Identifying and avoiding design-related biases in observational studies using the target trial framework

Harrison J Hansford[1,2], Nazrul Islam[3], Hopin Lee[4,5], Barbra A Dickerman[6], Aidan G Cashin[1,2]

1. School of Health Sciences, Faculty of Medicine and Health, UNSW Sydney, Sydney, Australia
2. Centre for Pain IMPACT, Neuroscience Research Australia, Sydney, Australia
3. Faculty of Medicine, University of Southampton, Southampton, United Kingdom
4. University of Exeter Medical School, Exeter, United Kingdom
5. EMEA Methods & Evidence Generation, IQVIA, London, United Kingdom
6. CAUSALab, Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts

Disclosures

HJH, HL, AGC are leading the development of the TARGET reporting guideline for studies emulating a target trial. NI and BAD are members of the consortium of experts involved in developing the TARGET guideline. NI is an Editor at The BMJ, and a member of the Editorial Advisory Board at BMJ Medicine. However, he will not be involved in the decision-making process of the manuscript.

Abstract

Observational studies are necessary to provide evidence to inform decision-making in the absence of a relevant randomised trial. Although commonly criticised for potential problems due to confounding bias, design-related biases in observational studies are often overlooked yet highly prevalent. Design-related biases occur due to decisions made by researchers during analyses of observational data. Common design-related biases include bias related to selection and treatment misclassification, resulting from misalignment of eligibility ascertainment, treatment strategy assignment, and start of follow-up. Conceptualising the analysis of observational data to identify the causal effects of interventions as an attempt to explicitly emulate a target trial can help avoid design-related biases, so that investigators can instead and focus on data-related biases (e.g., confounding, measurement error) not directly addressed by the framework. Target trial emulation may also help readers appraise an observational study when transparently reported. This article aims to help readers of observational studies identify, and investigators avoid, design-related biases to support the use of observational evidence to inform clinical and policy decision-making.

Standfirst

Design-related biases are common in observational studies investigating comparative effectiveness of interventions; this manuscript outlines how to identify them and how target trial emulation can help to avoid these biases.

Background

Randomised trials are the preferred approach to estimate the causal effects of medical interventions.[1,2] However, it is often not feasible to conduct randomised trials to investigate the long-term safety and effectiveness of interventions,[3] identify specific groups who benefit (or are harmed) most from interventions,[4] or to provide evidence when timely information is needed for decision making.[5] Routinely collected, observational data (e.g., in the form of electronic health records and claims data) have become increasingly available for research[6] and are frequently used to inform decision-making where randomised trials are not available.[7,8] Because of the ability to complement evidence from randomised trials, regulators such as the European Medicines Agency are reviewing the use of observational data to inform regulatory decision-making around the safety and effectiveness of medicines.[9] Despite the necessity of using observational data to inform health decision-making in the absence of a relevant randomised trial, there are clear challenges in drawing valid causal inferences from observational data.[10]

Clinicians and policy makers are often hesitant to rely on evidence from observational studies due to concerns of confounding bias due to the lack of randomisation.[10,11] However, design-related biases[12] can explain discrepancies between findings in observational data and randomised trials, even more so than confounding.[13–17] These biases are introduced when observational studies are designed in a manner which deviates from the design principles of a randomised trial.[18] For example, well-designed randomised trials will follow participants from the time of randomisation (treatment strategy assignment) immediately after meeting eligibility criteria. In observational studies of longitudinal data, the timing of eligibility, treatment assignment and start of follow-up are not naturally aligned, and analysts must therefore define these timepoints retrospectively. The choice of these timepoints is where errors commonly occur.[13,19] In a cross-sectional study of 200 observational studies of drug interventions,[20] Yaacoub et al., highlighted that over 75% of studies were subject to at least one avoidable design-related bias related to the timing of eligibility ascertainment, treatment strategy assignment, and start of follow-up.

It is critical that readers and investigators alike can identify and avoid these often-overlooked design-related biases to maximise the usefulness of observational data for decision-making. This Research, Methods and Reporting article aims to provide a detailed outline of how design-related biases may occur and how target trial emulation can help avoid these biases, with reference to the growing literature on this topic in health and medical

77 research.[12,13,17,21–27,27–33] Key terms used in this article are defined in Box 1, with

78 references to articles with more detail.

79

80 **Box 1: Key terms**

Term	Definition
Target trial	The hypothetical randomised trial that would be conducted to answer the question of interest, with respect to the variables available in the observational data.[22]
Target trial emulation	The process of mapping the analysis of observational data to the components of the target trial (eligibility criteria, treatment strategies, assignment procedures, follow-up, outcomes, causal contrast, data analysis plan) with specification of the assumptions necessary for causal inference.[22]
Time zero	The time at which follow-up begins, which should align with the time of eligibility ascertainment and treatment strategy assignment.[12]
Design-related bias	Biases introduce by investigator decisions about the design of an observational study, often occurring when investigators deviate from the design principles of a randomised trial.[13]

Selection bias	The systematic inclusion or exclusion of individuals into an analysis resulting in a biased treatment effect estimate.[23]
Misclassification of treatment assignment	Errors when classifying individuals into a treatment strategy caused by the use of information that emerges after the beginning of follow-up to assign individuals to a treatment strategy. This often occurs when baseline information alone is insufficient to classify individuals into a treatment strategy.[23]
Clone-censor-weight	An analytic approach where the study dataset is replicated (cloned) and replicates (clones) are assigned to each treatment strategy with which their data are compatible with at baseline and censored when they deviate from their assigned strategy, and inverse probability of censoring weights are applied to overcome biases introduced by the censoring. This approach is used to handle scenarios in which baseline information alone is insufficient to classify individuals into a treatment strategy.[30,32]
Sequential trial emulation	An analytic approach where a series of hypothetical target trials are conceptualised and emulated where individuals who are eligible at multiple times are included at each time of eligibility. This approach is used to handle scenarios where individuals may

	meet the eligibility criteria at multiple times over follow-up and is more statistically efficient than selecting only one of those times as time zero.[34]

Identifying common design-related biases in observational studies

Deviation from the design principles of a randomised trial in an observational study will often introduce biases that occur primarily due to the misalignment of 'time-zero' – that is, the time of eligibility ascertainment, treatment strategy assignment, and start of follow-up.[23] For example, selection bias can be introduced when prevalent users of the treatment(s) are included in an analysis,[35] and immortal time may appear when a) eligibility criteria are ascertained using information that emerges after the start of follow-up, or b) information that emerges after baseline is used to classify individuals into a treatment strategy.[20,36,37] In addition to the design-related biases highlighted in the following sections, we note that other biases such as confounding may also be present and require appropriate data and analytic approaches to address.[38]

Prevalent user bias

Studies that include prevalent users of treatment are prone to selection bias (sometimes referred to as 'prevalent-user' bias, or 'depletion of susceptibles').[37,39] This bias arises when the 'treated' group includes individuals who had been taking treatment for some time

before the start of follow-up and who remained alive and event-free until the start of follow-up (Figure 1). This results in a systematic 'depletion of susceptibles' in the 'treated' group, which may therefore appear safer or more effective.

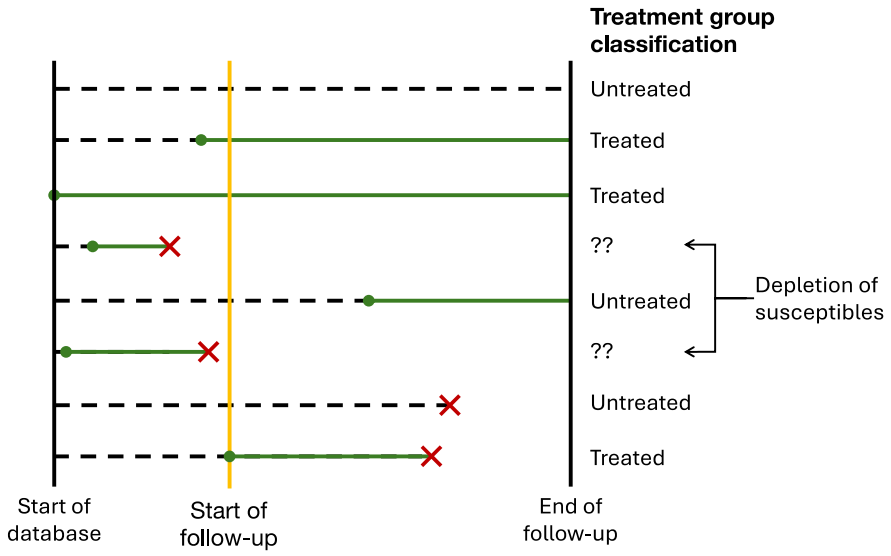

Figure 1: Graphical depiction of prevalent user bias. Green lines indicate treatment, dotted lines indicate no treatment. Red crosses depict people who died, experienced the outcome of interest, or were lost to follow-up. The yellow line indicates the start of follow up (time zero). Depletion of susceptibles occurs when individuals who experience outcomes during treatment prior to the start of follow-up are systematically excluded. Figure adapted from Fu et al., 2021, licensed under CC-BY-NC 4.0.[21]

Immortal time

Immortal time refers to a period of time included in the analysis where individuals cannot experience the outcome of interest due to choices about the study design.[12,26,40] In a

randomised trial, all outcomes that occur after eligibility ascertainment, treatment strategy assignment, and the start of follow-up are attributed to the appropriate treatment strategy. However, in observational analyses, when information that is ascertained after the start of follow up is used to inform eligibility (or classification into treatment strategies), immortal time may be introduced. Below are examples of specific design decisions that can introduce immortal time.

Using information emerging after treatment assignment to define eligibility (selection bias)

Using information that emerges after treatment assignment (e.g., adherence to treatment strategy, duration of treatment use, providing follow-up data) to define the eligibility criteria induces selection bias when the analysis starts the follow-up at treatment assignment. In a study investigating the effect of unilateral total knee replacement compared to staged bilateral total knee replacement for the outcome of functional impairment, investigators only included participants who provided follow-up data at 12 months post-surgery.[27] This decision will result in a biased treatment effect estimate if response rates differ based on the intervention and if there are common causes of response and the outcome of interest (e.g., persistent pain). In general, if information measured after treatment strategy assignment and the start of follow-up is used to determine eligibility, the treatment differentially influences those eligibility criteria, and there are common causes of the eligibility criteria and the outcome, then the treatment effect can be biased (Figure 2).[20] In a well-designed and

executed randomised trial, the time of eligibility ascertainment is anchored to the time of treatment strategy assignment (randomisation) and the start of follow-up, avoiding this issue

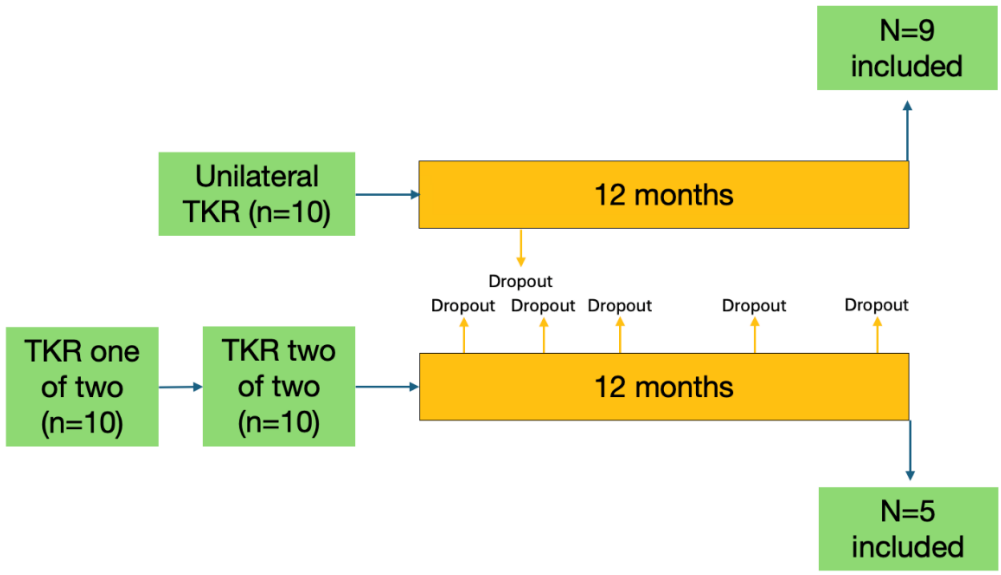

Figure 2: Immortal time due to selection bias. This can arise when eligibility criteria (e.g., 12-month follow-up response) are applied after treatment assignment (e.g., unilateral vs. bilateral total knee replacement [TKR]) and the start of follow-up, thereby creating a period of immortal time between baseline and eligibility ascertainment where the outcome cannot occur for included individuals. Whilst not the focus, this example also includes bias due to treatment misclassification (see errors in classifying individuals into treatment strategies).

Errors in classifying individuals into treatment strategies (misclassification bias)

Immortal time can also arise when individuals are classified into a treatment strategy that differs from the one they should have been assigned to. This often occurs when treatment

strategies under comparison cannot be distinguished at time zero. In a study that compared surgery to no surgery for lung cancer in the elderly, it was necessary to have a grace period of 6 months to allow individuals time from their diagnosis to receive surgery, as surgery often does not occur immediately after diagnosis.[28] Using only information at cancer diagnosis, it is unclear to which treatment strategy individuals should be classified as their data at baseline may be compatible with both strategies (surgery within 6 months or no surgery). If investigators 'look forward' and classify individuals who undergo surgery within 6 months to the surgery strategy, all individuals in that strategy would have survived until their surgery. Therefore, in this scenario, events (e.g., death) that occur before surgery will be attributed to the no surgery strategy. This may result in up to 6 months of immortal time for participants classified to the surgery strategy (Figure 3B).[28] In a randomised trial, participants would be allocated prospectively to either strategy and events are attributed to the strategy assigned at baseline, avoiding this bias. To avoid introducing immortal time, analysts of observational data can use approaches which overcome the lack of knowledge on what treatment strategy individuals were intended to be allocated, without using information that is measured after baseline.

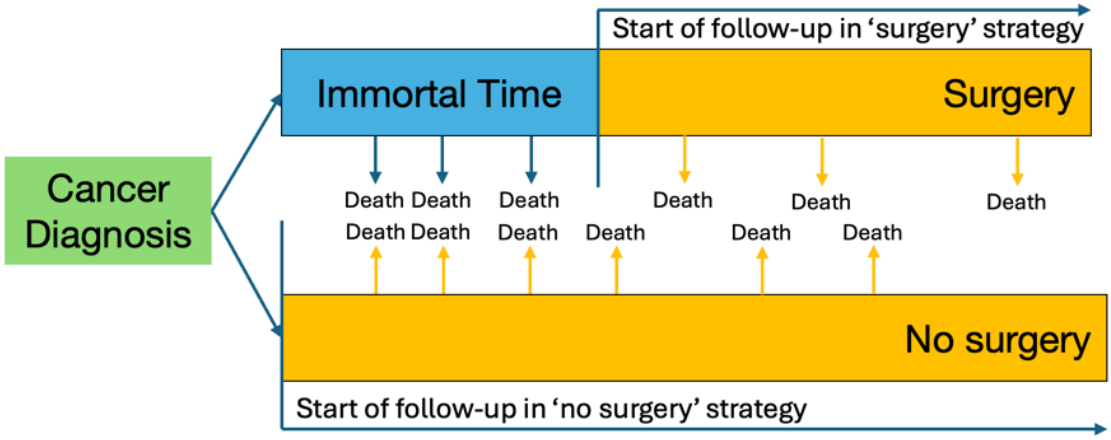

Figure 3: Immortal time due to misclassification of treatment assignment. This can arise when treatment strategies (e.g., surgery within 6 months or no surgery) are assigned using information after the start of follow up (e.g., cancer diagnosis), which can result in events being misattributed to one of the treatment groups (e.g., above, early deaths marked by blue arrows are misattributed to the no surgery strategy).

Avoiding design-related biases with target trial emulation

Carefully mapping the analysis of an observational study to the framework of a randomised trial that would answer the question of interest, known as target trial emulation, can help investigators avoid design-related biases.[22,41,42] To emulate a target trial there are two steps. The first step is to design the hypothetical pragmatic trial – the target trial – that a) would answer the question of interest and b) uses only variables that are available in the observational data source (i.e., the target trial is able to be emulated with the available data).[22] The target trial can be considered a detailed description of the causal question (or

estimand) that is being targeted.[22] The critical components to be considered when specifying the target trial are: eligibility criteria, treatment strategies, treatment assignment, outcomes, follow-up, and causal contrasts of interest (e.g., intention to treat effect), as well as any identifying assumptions and the data analysis plan.[14,22] After the target trial is specified, the second step is to map the components of the target trial to the observational data (i.e., emulate the target trial) and apply appropriate analytic techniques to adjust for confounding.[25,43,44] There may be instances where data on all variables needed to emulate the target trial are not available; however, still specifying the target trial can help readers understand these limitations.[45]

There are several approaches under the target trial framework to help investigators avoid inducing design-biases common in observational studies, such as to avoid inclusion of individuals already taking treatment at baseline and to avoid the use of information that is available after baseline to assign individuals to treatment strategies or to define eligibility.

Avoiding prevalent user bias

Prevalent user bias can be avoided by applying eligibility criteria in the target trial that restricts entry into the hypothetical trial to individuals who have not used the treatment for a defined period. This same criterion must be applied in the analysis of the observational data (emulation) and is often implemented using a 'look-back' period where individuals cannot have used the treatment for some defined period prior to baseline (time-zero). Only

including individuals who are just starting treatment has also been termed a new- or incident-user design in pharmacoepidemiology.[46] There are also occasions where prevalent users are of interest, and designs are emerging to handle the depletion of susceptibles appropriately;[47,48] however, such designs can be challenging to apply and interpret.[49]

An example of prevalent-user bias occurred in observational studies investigating hormone replacement therapy for the prevention of coronary heart disease.[50,51] Grodstein et al. (1996) compared women using hormone therapy (prevalent users) to women who did not use hormone therapy. The observational study suggested a reduced risk of coronary heart disease among users of hormone therapy compared to non-users,[51] a result that conflicted with a later randomised trial which showed an increased risk of coronary heart disease for those assigned to initiate hormone therapy compared to placebo.[52] The same data in Grodstein et al. (1996) was used by Hernán et al., (2008) although these investigators adhered more closely to the principles of the randomised trial by comparing initiators to non-initiators and found a similarly increased risk of coronary heart disease to the trial.[50] In the original observational study,[51] the inclusion of prevalent users introduced bias as individuals who were susceptible to adverse events of hormone therapy or who experienced coronary heart disease were less likely to be included in the study (depletion of susceptibles), resulting in an underestimation of the true risk of hormone therapy. The comparison between initiators and non-initiators of treatment better aligns with common clinical

decisions about whether to start a new treatment or not, and how a randomised trial would be conducted.[50]

Avoiding immortal time

Immortal time is introduced by selection of individuals based on eligibility criteria defined after treatment strategy assignment or misclassification of individuals into treatment strategies based on information available after eligibility ascertainment. These biases do not exist in well-designed randomised trials where there are negligible gaps between eligibility ascertainment, treatment strategy assignment, and the start of follow-up. Specifying a target trial that similarly aligns these features will avoid the introduction of immortal time.[12,34]

Immortal time can have substantial impacts on study results. To illustrate this, Kuehne et al., systematically applied different approaches to estimate the effect of LOT2 (a type of chemotherapy for ovarian cancer), including designs subject to immortal time, to highlight how the estimates change with design decisions.[17] With the 'naïve' approach subject to immortal time, they estimated a hazard ratio of 0.56 (95% CI 0.49 to 0.64). By contrast, when they emulated the design of an index trial, they estimated a hazard ratio of 1.12 (95% CI 0.96 to 1.28), which was compatible with the results of the trial.[17]

This alignment is straightforward when estimating the effect of receiving an intervention that occurs at a single timepoint such as a surgery or single-dose vaccine. However, this

alignment is more complex when estimating the effect of treatment strategies that are indistinguishable at baseline, such as when the effect of the timing of treatment is of interest (e.g., start treatment within 6 months of diagnosis or after 6 months). Several strategies have been developed to handle these situations.[30]

Clone-censor-weight approach

An approach to handle questions where treatment strategies are indistinguishable at time-zero, in conjunction with the target trial framework, is the clone-censor-weight approach.[30,32] To implement this, investigators can clone participants, that is, replicate individuals, and assign replicates to as many treatment strategies as their data are compatible with at baseline. As a person's observed treatment pattern becomes incompatible with their assigned strategy, their follow up in that treatment strategy is artificially stopped or 'censored' by the investigator at the time of deviation from their assigned strategy. This artificial censoring ensures individuals classified into each strategy are only included in follow-up whilst they are adherent to their assigned treatment strategy. However, if adherence is differential (non-random) between the treatment strategies, the artificial censoring may induce selection bias,[53] requiring a third step: weighting. Inverse probability of censoring weights are estimated and applied to eliminate the selection bias induced by the artificial censoring.[30,32] For interested readers, Hernán (2018) and Gaber et al., (2024) provide in depth explanations of how to conduct a clone-censor-weight analysis.[30,32]

An example of where the clone-censor-weight approach may be helpful is when comparing treatment strategies that involve a grace period for initiating treatment (that is, a predefined time after baseline when treatment can begin and the individual would be considered adherent to the treatment strategy throughout this period, reflecting delays common in clinical practice) (**Figure 4**).[32] Consider the example of comparing receipt of surgery within 6 months of baseline (the grace period) with no surgery over follow-up, among individuals with lung cancer (**Figure 3**). Since it is unclear which treatment strategy individuals should be assigned to at baseline (cancer diagnosis), all individuals who have baseline data compatible with both treatment strategies (e.g., have not had surgery on the day of cancer diagnosis) are replicated. Then, as replicates classified at baseline to the no surgery strategy have surgery over the follow-up, they are censored at that time as their data are no longer compatible with their strategy assigned at baseline. Similarly, for replicates classified into the surgery strategy who do not have surgery by the end of the grace period, they are censored at 6 months. Inverse probability of censoring weights are applied throughout the follow-up period to account for the censoring.

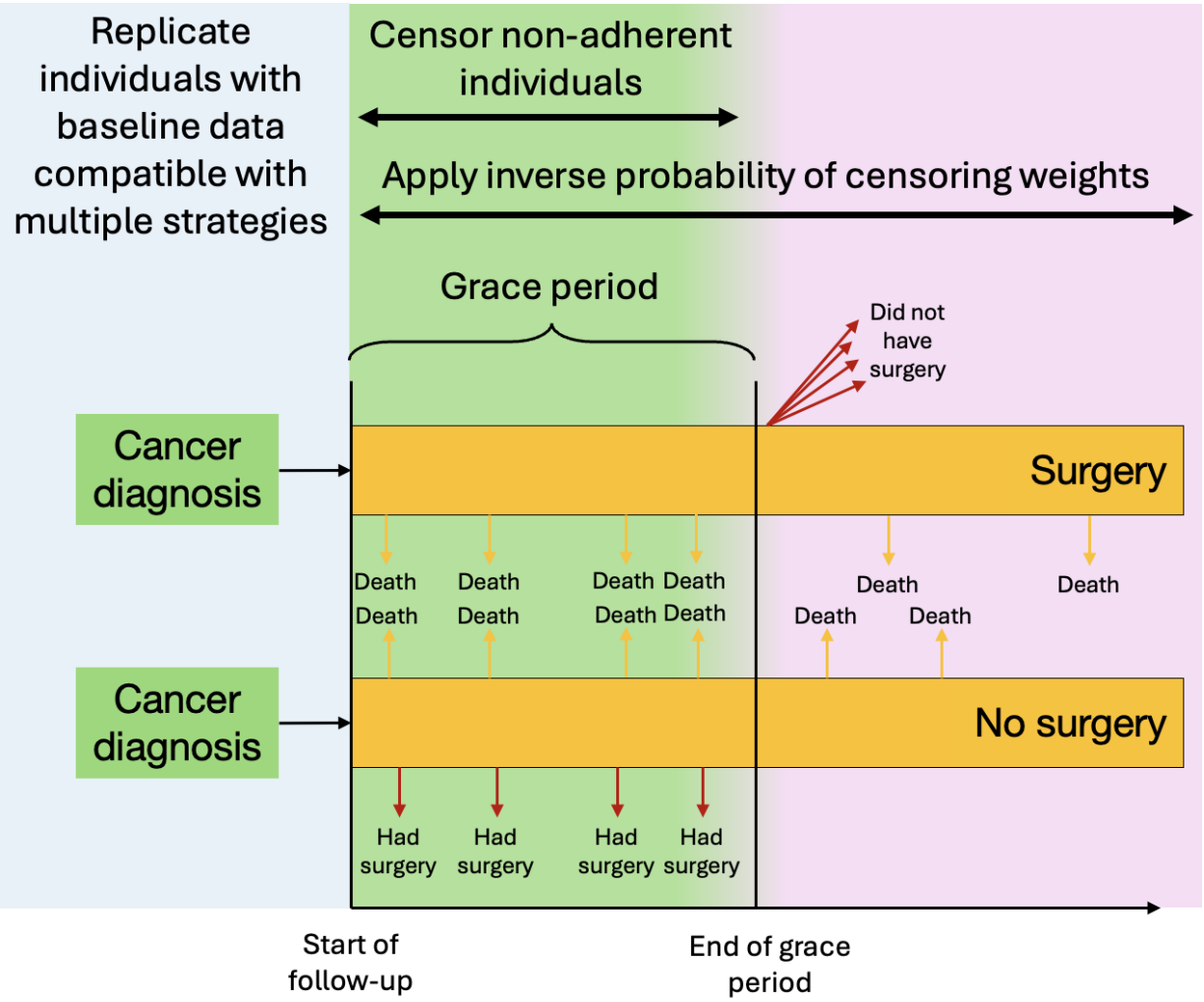

Figure 4: Illustration of clone-censor-weight approach. Red arrows indicate deviation from treatment strategies leading to censoring (e.g., when individuals classified to the no surgery strategy receive surgery, or when individuals classified to the surgery within a certain grace period strategy do not receive surgery by the end of the grace period).

Applying target trial emulation with cloning, censoring and weighting prevented immortal time in a recent example by Boyne et al. investigating survival following short or longer duration of adjuvant chemotherapy for colon cancer.[54] Prior to this study, findings from a

randomised trial (the IDEA trial) [55] conflicted with those from observational studies. In these previous observational studies, information emerging after baseline (achieved duration of treatment) was used to classify individuals into treatment strategies, leading to immortal time. Boyne et al. 2021, demonstrated that with the 'naïve' approach taken in previous observational studies, estimates suggested that shortener durations of treatment substantially worsened survival (hazard ratio = 3.33 [95% CI 1.04 to 10.65]).[54] However, when information emerging after baseline was not used to classify individuals into treatment strategies, and instead cloning, censoring and weighting was applied, the estimates were compatible with those from the IDEA trial (emulation hazard ratio = 0.96 [95% CI 0.43 to 2.14], trial hazard ratio = 0.96 [95% CI 0.85 to 1.08]).[54]

Emulating sequential trials

Another approach to avoid immortal time where treatment strategies are not distinguishable at baseline is reformulating the causal question to make the strategies distinguishable at baseline.[31] Consider the example outlined previously, comparing receipt of surgery within 6 months of baseline with no surgery over follow-up. Investigators could reformulate the first strategy as "receipt of surgery at baseline," that is, without a grace period. Therefore, individuals who initiate surgery at baseline will be classified into the surgery strategy, and into the no surgery strategy otherwise. This avoids the introduction of immortal time due to treatment strategy misclassification; however, whether the new causal question (e.g., without

grace periods, which may be realistic in many real-world treatment strategies) is still of interest should be considered.

In principle, using only a single eligible time (e.g., the first, or a random eligible time) would be an unbiased approach, as it relies only on baseline information to classify individuals into a treatment strategy.[12,23] However, a large amount of information could be lost, leading to less precise estimates.[12] A more efficient approach could be to instead emulate a sequence of nested trials, each initiated at a certain calendar period (e.g., each day, week, or month), which becomes time zero of a given nested trial.[31] For example, eligible individuals who initiate surgery in the first week of each nested trial could be compared to eligible individuals who do not have surgery in that week (Figure 5). Information across these trials can then be pooled to provide a single estimate of the treatment effect.

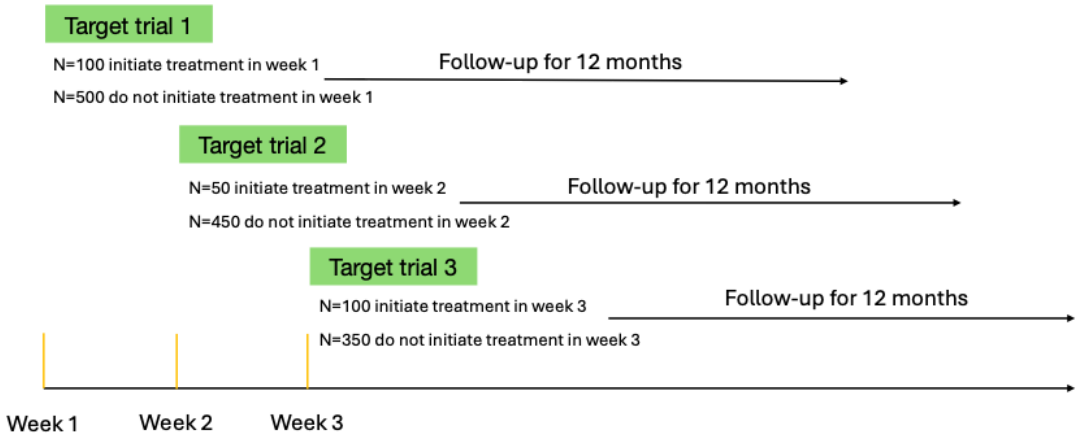

Figure 5: Emulation of sequential trials. Each target trial 'recruits' each week (or other relevant time-period, where follow-up begins [time zero]), where investigators classify

individuals into the treatment strategy their data is compatible with during that week, then follows individuals for 12 months.

Applying target trial emulation with sequential trial emulation prevented design-related biases in an evaluation of statins and risk of cancer.[56] In earlier observational analyses, investigators reported a substantially lower risk of cancer among statin users compared with non-users (odds ratio for lung cancer of 0.23 (95% CI 0.20 to 0.26), comparing long-term (>4 years) statin users with non-users).[57] These analyses had two flaws previously discussed:[29,50] post-baseline information on achieved duration of statin use was used to classify individuals into treatment strategies, generating immortal time, and prevalent users of treatment were included, leading to selection bias. Dickerman et al. reframed this analysis as a target trial with sequential trial emulation and classified individuals into one of two strategies (statin initiators vs. non-initiators) based on information available at the baseline of each sequential target trial, finding a null effect (hazard ratio = 1.02 [95% CI 0.99 to 1.05]), consistent with meta-analyses of randomized trials.[58,59]

Discussion

With the increasing use of observational studies of large healthcare databases for health decision-making, it is crucial that clinicians, policymakers and researchers are able to identify common design-related biases. Explicit emulation of a target trial supports readers' appraisal of observational studies investigating the comparative effectiveness of interventions

[42] and helps investigators overcome design-related biases. Cloning, censoring and weighting and sequential trial emulation are two approaches that help investigators avoid design-related biases when treatment assignment is unclear at baseline. The parametric g-formula is another approach available to analysts, however, is not described in this article with further reading available elsewhere.[60,61] Appropriate variance estimation is needed to handle the potential repeated inclusion of some individuals for many approaches used in target trial emulation.[12,50,62] Despite the potential benefits of these approaches, they require rich, longitudinal data and may become computationally intensive in large datasets.[62]

Target trial emulation facilitates more appropriate comparisons between observational studies and trials,[63] improving our understanding of where differences may occur.[16,64] Target trial emulation can also support the education of clinicians and epidemiologists in observational research design by communicating the study design in the form of a hypothetical trial.[65] However, it is important to note that target trial emulation does not eliminate the risk of bias,[66] and does not reduce the likelihood of biases within the data themselves, such as confounding bias or measurement error.[44] Such biases must be addressed through careful measurement and adjustment for key variables via appropriate analytic techniques, in combination with unverifiable assumptions.[67] Confounding by indication may be more of a concern when comparing an active intervention (e.g., a drug) to no intervention; the use of an active comparator (e.g., another drug) may reduce the

magnitude of confounding by indication.[68] Similarly, as with all analyses of observational data aiming to identify causal effects of interventions, time-varying confounding should be addressed using appropriate methods (such as inverse probability weighting[69]) when necessary.

Transparent reporting of target trial emulations is essential to allow readers to appraise, interpret and implement findings in their decision-making. The TARGET guideline [70–72] aims to assist investigators to report relevant information, and support readers to identify key aspects of the target trial emulation. TARGET was not designed to be, and should not be, used as a tool to assess risk of bias or quality of a study. However, although not the intention, the TARGET guideline may help improve study conduct by serving as an educational tool and clarifying key methodologic issues to be addressed. Study design diagrams, such as those described by Schneeweiss et al., (2019), can also assist investigators to communicate the timing of different aspects of the study design.[73]

Conclusion

In settings where trials are unavailable or infeasible, observational studies may provide evidence to address the relevant clinical question. Design-related biases are common in observational studies of comparative effectiveness, but explicit emulation of a target trial can help investigators avoid these biases and enable readers to identify them when present. The

use of target trial emulation may support sound policy and clinical decisions from observational studies.

Contributorship statement

HJH, NI and AGC conceived the project. HJH wrote the first draft. All authors provided extensive feedback on the manuscript. All authors contributed to the intellectual content of the manuscript and reviewed the final version.

Funding

AGC was supported by an Australian NHMRC Investigator Grant (ID 2010088). HJH was supported by an Australian National Health and Medical Research Council (NHMRC) Postgraduate Scholarship (ID 2021950) and a Neuroscience Research Australia PhD Pearl sponsored by Sandra Salteri. BAD was supported by National Cancer Institute grant R00 CA248335.

Data sharing statement

No data were generated for this project.

Patient and Public Involvement

No patients or members of the public were involved in the writing of this article.

Declaration of Interest Statement

References

1 Cochrane AL. Effectiveness and efficiency: random reflections on health services. 1972.

2 Concato J, Shah N, Horwitz RI. Randomized, Controlled Trials, Observational Studies, and the Hierarchy of Research Designs. *New England Journal of Medicine*. 2000;342:1887–92. doi: 10.1056/NEJM200006223422507

3 Matthews AA, Dahabreh IJ, Fröbert O, *et al.* Benchmarking Observational Analyses Before Using Them to Address Questions Trials Do Not Answer: An Application to Coronary Thrombus Aspiration. *American Journal of Epidemiology*. 2022;191:1652–65. doi: 10.1093/aje/kwac098

4 Moler-Zapata S, Hutchings A, O'Neill S, *et al.* Emulating Target Trials With Real-World Data to Inform Health Technology Assessment: Findings and Lessons From an Application to Emergency Surgery. *Value in Health*. 2023;26:1164–74. doi: 10.1016/j.jval.2023.04.010

5 Hulme WJ, Williamson EJ, Green ACA, *et al.* Comparative effectiveness of ChAdOx1 versus BNT162b2 covid-19 vaccines in health and social care workers in England: cohort study using OpenSAFELY. *The BMJ*. 2022;(no pagination). doi: 10.1136/bmj-2021-068946

6 Burns L, Roux NL, Kalesnik-Orszulak R, *et al.* Real-World Evidence for Regulatory Decision-Making: Guidance From Around the World. *Clinical Therapeutics*. 2022;44:420–37. doi: 10.1016/j.clinthera.2022.01.012

7 Hansford HJ, Cashin AG, Jones MD, *et al.* Reporting of Observational Studies Explicitly Aiming to Emulate Randomized Trials: A Systematic Review. *JAMA Network Open*. 2023;6:e2336023. doi: 10.1001/jamanetworkopen.2023.36023

8 Benchimol EI, Smeeth L, Guttmann A, *et al.* The REporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. *PLoS medicine*. 2015;12:e1001885.

9 European Medicines Agency. Real-world evidence framework to support EU regulatory decision-making. 2024.

10 Collins R, Bowman L, Landray M, *et al.* The Magic of Randomization versus the Myth of Real-World Evidence | New England Journal of Medicine. *New England Journal of Medicine*. 2020;674–8. doi: 10.1056/NEJMsb1901642

11 Califf RM, Hernandez AF, Landray M. Weighing the Benefits and Risks of Proliferating Observational Treatment Assessments: Observational Cacophony, Randomized Harmony. *JAMA*. 2020;324:625–6. doi: 10.1001/jama.2020.13319

12 Hernán MA, Sauer BC, Hernández-Díaz S, *et al.* Specifying a target trial prevents immortal time bias and other self-inflicted injuries in observational analyses. *J Clin Epidemiol*. 2016;79:70–5. doi: 10.1016/j.jclinepi.2016.04.014

13 Dickerman BA, García-Albéniz X, Logan RW, *et al.* Avoidable flaws in observational analyses: an application to statins and cancer. *Nat Med*. 2019;25:1601–6. doi: 10.1038/s41591-019-0597-x

14 Hernán MA. Methods of Public Health Research — Strengthening Causal Inference from Observational Data. *New England Journal of Medicine*. 2021;385:1345–8. doi: 10.1056/NEJMp2113319

15 Wang SV, Schneeweiss S, Initiative R-D. Emulation of Randomized Clinical Trials With Nonrandomized Database Analyses: Results of 32 Clinical Trials. *JAMA*. 2023;329:1376–85. doi: 10.1001/jama.2023.4221

16 Heyard R, Held L, Schneeweiss S, *et al.* Design differences and variation in results between randomised trials and non-randomised emulations: meta-analysis of RCT-DUPLICATE data. *BMJ Medicine*. 2024;3. doi: 10.1136/bmjmed-2023-000709

17 Kuehne F, Arvandi M, Hess LM, *et al.* Causal analyses with target trial emulation for real-world evidence removed large self-inflicted biases: systematic bias assessment of ovarian cancer treatment effectiveness. *Journal of Clinical Epidemiology*. 2022;152:269–80. doi: 10.1016/j.jclinepi.2022.10.005

18 Hernán MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *JAMA*. 2022;328:2446–7. doi: 10.1001/jama.2022.21383

19 McIntyre M, Yaacoub S, Perrodeau E, *et al.* Review of methods to deal with the misalignment of times of eligibility, start of follow up and treatment assignment in studies explicitly aimed at emulating target trials. *Journal of Clinical Epidemiology*. 2025;111898. doi: 10.1016/j.jclinepi.2025.111898

20 Yaacoub S, Porcher R, Pellat A, *et al.* Characteristics of non-randomised studies of drug treatments: cross sectional study. *BMJ Med*. 2024;3. doi: 10.1136/bmjmed-2024-000932

21 Fu EL, van Diepen M, Xu Y, *et al.* Pharmacoepidemiology for nephrologists (part 2): potential biases and how to overcome them. *Clinical Kidney Journal*. 2021;14:1317–26. doi: 10.1093/ckj/sfaa242

22 Hernán MA, Dahabreh IJ, Dickerman BA, *et al.* The Target Trial Framework for Causal Inference From Observational Data: Why and When Is It Helpful? *Ann Intern Med*. Published Online First: 18 February 2025. doi: 10.7326/ANNALS-24-01871

23 Hernán MA, Sterne JAC, Higgins JPT, *et al.* A structural description of biases that generate immortal time. *Epidemiology*. 2024.

24 Fu EL. Target Trial Emulation to Improve Causal Inference from Observational Data: What, Why, and How? *Journal of the American Society of Nephrology*. 2023;34:1305. doi: 10.1681/ASN.0000000000000152

25 Desai RJ, Wang SV, Sreedhara SK, *et al.* Process guide for inferential studies using healthcare data from routine clinical practice to evaluate causal effects of drugs (PRINCIPLED): considerations from the FDA Sentinel Innovation Center. *BMJ*. 2024;384:e076460. doi: 10.1136/bmj-2023-076460

26 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. doi: 10.1136/bmj.b5087

27 Larson DR, Crowson CS, Devick KL, *et al.* Immortal Time Bias in the Analysis of Time to Event Data in Orthopaedics. *J Arthroplasty*. 2021;36:3372–7. doi: 10.1016/j.arth.2021.06.012

28 Maringe C, Benitez Majano S, Exarchakou A, *et al.* Reflection on modern methods: trial emulation in the presence of immortal-time bias. Assessing the benefit of major surgery for elderly lung cancer patients using observational data. *International Journal of Epidemiology*. 2020;49:1719–29. doi: 10.1093/ije/dyaa057

29 Suissa S. Immortal time bias in pharmaco-epidemiology. *Am J Epidemiol*. 2008;167:492–9. doi: 10.1093/aje/kwm324

30 Hernán MA. How to estimate the effect of treatment duration on survival outcomes using observational data. *BMJ*. 2018;360:k182. doi: 10.1136/bmj.k182

31 García-Albéniz X, Hsu J, Hernán MA. The value of explicitly emulating a target trial when using real world evidence: an application to colorectal cancer screening. *Eur J Epidemiol*. 2017;32:495–500. doi: 10.1007/s10654-017-0287-2

32 Gaber CE, Ghazarian AA, Strassle PD, *et al.* De-Mystifying the Clone-Censor-Weight Method for Causal Research Using Observational Data: A Primer for Cancer Researchers. *Cancer Med*. 2024;13:e70461. doi: 10.1002/cam4.70461

33 Gaber CE, Hanson KA, Kim S, *et al.* The Clone-Censor-Weight Method in Pharmacoepidemiologic Research: Foundations and Methodological Implementation. *Curr Epidemiol Rep*. 2024;11:164–74. doi: 10.1007/s40471-024-00346-2

34 Caniglia EC, Zash R, Fennell C, *et al.* Emulating Target Trials to Avoid Immortal Time Bias - An Application to Antibiotic Initiation and Preterm Delivery. *Epidemiology*. 2023;34:430–8. doi: 10.1097/EDE.0000000000001601

35 Cribb L, Moreno-Betancur M, Wu Z, *et al.* Moving beyond the prevalent exposure design for causal inference in dementia research. *The Lancet Healthy Longevity*. 2025;6. doi: 10.1016/j.lanhl.2024.100675

36 Anes AM, Arana A, Blake K, *et al. The European network of centres for pharmacoepidemiology and pharmacovigilance (ENCePP). Guide on methodological standards in pharmacoepidemiology*. 2023.

37 Sterne JA, Hernán MA, McAleenan A, *et al.* Assessing risk of bias in a non-randomized study. *Cochrane handbook for systematic reviews of interventions*. 2019;621–41.

38 Desai RJ, Franklin JM. Alternative approaches for confounding adjustment in observational studies using weighting based on the propensity score: a primer for practitioners. *BMJ*. 2019;367:l5657. doi: 10.1136/bmj.l5657

39 Sackett DL. Bias in analytic research. *J Chronic Dis*. 1979;32:51–63. doi: 10.1016/0021-9681(79)90012-2

40 Lee H, Nunan D. Immortal time bias. Catalog of Bias. 2020. https://catalogofbias.org/biases/immortal-time-bias/ (accessed 13 November 2024)

41 Hernán MA. With great data comes great responsibility:Publishing comparative effectiveness research in EPIDEMIOLOGY. *Epidemiology (Cambridge, Mass)*. 2011;22:290. doi: 10.1097/EDE.0b013e3182114039

42 Sterne JA, Hernán MA, Reeves BC, *et al.* ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. *BMJ*. 2016;355:i4919. doi: 10.1136/bmj.i4919

43 Matthews AA, Danaei G, Islam N, *et al.* Target trial emulation: applying principles of randomised trials to observational studies. *BMJ*. 2022;378:e071108. doi: 10.1136/bmj-2022-071108

44 Hernán MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. *American Journal of Epidemiology*. 2016;183:758–64. doi: 10.1093/aje/kwv254

45 Swanson SA. The Causal Effects of Causal Inference Pedagogy. *Epidemiology*. 2023;34:611–3. doi: 10.1097/EDE.0000000000001639

46 Ray WA. Evaluating medication effects outside of clinical trials: new-user designs. *Am J Epidemiol*. 2003;158:915–20. doi: 10.1093/aje/kwg231

47 Suissa S, Dell'Aniello S, Renoux C. The Prevalent New-user Design for Studies With no Active Comparator: The Example of Statins and Cancer. *Epidemiology*. 2023;34:681–9. doi: 10.1097/EDE.0000000000001628

48 Suissa S, Moodie EEM, Dell'Aniello S. Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. *Pharmacoepidemiology and Drug Safety*. 2017;26:459–68. doi: 10.1002/pds.4107

49 Luijken K, Spekreijse JJ, van Smeden M, *et al.* New-user and prevalent-user designs and the definition of study time origin in pharmacoepidemiology: A review of reporting practices. *Pharmacoepidemiol Drug Saf*. 2021;30:960–74. doi: 10.1002/pds.5258

50 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. doi: 10.1097/EDE.0b013e3181875e61

51 Grodstein F, Stampfer MJ, Manson JE, *et al.* Postmenopausal Estrogen and Progestin Use and the Risk of Cardiovascular Disease. *New England Journal of Medicine*. 1996;335:453–61. doi: 10.1056/NEJM199608153350701

52 Manson JE, Hsia J, Johnson KC, *et al.* Estrogen plus Progestin and the Risk of Coronary Heart Disease. *New England Journal of Medicine*. 2003;349:523–34. doi: 10.1056/NEJMoa030808

53 Hernán MA, Robins JM. Per-Protocol Analyses of Pragmatic Trials. *New England Journal of Medicine*. 2017;377:1391–8. doi: 10.1056/NEJMsm1605385

54 Boyne DJ: C WY: Hilsden, RJ: Sajobi, TT: Batra, A: Friedenreich, CM: Brenner, DR. Association of a Shortened Duration of Adjuvant Chemotherapy with Overall Survival among Individuals with Stage III Colon Cancer. *JAMA Network Open*. Published Online First: 30 March 2021. doi: 10.1001/jamanetworkopen.2021.3587

55 Grothey A, Sobrero AF, Shields AF, *et al.* Duration of Adjuvant Chemotherapy for Stage III Colon Cancer. *New England Journal of Medicine*. 2018;378:1177–88. doi: 10.1056/NEJMoa1713709

56 Dickerman BA: G-A Xabier: Logan, Roger: Denaxas, Spiros: Hernán, Miguel A. Avoidable flaws in observational analyses: an application to statins and cancer. *Nat Med*. Published Online First: 2019. doi: 10.1038/s41591-019-0597-x

57 Khurana V, Bejjanki HR, Caldito G, *et al.* Statins reduce the risk of lung cancer in humans: a large case-control study of US veterans. *Chest*. 2007;131:1282–8. doi: 10.1378/chest.06-0931

58 Dale KM, Coleman CI, Henyan NN, *et al.* Statins and cancer risk: a meta-analysis. *JAMA*. 2006;295:74–80. doi: 10.1001/jama.295.1.74

59 Cholesterol Treatment Trialists' (CTT) Collaboration, Emberson JR, Kearney PM, *et al.* Lack of effect of lowering LDL cholesterol on cancer: meta-analysis of individual data from 175,000 people in 27 randomised trials of statin therapy. *PLoS One*. 2012;7:e29849. doi: 10.1371/journal.pone.0029849

60 Wen L, Young JG, Robins JM, *et al.* Parametric g-formula implementations for causal survival analyses. *Biometrics*. 2021;77:740–53. doi: 10.1111/biom.13321

61 Wanis KN, Sarvet AL, Wen L, *et al.* Grace periods in comparative effectiveness studies of sustained treatments. *J R Stat Soc Ser A Stat Soc*. 2024;187:796–810. doi: 10.1093/jrsssa/qnae002

62 Danaei G, Rodríguez LAG, Cantero OF, *et al.* Observational data for comparative effectiveness research: an emulation of randomised trials to estimate the effect of statins on primary prevention of coronary heart disease. *Statistical methods in medical research*. 2011;22:70. doi: 10.1177/0962280211403603

63 Lodi S, Phillips A, Lundgren J, *et al.* Effect Estimates in Randomized Trials and Observational Studies: Comparing Apples With Apples. *Am J Epidemiol*. 2019;188:1569–77. doi: 10.1093/aje/kwz100

64 Forbes SP, Dahabreh IJ. Benchmarking Observational Analyses Against Randomized Trials: a Review of Studies Assessing Propensity Score Methods. *J Gen Intern Med*. 2020;35:1396–404. doi: 10.1007/s11606-020-05713-5

65 Labrecque JA, Swanson SA. Target trial emulation: teaching epidemiology and beyond. *Eur J Epidemiol*. 2017;32:473–5. doi: 10.1007/s10654-017-0293-4

66 Simon-Tillaux N, Martin GL, Hajage D, *et al.* Conducting observational analyses with the target trial emulation approach: a methodological systematic review. *BMJ Open*. 2024;14:e086595. doi: 10.1136/bmjopen-2024-086595

67 Hernán MA, Robins JM. *Causal inference: What if?* CRC Boca Raton, FL 2010.

68 Lund JL, Richardson DB, Stürmer T. The active comparator, new user study design in pharmacoepidemiology: historical foundations and contemporary application. *Curr Epidemiol Rep*. 2015;2:221–8. doi: 10.1007/s40471-015-0053-5

69 Mansournia MA, Etminan M, Danaei G, *et al.* Handling time varying confounding in observational research. *BMJ*. 2017;359:j4587. doi: 10.1136/bmj.j4587

70 Hansford HJ, Cashin AG, Jones MD, *et al.* Development of the TrAnsparent ReportinG of observational studies Emulating a Target trial (TARGET) guideline. *BMJ Open*. 2023;13:e074626. doi: 10.1136/bmjopen-2023-074626

71 Cashin AG, Hansford HJ, Hernán MA, *et al.* Transparent reporting of observational studies emulating a target trial: the TARGET Statement. *BMJ*. 2025;390:e087179. doi: 10.1136/bmj-2025-087179

72 Cashin AG, Hansford HJ, Hernán MA, *et al.* Transparent Reporting of Observational Studies Emulating a Target Trial—The TARGET Statement. *JAMA*. Published Online First: 3 September 2025. doi: 10.1001/jama.2025.13350

73 Schneeweiss S, Rassen JA, Brown JS, *et al.* Graphical Depiction of Longitudinal Study Designs in Health Care Databases. *Ann Intern Med*. 2019;170:398–406. doi: 10.7326/M18-3079