Immortal time bias in critical care research: application of time-varying Cox regression for observational cohort studies

Ayumi K. Shintani, Ph.D., M.P.H., **Timothy D. Girard, M.D., M.S.C.I.**, **Patrick G. Arbogast, Ph.D.**, **Karel G.M. Moons, M.Sc., Ph.D.**, and **E. Wesley Ely, M.D., M.P.H.**
Department of Biostatistics (A.K.S. and P.G.A.) and Department of Medicine, Division of Allergy, Pulmonary, and Critical Care Medicine and Center for Health Services Research (T.D.G. and E.W.E.), Vanderbilt University School of Medicine, Nashville, TN; Julius Center for Health Sciences and Primary Care (K.G.M.M.), University Medical Center Utrecht, Utrecht, The Netherlands; and VA Tennessee Valley Geriatric Research, Education and Clinical Center (GRECC) (T.D.G. and E.W.E.), VA Service, Department of Veterans Affairs Medical Center, Nashville, TN

Abstract

Objectives—To determine the magnitude of bias when using time-<u>fixed</u> methodology to analyze the effect of a time-<u>varying</u> exposure incurred in the ICU on ICU length of stay.

Design—Prospective cohort study (Part 1) and Monte Carlo stimulation study (Part 2)

Setting—Medical and coronary intensive care units (ICUs) in a university hospital

Patients—224 mechanically ventilated patients

Measurements and Main Results—This was a two-part investigation, with Part I being a case-study analyzing the association between delirium in the ICU (exposure variable/risk factor) and ICU length of stay (outcome) in a prospective cohort study and Part II being a Monte Carlo simulation generating 6,000 simulated datasets with a known absence of association between delirium and ICU length of stay. In both parts, we assessed the associations between delirium in the ICU and ICU length of stay using time-fixed versus time-varying Cox regression methodology. Both Kaplan-Meier survival curves and Cox regression indicated a strong association between delirium and ICU stay when analyzed as a time-fixed variable. There was no association when appropriate time-varying analyses were used. Adjusted hazard ratios (HRs) for the effect of delirium on ICU stay were 1.9 (95% CI, 1.3–2.7, p<0.001) for time-fixed and 1.1 (0.7–1.6, p=0.70) for time-varying approach controlling for an a priori chosen set of clinically relevant covariates. Based on our Monte Carlo simulation study, we found the median HR was 12.7 (7.7–24.9) from typically adopted time-fixed Cox regression and 1.0 (0.6–1.6) from time-varying Cox regression model.

Conclusions—Large errors and degrees of bias can result from studies using a time-fixed analytic approach to understand relationships between risk factors and clinical outcomes when the variables overlap temporally in occurrence. Those conducting such studies, and clinicians reading them, should be sure that time-varying covariates be correctly handled to avoid mistaken conclusions.

Address correspondence to: Dr. Shintani at the Department of Biostatistics, Vanderbilt University Medical Center, S-2323 Medical Center North, Nashville, TN 37232-2158, ayumi.shintani@vanderbilt.edu.

Introduction

Though randomized clinical trials are the ideal study design used to assess the effect of a treatment on disease, researchers often need to study an exposure to which patients cannot be randomly assigned, e.g., potentially harmful exposures such as cigarette smoking or delirium in an intensive care unit (ICU) patient. Thus, observational cohort studies remain an important method for evaluating associations between exposure variables (i.e., risk factors) and clinical outcomes (1). Determining the truth about relationships between exposures on outcomes using observational studies, however, is challenging and precarious. Indeed, results are often biased by multiple factors, including confounding, selection bias, and other biases that arise during the design, conduct, and analysis of observational cohort studies.

One challenge frequently encountered in observational studies that is often improperly managed is the analysis of exposure variables whose status can change during the follow-up period (i.e., time-varying or time-dependent variables) (2). E.g., in a study of patients with chronic obstructive pulmonary disease (COPD), corticosteroid use may be a time-varying exposure; a patient may begin the study as a non-user but become a user during the period of follow-up. An analysis that simply categorizes the exposure variable according to its final status rather than considering the timing of the change in status can create bias. Patients who experience an early outcome (e.g., a COPD patient who dies early during follow-up) will more likely be categorized as unexposed (i.e., a non-user of steroids) because they had less opportunity for exposure. Such patients will contribute worse outcomes to the non-user group, thereby artificially inflating the benefit of the exposure. A survival analysis of the association between corticosteroid use in COPD that does not consider the exposure variable as a time-varying factor yields erroneous results due to such misclassification of the exposure variable and resultant "immortal time bias" (3). Immortal time is the period of time prior to initiation of treatment (i.e., change in exposure status to corticosteroid user). In time-fixed analyses, this period of time is inappropriately attributed to treatment effect even though the treatment had not yet started and could not have truly influenced this period of time. This period is called "immortal time" since the subject could not have died prior to receiving the treatment and thus was immortal prior to exposure. Bias resulting from inclusion of such time in survival analysis has also been termed "time-dependent bias" (2) or "survivor treatment selection bias" because longer survival provides a greater opportunity to be selected for treatment (4) (5). Proper analysis must consider the occurrence of time-varying exposure, by allowing patients to transfer from unexposed to exposed group at the onset of exposure during the follow-up period to avoid "immortal time" bias.

Though examples of time-varying exposures abound in the medical literature, appropriate analysis of such variables is much less common. A recent systematic review estimated that up to 1 in every 10 observational studies using survival analysis were susceptible to immortal time bias (2). Such bias is not limited to survival analyses, but may affect any analysis that applies a time-fixed definition to an exposure that actually varies over time. Importantly, conclusions based on biased analyses of time-varying exposures (e.g., regarding the efficacy of treatment with drugs or procedures as well as outcomes associated with acute or chronic illnesses) can misinform researchers and clinicians alike, potentially resulting in inappropriate allocation of resources, inflation of medical costs, and even direct harm to patients.

In this two-part investigation, we sought to describe and quantify the bias introduced by using time-fixed methods of analysis to analyze relationships from an observational study with a time-dependent exposure variable. We first conducted a case study, analyzing a previously reported prospective cohort study of mechanically ventilated ICU patients for an

outcome (ICU length of stay) that temporally overlapped with the exposure variable (ICU delirium) using both time-fixed and time-varying methods of multivariable survival analysis and compared the results. Next, we objectively quantified the bias introduced by time-fixed analysis (i.e., immortal-time bias) using a Monte Carlo simulation study.

CASE STUDY

Case Study Methods

To determine whether immortal-time bias can influence the results of a prospective observational study and to describe the bias if observed, we analyzed the association between delirium and ICU length of stay in a cohort of mechanically ventilated medical ICU patients using both time-fixed and time-varying methods. Delirium, the exposure variable of interest, can occur at any time during the ICU stay; it is a time-dependent exposure. Thus, we hypothesized that a time-fixed analysis would conclude that delirium is associated with prolonged stays in the ICU whereas time-varying analysis would show that no such association exists.

The cohort consisted of 275 adult, mechanically ventilated patients admitted to the medical or coronary ICUs at Vanderbilt University Medical Center between February 2000 and May 2001. As described previously (6), all patients were prospectively enrolled and evaluated daily for delirium using the Confusion Assessment Method for the ICU (CAM-ICU) (7, 8). Fifty-one patients who died were never evaluated for delirium due to persistent coma. The remaining 224 patients were analyzed as described herein. Length of stay in the ICU was prospectively determined for all patients.

In the time-fixed analyses, patients were classified according to presence or absence of the exposure variable during the study period, i.e., each patient was grouped into the "ever delirious" group or the "never delirious" group. To evaluate the effect of ever being delirious vs. never being delirious on ICU length of stay, we used time-to-event analyses. Patients were censored at time of death. Median and interquartile range (IQR) of time to ICU discharge were obtained via Kaplan-Meier analyses, and a log-rank test was used to assess the effect of delirium. Additionally, Cox proportional hazards regression was used to evaluate the association between delirium as a time-fixed exposure and ICU length of stay after adjustment for the following a priori-chosen covariates: age, severity of illness (measured by the Acute Physiology and Chronic Health Evaluation II), comorbid illnesses (measured by the Charlson Comorbidity Index), coma (as time-dependent covariate), and treatment with sedative and analgesic medications (lorazepam, propofol, morphine, and fentanyl).

In the time-varying analyses, all patients were considered unexposed (i.e., without delirium) from the time of study entry until the onset of delirium, defined as the first day delirium was diagnosed using the CAM-ICU. Patients who developed delirium were considered delirious from the time of delirium onset until death or discharge from the ICU, whichever occurred first. To graphically display this definition of exposure and compare it to that used in the time-fixed method, we created event history graphs (9) using the data from 10 randomly selected patients. Event history graphs can overlay the actual shift of time-varying exposure status within a subject in Kaplan-Meier curves. Because no formal statistical inferences can be made using event history graphs, however, we also graphed survival functions using the method of Simon and Makuch (10), who modified Kaplan-Meier survival curves to account for time-varying exposures. When using the standard Kaplan-Meier method to create survival curves, exposure status for each patient is fixed for the entire period of follow-up. The Simon-Makuch method, alternatively, can graphically represent survival curves for time-dependent exposure variables. We used the log-rank test to assess the effect of

delirium. Finally, Cox proportional hazards regression was used to evaluate the association between delirium as a time-varying exposure and ICU length of stay after adjustment for the previously listed covariates. The inverse of each hazard ratio (HR) is reported to facilitate comparisons and graphical display; i.e., a HR>1 indicates that delirium is associated with delayed ICU discharge. We used SAS version 9.0 (SAS Institutes, Cary, NC) in the case-study.

Case Study Results

As previously reported (6), 183 (82%) of 224 patients studied developed ICU delirium. The median time [interquartile range] to delirium onset was 2 days [1, 3] from ICU admission. Twenty-eight (13%) patients died in the ICU.

The event history graphs in Figure 1 indicate exposure status prior to ICU discharge or death, whichever came first, for 10 randomly selected patients according to a time-fixed (Figure 1A) or time-varying (Figure 1B) method of analysis. Figure 1A suggests that patients without delirium are likely to be discharged early, whereas patients with delirium appear to have longer ICU stays. Figure 1B, revealing that some patients who ultimately develop delirium do not do so until later in their ICU course, does not suggest an association between delirium and delayed discharge from the ICU. That is, normal neurologic status (as indicated by the green bars) is not exclusively experienced by patients who are discharged early, and delirium (as indicated by the red bars) is not exclusively experienced by patients who have delayed discharge.

The Kaplan-Meier curves in Figure 2A, created using a time-fixed definition of exposure, suggest that delirium is associated with delayed discharge from the ICU (p<0.0001). The curves in Figure 2B, alternatively, which are modified to reflect the time-varying nature of the exposure, show that delirium does not prolong the period of time spend in the ICU (p=0.97). Similarly, Cox proportional hazards regression found a significant association between delirium as a time-fixed exposure and time to ICU discharge; the adjusted HR was 1.9 (95% CI, 1.3–2.7; p<0.001) using the time-fixed method. No association was found, however, between delirium and time to ICU discharge (adjusted HR, 1.1; 95% CI, 0.7–1.6; p=0.70) using the time-varying method of analysis.

MONTE CARLO SIMULATION STUDY

Monte Carlo Study Methods

Though the analyses conducted in the case study can demonstrate that contradictory results can be obtained when different analytical methods are applied to an observational study with a time-dependent exposure, these analyses are limited for two reasons. First, the magnitude of the probable bias introduced by time-fixed analysis of such data cannot be estimated because the study analyzed is only one example (e.g., one sample of the true population of ICU patients at risk for delirium and the outcomes potentially associated with it). And second—perhaps more important—the true association between exposure and outcome (in this case, between ICU delirium and length of stay in the ICU) is not known apart from the inferences made by our statistical analyses. Thus, if the reader does not trust that the results of the time-varying methods of analysis are valid compared with those of the time-fixed analysis, the results of the case-study may be simply disregarded.

To address these limitations, we conducted a computer simulation study in which both time-fixed and time-varying methods were used to analyze numerous datasets wherein, by design, no association between ICU delirium and length of stay in the ICU existed. First, we created 1,000 datasets—each including 224 simulated patients—in which the time to delirium onset was randomly set along an exponential distribution with a median onset of 1 day, and the

time to ICU discharge was randomly set along an independent exponential distribution with a median time to discharge of 6 days. Because the two variables were set independently, the null hypothesis was known to be true— i.e., no association existed between ICU delirium and length of stay in the ICU in any simulated dataset except for that which occurred due to chance alone. A set of 1000 simulated datasets were generated for 6 varying frequency rate of delirium (0.01, 0.05, 0.1, 0.1, 0.33, 0.5, 1 event per person-day) the corresponding median time to delirium onset set are 1, 2, 3, 10, 20, and 100 days, respectively, so that a total of 6 sets of 1,000 simulated datasets were created. For each set of 1000 simulated datasets, we determined the unadjusted association between delirium and ICU length of stay using two Cox proportional hazards regression models—one in which delirium was a time-fixed variable and one in which it was a time-varying variable. The each set of 1,000 HRs generated from the time-fixed analyses were ranked, and the 50th, 2.5th, and 97.5th percentile HRs were identified as representing the point estimate and 95% confidence interval (CI) of the true HR. The same was done using the 6,000 HRs generated from the time-varying analyses. As in Part I, HR was computed by taking reciprocating each HR so that greater HR reflects the greater likelihood of delirium prolonging ICU stay.

R (version 2.4 Patched, www.r-project.org) for all statistical analyses conducted for the Monte Carlo Study (11).

Monte Carlo Study Results

Analyses of the simulated datasets—which were generated in such a manner that the null hypothesis was known a priori to be true (i.e., this simulated dataset was known to have no association between delirium and ICU length of stay)—revealed results similar to those described in the case study, supporting the validity of the time-varying of analysis and the pitfall of the time-fixed analysis.

The results of the Monte Carlo study are displayed in Figure 3A and Figure 3b. Both of these components of Figure 3 show similar findings, with the X-axes the same for both graphs indicating varying frequency of delirium, but the Y-axis representing different "views" of the statistical results using time-fixed vs. time-varying approaches to determining the relationship between delirium and the hazard of a delayed ICU discharge. In the time-fixed analyses, delirium was found to be associated with delayed ICU discharge in more than 95% of the 6,000 simulated datasets, and the association was strongest in datasets where delirium occurred more frequently. The median HRs were 2.1 (95% CI, 1.3–3.7; 74% of p values<0.05), 3.0 (95% CI, 2.3–4.2; all p values<0.001), and 12.7 (95% CI, 7.7–24.9; all p values<0.001) when delirium rates were 0.01, 0.5, and 1.0 per patient-day, respectively. Thus, in populations where delirium occurs frequently, such as in the ICU, the bias that results from a time-fixed analysis of a time-varying exposure is inflated.

In the time-varying analyses, the null hypothesis was correctly detected regardless of the frequency of delirium. The median HRs were 1.0 (95% CI, 0.5–2.0; 6.9% of p values<0.05), 1.0 (95% CI, 0.7–1.; 5.6% of p values<0.05), and 1.0 (95% CI, 0.6–1.6; 4.5% of p values<0.05) when delirium rates were 0.01, 0.5, and 1.0 per patient-day, respectively.

Discussion

In this investigation, we have used the example of delirium in the ICU as a template to expose and quantify the bias that can result from inappropriately analyzing the relationships between time-varying exposures (risk factors) and clinical outcomes in an observational cohort study using time-fixed methods of analysis rather than time-varying techniques. The main findings of this investigation were that ICU delirium was erroneously shown to be was associated with delayed discharge from the ICU when time-fixed methods were used to

analyze both real and simulated data, whereas a method of analysis that accounted for the time-varying nature of the delirium exposure correctly determined there was no association between delirium and length of stay in the ICU (a fact known a priori to be true in the Monte Carlo simulation study).

Mistakenly choosing the wrong statistical approach in such circumstances is easily correctable as shown by our data. Though statisticians and other investigators have stressed the importance of appropriate analysis of time-dependent variables (2–4, 12, 13) (5), observational studies that are potentially biased by time-fixed analyses continue to abound in the literature. Numerous examples can be found in the field of critical care research alone (14–20), especially in studies evaluating the effect of an exposure in the ICU on length of ICU stay or ICU mortality. In such studies, the period of potential exposure may overlap with the period of time during which the outcome of interest occurs, creating the possibility of immortal time bias. Of course, this type of bias is pervasive in multiple fields of medical research (2). This has occurred despite the avail ability of several statistical methods that can account for the time-varying nature of some exposure variables, including Cox proportional hazards regression with time-varying covariates (21, 22), modified Kaplan-Meier survival functions (10, 23, 24), and Poisson regression (25). We chose to use Cox regression in the current investigation because Cox regression is more flexible than Poisson regression; in Cox regression, a change in the hazard of an outcome may change over time as long as the hazard ratio remains constant (proportional hazards assumption), whereas in Poisson regression, hazards must be constant over time. We did analyze the simulated datasets created in this investigation using Poisson regression, and the results mirrored those produced by Cox regression with time-varying covariates.

Barriers to the use of time-varying methods of analysis have not been formally studied, but several possible explanations warrant discussion. Many investigators may simply be unaware of the bias introduced when time-fixed methods of analysis are inappropriately applied or may not know of the availability of appropriate statistical methods, i.e., methods that can account for time-varying variables. Some statistical software packages cannot perform time-varying methods of analysis, but others can, including SAS, STATA (Stata Corp, College Station, TX), and R. There are currently few packages (e.g., STATA and R) that can create modified Kaplan-Meier curves such as those presented in this investigation. This may explain why, from the time they described their method in 1984 through 2006, Simon and Makuch's original publication (10) has only been cited 54 times.

Even when time-fixed methods of analysis are used, immortal-time bias can be avoided through the use of appropriate study designs. Pittet D et al. (26), for example, examined the effect of nosocomial bloodstream infection on multiple outcomes among critically ill patients in a matched cohort study and dealt with immortal-time bias by using proper controls in the analysis; each patient with a bloodstream infection was matched to a control patient according to the number of days spent in the ICU before the onset of the bloodstream infection, and the remaining days in ICU were used as the outcome. Although it avoids immortal time bias, this design results in a loss of statistical power because a large portion of the control population is not used in matching.

Our investigation is the first to describe the potential magnitude of immortal time bias in ICU studies using Monte Carlo Simulation. We clearly demonstrate that inappropriate analyses of a time-varying exposure can result in significant inflation of type I error— rejection of the null hypothesis when it is actually true, i.e., no association exists between exposure and outcome. Our findings can apply to many scenarios in which patients are at risk for both the exposure and the outcome during the same period of time. Immortal time bias has previously been discussed in the context of survival analyses (3, 4) (5) with the

outcome being time to death. Austin et al., for example, used Monte Carlo simulations to analyze the effect of a treatment on survival and showed that immortal time bias (or survivor treatment selection bias) increased as the time required for exposed patients to receive treatment increased. In this paper, we have extended the discussion to another important outcome, length of stay.

In many situations, length of stay is analyzed with linear regression, and bias is quickly introduced due to the lack of a time-varying method of analysis. We also assessed the impact of immortal time bias resulting from use of linear regression in our simulated datasets, and the results were similar to those generated using Cox regression with time-fixed covariates. Several limitations of the current study warrant discussion. First, our simulated data were intended to mimic an ICU cohort, and it does not reflect the nature of many other clinical scenarios in which the overlap between periods of potential exposure and occurrence of the outcome is different. The magnitude of potential immortal time bias in other observational studies depends, in part, upon the percentage of the study period during which patients are at risk for both exposure and outcome. More exhaustive simulation studies, which model a broader range of scenarios, could determine the magnitude of potential bias introduced in these situations as well as examine the effect of exposures rates that vary over time and of censoring due to death, which were not considered in our investigation. Because delirium is associated with increased mortality (6), the magnitude of immortal time bias may be attenuated in the current study, which did not include patients who died. Lastly, our investigation only examines immortal time bias within the context of one type of exposure (harmful, in this case) and one type of outcome (one in which an increase is bad). The nature of the exposure and outcome variable are important when determining the directionality of immortal time bias. E.g., an ineffective drug may be found effective in prolonging life, or an unharmful exposure may be claimed harmful in delayed ICU stay (Table 1.)

In conclusion, failure to properly account for the time-varying nature of an exposure variable can have severe consequences, resulting in profoundly biased study results. Statistical methods that properly account for time-varying exposures, such as time-dependent Cox regression and Simon and Makuch's modified Kaplan-Meier curves can easily be implemented in commercial statistical software packages such as Stata (Stata Corp, College Station, TX) and R (www.r-project.org). Clinicians reading results of well done cohort studies should also check methods sections of manuscript to ensure that appropriate time-varying analytical techniques have been used in order to avoid immortal time bias.

Acknowledgments

REFERENCES

1. Rosenbaum, PR. Observational Studies. 2nd Edition. New York: Springer; 2002.

2. 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 672682.

3. 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. [PubMed: 12663327]

4. Glesby MJ, Hoover DR. Survivor treatment selection bias in observational studies: examples from the AIDS literature. Ann Intern Med. 1996; 124:999–1005. [PubMed: 8624068]

5. Austin PC, Mamdani MM, van Walraven C, et al. Quantifying the impact of survivor treatment bias in observational studies. J Eval Clin Pract. 2006; 12:601–612. [PubMed: 17100859]

6. Ely EW, Shintani A, Truman B, et al. Delirium as a predictor of mortality in mechanically ventilated patients in the intensive care unit. JAMA. 2004; 291:1753–1762. [PubMed: 15082703]

7. Ely EW, Margolin R, Francis J, et al. Evaluation of delirium in critically ill patients: validation of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU). Crit Care Med. 2001; 29:1370–1379. [PubMed: 11445689]

8. Ely EW, Inouye SK, Bernard GR, et al. Delirium in mechanically ventilated patients: validity and reliability of the confusion assessment method for the intensive care unit (CAM-ICU). JAMA. 2001; 286:2703–2710. [PubMed: 11730446]

9. Dubin JA, Muller HG, Wang JL. Event history graphs for censored survival data. Stat Med. 2001; 20:2951–2964. [PubMed: 11568951]

10. Simon R, Makuch RW. A non-parametric graphical representation of the relationship between survival and the occurrence of an event: application to responder versus non-responder bias. Stat Med. 1984; 3:35–44. [PubMed: 6729287]

11. Ihaka R, Gentleman R. R: a language for data analysis and graphics. J Comput Graph Stat. 1996; 5:299–314.

12. Sylvestre MP, Huszti E, Hanley JA. Do Oscar winners live longer than less successful peers? A reanalysis of the evidence. Ann Intern Med. 2006; 145:361–363. [PubMed: 16954361]

13. Girard TD, Shintani AK, Ely EW. Comment on "Incidence, risk factors and consequences of ICU delirium" by Ouimet et al. Intensive Care Med. 2007; 33:1479–1480. [PubMed: 17549453]

14. Rello J, Lorente C, Diaz E, et al. Incidence, etiology, and outcome of nosocomial pneumonia in ICU patients requiring percutaneous tracheotomy for mechanical ventilation. Chest. 2003; 124:2239–2243. [PubMed: 14665506]

15. Laupland KB, Gregson DB, Zygun DA, et al. Severe bloodstream infections: a population-based assessment. Crit Care Med. 2004; 32:992–997. [PubMed: 15071391]

16. Lin SM, Liu CY, Wang CH, et al. The impact of delirium on the survival of mechanically ventilated patients. Crit Care Med. 2004; 32:2254–2259. [PubMed: 15640638]

17. Jaber S, Chanques G, Borry J, et al. Cytomegalovirus infection in critically ill patients: associated factors and consequences. Chest. 2005; 127:233–241. [PubMed: 15653989]

18. Milbrandt EB, Kersten A, Kong L, et al. Haloperidol use is associated with lower hospital mortality in mechanically ventilated patients. Crit Care Med. 2005; 33:226–229. [PubMed: 15644675]

19. Ouimet S, Kavanagh BP, Gottfried SB, et al. Incidence, risk factors and consequences of ICU delirium. Intensive Care Med. 2007; 33:66–73. [PubMed: 17102966]

20. Barrantes F, Tian J, Vazquez R, et al. Acute kidney injury criteria predict outcomes of critically ill patients. Crit Care Med. 2008; 36:1397–1403. [PubMed: 18434915]

21. Cox DR. Regression models and life-tables. J R Stat Soc Ser B. 1972; 34:187–220.

22. Crowley J, Hu M. Covariance analysis of heart-transplant survival data. J Am Stat Assoc. 1977; 72:27–36.

23. Mantel N, Byar DP. Evaluation of response-time data involving transient states Illustration ssing heart-transplant data. J Am Stat Assoc. 1974; 69:81–86.

24. Snapinn SM, Jiang Q, Iglewicz B. Illustrating the impact of a time-varying covariate with an extended Kaplan-Meier estimator. Am Stat. 2005; 59:301–307.

25. Suissa S. Immortal time bias in observational studies of drug effects. Pharmacoepidemiol Drug Saf. 2007; 16:241–249. [PubMed: 17252614]

26. Pittet D, Tarara D, Wenzel RP. Nosocomial bloodstream infection in critically ill patients. Excess length of stay, extra costs, and attributable mortality. JAMA. 1994; 271:1598–1601. [PubMed: 8182812]
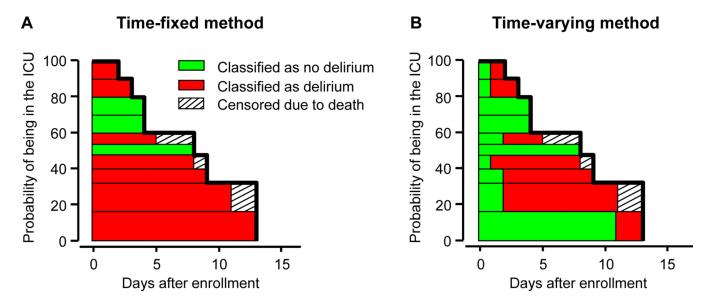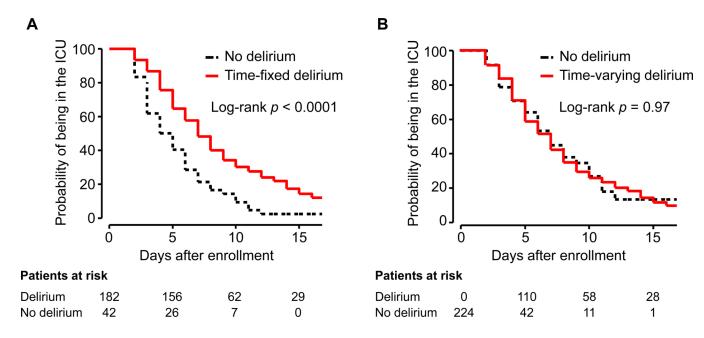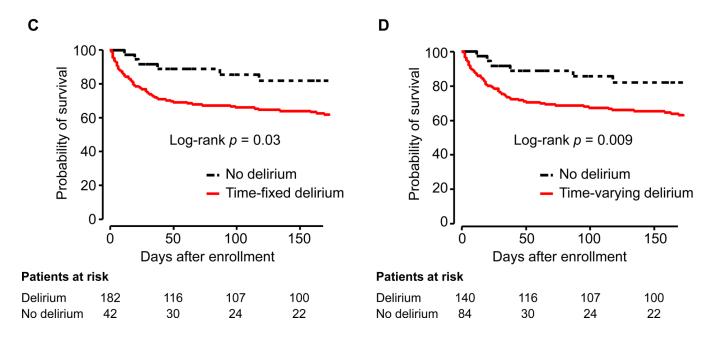
Figure 1.
1A and 1B. Combined Delirium Exposure Status and Kaplan Meier (K-M) Curve for 10 Randomly Selected Patients using Event History Graphs for 10 randomly selected patients from among 224 patients in the example cohort study. Reader should interpret the upper boundary and the right boundary of the bars as creating a K-M plot (bolded line), with details of the "event history" displayed via green, red, or diagonal shading of the bars as indicated period of exposure or censoring used in each analysis. Seven subjects were discharged alive from the ICU (bars ending without diagonal shading), and 3 subjects were censored due to death (diagonally shaded bars). In (A) exposure status does not change for the entire follow-up period "time-fixed", and according to retrospective classification in ever-or-never having delirium. This graph makes it appear that patients without delirium (green) are likely to be discharged early, whereas patients with delirium (red) appear to have longer ICU stays. Event-history graph (B) conveys information regarding how a patient's status of exposure changed during the period of study (i.e., "time-varying") according to the "actual" onset of delirium, showing that some patients who ultimately develop delirium do not do so until later in their ICU course, thus, suggesting less of an association between delirium and delayed discharge from the ICU. For example, the top horizontal bar indicates a patient who was delirium free at enrollment developed delirium on day 1, and was discharged alive from ICU on day 2. The fifth bar indicates a patient who developed delirium for the first time on day 2 died at day 5 in ICU.

A

Probability of being in the ICU

- No delirium
- Time-fixed delirium

Log-rank *p* < 0.0001

Days after enrollment

Patients at risk

Delirium	182	156	62	29
No delirium	42	26	7	0

B

Probability of being in the ICU

- No delirium
- Time-varying delirium

Log-rank *p* = 0.97

Days after enrollment

Patients at risk

Delirium	0	110	58	28
No delirium	224	42	11	1

C

Probability of survival

Log-rank *p* = 0.03

- No delirium
- Time-fixed delirium

Days after enrollment

Patients at risk

Delirium	182	116	107	100
No delirium	42	30	24	22

D

Probability of survival

Log-rank *p* = 0.009

- No delirium
- Time-varying delirium

Days after enrollment

Patients at risk

Delirium	140	116	107	100
No delirium	84	30	24	22

Figure 2.
2A and 2B. K-M analysis of assessing the effect of delirium on time to ICU discharge
Figure 2A plots the percentage of patients remaining in the ICU from study enrollment to
completion using traditional Kaplan-Meier curves, one for patients who were ever delirious
and one for those who were never delirious. Alternatively, Figure 2B plots the percentage of
patients remaining in the ICU from study enrollment to completion using Kaplan-Meier
survival curves modified according to the Simon-Makuch method (see reference below *).
On any given study day, one curve represents patients with delirium and the other represents
those without delirium. Because the Simon-Makuch method accounts for a change in
exposure status over time, the number of patients represented by each curve changes as the
curve moves from left to right on the x-axis. In the time-varying approach, all 224 patients

start off as not-having delirium and transfer to delirium category at its onset, and are classified as delirium thereafter. Hence, the number of patients in delirium category increased as this transfer occurs and decreased as censoring or ICU discharge took place.

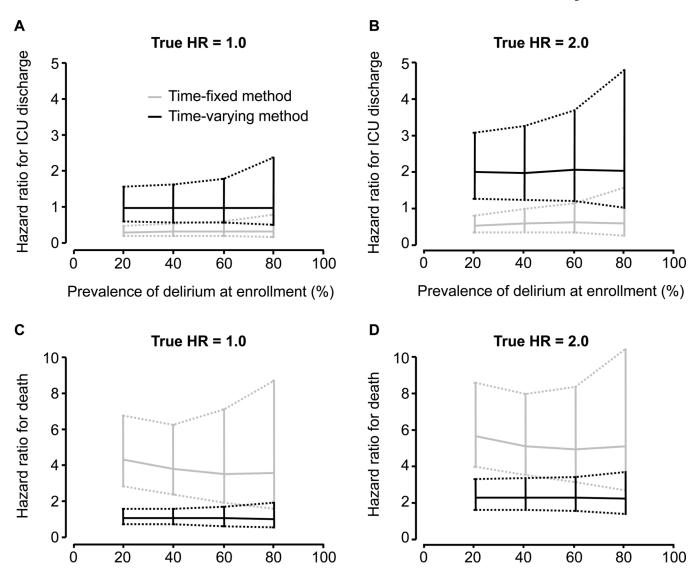<u>Reference</u> *: Simon R and Makuch R. Statistics in Medicine, Vol 3, pp. 35–44 (1984).

Figure 3.
3A and 3B. Line Graph (with 95% CI) showing relationship between delirium rate (X-axes) vs. Y-axes values for the (A) Hazard Ratio and (B) Chi Square Value of a delayed ICU discharge according to analysis by Time-fixed and Time-varying Cox models. These data represent the results of Part II of this investigation, which used over the 1000 simulation datasets with a known lack of association between delirium rate and ICU length of stay. The Y-axis Chi-square test statistics shown in Figure 3B correspond with the hazard ratios shown on the Y-axis in Figure 3A. Each test statistic value plotted in Figure 3B corresponds with p-value. A larger value of test statistics indicates a smaller p-value. For example, chi-square of 10 is equivalent with p-value = 0.0016, chi-square of 50 is equivalent with $P = 10^{-12}$. P-values approach to negative infinity for corresponding chi-square test statistics being larger than 50. Both Figure 3A and 3B reinforce the findings of Part I of this investigation, showing marked differences in the results of time-fixed vs. time-varying approaches to determining the relationship between delirium and the hazard of a delayed ICU discharge. See the text in Results for a more detailed description of the numerical interpretations of these data, which mainly emphasize that the higher the delirium rates from

left to right on the X-axes, the more discrepant the results between the time-fixed vs. time-varying Cox models, and thus the greater the error found when using a time-fixed analytical approach.

Table 1

Types of error which can be inflated by immortal time bias depending on outcomes and exposures analyzed.

	Exposure	
Outcome	Harmful exposure	Beneficial exposure
Length of stay	Type I error	Type II error
Mortality	Type II error	Type I error

The direction of immortal time bias depends on the type of outcome assessed and the nature of the exposure. As follow-up continues, the opportunity for exposure increases. Thus, if the exposure is beneficial, immortal time bias works to inflate the true survival benefit while simultaneously falsely showing prolongation of length of stay (instead of showing a reduced length of stay, which would be expected from a beneficial exposure). When the exposure is harmful, immortal time bias masks the increased risk of mortality but falsely increases length of stay.