International Epidemiological Association

International Journal of Epidemiology, 2021, 1–5
doi: 10.1093/ije/dyab157
Editorial [Invited Only]

Editorial [Invited Only]

Causal diagrams for immortal time bias

Mohammad Ali Mansournia,[1] Maryam Nazemipour[2,*] and Mahyar Etminan[3]

[1]Department of Epidemiology and Biostatistics, School of Public Health, Tehran University of Medical Sciences, Tehran, Iran, [2]Osteoporosis Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, Tehran, Iran and [3]Department of Ophthalmology and Visual Sciences and Medicine, University of British Columbia, Vancouver, Canada

*Corresponding author. Osteoporosis Research Center, Endocrinology and Metabolism Clinical Sciences Institute, Tehran University of Medical Sciences, PO Box: 1411713119, Tehran, Iran. E-mail: ma_nazemipour@yahoo.com

Editorial decision 7 July 2021; Accepted 9 July 2021

Introduction

Well-designed observational cohort studies are frequently used for estimating the causal effects of exposures when randomized trials are not feasible or ethical. However, these studies are subject to several biases in the taxonomy of confounding, selection bias and measurement bias.[1] In particular, incorrect handling of follow-up times in terms of exposure status in the analysis of such studies may introduce immortal time bias (ITB) in favour of the exposed group.[2,3] Immortal time refers to a period of time in which, by design, participants in the exposed group cannot experience the outcome. This often happens in pharmacoepidemiologic studies in which treatment is prescribed at variable times (with delay) after disease diagnosis. The bias occurs when the exposed group is considered to be exposed during their entire follow-up time (even during periods in which they are theoretically unexposed) or their unexposed follow-up times are discarded.[2,3]

In recent years, ITB has been identified in a large body of literature in different areas of medicine including cancer,[4] diabetes,[3] heart disease[5] and COPD[2] among others. Biased results from these studies can lead to inappropriate prescribing of drug therapies or even harmful outcomes in patients. Despite the publication of many examples and tutorials on this topic, ITB still seems to be prevalent in many cohort studies.[6,7] Perhaps one potential reason might be that the structure of ITB is still poorly understood.

Better understanding of ITB requires appreciation for its structure. Causal diagrams have been extensively used to represent several biases in epidemiology.[8–13] A detailed description of causal diagrams is beyond the scope of this paper. In brief, directed acyclic graphs (DAGs) include nodes (measured and unmeasured variables) linked by directed edges (arrows). A causal DAG (cDAG) is one in which the absence of an arrow between two variables implies the absence of a direct causal effect and in which it is assumed that all shared causes of any pair of variables (i.e. confounders) are included in the graph. Two variables A and B are connected with an arrow that starts from A and ends on B, meaning that A causes B. Variable A cannot cause itself as cDAGs are acyclic. Adjacent arrows, irrespective of their direction, form paths in a cDAG. Conditioning on a variable means restricting that variable to a subset of its values and can be undertaken at the data-analysis stage (e.g. regression adjustment) or study-design stage (e.g. restriction).

Causal diagrams provide a mathematically rigorous yet intuitive tool for identifying structural biases. To our knowledge, a detailed discussion of the ITB structure using causal diagrams has not been presented to date. In this paper, we use causal diagrams to represent the structure of ITB. We graphically explore the structure of ITB using causal diagrams and demonstrate how ITB can lead to selection bias and measurement bias, considered as two of

© The Author(s) 2021; all rights reserved. Published by Oxford University Press on behalf of the International Epidemiological Association

Downloaded from https://academic.oup.com/ije/advance-article/doi/10.1093/ije/dyab157/6333514 by University of British Columbia user on 02 August 2021

the three (along with confounding bias) fundamental types of bias in epidemiology. We will also discuss mitigation strategies that researchers can use to prevent ITB in their studies.

An illustrative example

We will use the first description of ITB using a land mark study undertaken in the late 1960s in the context of two cohort studies that assessed the effect of heart transplantation on mortality (compared with medical therapy).[14] For example, imagine a group of patients with end-stage heart failure who are actively receiving medical therapy while also registered on the waiting list of receiving a heart transplant. For simplicity, we assume a large sequentially randomized experiment without censoring, random confounding and non-compliance with treatment,[11,15] i.e. the participants randomly received heart transplant at monthly visits, and all were followed until death or a priori end date, whichever came first (Figure 1). The waiting time for patients who survived to receive the transplant is immortal, i.e. patients could not die during the time they waited for a transplant, leading to ITB if this immortal time is not appropriately accounted for in the analysis.

ITB I: misclassification

One can incorrectly assign the waiting time of subjects who receive heart transplant to the exposed group in the analysis even though they had received medical therapy during that period. ITB can occur if the immortal time (unexposed time) of those who receive a heart transplant is incorrectly considered as exposed person-time. For example, if a subject is on medical therapy since cohort entry (month 0) and receives a heart transplant on day 90 (third month), this subject's time for the first 3 months (days 0–90) is incorrectly counted as exposed.

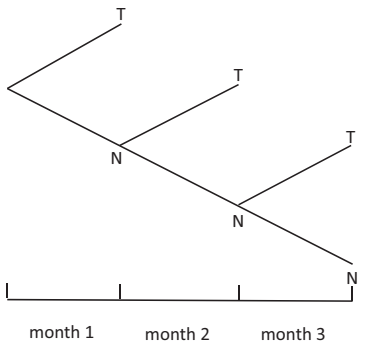

Figure 1 A sequentially randomized experiment in which patients with end-stage heart failure who are on medical therapy randomly receive heart transplant at monthly visits. T denotes heart transplant and N represents no heart transplant. For simplicity, the diagram represents only the first three visits.

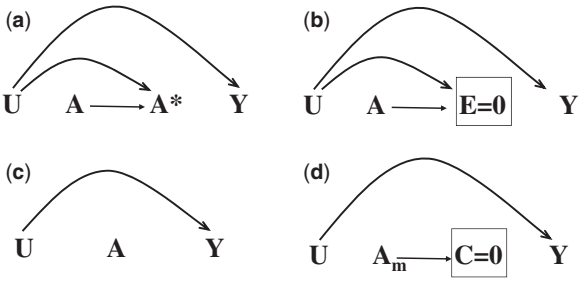

Figure 2 Causal diagrams representing different approaches for handling immortal times: (A) misclassification or measurement bias due to the assignment of the immortal times to the exposed, (B) selection bias due to the exclusion of the immortal times, (C) no immortal time bias in the time-dependent analysis and (D) no immortal time bias in the sequential approach.

The causal diagram in Figure 2A, representing the subjects' person-time data so that each exposed subject effectively contributes as two persons: (i) unexposed before receiving a heart transplant and (ii) exposed after receiving a transplant, demonstrates that the ITB arising from this error is an example of misclassification or measurement bias.[10] The variables A and Y denote a subject's exposure status (1: heart transplantation plus medical therapy, 0: medical therapy alone) and outcome (1: died, 0: survived), respectively. The variable U represents an unmeasured protective cause of the outcome, e.g. a cardioprotective haplotype (1: yes, 0: no). Variable A* is a misclassified version of A, i.e. $A^* = 1$ during the immortal time period ($A = 0$); otherwise $A^* = A$. The arrows from A and U to A* reflect that unexposed subjects with the haplotype ($A = 0$ and $U = 1$) are more likely to be misclassified (from $A = 0$ to $A^* = 1$) than those without the haplotype ($A = 0$ and $U = 0$) because the former has a higher probability of surviving the waiting time and hence receiving a heart transplant than the latter.

In Figure 2A, the absence of an arrow from A to Y represents the causal null hypothesis of no effect of heart transplantation on death. However, the misclassified exposure (A*) is associated with Y through U as the path $A^* \leftarrow U \rightarrow Y$ is open. The variables A and U are independent since the (true) exposure was randomly assigned to the patients, and in fact A* is a collider on the path $A \rightarrow A^* \leftarrow U$. However, the misclassified exposure (A*) is positively associated with U because, as previously mentioned, the subjects with the haplotype ($U = 1$) are more likely to receive a heart transplant ($A^* = 1$) than those without the haplotype ($U = 0$). The ITB presented in Figure 2A will be in favour of heart transplantation because the sign of the association between A* and U is positive, and also U and Y are negatively associated.

Downloaded from https://academic.oup.com/ije/advance-article/doi/10.1093/ije/dyab157/6333514 by University of British Columbia user on 02 August 2021

ITB II: selection bias

The second ITB type is through inappropriate exclusion of immortal times of the exposed participants (those who received heart transplant) from the analysis, which can lead to selection bias.[12,16] The causal diagram in Figure 2B represents the structure of this selection bias. Variable E indicates the exclusion of immortal times: E = 1 for the excluded (immortal) person-times and E = 0 otherwise. The square around E = 0 indicates that the analysis is limited and subsequently conditioned to person-times that were not excluded. The rationale behind arrows from A and U to E is that the unexposed group with haplotype (A = 0 and U = 1) is more likely to be excluded from the analysis (E = 1) than those without haplotype (A = 0 and U = 0), because the former have a higher probability of receiving a heart transplant than the latter, as they are more likely to survive the waiting time.

The causal diagram in Figure 2B displays the null hypothesis of no effect of heart transplantation on death as there is no arrow from A to Y. However, A and Y become associated as, by exclusion of the immortal times, the collider E on the path A→E←U→Y has been conditioned on, creating an open path between A and Y. As a result of this conditioning, the variables A and U become positively associated as the unexposed subjects without haplotype (A=0 and U = 0) are more likely to be included in the analysis than those with haplotype (A=0 and U = 1). Again, the ITB presented in Figure 2B is in favour of the heart-transplantation group because the sign of the association between A and U is positive, and also U and Y are negatively associated.

In an attempt to correct the ITB mentioned above, one can emulate immortal times for the unexposed subjects from the distribution of the immortal times of the exposed group. Using this approach, similar exclusions are applied to the unexposed subjects by randomly sampling (with replacement) from the list of immortal times for each unexposed subject and then subtracting this immortal time from his/her follow-up time; the subject will be excluded if the result of the subtraction is negative (Figure 3).[17,18] This approach, which effectively recalibrates the follow-up times for the unexposed subjects, can partly remedy the selection-bias problem in the exclusion scenario by discarding the unexposed subjects who are more likely 'frail' and deemed to die before the assigned immortal time (those designated as A = 0 and U = 0) from the analysis (Figure 3B). However, the danger with this method can be observed for unexposed subjects whose follow-up times are longer than the matched immortal time (Figure 3C). In fact, the bias can be further amplified due to elimination of the immortal time twice: once from the exposed subjects and once from the unexposed subjects (Figure 3A and C). Specifically, it involves the probable exclusion of the healthy unexposed follow-up times (A = 0 and U = 1) from unexposed subjects. Thus, consistent with the previous theoretical derivation and simulation studies,[18] the net result of matched exclusions is selection bias and the introduction of a positive association between A and U, which in turn induces a biased negative association between A and Y. This approach has been called the 'prescription time-distribution matching method',[17,18] although, unlike its name, it does not really involve matched sampling.

Mitigation strategies for avoiding ITB

Time-dependent analysis

The optimal approach for avoiding ITB is correctly allocating each subject's follow-up time contribution during exposed and unexposed periods, i.e. the follow-up time of a subject should be analysed as unexposed before he/she receives a heart transplant, and exposed thereafter.[2,3,17,18] Figure 2C represents the causal diagram for this time-dependent analysis and suggests that under the causal null hypothesis of no effect of heart transplantation on death, A and Y are independent of each other as there is no path between them. Thus, there is no bias with a time-dependent analysis. An example of a time-dependent analysis is using a time-dependent Cox regression model. However, other methods such as time-dependent parametric survival models can also be used.[19]

Sequential approach

Another potential remedy to control for ITB is the use of a sequential approach that mimics mini-randomized trials at monthly intervals.[20] Using this approach, in each month, only patients who have not yet received a heart transplant are included. Moreover, those who receive a heart transplant in the subsequent months are artificially censored at the time of receiving a heart transplant. Thus patients are either exposed or unexposed in each mini-randomized trial, circumventing the immortal time.[18] The analysis results of mini-randomized trials can be pooled to estimate the causal effect of the exposure, e.g. using a monthly stratified Cox model, assuming that baseline hazard functions vary over mini-randomized trials. Within-subject correlation, introduced as the same subject may be included in several mini-randomized trials, should be taken into account using e.g. cluster-robust standard errors.[21,22] The

Downloaded from https://academic.oup.com/ije/advance-article/doi/10.1093/ije/dyab157/6333514 by University of British Columbia user on 02 August 2021

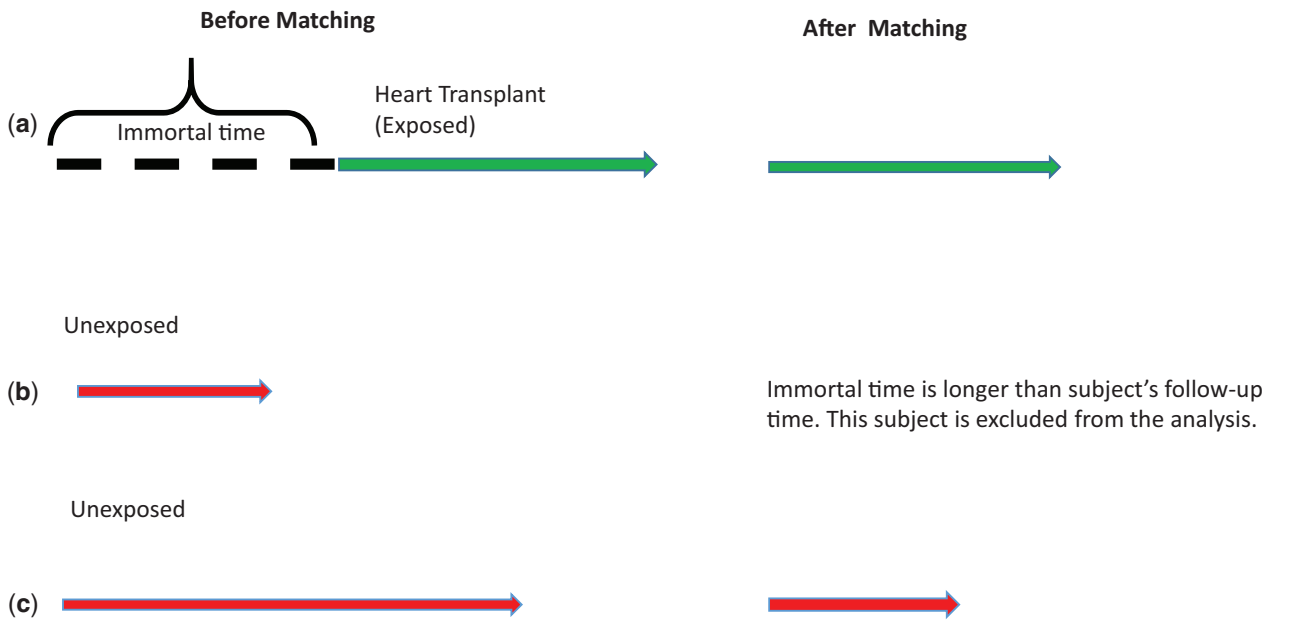

Figure 3 Immortal time matched exclusion method to control for immortal time bias assuming that subject A (exposed) is matched to subject B or C (unexposed). Subject B's follow-up time is less than subject A's immortal time, so subject B is excluded from the analysis. Subject C's updated follow-up time (after matching) equals his/her follow-up time (before matching) minus subject A's immortal time. Although this approach reduces selection bias in the exclusion scenario by eliminating some of the frail unexposed like subject B, it reintroduces bias due to omitting the early unexposed times from the rigorous unexposed like subject C. Thus the net effect of immortal time matching can be selection bias.

main idea of the sequential approach stems from the emulation of randomized trials from observational studies in which an observational cohort is considered as a sequence of non-randomized nested trials and subjects are censored when they have discontinued their baseline treatment and then weighted using inverse probability-of-censoring weighing (IPCW).[23,24]

The causal diagram in Figure 2D shows the mini-randomized trial in the sequential approach at month 'm'. The variable A_m denotes heart transplant at month 'm' in those who have not yet received a transplant. Variable C represents the artificial censoring of patients receiving a transplant at a later month, hence the arrow from A_m to C. The square around $C = 0$ suggests that the analysis is limited to those who are not censored. Figure 2D suggests that A_m and Y are independent under the causal null hypothesis of no effect of heart transplantation on death and so there is no bias.

Discussion

We have used causal diagrams to demonstrate the structure of immortal time bias and have provided diagrammatic explanations of the two biases inherent in ITB. Allocation of the immortal times to exposed subjects introduces misclassification or measurement bias in favour of the exposed subjects. This bias is proportional to the ratios of immortal person-time (sum of waiting times) to exposed and unexposed person-times. Similarly, exclusion of the immortal times from the analysis results in selection bias in favour of the exposed, which is proportional to the ratio of immortal person-time to unexposed person-time.[2,18] Of note, the selection bias remains in effect even if one resorts to matched exclusions in the unexposed group. Thus the ITB would be away from the null if exposure is protective or without any effect, and towards the null if it is a risk factor.

The magnitude of the misclassification bias is generally higher than the selection bias as the error involves two folds. The first is removal of the (unexposed) immortal times from the unexposed person-times and the second is adding them to the exposed person-times. However, in the case of the exclusion of immortal times, the error is only one-fold, mainly only the exclusion of immortal times from the unexposed person-times. Both biases can be remarkable if the prevalence of exposure is high or the outcome incidence is low.[18]

Suissa has described how different variations of cohort studies including time-based cohorts, event-based cohorts, exposure-based cohorts, multiple-event-based cohorts and event-exposure-based cohorts can lead to ITB.[2] We note that these designs are simply variations of cohort-study scenarios in which immortal time is observed. The ITB in these situations still falls under the categories of either (i) immortal time is misclassified to the exposed (misclassification) or (ii) excluded from the analysis (selection bias).

Downloaded from https://academic.oup.com/ije/advance-article/doi/10.1093/ije/dyab157/6333514 by University of British Columbia user on 02 August 2021

The optimal solution to the ITB is a time-dependent analysis, i.e. the immortal times should be considered as unexposed so that the exposed subjects contribute to both exposed and unexposed person-times.[2,3,17,18] Another remedy is the sequential approach in which the study is considered as a series of nested mini-trials with artificial censoring at the time of receiving exposure.[18] In practice, the observational studies are subject to confounding as well as selection bias due to censoring, which should be appropriately adjusted for in the analysis. In particular, confounders would be the predictors of artificial censoring in the sequential approach and the resulting selection bias should be adjusted for using IPCW in the analysis.[25,26] Finally, it is important to note that heart transplant and unmeasured haplotype in our example are independent by design, i.e. a sequentially randomized experiment, but as survival is affected by both of them, there will be a built-in selection bias in the hazard ratio generated from survival analysis. Therefore, it is preferred to perform risk (survival curve) comparisons based on time-dependent Cox regression or the stratified Cox model when undertaking the sequential approach.[1]

ITB is a complex bias that can present itself as either exposure misclassification or selection bias. We have demonstrated the mechanics of this bias using casual diagrams in the hope that researchers can have a better appreciation for this bias and use appropriate mitigation strategies to control for this bias in their studies.

Conflict of interest

None declared.

References

1. Hernán MA, Robins JM. *Causal Inference: What If*. Boca Raton: Chapman & Hall/CRC, 2021.
2. Suissa S. Immortal time bias in pharmacoepidemiology. *Am J Epidemiol* 2008;**167**:492–99.
3. Lévesque LE, Hanley JA, Kezouh A, Suissa S. Problem of immortal time bias in cohort studies: example using statins for preventing progression of diabetes. *BMJ* 2010;**340**:b5087.
4. Suissa S, Azoulay L. Metformin and the risk of cancer: time-related biases in observational studies. *Diabetes Care* 2012;**35**: 2665–73.
5. Suissa S. Risk-treatment mismatch for heart failure. *JAMA* 2005; **294**:3087–88.
6. Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf* 2007;**16**:241–49.
7. Etminan M, Nazemipour M, Candidate MS, Mansournia MA. Potential biases in studies of acid-suppressing drugs and COVID-19 infection. *Gastroenterology* 2021;**160**:1443–46.
8. Mansournia MA, Hernán MA, Greenland S. Matched designs and causal diagrams. *Int J Epidemiol* 2013;**42**:860–69.
9. Etminan M, Collins GS, Mansournia MA. Using causal diagrams to improve the design and interpretation of medical research. *Chest* 2020;**158**:S21–28.
10. Hernán MA, Cole SR. Invited commentary: causal diagrams and measurement bias. *Am J Epidemiol* 2009;**170**:959–62.
11. Mansournia MA, Higgins JP, Sterne JA, Hernán MA. Biases in randomized trials: a conversation between trialists and epidemiologists. *Epidemiology* 2017;**28**:54–59.
12. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;**15**:615–25.
13. Etminan M, Brophy JM, Collins G, Nazemipour M, Mansournia MA. To adjust or not to adjust: the role of different covariates in cardiovascular observational studies. *Am Heart J* 2021;**237**:62–67.
14. Messmer B, Leachman R, Nora J, Cooley D. Survival-times after cardiac allografts. *Lancet* 1969;**293**:954–56.
15. Greenland S, Mansournia MA. Limitations of individual causal models, causal graphs, and ignorability assumptions, as illustrated by random confounding and design unfaithfulness. *Eur J Epidemiol* 2015;**30**:1101–10.
16. Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology* 2003;**14**: 300–06.
17. Zhou Z, Rahme E, Abrahamowicz M, Pilote L. Survival bias associated with time-to-treatment initiation in drug effectiveness evaluation: a comparison of methods. *Am J Epidemiol* 2005; **162**:1016–23.
18. Karim ME, Gustafson P, Petkau J *et al.* Long-Term Benefits and Adverse Effects of Beta-Interferon for Multiple Sclerosis (BeAMS) Study Group. Comparison of statistical approaches for dealing with immortal time bias in drug effectiveness studies. *Am J Epidemiol* 2016;**184**:325–35.
19. Kleinbaum DG, Klein M. *Survival Analysis: A Self-Learning Text*, 3rd edn. New York: Spinger, 2012.
20. 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**:2757–68.
21. Mansournia MA, Nazemipour M, Naimi AI, Collins GS, Campbell MJ. Reflection on modern methods: demystifying robust standard errors for epidemiologists. *Int J Epidemiol* 2021; **50**:346–51.
22. Mansournia MA, Collins GS, Nielsen RO *et al.* A CHecklist for statistical Assessment of Medical Papers (the CHAMP statement): explanation and elaboration. *Br J Sports Med* 2021.
23. 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**:766–79.
24. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol* 2016; **183**:758–64.
25. Mansournia MA, Altman DG. Inverse probability weighting. *BMJ* 2016;**352**:i189.
26. Mansournia MA, Etminan M, Danaei G, Kaufman JS, Collins G. Handling time varying confounding in observational research. *BMJ* 2017;**359**:j4587.

Downloaded from https://academic.oup.com/ije/advance-article/doi/10.1093/ije/dyab157/6333514 by University of British Columbia user on 02 August 2021