Practice of Epidemiology

Comparison of Statistical Approaches for Dealing With Immortal Time Bias in Drug Effectiveness Studies

Mohammad Ehsanul Karim*, Paul Gustafson, John Petkau, Helen Tremlett, and the Long-Term Benefits and Adverse Effects of Beta-Interferon for Multiple Sclerosis (BeAMS) Study Group

* Correspondence to Dr. Mohammad Ehsanul Karim, Department of Epidemiology, Biostatistics and Occupational Health, Faculty of Medicine, McGill University, Purvis Hall, 1020 Pine Avenue West, Montreal, Quebec H3A 1A2, Canada (e-mail: ehsan.karim@mail.mcgill.ca).

In time-to-event analyses of observational studies of drug effectiveness, incorrect handling of the period between cohort entry and first treatment exposure during follow-up may result in immortal time bias. This bias can be eliminated by acknowledging a change in treatment exposure status with time-dependent analyses, such as fitting a time-dependent Cox model. The prescription time-distribution matching (PTDM) method has been proposed as a simpler approach for controlling immortal time bias. Using simulation studies and theoretical quantification of bias, we compared the performance of the PTDM approach with that of the time-dependent Cox model in the presence of immortal time. Both assessments revealed that the PTDM approach did not adequately address immortal time bias. Based on our simulation results, another recently proposed observational data analysis technique, the sequential Cox approach, was found to be more useful than the PTDM approach (Cox: bias = −0.002, mean squared error = 0.025; PTDM: bias = −1.411, mean squared error = 2.011). We applied these approaches to investigate the association of β-interferon treatment with delaying disability progression in a multiple sclerosis cohort in British Columbia, Canada (Long-Term Benefits and Adverse Effects of Beta-Interferon for Multiple Sclerosis (BeAMS) Study, 1995–2008).

bias (epidemiology); confounding factors (epidemiology); epidemiologic methods; immortal time bias; longitudinal studies; models; survival analysis

Abbreviations: BeAMS, Long-Term Benefits and Adverse Effects of Beta-Interferon for Multiple Sclerosis; CI, confidence interval; HR, hazard ratio; log HR, log hazard ratio; MS, multiple sclerosis; PTDM, prescription time-distribution matching; RR, rate ratio.

In many observational studies of drug effectiveness, there may be a delay or waiting period before a subject begins to receive a treatment. Let us use multiple sclerosis (MS) as an example, with the disease-modifying drug β-interferon being the study drug of interest. If treated subjects entered a cohort at the time of their first β-interferon administration but untreated subjects entered the cohort at an earlier time, such as their first assessment at an MS clinic or study center conducting the observational study (as in the paper by Trojano et al. ([1])), then the time period between the first clinic visit (baseline) and the first administration of β-interferon corresponds to "immortal time" ([2], [3]). This delay period is labeled "immortal" because, by design, subjects in the treated group cannot develop an outcome before receiving the treatment. Therefore, if some subjects develop the outcome before getting the opportunity to initiate treatment, they will be assigned to the untreated group.

If treatment status is misclassified during this delay period or if this delay period is excluded from the analysis, "immortal time bias" is introduced. Classifying subjects as "treated" or "untreated" with a time-invariant variable results in a spurious survival advantage (protective association) in favor of the treated group. This can considerably distort the estimated hazard ratio if a large number of failures (events or reaching outcome) occur before the initiation of treatment or if the duration of the immortal time is large ([4]). Although this bias was first identified in the 1970s ([5]), many studies still fail to account for this source of bias ([2], [6], [7]). Recently, the issue of immortal time bias has resurfaced in pharmacoepidemiology studies in the context of propensity score adjustment ([3], [8]–[11]), where treatment status must be defined at baseline.

The most common approach to accounting for immortal time bias is the adoption of a proper treatment exposure

definition via time-dependent analyses (12), such as time-dependent Cox proportional hazards models (13–16). Prescription time-distribution matching (PTDM) (13) is a simple approach to adjusting for immortal time bias that has been frequently cited (17–21) and used (22–28) in recent years. The performance and statistical efficiency of this approach have been claimed to be similar to those of the Cox model with a binary time-dependent exposure (hereafter called the "time-dependent Cox model") (13).

Several studies have quantified the amount and direction of bias due to "misclassifying" or "ignoring" immortal time by means of simulation (14, 29–31) or theoretically (15, 32–34); however, to the best of our knowledge, no attempt has been made to explore the appropriateness of the PTDM approach in minimizing immortal time bias. Our primary focus in this study was to assess the performance of the PTDM for dealing with this bias. We quantified the bias due to PTDM via simulation studies as well as theoretically (see Web Appendix 1 and Web Figure 1, available at http://aje.oxfordjournals.org/) For comparison, we considered a simplified version of the sequential Cox approach (35), which was originally proposed for dealing with time-dependent confounding (36). Additionally, misclassified and excluded immortal time approaches were also compared. We applied these methods to investigate the association of time-varying β-interferon treatment with delaying disability progression in subjects from the database of a British Columbia, Canada, MS study, the Long-Term Benefits and Adverse Effects of Beta-Interferon for Multiple Sclerosis (BeAMS) Study (1995–2008) (37, 38).

METHODS

Notation

Consider a hypothetical longitudinal study consisting of n subjects ($i = 1, 2, . . ., n$). Let $t_0 = 0$ be the start of follow-up (i.e., the time of the baseline visit). Data on baseline covariates L_0 (binary) are recorded at this time. Follow-up continues until the time of failure (reaching outcome) T or censoring T^C. Regular measurements of the binary treatment status A_{t_m} ($A_{t_m} = 1$ for treated and 0 otherwise) are recorded at intervals $m = 0, 1, 2, . . ., K$. Let $[t_m, t_{m+1}]$ denote the mth interval. Because this study is focusing on the implications of immortal time, we assume that the subjects may initiate treatment no more than once and continue taking the treatment thereafter until the study ends (i.e., the value for treatment exposure changes only once, from 0 to 1). Let treatment initiation occur at time T^A.

For group-based comparison, suppose $k = 1$ indicates the ever-treated group, whereas $k = 0$ indicates the never-treated group. Let n_k indicate the size of these groups. Further, let N_k and T_k ($k = 0, 1$) indicate the observed number of failures and follow-up person-time in these groups, respectively. Let $r = T_0 : T_1$, the ratio of the observed person-times in the never-treated and ever-treated subjects. Denote T_{IT} as the observed immortal time—that is, the aggregated follow-up time not under treatment in the ever-treated group—and set $f = T_{IT}/T_1$. We describe the analytical approaches in detail in the following sections.

Cox models with a time-invariant treatment variable

Two imprecise treatment exposure definitions are often used to estimate the log hazard ratio (log HR), where a Cox model with time-invariant treatment variable A (which does not depend on the time interval) is fitted while adjusting for the potential confounders measured at the original baseline as follows:

$$\lambda_T(t_m|L_0) = \lambda_0(t_m)\exp(\psi_1 A + \psi_2 L_0),$$

where m is the visit index, $\lambda_0(t_m)$ is the unspecified baseline hazard function, ψ_1 is the log HR of the treatment status (A), and ψ_2 is the vector of log HRs for the baseline covariates L_0.

Included immortal time. In this approach, subjects who were ever exposed to treatment ("ever-treated") are classified as treated for their whole duration of follow-up. Under the assumption of constant hazard, comparing the failure rate ratio from this approach (RR′) with the correct rate ratio (RR) obtained from a time-dependent analysis (see Suissa (15) and Web Appendix 2) yields

$$\frac{RR'}{RR} = \frac{\dfrac{N_1/T_1}{N_0/T_0}}{\dfrac{N_1/(T_1 - T_{IT})}{N_0/(T_0 + T_{IT})}} = (1 - f) \times \frac{r}{(r+f)}.$$

From this equation, we can see that this approach always underestimates the correct failure rate ratio (also see Web Figure 2).

Excluded immortal time. In this approach, the immortal time—that is, time from cohort entry to the initiation of treatment—is excluded from the ever-treated subjects' follow-up, and time zero for these subjects is taken to be the time of treatment initiation, T^A. However, the follow-up period for the subjects who were never exposed to treatment ("never-treated") remains the same; that is, time zero is the time of cohort entry $t_0 = 0$. Comparing the failure rate ratio from this approach (RR″) with the correct rate ratio obtained from a time-dependent analysis (see Suissa (15) and Web Appendix 2) yields

$$\frac{RR''}{RR} = \frac{\dfrac{N_1/(T_1 - T_{IT})}{N_0/T_0}}{\dfrac{N_1/(T_1 - T_{IT})}{N_0/(T_0 + T_{IT})}} = \frac{r}{(r+f)}.$$

From this equation, we can see that if the person-times in the ever-treated cohort are much smaller than the person-times in the never-treated cohort (i.e., if r is large), the bias from this approach may be negligible, even for large fractions of immortal time f (also see Web Figure 3).

Cox model with a time-dependent treatment variable

In the presence of the baseline covariates L_0, the hazard function can be expressed as the following time-dependent Cox model:

$$\lambda_T(t_m|L_0) = \lambda_0(t_m) \exp(\psi_1 A_{t_m} + \psi_2 L_0), \qquad (1)$$

where ψ_1 is the log HR of the time-dependent treatment status (A_{t_m}).

The PTDM approach

The essence of the PTDM approach is to redefine time zero in both the ever-treated and the never-treated groups (13) by shifting the start of follow-up to the time of treatment initiation (the end of the immortal time period) T^A for the ever-treated subjects. First the immortal time (wait) periods for all of the ever-treated subjects are listed. For each never-treated subject, a time zero is selected at random (with replacement) from this list and assigned as the new time zero for the corresponding never-treated subjects. If a never-treated subject has an event before the end of the assigned immortal time (new time zero), he or she is excluded from further analysis. The analysis is performed on the new time zero points—that is, the newly defined baselines after exclusion of the observed or assigned immortal time from the follow-up for the ever-treated and never-treated groups, respectively (see Web Figure 1). This is expected to eliminate the imbalance in the excluded time distribution between the two treatment groups.

Let us denote the immortal time for the jth treated subject as $T_{j,\mathrm{IT}}$. Similarly, the matched-immortal time for the j'th untreated subject is denoted by $T_{j',\mathrm{IT}}$. The aggregated immortal time is denoted by $T_{\mathrm{IT}} \equiv \sum_{j=1}^{n1} T_{j,\mathrm{IT}}$, and we denote the aggregated matched-immortal time by $T'_{\mathrm{IT}} \equiv \sum_{j'=1}^{n0} T_{j',\mathrm{IT}}$. The number of failures observed during the aggregated immortal time T_{IT} is zero. Let N'_{IT} be the number of failures observed during the aggregated matched-immortal time T'_{IT}. For $N'_{\mathrm{IT}} \geq 0$, the total untreated person-time T_x of the N'_{IT} patients who had failures within their assigned immortal times is excluded. Set $x = T_x/T_1 \geq 0$. In addition, set $T'_{\mathrm{IT}} = q \times T_{\mathrm{IT}}$; that is, q is the ratio of assigned and observed immortal times. Here, $q > 1$ for the setting in which the aggregated matched-immortal time T'_{IT} is larger than the aggregated observed immortal time T_{IT}; otherwise, $0 < q < 1$. Comparing the failure rate ratio from PTDM (RR''') with the correct rate ratio obtained from a time-dependent analysis (detailed derivation shown in Web Appendix 1) yields

$$\frac{RR'''}{RR} = \frac{\dfrac{N_1/(T_1 - T_{\mathrm{IT}})}{(N_0 - N'_{\mathrm{IT}})/(T_0 - T'_{\mathrm{IT}} - T_x)}}{\dfrac{N_1/(T_1 - T_{\mathrm{IT}})}{N_0/(T_0 + T_{\mathrm{IT}})}} = \frac{N_0}{N_0 - N'_{\mathrm{IT}}} \frac{r - q \times f - x}{r + f}.$$

Varying r, f, x, and q yields Figure 1. This approach shows a downward bias for increasing values of f, the fraction of the immortal person-time in the ever-treated subjects. The level of bias increases with both q and x. As with the excluded immortal time approach (described in Web Appendix 2), the bias from this approach is also small if the ever-treated cohort (or its person-time contribution) is much smaller than the never-treated cohort (i.e., r is large), even for large fractions of immortal time f (see Figure 1).

Further theoretical assessment of this approach is provided in Web Appendix 3.

The sequential Cox approach

The sequential Cox approach (35) mimics a randomized clinical trial for each of the intervals m ($m = 0, 1, 2, \ldots, K$) during which subjects initiate treatment. Based on treatment initiation at the mth interval, the mth minitrial is created as

follows: Only subjects who have not received any treatment before the mth interval are considered. Among these subjects, those initiating treatment during the mth interval ($t_m < T^A \leq t_{m+1}$) are considered the treated group, while the remaining subjects are considered the control group (for an illustration, see Web Figure 4 in Web Appendix 4). This mode of data restructuring closely resembles the "sequence of trials" approach (39) but differs with regard to some details; for example, in the current approach, treatment weights are not used to avoid potential instability in the estimated weights (35, 36) and the control subjects are artificially censored at the time of later treatment initiation ($T^A > t_{m+1}$) to avoid confounding due to misclassification of treatment status. To address such artificial censoring, the analysis results are adjusted using inverse-probability-of-censoring weights (36, 40).

We create the pseudodata by aggregating the data from all minitrials (see Web Appendix 4 for implementation details). Under the assumption that the different minitrials may have different baseline hazard functions but all subjects in the same minitrial will have the same baseline hazard function, we can fit a weighted stratified (time-invariant) Cox model to the pseudodata, stratified by the treatment initiation interval and weighted by the inverse-probability-of-censoring weights. Because the same subject may be included in more than 1 minitrial, we use the robust (sandwich) estimator to obtain the standard error (41, p. 170; 42, 43).

Design of simulation

To simulate survival times, we adapt the permutation algorithm (44), which has been validated for generating survival times conditional on time-dependent variables (45, 46) (see Web Appendix 5).

Simulation specifications. In all of our simulations, we generated $N = 1,000$ data sets, each with $n = 2,000$ subjects followed for up to $m = 30$ subsequent visits. For a given percentage of subjects, treatment initiation time T^A is generated from a uniform distribution $U(0, 30)$ (in months). In our simulations 1–3, the percentages of subjects who remain unexposed to treatment for the entire duration of follow-up (infinite T^A) are set to 0%, 25%, and 75%, respectively, for the case in which there is no failure or censoring during follow-up. When subjects fail (reach the outcome) or become censored before they get a chance to initiate treatment, the percentages of never-exposed subjects increase. In all of these simulation settings, we assume an exponential distribution for generating failure times T with the constant $\lambda_0 = 0.01$ rate of monthly events throughout the follow-up period. A uniform distribution $U(1, 60)$ months is assumed to generate censoring times T^C. Administrative censoring is set at 30 months of follow-up for all subjects who remain at risk at that time. For each simulated data set, we can calculate the exact percentages of ever-treated and never-treated subjects. For any subject, treatment status is binary for any given month, either exposed or unexposed, where a person is considered exposed throughout the month of the treatment initiation. To focus on the immortal time issue, we assume there are no discontinuations or interruptions for persons who initiate treatment. Additionally, we consider sex as a baseline confounder. A subject's sex is generated based on a Bernoulli
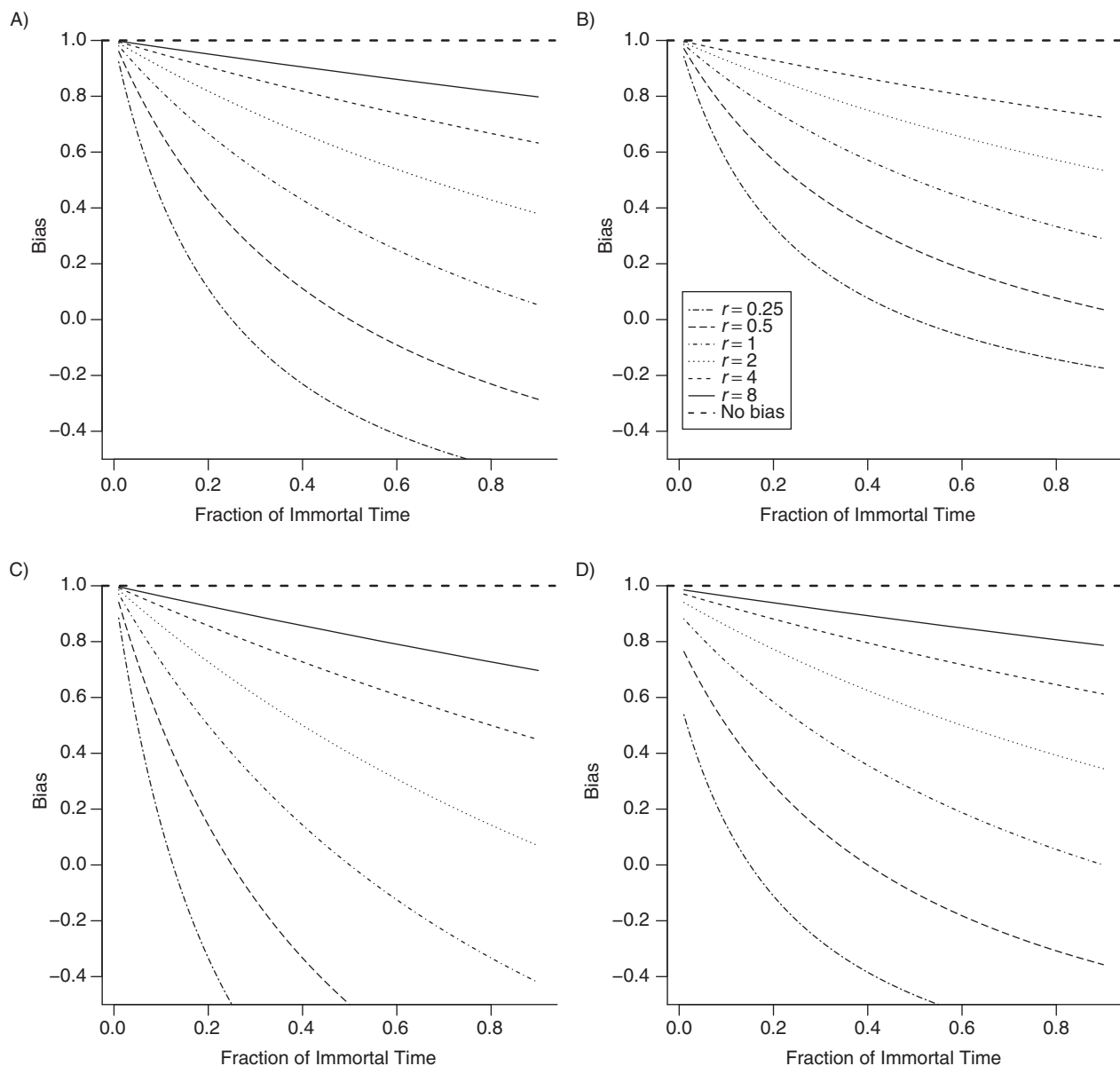
Figure 1.   Rate ratios for the prescription time-distribution matching method (RR″′) of dealing with immortal time bias in drug effectiveness studies as compared with rate ratios for a time-dependent analysis (RR) as a function of the fraction of immortal time *f* and for various ratios of the observed person-times in never-treated and ever-treated subjects, *r*. Specific cases: A) $q = 1$ (where q is the ratio of assigned and observed immortal times) and $x = 0$ (i.e., number of failures (reaching outcome) observed during the aggregated matched-immortal time, $N'_{IT} = 0$); B) $q = 0.5$ and $x = 0$; C) $q = 2$ and $x = 0$; D) $q = 1$ (assuming that N'_{IT} is very small compared with N_0) and $x = 0.5$. The "bias" is the deviation of RR″′/RR from the null value 1.

distribution where the probability of being male is 0.3 (as might occur in an MS population).

After generating values for the survival time T_i, the censoring time T_i^C, and the treatment and covariate matrix $X_{im} = (A_{im}, L_{i0})$ for each subject $i = 1, 2, \ldots, n$ for up to $m = 30$ time periods, the permutation algorithm (44) is used to generate survival data, where treatment A_{t_m} is time-dependent but the confounder L_0 is fixed at the baseline value. Arbitrarily, the parameters for the effects of treatment and sex on the survival outcome are set such that the treatment is associated

with the outcome in a harmful direction (a log HR of $\psi_1 = 0.5$) and males are at a lower risk than females (a log HR of $\psi_2 = -0.7$).

In order to consider more frequent events, we repeat the Monte Carlo study (similar to simulation 1) with $\lambda_0 = 0.10$ (on a monthly scale) (simulation 4). We further assume a gamma distribution (scale = 1/0.01, shape = 0.4) and a Weibull distribution (scale = 1/0.01, shape = 2) for generating failure times T (simulations 5 and 6, respectively; all other settings are similar to simulation 1).
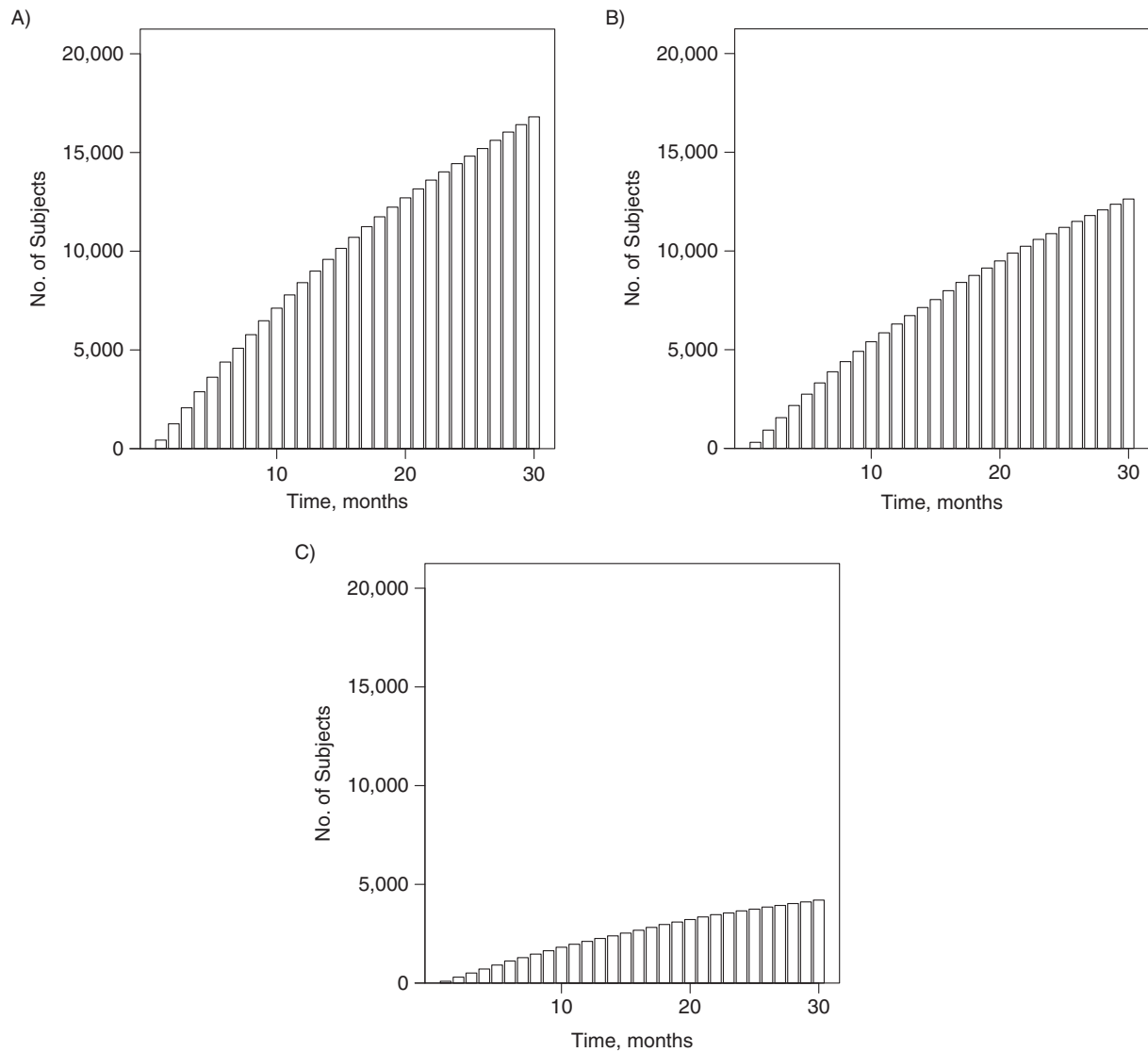
Figure 2. Cumulative percentages of a simulated cohort of 25,000 multiple sclerosis patients having initiated treatment over 30 follow-up periods. A) Simulation 1, with percentages of "failure," "never-treated," and "ever-treated" patients being 0.96%, 32.78%, and 67.22%, respectively; B) simulation 2, with the above percentages being 0.99%, 49.48%, and 50.52%, respectively; C) simulation 3, with the above percentages being 0.99%, 83.19%, and 16.81%, respectively. These percentages refer to percentages of subjects who fail (reach the outcome), remain unexposed to treatment, and initiate treatment, respectively, over the 30-month period.

Performance metrics. We assess the performance of the various approaches by means of the following measures:

- Bias $(\widehat{\psi_1}) = \sum_{i=1}^{N} (\widehat{\psi_{1i}} - \psi_1)/N$: the average difference between the true parameters and $N = 1,000$ estimated parameters (log HR).
- Standard deviation $(\widehat{\psi_1}) = \sqrt{\sum_{i=1}^{N} (\widehat{\psi_{1i}} - \overline{\psi_1})^2/(N-1)}$, where $\overline{\psi_1} = \sum_{i=1}^{N} \frac{\hat{\psi}_{1i}}{N}$.
- Model-based (average) standard error: The average of $N = 1,000$ estimated standard errors of the estimated causal association parameter.
- Coverage probabilities of model-based nominal 95% confidence intervals: proportion of $N = 1,000$ data sets in which

the true parameter is contained in the nominal 95% confidence interval.
- Mean squared error $(\widehat{\psi_1}) = \sqrt{\sum_{i=1}^{N} (\widehat{\psi_{1i}} - \psi_1)^2/N}$.

SIMULATION RESULTS

Description of the simulated data

To describe the data sets generated from each of the simulations, we generated 1 data set for each simulation with a larger number of subjects ($n = 25,000$). Figure 2 shows the cumulative percentages of the cohort having initiated treatment over the 30 follow-up periods.

Rare event condition

The results from the simulations of the rare event condition ($\lambda_0 = 0.01$ on a monthly time scale), simulations 1–3, are given in Tables 1–3, respectively.

When the immortal time is misclassified as exposed time, we see a substantial downward bias. The bias improves slightly when excluded immortal time and PTDM approaches are used, but these estimate are still off the target. As the ratio for the subjects in the never-treated group versus the ever-treated group increases, the bias monotonically goes down (see Tables 2 and 3). The variability of the estimator is larger for PTDM than for the time-dependent Cox results. Interestingly, the standard errors of the estimates from both the misclassified and excluded immortal-time approaches are smaller than those for the time-dependent Cox results, but the mean squared errors are generally higher. When the sequential Cox approach is applied, in all 3 of these simulations, the biases are negligible and the average coverage probabilities of the model-based nominal 95% confidence intervals are close to 0.95; both are comparable to the time-dependent Cox results. This estimator is generally more variable than that from the time-dependent Cox model fit (from simulations 1 and 2). However, from simulation 3 (Table 3), we see that exceptions may be possible.

When more events are available

Results from the more frequent event condition ($\lambda_0 = 0.10$ on a monthly time scale) are presented in Table 4. The bias for misclassified, excluded immortal-time, and PTDM approaches exhibiting large bias in Table 1 is considerably

Table 1. Results Derived From Analytical Approaches to Adjusting for Immortal Time Bias in a Simulation of 1,000 Data Sets, Each Containing 2,000 Multiple Sclerosis Patients Followed for Up to 30 Time Intervals (Rare Event Case $\lambda_0 = 0.01$) (Simulation 1)

Approach[a]	Bias ($\widehat{\psi_1}$)	SD ($\widehat{\psi_1}$)	SE ($\widehat{\psi_1}$)	CP	MSE ($\widehat{\psi_1}$)
TD Cox[b]	−0.004	0.122	0.121	0.955	0.015
Included IT	−2.307	0.109	0.108	0.000	5.332
Excluded IT	−1.360	0.102	0.102	0.000	1.859
PTDM	−1.411	0.144	0.143	0.000	2.011
Sequential Cox[c,d,e]	−0.002	0.157	0.150	0.941	0.025

Abbreviations: CP, coverage probability; IT, immortal time; MSE, mean squared error; PTDM, prescription time-distribution matching; SD, standard deviation; SE, standard error; TD, time-dependent.

[a] When estimating the treatment effect, the baseline covariate L_0 is included in all of the models under consideration. The results from the TD Cox model are considered the standard for comparison purposes. Results are provided on the log hazard ratio scale.

[b] Cox proportional hazards model with time-dependent exposure.

[c] In the sequential Cox approach, the corresponding inverse-probability-of-censoring weights model is fitted using Aalen's additive regression model (58, 59), adjusting for A_{t_m} and L_0 to predict future censoring status.

[d] A robust (sandwich) estimator was used to obtain the SE.

[e] Event times were generated from an exponential distribution with rate 0.01.

Table 2. Results Derived From Analytical Approaches to Adjusting for Immortal Time Bias in a Simulation of 1,000 Data Sets, Each Containing 2,000 Multiple Sclerosis Patients Followed for Up to 30 Time Intervals (Rare Event Case $\lambda_0 = 0.01$) (Simulation 2)[a]

Approach[b]	Bias ($\widehat{\psi_1}$)	SD ($\widehat{\psi_1}$)	SE ($\widehat{\psi_1}$)	CP	MSE ($\widehat{\psi_1}$)
TD Cox[c]	−0.000	0.117	0.116	0.949	0.014
Included IT	−1.394	0.099	0.100	0.000	1.953
Excluded IT	−0.698	0.104	0.104	0.000	0.499
PTDM	−0.558	0.134	0.131	0.012	0.329
Sequential Cox[d,e,f]	−0.001	0.127	0.126	0.950	0.016

Abbreviations: CP, coverage probability; IT, immortal time; MSE, mean squared error; PTDM, prescription time-distribution matching; SD, standard deviation; SE, standard error; TD, time-dependent.

[a] At least 25% of subjects unexposed at baseline remained untreated during the entire follow-up period.

[b] When estimating the treatment effect, the baseline covariate L_0 is included in all of the models under consideration. The results from the TD Cox model are considered the standard for comparison purposes. Results are provided on the log hazard ratio scale.

[c] Cox proportional hazards model with time-dependent exposure.

[d] In the sequential Cox approach, the corresponding inverse-probability-of-censoring weights model is fitted using Aalen's additive regression model (58, 59), adjusting for A_{t_m} and L_0 to predict future censoring status.

[e] A robust (sandwich) estimator was used to obtain the SE.

[f] Event times were generated from an exponential distribution with rate 0.01.

Table 3. Results Derived From Analytical Approaches to Adjusting for Immortal Time Bias in a Simulation of 1,000 Data Sets, Each Containing 2,000 Multiple Sclerosis Patients Followed for Up to 30 Time Intervals (Rare Event Case $\lambda_0 = 0.01$) (Simulation 3)[a]

Approach[b]	Bias ($\widehat{\psi_1}$)	SD ($\widehat{\psi_1}$)	SE ($\widehat{\psi_1}$)	CP	MSE ($\widehat{\psi_1}$)
TD Cox[c]	−0.001	0.146	0.145	0.956	0.021
Included IT	−0.832	0.138	0.135	0.000	0.711
Excluded IT	−0.136	0.146	0.142	0.850	0.040
PTDM	−0.101	0.152	0.151	0.912	0.033
Sequential Cox[d,e,f]	−0.008	0.146[g]	0.145	0.954	0.021

Abbreviations: CP, coverage probability; IT, immortal time; MSE, mean squared error; PTDM, prescription time-distribution matching; SD, standard deviation; SE, standard error; TD, time-dependent.

[a] At least 75% subjects unexposed at baseline remained untreated during the entire follow-up period.

[b] When estimating the treatment effect, the baseline covariate L_0 is included in all of the models under consideration. The results from the TD Cox model are considered the standard for comparison purposes. Results are provided on the log hazard ratio scale.

[c] Cox proportional hazards model with time-dependent exposure.

[d] In the sequential Cox approach, the corresponding inverse-probability-of-censoring weights model is fitted using Aalen's additive regression model (58, 59), adjusting for A_{t_m} and L_0 to predict future censoring status.

[e] A robust (sandwich) estimator was used to obtain the SE.

[f] Event times were generated from an exponential distribution with rate 0.01.

[g] The SD from the sequential Cox model (0.1456) was slightly smaller than the SD from the TD Cox model fit (0.1460).

Table 4. Results Derived From Analytical Approaches to Adjusting for Immortal Time Bias in a Simulation of 1,000 Data Sets, Each Containing 2,000 Multiple Sclerosis Patients Followed for Up to 30 Time Intervals (More Frequent Event $\lambda_0 = 0.10$) (Simulation 4)[a,b]

Approach[c]	Bias ($\widehat{\psi_1}$)	SD ($\widehat{\psi_1}$)	SE ($\widehat{\psi_1}$)	CP	MSE ($\widehat{\psi_1}$)
TD Cox[d]	−0.004	0.061	0.060	0.946	0.004
Included IT	−1.475	0.062	0.057	0.000	2.179
Excluded IT	−0.412	0.066	0.057	0.000	0.174
PTDM	−0.522	0.071	0.065	0.000	0.278
Sequential Cox[e,f]	−0.037	0.077	0.072	0.905	0.007

Abbreviations: CP, coverage probability; IT, immortal time; MSE, mean squared error; PTDM, prescription time-distribution matching; SD, standard deviation; SE, standard error; TD, time-dependent.

[a] $\lambda_0 = 0.10$ on a monthly time scale implies that almost all subjects should reach the outcome within the first 10 months.

[b] Event times were generated from an exponential distribution with rate 0.1.

[c] When estimating the treatment effect, the baseline covariate L_0 is included in all of the models under consideration. The results from the TD Cox model are considered the standard for comparison purposes. Results are provided on the log hazard ratio scale.

[d] Cox proportional hazards model with time-dependent exposure.

[e] In the sequential Cox approach, the corresponding inverse-probability-of-censoring weights model is fitted using Aalen's additive regression model ([58], [59]), adjusting for A_{t_m} and L_0 to predict future censoring status.

[f] A robust (sandwich) estimator was used to obtain the SE.

reduced but is still large, whereas the time-dependent Cox and sequential Cox approaches again exhibit negligible bias. For the PTDM approach, less bias is expected when the event rate is higher, as is discussed via an example of the analytical results in Web Appendix 3. In addition, the standard errors are much larger when failure rates are low.

When event times are generated from gamma and Weibull distributions

When event times are generated from a gamma distribution and a Weibull distribution, the results (Web Tables 1 and 2) closely resemble those obtained when event times are generated from an exponential distribution (Table 1).

Application to MS

We applied the methods described in this paper to data on the British Columbia MS cohort used earlier in the BeAMS Study (1995–2008) ([38], [47–51]). Web Appendix 6 shows the baseline characteristics of this MS cohort (see Web Table 3) and gives the study eligibility and exclusion criteria. In this retrospective observational study, the outcome measure was time from β-interferon eligibility to a "confirmed and sustained" Expanded Disability Status Scale score of 6, considered irreversible disability when all subsequent such scores were at least 6, with 1 or more measurements being taken at least 150 days later ([52], [53]). The goal was to assess the association of β-interferon with irreversible disease progression in MS patients.

Table 5. Estimated Parameters (Using the Hazard Scale) for Multiple Sclerosis Patients From British Columbia, Canada, 1995–2008

Approach[a]	$\widehat{HR}$	SE ($\widehat{HR}$)	95% CI
TD Cox[b]	1.29	0.23	0.91, 1.82
Included IT	0.98	0.19	0.67, 1.42
Excluded IT	1.35	0.26	0.93, 1.96
PTDM	1.12	0.21	0.77, 1.63
PTDM (average)[c]	1.26	0.25	0.86, 1.85
Sequential Cox[d]	1.22	0.33	0.72, 2.08

Abbreviations: CI, confidence interval; HR, hazard ratio; IPCW, inverse-probability-of-censoring weights; IT, immortal time; PTDM, prescription time-distribution matching; SD, standard deviation; SE, standard error; TD, time-dependent.

[a] To demonstrate the impact of the various immortal or immune time (pertains to nonfatal outcomes) adjustment methods, we assume here that once the subjects initiate β-interferon use, they continue taking the drug without any discontinuation until they develop the outcome (time to irreversible progression of disability) or are censored. All of the models under consideration adjust for the baseline confounders L_0 (age, sex, disease duration, and Expanded Disability Status Scale score) when estimating the treatment effect. The Cox model is used in all of these approaches.

[b] Cox proportional hazards model with time-dependent exposure.

[c] The HR was estimated from British Columbia multiple sclerosis data ([38], [47]) 1,000 times and averaged. The mean HR was 1.26 (ranging from 1.07 to 1.52; SD, 0.08), the average SE was 0.25 (SD, 0.02), and the average 95% CI was 0.86, 1.85.

[d] To create the IPCW for the sequential Cox approach, Aalen's additive regression model ([58], [59]) is used. The IPCW denominator model adjusts for A_{t_m} and L_0. The average weight is 1.00 (SD, 0.13), ranging from 0.66 to 3.45. Estimated IPCW distributions are unimodal and symmetrical.

Results are reported in Table 5. None of the approaches showed a significant hazard ratio. The hazard ratio estimated from the included immortal time approach was the lowest, and this was expected from the simulation. The hazard ratio estimated from the sequential Cox approach was slightly lower than that estimated from the time-dependent Cox analysis, while that estimated from the PTDM was even lower. The distributions of excluded times from both groups when using the PTDM method looked reasonably similar (Figure 3). Because the PTDM approach produces different estimates from the same data due to the random sampling of the control subjects, we estimated the hazard ratio from the British Columbia MS data 1,000 times (average hazard ratio = 1.26; listed in Table 5). The distribution of the estimated hazard ratios was moderately symmetric (see Web Figure 5). Density plots of the estimated inverse-probability-of-censoring weights are shown in Web Figure 6.

To better understand the possible level of bias from each approach in this context, we conducted another Monte Carlo study, choosing data simulation settings similar to those of the MS data set. Simulation specifications are shown in Web Appendix 7 (simulated cohort characteristics are shown in Web Table 4 and Web Figure 7). The results are reported in Table 6. Here, the level of bias from each approach is much
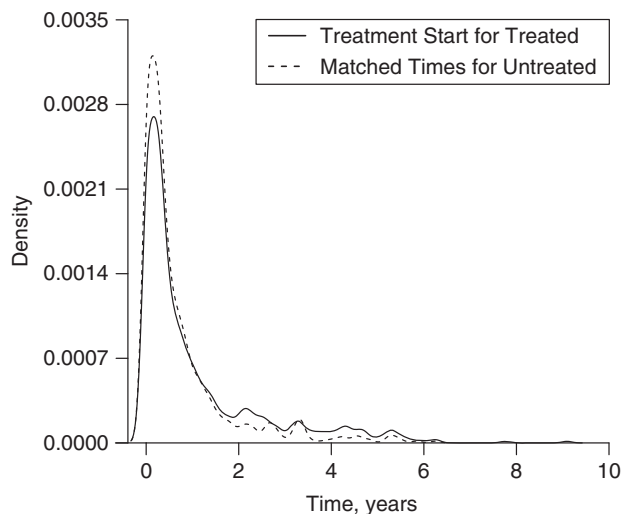
Figure 3. Matched wait periods (or matched immortal times), in years, from the prescription time-distribution matching approach in a simulated cohort of 25,000 multiple sclerosis patients initiating β-interferon treatment, British Columbia, Canada, 1995–2008.

less than that reported for the previous simulations (Table 6 vs. Tables 1 and 2) but similar to that in simulation 3 (Table 3).

DISCUSSION

In observational studies of drug effectiveness, subjects may not be exposed to the treatment at the beginning of follow-up. Statistical procedures, such as the time-dependent Cox model, are known to deal with time-dependent treatment. However, in the medical literature, it is common to see the Cox model applied on the basis of baseline treatment information, likely for the convenience of model-fitting and group-based interpretation (54). Therefore, there is a need for methods that are capable of handling the time-dependent nature of longitudinal data and will also help us to better understand the treatment-outcome mechanism. To this end, the

PTDM approach has been proposed for the situation where treatment initiation occurs later than cohort entry (13). The sequential Cox approach has been proposed for more complex longitudinal settings (35), but we considered a simplified version of this approach to deal with the immortal time bias.

In the PTDM approach, treatment exposure is converted into a time-independent variable so that a simple Cox model for treatment-group comparison can be applied. New time zeros are defined after excluding the observed and assigned immortal times in the ever- and never-treated groups. The excluded times (wait periods) for the ever-treated and never-treated groups follow the same distribution. However, the baseline is not the same for all subjects. The question of whether this approach (assigning the baseline borrowed from the ever-treated cohort) adequately addresses the immortal time bias has not previously been investigated in a systematic fashion, to our knowledge. From our bias-quantifying equations (Web Appendices 1 and 3) and the results of our simulations, we can see that the bias is still substantial in the PTDM analysis and the direction of the bias is negative, highlighting the value of setting a well-defined time zero or baseline. The results of simulations 1–3 (also in Web Appendix 1) identify an increase in the ratio of subjects in the never-treated group to subjects in the ever-treated group as an important design parameter for reducing bias associated with the PTDM method (biases were −1.411, −0.558, and −0.101 for the ratios 0.49, 0.98, and 4.95, respectively). From simulation 4, we learned that less bias is expected for the PTDM approach when the event rate is higher. We also considered 2 other popular event-time-generating distributions, namely the gamma and Weibull distributions (simulations 5 and 6, respectively), and the results were consistent with those from simulation 1.

In contrast, the sequential Cox approach allows a consistent definition of baseline and utilizes all subjects. The baseline is set on the basis of treatment initiation, and the subjects are stratified accordingly. This approach recreates the covariate process at each treatment start using the minitrial approach (36) and works very well in our simulation studies in terms of bias. In practice, as the time-dependent Cox model approach

Table 6. Comparison of Analytical Approaches to Adjusting for Immortal Time Bias in a Simulation Inspired by Multiple Sclerosis Data From 1,000 Data Sets, Each Containing 1,700 Subjects Followed for Up to 150 Months[a]

Approach[b]	Bias ($\widehat{\psi_1}$)	SD ($\widehat{\psi_1}$)	SE ($\widehat{\psi_1}$)	CP	MSE ($\widehat{\psi_1}$)	95% CI for Mean Bias[c]
TD Cox[d]	0.000	0.136	0.135	0.949	0.018	−0.008, 0.008
Included IT	−0.833	0.122	0.122	0.000	0.709	−0.841, −0.825[e]
Excluded IT	−0.027	0.134	0.129	0.942	0.019	−0.035, −0.019[e]
PTDM	−0.105	0.139	0.139	0.890	0.030	−0.114, −0.096[e]
Sequential Cox	−0.001	0.231	0.207	0.928	0.053	−0.015, 0.013

Abbreviations: CI, confidence interval; CP, coverage probability; IT, immortal time; MSE, mean squared error; PTDM, prescription time-distribution matching; SD, standard deviation; SE, standard error; TD, time-dependent.

[a] The simulation included 4 baseline covariates and a time-dependent treatment exposure.
[b] Results are provided on the log hazard ratio scale.
[c] The 95% CI for the mean bias was calculated as bias $\pm 1.96 \times SD/\sqrt{(N)}$ across the 1,000 simulations.
[d] Cox proportional hazards model with time-dependent exposure.
[e] Indicative of a finite-sample bias across 1,000 simulations.

can easily deal with the immortal time bias and is simpler to implement, one may not be inclined to use the sequential Cox approach, because additional steps are involved in the data-restructuring for the latter approach (i.e., to create minitrials and then pseudodata). The data associated with a given minitrial in a sequential Cox approach, however, can be extracted and separated quite easily from the pseudodata, and it is straightforward to compare the treatment associations (of early versus late treatment initiation) with the outcome. It is also possible to estimate the treatment association for patients with a specific level of a covariate at treatment initiation (35). Such detailed scrutiny using this minitrial approach could yield insights about the data which may be hard to gain using the time-dependent Cox approach.

We applied the methods under consideration to estimate the association of β-interferon with disease progression in the British Columbia MS cohort. In comparison with results from the time-dependent Cox model (hazard ratio (HR) = 1.29, 95% confidence interval (CI): 0.91, 1.81), the sequential Cox approach provided a slightly lower association estimate with a wider confidence interval (HR = 1.22, 95% CI: 0.72, 2.08). Repeating the PTDM analysis 1,000 times, the average results (HR = 1.26, average 95% CI: 0.86, 1.85) were quite similar to those from the time-dependent Cox approach, although the average confidence interval for the estimated treatment association was slightly wider. We designed another Monte Carlo study using simulation settings reflecting our MS data set to explore further the level of bias from each approach. The smaller amount of bias in this simulation for each approach may have been a consequence of the fact that the ever-treated group was much smaller than the never-treated group in this simulation (19% vs. 81%; ratio: 4.26). This phenomenon also explains why the PTDM approach provided estimates close to those from the time-dependent Cox model fit (29% user vs. 71% nonuser) (13). Even in this simulation setting, the time-dependent Cox model and sequential Cox approach were still associated with less bias than the PTDM approach.

Similar to other simulation studies, we investigated only a few scenarios. However, the assumptions underlying the data simulation were consistent with patterns typical of epidemiologic observational survival studies, where treatment initiation may happen later for some subjects than others (7). Furthermore, our assumption of no discontinuations or interruptions in the treatment was restrictive and may not be suitable for more complex disease scenarios in which an investigator might wish to assess different treatment strategies (i.e., switching between therapies) over the course of time. Substantial immortal time bias was introduced by group-based Cox model analyses in the scenarios investigated (analytical expressions for PTDM approach in Web Appendix 3 also support this finding). However, even in these scenarios, the sequential Cox method appeared to estimate the target parameter adequately. Variance estimation is still a challenge with the sequential Cox model and similar methods (55, 56), and the resampling methods suggested for estimating the variance are computationally demanding (35, 57). Future research could focus on assessing the ability of the sequential Cox approach to adjust for time-dependent confounders (37).

ACKNOWLEDGMENTS

make presentations at conferences (2011, 2012) and has received financial support from the Pacific Institute for the Mathematical Sciences to attend a workshop (2013). P.G. has received consulting fees from Biogen to serve on an advisory board. Over the past 3 years, J.P. has received consulting fees and/or fees for service on Data Safety Monitoring Boards from EMD Serono, the Myelin Research Foundation, and Novartis. H.T. has received speaker honoraria and/or money for travel expenses to attend conferences of the Consortium of MS Centres (2013), the MS Society of Canada (2013), the National Multiple Sclerosis Society (2012, 2014), Bayer Pharmaceutical (speaker, 2010; honoraria declined), Teva Pharmaceutical (speaker, 2011), the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS) (2011, 2012, 2013, 2014), the Multiple Sclerosis Trust (2011), the Chesapeake Health Education Program, the US Department of Veterans Affairs (2012, honorarium declined), Novartis Canada (2012), Biogen Idec (2014; honorarium declined), and the American Academy of Neurology (2013, 2014, 2015). Unless otherwise stated, H.T. donates all speaker honoraria to either an MS charity or an unrestricted grant fund for use by her research group. A.S. has received travel grants to attend and make presentations at conferences of the ERTN (2010, 2011), ECTRIMS (2010, 2011), and the Consortium of MS Centres (2012). E.K. has had travel and accommodation costs covered to attend and make presentations at conferences of the ERTN (2008, 2011), the International Society for Pharmacoepidemiology (2010, 2013), and Bayer Schering Pharma (2010). C.E. has received travel grants to attend and make presentations at conferences of the ERTN (2011) and ECTRIMS (2011). Over the past 5 years, J.O. has received speaker honoraria, consulting fees, travel grants, research grants, and/or educational grants from Aventis, Bayer, Biogen Idec, BioMS, Corixa, Genentech, GlaxoSmithKlein, Merck-Serono, Novartis, Serono, Shering, Talecris, and Teva Neuroscience. J.O. has received fees for services from Bayer, Novartis, and Biogen Idec to serve on advisory committees.

REFERENCES

1. Trojano M, Pellegrini F, Fuiani A, et al. New natural history of interferon-beta-treated relapsing multiple sclerosis. *Ann Neurol.* 2007;61(4):300–306.
2. Lévesque LE, Hanley JA, Kezouh A, et al. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ.* 2010;340:b5087.
3. Renoux C, Suissa S. Immortal time bias in the study of effectiveness of interferon-beta in multiple sclerosis [letter]. *Ann Neurol.* 2008;64(1):109–110.
4. Wolkewitz M, Allignol A, Harbarth S, et al. Time-dependent study entries and exposures in cohort studies can easily be sources of different and avoidable types of bias. *J Clin Epidemiol.* 2012;65(11):1171–1180.
5. Gail MH. Does cardiac transplantation prolong life? A reassessment. *Ann Intern Med.* 1972;76(5):815–817.
6. van Walraven C, Davis D, Forster AJ, et al. Time-dependent bias was common in survival analyses published in leading clinical journals. *J Clin Epidemiol.* 2004;57(7):672–682.
7. Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf.* 2007;16(3):241–249.
8. Kiri VA, Pride NB, Soriano JB, et al. Inhaled corticosteroids in chronic obstructive pulmonary disease: results from two observational designs free of immortal time bias. *Am J Respir Crit Care Med.* 2005;172(4):460–464.
9. Austin PC, Platt RW. Survivor treatment bias, treatment selection bias, and propensity scores in observational research. *J Clin Epidemiol.* 2010;63(2):136–138.
10. Ravi B, Croxford R, Austin PC, et al. The relation between total joint arthroplasty and risk for serious cardiovascular events in patients with moderate-severe osteoarthritis: propensity score matched landmark analysis. *BMJ.* 2013;347:f6187.
11. Karim ME. Can joint replacement reduce cardiovascular risk? [editorial]. *BMJ.* 2013;347:f6651.
12. 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(1):49–53.
13. Zhou Z, Rahme E, Abrahamowicz M, et al. Survival bias associated with time-to-treatment initiation in drug effectiveness evaluation: a comparison of methods. *Am J Epidemiol.* 2005;162(10):1016–1023.
14. Austin PC, Mamdani MM, Van Walraven C, et al. Quantifying the impact of survivor treatment bias in observational studies. *J Eval Clin Pract.* 2006;12(6):601–612.
15. Suissa S. Immortal time bias in pharmacoepidemiology. *Am J Epidemiol.* 2008;167(4):492–499.
16. Wolkewitz M, Allignol A, Schumacher M, et al. Two pitfalls in survival analyses of time-dependent exposure: a case study in a cohort of Oscar nominees. *Am Stat.* 2010;64(3):205–211.
17. Sylvestre MP, Huszti E, Hanley JA. Do OSCAR winners live longer than less successful peers? A reanalysis of the evidence. *Ann Intern Med.* 2006;145(5):361–363.
18. Ho PM, Fihn SD, Wang L, et al. Clopidogrel and long-term outcomes after stent implantation for acute coronary syndrome. *Am Heart J.* 2007;154(5):846–851.
19. Karp I, Behlouli H, Lelorier J, et al. Statins and cancer risk. *Am J Med.* 2008;121(4):302–309.
20. Ho PM, Maddox TM, Wang L, et al. Risk of adverse outcomes associated with concomitant use of clopidogrel and proton pump inhibitors following acute coronary syndrome. *JAMA.* 2009;301(9):937–944.
21. Snyder CW, Weinberg JA, McGwin G Jr, et al. The relationship of blood product ratio to mortality: survival benefit or survival bias? *J Trauma.* 2009;66(2):358–362.
22. Keyhan G, Chen SF, Pilote L. The effectiveness of beta-blockers in women with congestive heart failure. *J Gen Intern Med.* 2007;22(7):955–961.
23. Keyhan G, Chen SF, Pilote L. Angiotensin-converting enzyme inhibitors and survival in women and men with heart failure. *Eur J Heart Fail.* 2007;9(6-7):594–601.
24. Morin S, Rahme E, Behlouli H, et al. Effectiveness of antiresorptive agents in the prevention of recurrent hip fractures. *Osteoporos Int.* 2007;18(12):1625–1632.
25. Walkey AJ, Wiener RS. Utilization patterns and patient outcomes associated with use of rescue therapies in acute lung injury. *Crit Care Med.* 2011;39(6):1322–1328.
26. Manari A, Varani E, Guastaroba P, et al. Long-term outcome in patients with ST segment elevation myocardial infarction and multivessel disease treated with culprit-only, immediate, or staged multivessel percutaneous revascularization strategies: insights from the REAL registry. *Catheter Cardiovasc Interv.* 2014;84(6):912–922.
27. Prieto-Alhambra D, Javaid MK, Judge A, et al. Hormone replacement therapy and mid-term implant survival following knee or hip arthroplasty for osteoarthritis: a population-based cohort study. *Ann Rheum Dis.* 2015;74(3):557–563.

28. Li G, Holbrook A, Delate T, et al. Prediction of individual combined benefit and harm for patients with atrial fibrillation considering warfarin therapy: a study protocol. *BMJ Open*. 2015;5(11):e009518.

29. Shintani AK, Girard TD, Eden SK, et al. Immortal time bias in critical care research: application of time-varying cox regression for observational cohort studies. *Crit Care Med*. 2009;37(11):2939–2945.

30. Liu J, Weinhandl ED, Gilbertson DT, et al. Issues regarding 'immortal time' in the analysis of the treatment effects in observational studies. *Kidney Int*. 2012;81(4):341–350.

31. Ho AM, Dion PW, Yeung JH, et al. Simulation of survivorship bias in observational studies on plasma to red blood cell ratios in massive transfusion for trauma. *Br J Surg*. 2012;99(suppl 1): 132–139.

32. Buyse M, Piedbois P. On the relationship between response to treatment and survival time. *Stat Med*. 1996;15(24): 2797–2812.

33. Beyersmann J, Gastmeier P, Wolkewitz M, et al. An easy mathematical proof showed that time-dependent bias inevitably leads to biased effect estimation. *J Clin Epidemiol*. 2008; 61(12):1216–1221.

34. Beyersmann J, Wolkewitz M, Schumacher M. The impact of time-dependent bias in proportional hazards modelling. *Stat Med*. 2008;27(30):6439–6454.

35. 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;29(26):2757–2768.

36. Aalen O. Armitage Lecture 2010. Understanding treatment effects: the value of integrating longitudinal data and survival analysis. *Stat Med*. 2012;31(18):1903–1917.

37. Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology*. 2000;11(5):561–570.

38. Shirani A, Zhao Y, Karim ME, et al. Association between use of interferon beta and progression of disability in patients with relapsing-remitting multiple sclerosis. *JAMA*. 2012;308(3): 247–256.

39. Hernán MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology*. 2008;19(6):766–779.

40. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11(5):550–560.

41. Therneau T. *Modeling Survival Data: Extending the Cox Model*. New York, NY: Springer-Verlag New York; 2000.

42. Xiao Y, Abrahamowicz M, Moodie EE. Accuracy of conventional and marginal structural Cox model estimators: a simulation study. *Int J Biostat*. 2010;6(2):Article 13.

43. Taylor JM, Shen J, Kennedy EH, et al. Comparison of methods for estimating the effect of salvage therapy in prostate cancer when treatment is given by indication. *Stat Med*. 2014;33(2): 257–274.

44. Abrahamowicz M, Mackenzie T, Esdaile J. Time-dependent hazard ratio: modeling and hypothesis testing with application in lupus nephritis. *J Am Stat Assoc*. 1996;91(436): 1432–1439.

45. Mackenzie T, Abrahamowicz M. Marginal and hazard ratio specific random data generation: applications to semi-parametric bootstrapping. *Stat Comput*. 2002;12(3):245–252.

46. Sylvestre MP, Abrahamowicz M. Comparison of algorithms to generate event times conditional on time-dependent covariates. *Stat Med*. 2008;27(14):2618–2634.

47. Karim ME, Gustafson P, Petkau J, et al. Marginal structural Cox models for estimating the association between β-interferon exposure and disease progression in a multiple sclerosis cohort. *Am J Epidemiol*. 2014;180(2):160–171.

48. Karim M. *Causal Inference Approaches for Dealing With Time-Dependent Confounding in Longitudinal Studies, With Applications to Multiple Sclerosis Research* [doctoral thesis]. Vancouver, BC, Canada: University of British Columbia; 2015.

49. Shirani A, Zhao Y, Karim ME, et al. Investigation of heterogeneity in the association between interferon beta and disability progression in multiple sclerosis: an observational study. *Eur J Neurol*. 2014;21(6):835–844.

50. Zhang T, Shirani A, Zhao Y, et al. Beta-interferon exposure and onset of secondary progressive multiple sclerosis. *Eur J Neurol*. 2015;22(6):990–1000.

51. Shirani A, Zhao Y, Petkau J, et al. Multiple sclerosis in older adults: the clinical profile and impact of interferon beta treatment. *Biomed Res Int*. 2015;2015:451912.

52. Tremlett H, Paty D, Devonshire V. Disability progression in multiple sclerosis is slower than previously reported. *Neurology*. 2006;66(2):172–177.

53. Tremlett H, Yousefi M, Devonshire V, et al. Impact of multiple sclerosis relapses on progression diminishes with time. *Neurology*. 2009;73(20):1616–1623.

54. Leffondré K, Abrahamowicz M, Siemiatycki J. Evaluation of Cox's model and logistic regression for matched case-control data with time-dependent covariates: a simulation study. *Stat Med*. 2003;22(24):3781–3794.

55. 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;62(3):910–917.

56. 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; 104(485):49–59.

57. Røysland K, Gran JM, Ledergerber B, et al. Analyzing direct and indirect effects of treatment using dynamic path analysis applied to data from the Swiss HIV Cohort Study. *Stat Med*. 2011;30(24):2947–2958.

58. Aalen OO. A linear regression model for the analysis of life times. *Stat Med*. 1989;8(8):907–925.

59. Aalen OO. Further results on the non-parametric linear model in survival analysis. *Stat Med*. 1993;12(17):1569–1588.