistical principle of treatment switching and as such, it is susceptible to the problems associated with inappropriate adoption of that statistical principle- namely, that the reason for switching should be independent of the risk of the event of interest. In other words, it assumes that in routine clinical practice, the reason for initiation of treatment by the physician is unconnected with the guidelines that inform appropriate prescribing of the treatment, as if the decision to initiate treatment is not informed by any observed change(s) or even concern about possible change(s) in the patient's risk profile, but by something else that is quantifiable. Even if we accept that we can use a suitable statistical technique to adequately account for the repeated use of individuals, this assumption- which translates to a random process as responsible for the reason to initiate treatment- is unlikely to be tenable in the real-world setting of clinical practice involving prescription drugs because of the realistic possibility of time-varying (1) risk profiles, (2) unmeasured confounding and of course, (3) confounding by indication, to list but three. Furthermore, these models are generally being applied within the intention-to-treat principle, to address the problem of immortal time bias, whilst ignoring many of the problems associated with routine clinical practice- namely issues resulting from patient noncompliance, such as treatment discontinuation and ad hoc consumption, to list but two.

Our position is that the assumptions we adopt to address the problem of immortal time bias in real-world data, especially those about prescription drugs, should also take cognisant of the other sources of bias, some of which may be more impactful- such as exposure misclassification, confounding by indication and channelling bias, to name but only three. We are suggesting the issues associated with the use of unproven/untested assumptions are particularly relevant to the realm of assessment of the effectiveness and safety

of clinical interventions in real-world evidence generation. In this regard, the ETCR approach is particularly suitable for such studies because of the need to avoid or minimize the critical problems of exposure misclassification and confounding. It is designed to address the problem associated with arbitrary or unproven assumptions in the definition of an index date. The literature suggests none of the approaches involving sophisticated statistical models and designs, which have been proposed for addressing the problem of immortal time bias have been utilized in the generation of evidence for regulatory submission. We are not surprised! The difficulty with establishing the appropriateness and validity of the assumptions associated with these methodologies may account for the lack of their application in the regulatory space. Indeed, the much simple nested-case control design, with its inherent time-dependent nature is generally preferred in such settings even by some of the chief proponents of the use of the TDCR approach for addressing immortal time bias.[66-72]

Our experience suggests most observational treatment safety and effectiveness studies are based on the "no treatment" strategy- most are based on either single product registries or chart reviews on patients exposed to the treatment of interest. The problem of a suitable index date for patients either on standard care or without treatment may account for some of the reasons for the dominance of single-arm (i.e. single product) observational studies of the descriptive type. The simplicity and easy of application of the ETCR methodology ought to make it the method of choice for addressing immortal time bias in such studies. Indeed, we expect the new approach to enhance the value of not only real-word data in the design and conduct of externally controlled trials, but also facilitate enhanced utility of disease registries in general, and product registries in particular, for the conduct of comparative post-authorization safety studies, by the involvement of external and/or internal comparator groups with "no treatment" as the treatment strategy.

Matching is an inherently inefficient, limiting method and more so for application in a rare disease- a limitation which may also apply to 1:n greedy matching on the first set of propensity scores to construct the blocks of exposed and unexposed patients. Another limitation of the ETCR is the problem of unmeasured confounding which is likely in the absence of data on any key patient characteristic because of its reliance on propensity scores derived from observed/measured characteristics. We also acknowledge as a limitation, the assumption that the updated data on the relevant covariates obtained at the time of treatment initiation will be sufficiently reliable between the comparative groups, since information derived at treatment initiation are likely to be more up to date for the treated (exposed) patients than for their corresponding untreated (unexposed) counterparts. Nevertheless, it is reasonable to assume in the derivation of the second set of propensity scores, that the values obtained at cohort entry on the unexposed patient may remain valid until they are updated in the patient's subsequent records prior to or at the start of follow-up. We acknowledge that like the ETCR approach, the Landmark method is simple to implement, and its results are easy to describe and interpret. Indeed, the approach may be more suitable than the ETCR for settings where treatment is known to have started within a specific time-period from cohort entry[50,73-74] or where a suitable threshold for exposure definition can be defined.[67-68] We also acknowledge that there may be settings where the assumptions associated with the use of the time-dependent Cox regression approach may be realistic, such as for data from surgical interventions, including organ transplantation that are based on waiting lists. Indeed, the TDCR may be more effective than the ETCR for addressing the problem of immortal time bias in such settings, including vaccination programs where the drug is administered to every eligible, consenting candidate on a waiting list. However, in certain settings where the TDCR can be utilized appropriately, we may also need to take into account, some other possible implications of the time-dependent treatment effect. For example, inclusion of a separate parameter in the TDCR model to control for possible transient effect on the risk of interest by the initiation of treatment has been proposed for some settings.[23,75]

It is noteworthy that the problem of confounding can be more impactful than that of immortal time in many observational studies, and hence the need to validate the assumptions we introduce in our attempt at addressing the lesser problem. In this regard, the ETCR methodology offers a simplistic approach for handling both problems and although, we have used greedy matching at 1:n ratio in our illustrations, other propensity score-based matching or weighting methods can be used with the new method. The methodology is applicable to any clinical setting where patients may have to wait for the intervention of interest primarily for clinical reasons (e.g. compliance with clinical guidelines instead of non-availability of treatment). Indeed, the increasing involvement of machine learning techniques along with their associated computational power in the derivation of propensity scores should make the ETCR methodology readily scalable to larger datasets and facilitate improvement in the handling of immortal time bias.[76-82]

6. CONCLUSIONS

Immortal time bias, which occurs when the period prior to treatment initiation is unaccounted for in the assessment of treatment effect, is common in observational studies. However, current methods for resolving the problem have their limitations and the most preferred approach is based on assumptions, which may not necessarily hold true in certain real-world settings, and more so in observational safety studies. A new, simple-and-easy-to-use approach called the Entry-time Comparability Retention (ETCR) method is proposed that involves using two sets of the propensity scores-

firstly, as a design tool for the creation of patient blocks, to handle the source of the immortal time such that the unexposed patients (i.e. those on standard care) who are comparable to each exposed patient at cohort entry can constitute the same block and can be assigned the same start of follow-up time as that of the treatment initiator, and secondly in its conventional role as a balancing tool, to retain comparability and address the problem of confounding. The ETCR methodology allows for the use of any suitable analytical approach such as matching or weighting with its second set of propensity scores. In other words, although the approach does not add anything new to any of the current analytical methods for handling the problems of confounding, it does avoid most of the limitations associated with the current approaches. The results from simulated data that reflect some of the different possible real-world settings associated with observational studies, support the ETCR methodology when applied with the Inverse Probability of Treatment Weighting approach as superior to two of the leading alternative methodologies in the current literature. Although the literature on immortal time bias supports our use of simulated data for comparative assessment of the performance of the different methods, we expect and strongly recommend the conduct of similar comparative assessments of these methods based on real-life data.

AUTHORS' CONTRIBUTIONS

VK was involved in the conception of the methodology, management of the research, interpretation of the results and writing of the manuscript.

PM was involved in the management of the research, data analysis, interpretation of the results and writing of the manuscript.

ACKNOWLEDGMENTS

The authors wish to acknowledge the enormous support and encouragement provided by their colleague, Tyler Knight on this research, especially for the useful suggestions about the results and for the several rounds of review of the manuscript.

FUNDING

Research was funded by Fortrea Inc.

CONFLICT OF INTEREST

Both authors are employees of Fortrea Inc. and since their employer had no role in any aspect of the research, there are no actual or potential conflicts of interest to disclose.

AVAILABILITY OF DATA AND MATERIALS

The data that support the findings of this study were generated by simulation as described in the 'Methods' section and are available from the corresponding author on request.

REFERENCES

1. Suissa S: Effectiveness of inhaled corticosteroids in chronic obstructive pulmonary disease: immortal time bias in observational studies. Am J Respir.Crit Care med., 2003; 168: 49-53
2. Kiri VA. Mishandling of immortal time bias in observational studies- we can do better. Biom Biostat Int J., 2022; 11(4): 131-132. http://dx.doi.org/10.15406/bbij.2022.11.00366
3. Suissa S. Immortal time bias in observational studies of drug effects. Pharmacoepidemiol Drug Saf., 2007; 16: 241-249.
4. Rothman KJ, Greenland S. Cohort studies. In: Rothman KJ, Greenland S, eds. Modern epidemiology. 2nd ed. Lippincott-Raven, 1998; 79-91.
5. Suissa S. Inhaled steroids and mortality in COPD: bias from unaccounted immortal time. Eur Respir J., 2004; 23: 391-395.
6. Tyrer F, Bhaskaran K, Rutherford MJ. Immortal time bias for life-long conditions in retrospective observational studies using electronic health records. BMC Med Res Methodol., 2022 Mar 27; 22(1): 86. doi: 10.1186/s12874-022-01581-1
7. Yadav K, Lewis RJ. Immortal Time Bias in Observational Studies. JAMA, 2021 Feb 16; 325(7): 686-687. doi: 10.1001/jama.2020.9151
8. Daniel S, Koren G, Lunenfeld E, Levy A. Immortal time bias in drug safety cohort studies: spontaneous abortion following nonsteroidal antiinflammatory drug exposure. Am J Obstet Gynecol., 2015 Mar; 212(3): 307.e1-6. doi: 10.1016/j.ajog.2014.09.02
9. Liang C, Seeger JD, Dore DD. Implications of immortal person-time when outcomes are nonfatal. Ann Epidemiol., 2016 Mar; 26(3): 212-217. doi: 10.1016/j.annepidem.2015.12.004.
10. Targownik LE, Suissa S. Understanding and Avoiding Immortal-Time Bias in Gastrointestinal Observational Research. Am J Gastroenterol., 2015 Dec; 110(12): 1647-1650. doi: 10.1038/ajg.2015.210
11. Targownik LE. Immortal time bias and infliximab-related mortality and malignancy incidence. Gut., 2010 Mar; 59(3): 416; author reply 416. doi: 10.1136/gut.2009.186445
12. Lewis JD. Immortal time bias in estimates of mortality among infliximab-treated patients with Crohn's disease. Gut., 2010 Nov; 59(11): 1586-1587. doi: 10.1136/gut.2009.191015
13. Liu J, Weinhandl ED, Gilbertson DT, Collins AJ, St Peter WL. Issues regarding 'immortal time' in the analysis of the treatment effects in observational studies. Kidney Int., 2012 Feb; 81(4): 341-350. doi: 10.1038/ki.2011.388
14. Hemmelgarn BR, Zhang J, Manns BJ, Tonelli M. Multidisciplinary care and immortal time bias. Kidney Int., 2013 Nov; 84(5): 1052. doi: 10.1038/ki.2013.257
15. Nichol KL, Nordin JD, Nelson DB, Mullooly JP, Hak E. Effectiveness of influenza vaccine in the

community-dwelling elderly. N Engl J Med., 2007; 357: 1373-1381
16. Trojano M, Pellegrini F, Fuiani A, Paolicelli D, Zipoli V, Zimatore GB, et al. New natural history of interferon-beta-treated relapsing multiple sclerosis. Ann Neurol., 2007; 61: 300-306.
17. Jackevicius CA, Li P, Tu JV. Prevalence, predictors, and outcomes of primary non-adherence after acute myocardial infarction. Circulation, 2008; 117: 1028-1036.
18. Kiri V, Messina P, Knight T, Anastassopoulos, KP. A Novel Approach for Minimizing Immortal Time Bias in Observational Studies. Value in Health, 2023; 25(12): S350. (ISPOR Annual European Congress Meeting)
19. Harding BN, Weiss NS. Point: Immortal Time Bias-What Are the Determinants of Its Magnitude? Am J Epidemiol., 2019 Jun 1; 188(6): 1013-1015. doi: 10.1093/aje/kwz067
20. Messmer BJ, Nora JJ, Leachman RD, Cooley DA. Survival-times after cardiac allografts. Lancet., 1969; 1(7602): 954-956
21. Clark DA, Stinson EB, Griepp RB, Schroeder JS, Shumway NE, Harrison DC. Cardiac transplantation in man. VI. Prognosis of patients selected for cardiac transplantation. Ann Intern Med., 1971; 75(1): 15-21.
22. Kalbfleisch J and Prentice R. The statistical analysis of failure time data. New York: Wiley, 1980.
23. Cox DR and Oakes D. Analysis of survival data. New York: Chapman and Hall, 1984.
24. Sheehy O, LeLorier J, Rinfret S. Restrictive access to clopidogrel and mortality following coronary stent implantation. CMAJ., 2008; 178: 413-420.
25. Gail MH. Does cardiac transplantation prolong life? A reassessment. Ann Intern Med., 1972; 76(5): 815-817.
26. Glynn RJ, Knight EL, Levin R, Avorn J. Paradoxical relations of drug treatment with mortality in older persons. Epidemiology, 2001; 12(6): 682-689.
27. Schneeweiss S, Patrick AR, Stürmer T, et al. Increasing levels of restriction in pharmacoepidemiologic database studies of elderly and comparison with randomized trial results. Med Care., 2007; 45(10 suppl 2): S131-S142.
28. Dekkers OM, Groenwold RHH. When observational studies can give wrong answers: the potential of immortal time bias. Eur J Endocrinol., 2021; 184(1): E1-E4.
29. Kiri VA, McKenzie G. How Real is Intention-To-Treat (ITT) Analysis in Non-Interventional Post Authorization Safety Studies? We Can Do Better. Current Drug Safety, 2009; 4(2): 137-142.
30. Rosenbaum PR and Rubin DB. The central role of propensity score in observational studies for causal effects Biometrika, 1983; 70: 41-55.
31. Karim ME, Gustafson P, Petkau J, Tremlett H. Comparison of statistical approaches for dealing with immortal time bias in drug effectiveness studies. Am J Epidemiol., 2016; 184(4): 325–335.
32. 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 Methodol., 2017; 17(1): 126.
33. Buyse M, Piedbois P. On the relationship between response to treatment and survival time. Stat Med., 1996; 15: 2797–2812.
34. Snapinn S, Jiang Q, Iglewicz B. Statistical practice illustrating the impact of a time-varying covariate with an extended Kaplan–Meier estimator. Am Stat., 2005; 59: 301–307.
35. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response and other comparisons of time-to-event by outcome variables. J Clin Oncol., 2008; 26: 3913–3915.
36. Dafni U. Landmark analysis at the 25-year landmark point. Circ Cardiovasc Qual Outcomes, 2011 May; 4(3): 363-371. doi: 10.1161/CIRCOUTCOMES.110.957951
37. Mi X, Hammill BG, Curtis LH, Lai EC, Setoguchi S. Use of the landmark method to address immortal person-time bias in comparative effectiveness research: a simulation study. Stat Med., 2016 Nov 20; 35(26): 4824-4836. doi: 10.1002/sim.7019
38. Duchesneau ED, Jackson BE, Webster-Clark M, Lund JL, Reeder-Hayes KE, Nápoles AM, Strassle PD. The Timing, the Treatment, the Question: Comparison of Epidemiologic Approaches to Minimize Immortal Time Bias in Real-World Data Using a Surgical Oncology Example. Cancer Epidemiol Biomarkers Prev, 2022 Nov 2; 31(11): 2079-2086. doi: 10.1158/1055-9965.EPI-22-0495
39. Zhou Z, Rahme E, Abrahamowicz M, Pilote L. Survival bias associated with time-to-treatment initiation in drug effectiveness evaluation: A comparison of methods. Am J Epidemiol., 2005; 162(10): 1016-1023.
40. Agarwal P, Moshier E, Ru M, Ohri N, Ennis R, Rosenzweig K, Mazumdar M. 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 Jan-Dec; 25(1): 1073274818789355. doi: 10.1177/1073274818789355
41. Jones M, Fowler R. Immortal time bias in observational studies of time-to-event outcomes. J Crit Care., 2016; 36: 195–199
42. Seeger JD, Walker AM, Williams PL, Saperia GM, Sacks FM. A propensity score-matched cohort study of the effect of statins, mainly fluvastatin, on the occurrence of acute myocardial infarction. Am J Cardiol., 2003 Dec 15; 92(12): 1447-51. doi: 10.1016/j.amjcard.2003.08.057
43. Bender R, Augustin T, Blettner M. Generating survival times to simulate Cox proportional hazards models. Stat Med., 2005 Jun 15; 24(11): 1713-1723.
44. Kiri VA, Mackenzie G. Re: "Immortal time bias in pharmacoepidemiology" [letter]. Am J Epidemiol., 2009 Sep 1; 170(5): 667-668

45. Fisher LD, Lin DY. Time-dependent covariates in the cox proportional-hazards regression model. Annu. Rev. Public Health, 1999; 20: 145–157.
46. Suissa S. Inhaled steroids and mortality in COPD: bias from unaccounted immortal time. Eur Respir J., 2004; 23: 391-395.
47. Suissa S. Immortal time bias in pharmaco-epidemiology. Am J Epidemiol., 2008; 167(4): 492-499.
48. Kraemer HC. Events per person-time (incidence rate): a misleading statistic? Stat Med., 2009; 28(6): 1028-1039.
49. FDA Guidance for Industry: Considerations for the Design and Conduct of Externally Controlled Trials for Drug and Biological Products. https://www.fda.gov/regulatory-information/search-fda-guidance. February, 2023.
50. Anderson JR, Cain KC, Gelber RD. Analysis of survival by tumor response. J Clin Oncol., 1983; 1: 710-719.
51. Lu B. Propensity score matching with time-dependent covariates. Biometrics, 2005; 61: 721–728.
52. Suissa S, Dell'Aniello S. Time-related biases in pharmacoepidemiology. Pharmacoepidemiol Drug Saf., 2020; 29(9): 1101-1110. DOI: 10.1002/pds.5083.
53. Suissa S, Moodie EE, Dell'Aniello S. Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. Pharmacoepidemiol Drug Saf., 2017; 26: 459–468. doi: 10.1002/pds.4107
54. Suissa S, Dell'Aniello S, Renoux C. The Prevalent New-user Design for Studies With no Active Comparator: The Example of Statins and Cancer. Epidemiology, 2023 Sep 1; 34(5): 681-689. doi: 10.1097/EDE.0000000000001628.
55. Rhodes A., Evans L.E., Alhazzani W., Levy M.M., Antonelli M., Ferrer R., Kumar A., Sevransky J.E., Sprung C.L., Nunnally M.E., et al. Surviving Sepsis Campaign: International Guidelines for Management of Sepsis and Septic Shock: 2016. Crit. Care Med., 2017; 45: 486–552. doi: 10.1097/CCM.0000000000002255
56. Levitov A., Frankel H.L., Blaivas M., Kirkpatrick A.W., Su E., Evans D., Summerfield D.T., Slonim A., Breitkreutz R., Price S., et al. Guidelines for the Appropriate Use of Bedside General and Cardiac Ultrasonography in the Evaluation of Critically Ill Patients-Part II: Cardiac Ultrasonography. Crit. Care Med., 2016; 44: 1206–1227. doi: 10.1097/CCM.0000000000001847
57. Zheng J, Yang Q, Kong T, Chen X, Wang R, Huo J, Huang W, Wen D, Xiong X, Zhang Z. Association between Wait Time for Transthoracic Echocardiography and 28-Day Mortality in Patients with Septic Shock: A Cohort Study. J Clin Med., 2022 Jul 16; 11(14): 4131. doi: 10.3390/jcm11144131
58. Hade EM, Nattino G, Frey HA, Lu B. Propensity score matching for treatment delay effects with observational survival data. Stat Methods Med Res., 2020 Mar; 29(3): 695-708
59. Zhang Z, Li X, Wu X, Qiu H, Shi H; written on behalf of AME Big-Data Clinical Trial Collaborative Group. Propensity score analysis for time-dependent exposure. Ann Transl Med., 2020 Mar; 8(5): 246.
60. Gran JM, Røysland K, Wolbers M, et al. A sequential Cox approach for estimating the causal effect of treatment in the presence of time-dependent confounding applied to data from the Swiss HIV Cohort Study. Stat Med., 2010 Nov 20; 29(26): 2757-2768. doi: 10.1002/sim.4048. PMID: 20803557
61. Williamson T, Ravani P. Marginal structural models in clinical research: when and how to use them? Nephrol Dial Transplant., 2017; 32(suppl_2): ii84–ii90. 10.1093/ndt/gfw341
62. Schaubel DE, Wolfe RA, Port FK. A sequential stratification method for estimating the effect of a time-dependent experimental treatment in observational studies. Biometrics, 2006; 623: 910–917.
63. Latimer NR, Henshall C, Siebert U, Bell H. Treatment switching: statistical and decision-making challenges and approaches. International Journal of Technology Assessment in Health Care., 2016; 32(3): 160-166.
64. Schaubel D, Wolfe R, Sima C et al. Estimating the effect of a time-dependent treatment by levels of an internal time-dependent covariate: application to the contrast between liver wait-list and posttransplant mortality. J Am Stat Assoc, 2009; 104485: 49–59.
65. García-Albéniz X, Hsu J, Hernán MA. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. Eur J Epidemiol., 2017 Jun; 32(6): 495-500. doi: 10.1007/s10654-017-0287-2
66. Sin DD, Tu JV. Inhaled corticosteroids and the risk of mortality and readmission in elderly patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med., 2001; 164: 580-584.
67. Kiri VA, Pride NB, Soriano JB, Vestbo J. *Authors' Reply* "Observational studies of inhaled corticosteroids in chronic obstructive pulmonary disease: misconstrued immortal time bias", Am J Respir Crit Care Med., 2006; 173: 464–465.
68. Kiri VA, Pride NB Soriano JB, Vestbo J. Inhaled Corticosteroids in COPD: results from two observational designs free of immortal-time bias, Am J Respir Crit Care Med., 2005; 172: 460-464.
69. Liddell FDK, McDonald JC, Thomas DC. Methods of cohort analysis: appraisal by application to asbestos mining. J R Stat Soc Ser A., 1977; 140: 469-491.
70. Breslow NE, Day NE, eds. Fitting models to continuous data. Statistical methods in cancer research. Vol.2. The design and analysis of cohort

studies. IARC Scientific Publications, 1987; 178-229.
71. Suissa S. Novel approaches to pharmacoepidemiology study design and statistical analysis. In: Strom BL, ed. Pharmacoepidemiology. 4th ed. John Wiley, 2005; 811-829.
72. Lévesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. BMJ., 2010 Mar 12; 340: b5087. doi: 10.1136/bmj.b5087
73. Malenka DJ, Kaplan AV, Lucas FL, Sharp SM, Skinner JS. Outcomes following coronary stenting in the era of bare-metal vs the era of drug-eluting stents. JAMA, 2008; 299: 2868-2876.
74. Eisenstein EL, Anstrom KJ, Kong DF, Shaw LK, Tuttle RH, Mark DB, Kramer JM, Harrington RA, Matchar DB, Kandzari DE, Peterson ED, Schulman KA, Califf RM. Clopidogrel use and long-term clinical outcomes after drug-eluting stent implantation. JAMA, 2007; 297: 159-168.
75. Bunday, B.D. & Kiri, V.A. Analysis of Censored Recidivism Data using a Proportional Hazards-type Model, The Statist, 1992; 41(1): 85-96. https://doi.org/10.2307/2348639
76. Lee BK, Lessler J, Stuart EA. Improving propensity score weighting using machine learning. Stat Med., 2010 Feb 10; 29(3): 337-46. doi: 10.1002/sim.3782
77. Setoguchi S, Schneeweiss S, Brookhart MA, Glynn RJ, Cook EF. Evaluating uses of data mining techniques in propensity score estimation: a simulation study. Pharmacoepidemiol Drug Saf., 2008 Jun; 17(6): 546-55. doi: 10.1002/pds.1555
78. McCaffrey DF, Ridgeway G, Morral AR. Propensity score estimation with boosted regression for evaluating causal effects in observational studies. Psychol Methods., 2004 Dec; 9(4): 403-25. doi: 10.1037/1082-989X.9.4.403
79. 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 Oct 15; 17(19): 2265-81. doi: 10.1002/(sici)1097-0258(19981015)17:19<2265::aid-sim918>3.0.co;2-b
80. Drake C. Effects of misspecification of the propensity score on estimators of treatment effect. Biometrics, 1993; 49: 1231–1236.
81. Elith J, Leathwick JR, Hastie T. A working guide to boosted regression trees. J Anim Ecol., 2008 Jul; 77(4): 802-13. doi: 10.1111/j.1365-2656.2008.01390.x
82. Tsiatis AA, Davidian M. Comment: Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean from Incomplete Data. Stat Sci., 2007; 22(4): 569-573. doi: 10.1214/07-STS227