Canadian Journal of Cardiology 41 (2025) 996–1009

Review

Breaking Down Bias: A Methodological Primer on Identifying, Evaluating, and Mitigating Bias in Cardiovascular Research

Nicholas Grubic, MSc,[a] Amy Johnston, PhD,[b] Varinder K. Randhawa, MD, PhD,[c,d] Karin H. Humphries, DSc,[e] Laura C. Rosella, PhD,[a,f,g] and Katerina Maximova, PhD[a,h]

[a] *Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada*

[b] *Department of Obstetrics and Gynecology, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada*

[c] *Department of Critical Care Medicine, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada*

[d] *Department of Medicine, Temerty Faculty of Medicine, University of Toronto, Toronto, Canada*

[e] *Faculty of Medicine, University of British Columbia, Vancouver, British Columbia, Canada*

[f] *Department of Laboratory Medicine and Pathobiology, Temerty Faculty of Medicine, Toronto, Ontario, Canada*

[g] *Institute for Better Health, Trillium Health Partners, Mississauga, Ontario, Canada*

[h] *MAP Centre for Urban Health Solutions, St Michael's Hospital, Toronto, Ontario, Canada*

ABSTRACT

Systematic error, often referred to as bias is an inherent challenge in observational cardiovascular research, and has the potential to profoundly influence the design, conduct, and interpretation of study results. If not carefully considered and managed, bias can lead to spurious results, which can misinform clinical practice or public health initiatives and compromise patient outcomes. This methodological primer offers a concise introduction to identifying, evaluating, and mitigating bias in observational cardiovascular research studies that examine the causal association between an exposure (or treatment) and an outcome. Using high-profile examples from the cardiovascular literature, this review provides a theoretical overview of 3 main types of bias—selection bias, information bias, and confounding—and discusses the implications of specialized types of biases commonly encountered in longitudinal cardiovascular research studies, namely, competing risks, immortal time bias, and confounding by indication. Furthermore, strategies and tools that can be used to minimize and assess the influence of bias are highlighted, with a specific focus on using the target trial framework, directed acyclic graphs, quantitative bias analysis, and formal risk of bias assessments. This review aims to assist researchers and health care professionals in designing obser-

RÉSUMÉ

L'erreur systématique, souvent définie comme un biais, est un problème inhérent dans la recherche cardiovasculaire observationnelle et a la capacité de profondément influer sur la conception et l'exécution des études ainsi que sur l'interprétation de leurs résultats. S'il n'est pas soigneusement considéré et traité, ce biais peut entraîner des résultats erronés, qui peuvent mal orienter la pratique clinique ou les initiatives de santé publique et ainsi compromettre les issues pour les patients. Cet abécédaire méthodologique propose une brève introduction à la détection, à l'évaluation et à l'atténuation de ce biais dans les études en recherche cardiovasculaire observationnelle qui examinent le lien de causalité entre une exposition (ou un traitement) et un résultat. Utilisant des exemples notoires tirés de la littérature cardiovasculaire, cette revue propose un aperçu théorique de trois principaux types de biais — les biais de sélection, les biais de classement et les biais de confusion — et discute des implications de types de biais spécialisés fréquemment rencontrés dans des études en recherche cardiovasculaire longitudinales, soit les risques concurrents, le biais de temps immortel et le biais de confusion par indication. De plus, les stratégies et les outils qui peuvent servir à réduire au minimum et à évaluer l'influence des biais y sont mis en évidence, avec un

"The findings of this study might be subject to [insert name of bias here]…"—a routine placeholder statement made in almost all scientific articles. Although sources of bias are frequently acknowledged as a potential limitation in health research, the implications that bias might have on study results are often ignored, poorly explained, and rarely quantified, which can lead to spurious or misleading conclusions.[1] Bias can be introduced into various study designs, occur at any stage of the research process, and diminish the validity of the

Received for publication September 30, 2024. Accepted December 13, 2024.

Corresponding author: Nicholas Grubic, Dalla Lana School of Public Health, University of Toronto, 155 College Street, 6th Floor, Toronto, Ontario M5T 3M7, Canada.

E-mail: nicholas.j.grubic@gmail.com

@nickgrubic, @amydjohn, @VarinderKaurRa1, @LauraCRosella

See page 1006 for disclosure information.

https://doi.org/10.1016/j.cjca.2024.12.022

vational studies and selecting appropriate methodologies to reduce bias, ultimately enhancing the estimation of causal associations in cardiovascular research.

accent particulier sur le cadre de l'essai ciblé, les graphes acycliques orientés, l'analyse quantitative des biais et les évaluations formelles des risques de biais. Cette revue vise à aider les chercheurs et les professionnels de la santé à concevoir des études observationnelles et à choisir les bonnes méthodologies pour réduire les biais et, ultimement, améliorer l'estimation des liens de causalité dans la recherche cardiovasculaire.

results. These errors are often complex, subtle, and difficult to detect, which complicates interpretation, distorts scientific reasoning, and muddles knowledge dissemination.[2]

Cardiovascular research serves as the foundation for identifying risk factors, developing preventative measures, and improving treatments for cardiovascular disease (CVD). However, the introduction or failure to address bias in cardiovascular research studies can lead to spurious conclusions, which might result in the implementation of ineffective interventions, reduced generalizability across diverse patient populations, and misleading information that undermines clinical decision-making. To evaluate and reduce bias in cardiovascular research, it is essential for health care professionals and scientists to understand the theoretical basis of common biases and become familiar with strategies that can be used to limit their effects on study findings.

The main objectives of this methodological primer are to: (1) describe common types of bias encountered in observational epidemiologic studies within cardiovascular research; and (2) highlight strategies to minimize and detect bias in these studies, including methods for improving study design, data collection, analysis, and interpretation of results. This review is focused on bias in etiologic studies with causal goals, which primarily aim to quantify a causal association of ≥ 1 exposure (or treatment) on an outcome, expressed using a measure of association, such as the odds ratio, relative risk (RR), hazard ratio, or risk difference.[3] Interested readers should consult other resources to learn more about bias in prediction studies[4-6] and clinical trials.[7,8]

This article also uses causal language consistent with contemporary perspectives on drawing causal inferences from observational studies,[9,10] which assert that randomization alone should not be the sole criterion for determining whether causal conclusions can be drawn from a study. Indeed, observational studies are often conducted with causal aims, and authors should not be dissuaded from making this explicit in their work. However, interpreting observational findings as "causal"—though not impossible—requires researchers to make assumptions that cannot be directly tested.[9] As such, assessments about whether such an interpretation is feasible require: (1) subjective judgements that are grounded on background knowledge and research context; (2) insights informed from analyses carried out on different data sources and with different assumptions (ie, triangulation); and (3) careful consideration of the study's design, conduct, and statistical approaches (including the results of sensitivity analyses).[9,10]

Bias and Error

According to *A Dictionary of Epidemiology*, bias is defined as "an error in the conception and design of a study — or in the collection, analysis, interpretation, reporting, publication, or review of data — leading to results or conclusions that are systematically (as opposed to randomly) different from the truth."[11] It is important to distinguish systematic error, which is introduced by bias and relates to inaccuracy, from random error, which occurs by randomness or chance (due to sampling) and relates to imprecision.[12] To illustrate this distinction, consider a scenario in which a clinician is measuring a patient's systolic blood pressure (SBP) with a sphygmomanometer to obtain an accurate reading. Random error in SBP measurements can occur by chance if the patient makes slight movements or because of transient changes in the surrounding environment (eg, temperature fluctuations, background noises). Although each measurement might not exactly reflect their "true" SBP, calculating the average over a large number of repeated measurements could lead to an accurate reading.[13] However, if the number of repeated measurements is small, the average of these measurements might still differ substantially from their true SBP measure.

In comparison, systematic error in SBP measurements can occur because of a poorly calibrated device or the use of an incorrect cuff size, resulting in consistently underestimated or overestimated measurements. An important distinction between systematic error and random error is that, with systematic error, the direction of the error is often predictable.[13] For instance, using a regular-sized blood pressure cuff on an individual who requires a small cuff has been shown to underestimate SBP measurements.[14] Since each measurement would be underestimated, the average of any number of repeated measurements would deviate from the patient's true SBP.[13]

In every study, some degree of systematic and random error is inevitable. However, because systematic errors might lead to erroneous conclusions, preventing or minimizing them through robust design strategies and analytical methods is a focal point of scientific inquiry. The previous example demonstrates how bias can arise from measurement error; however, bias can also be introduced through selection, response, and other factors.

Classification of Bias

Systematic error can be classified into 3 main types: selection bias, information bias, and confounding. This section provides a brief theoretical overview of each type of bias and illustrates their potential effect on study findings if left unaddressed, using prominent examples from the cardiovascular literature. A detailed introduction to specific types of selection bias, information bias, and confounding can be found in online resources, such as the *Catalogue of Bias*,[15] as well as in other comprehensive articles.[16-18]

Selection bias

Selection bias, one of the most conceptually complex biases in epidemiology, has been the subject of considerable debate and definitional ambiguity.[19,20] This complexity stems from the fact that selection bias can occur at study entry and exit, follow various structural mechanisms, and affect different aspects of validity in observational studies.[19-21] A clear grasp of selection bias first requires familiarity with 5 key concepts including the target population, study population, and analytic sample, as well as internal and external validity.[20]

Populations, samples, and validity. A study's target population is the population to which the results of a study are intended to be generalized. The study population (or study sample) refers to the population that is included in a study. The analytic sample is the subset of the study population that is ultimately included in the analysis. For example, in a study on the effectiveness of a cardiac rehabilitation program for adults diagnosed with coronary artery disease (CAD), the target population might include all adult Canadians with CAD, the study population might consist of individuals with CAD from a specific hospital or in a given cohort study, and the analytic sample might include only those with complete data on key study variables and who were not lost to follow-up.[22]

Building on this framework, internal validity refers to how similar the estimated association generated from the analytic sample is to the true causal effect in the study population, whereas external validity refers to how similar the true causal effect in the study population is to the true causal effect in the broader target population.[20] External validity encompasses 2 main concepts: generalizability and transportability.[23] In brief, generalizability is concerned with extending causal inferences from the study population back to the target population (eg, generalizing results from a specific region to an entire country), whereas transportability is concerned with extending causal inferences to an external target population (eg, transporting results from a region in Canada to a region in Switzerland).[23,24] Before assessing a study's external validity, researchers must first be reasonably confident that their study is internally valid. The relationships between the target population, study population, and analytic sample in the context of selection bias, as well as their implications on internal and external validity, are illustrated in Figure 1.

Demystifying selection bias. Selection bias can occur when the participants in a study population are not representative of the target population.[20] It can result from errors in selecting participants either into the study (eg, recruitment) or out of the study (eg, losses to follow-up). From an empirical standpoint, selection bias is present when individuals in the target population have different probabilities of being included in the study population on the basis of their exposure and outcome status (or factors that causally affect the exposure and the outcome). This poses a threat to external validity.[21,25] A hypothetical situation in which unexposed individuals without the outcome (ie, unexposed "healthy" controls in a case-control study) have a lower probability of being selected into the study population than other individuals, thus leading to selection bias, is described in Figure 2. Selection bias in longitudinal studies can also occur when factors that influence an individual's retention in a study lead to differential attrition on the basis of their exposure and outcome status. This poses a threat to internal validity.[21,26] Selection bias in case-control studies can also arise because of the challenges of defining the study base and sampling the appropriate controls.[27,28] For this reason, care must be taken to ensure that controls are sampled from the same study base that gives rise to the cases and selected independent of their exposure status.[29]

If selection bias is not prevented or addressed statistically, it can lead to inaccurate estimates of association that differ systematically from what would be observed if the entire target population were included in the study. Although out of scope for this methodological primer, more elaborate structural definitions of the 2 primary subtypes of selection bias, namely, collider stratification bias (type 1) and generalizability bias (type 2), have been formally described using causal diagrams.[20,21,30-32]

A widely recognized example of selection bias relates to the reported association between postmenopausal hormone replacement therapy (HRT) and CAD risk. Although early observational studies suggested a cardioprotective effect of postmenopausal HRT in prevalent users compared with nonusers, several subsequent placebo-controlled randomized trials reached the opposite conclusion, which was particularly evident during the first year after initiating HRT.[33,34] These contradictory findings have been partially attributed to prevalent user selection bias. Specifically, in the observational studies conducted, most prevalent users of HRT had passed the window of increased short-term CAD risk and were in a phase of decreased incidence, thereby undersampling HRT users who would ultimately be diagnosed with CAD during follow-up (ie, box A in Fig. 2).[34] A subsequent reanalysis of observational data that compared incident HRT users with nonusers, thus mimicking a key element of previous trials (ie, HRT therapy begins at randomization or "time-zero"), showed an initial increase in risk of CAD for HRT users, although no overall benefit over the entire follow-up period—the same pattern observed in the aforementioned randomized trials.[35] This finding confirmed that selection bias had influenced the findings of earlier observational studies.[35]

The recruitment or inclusion of hospitalized patients—a common design choice in cardiovascular research studies—can also lead to selection bias. This practice has become more common because of the increasing availability of health administrative data over the past few decades. A notable example of selection bias arising from this recruitment strategy is that of Canto et al.,[36] which identified a counterintuitive protective association (ie, RR < 1) between a higher number of CVD risk factors and mortality among patients hospitalized for an acute myocardial infarction (MI). In this study, participant selection was different according to participants' exposure and outcome status, because individuals with a greater number of CVD risk factors (exposure) were more likely to die (outcome) from an incident MI before hospital admission, thereby preventing their inclusion in the analytic sample. Indeed, these participants did not survive long enough to be admitted to the hospital.[37] In a follow-up investigation, Banack et al.[38] used a methodological technique known as inverse probability of censoring weights to address this type of selection bias and showed that excluding individuals who died before hospitalization for an acute cardiac event severely attenuated the observed effect of CVD risk factors on

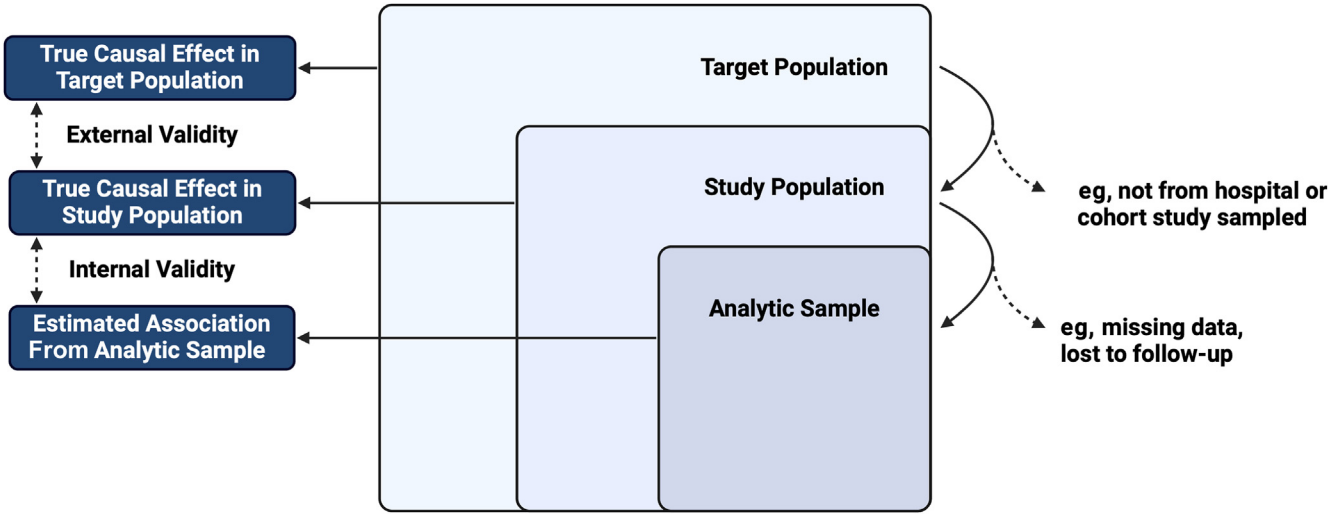

Figure 1. Relationships between the target population, study population, and analytic sample in the context of selection bias. Modified from Lu et al.[20] with permission from Wolter Kluwer Health, Inc.

mortality. Specifically, the RR of mortality for individuals with 5 risk factors compared to those with no risk factors (reference group) was 1.95 without accounting for this selection bias and 4.23 after accounting for this bias—a 117% percent change in RR![38] As illustrated by Canto et al.[36] and others, selection bias might be so severe that it reverses the direction of the effect estimate, creating a paradox in which harmful exposures mistakenly appear protective, such as seen in the "obesity paradox"[39] or "birth weight paradox."[40]

Information bias

Information bias occurs during the data collection phase of a study when an exposure or outcome is imperfectly measured or classified among individuals included in the analytic sample, thus affecting internal validity. In simple terms, information bias is a mismeasurement that affects the primary exposure-outcome association and often manifests in 2 forms on the basis of the type of variable that is mismeasured.[41,42] The inaccurate measurement of a continuous exposure or outcome is commonly referred to as measurement error.[43] In comparison, misclassification refers to the systematic tendency for study participants to be erroneously placed into different exposure or outcome categories.[25] Other types of information bias, such as ecological fallacy and regression to the mean, have been described elsewhere.[18]

Misclassification can be categorized into 2 subtypes, including nondifferential and differential misclassification. These subtypes are summarized in Figure 3, with their effect on an observed measure of association described using a simplistic scenario of a binary exposure or outcome. Nondifferential misclassification occurs when the degree of misclassification is the same across the study groups of interest, for example, when exposure status is similarly misclassified in cases and controls in a case-control study or when the accuracy of outcome ascertainment is the same between exposed and unexposed individuals in a cohort study. Nondifferential misclassification typically leads to measures of association that are biased toward the null, leading to a greater chance that real associations will be missed (ie, a type II error or false negative). In contrast, differential misclassification occurs when the degree of misclassification differs across the study groups of interest, which can lead to measures of associations that might be biased either toward or away from the null.

Importantly, nondifferential misclassification might not always lead to bias toward the null or, in other words, a reduction in the strength of a measure of association.[44] Specifically, it is important to recognize that this rule of thumb should not be inferred for polytomous variables (> 2 levels) and requires several conditions to be met (most of which cannot be empirically verified), such as exact nondifferentiality and the presence of misclassification errors that are independent of errors in other variables.[42,44-46] A detailed overview of the conditions required to assume that nondifferential misclassification produces bias toward the null has been summarized by others.[44,46]

In a case-control study, Oliveira et al.[47] evaluated how misclassification of self-reported body mass index (BMI) can affect estimates of the association between BMI (when categorized as underweight or normal weight, overweight, and obese) and the occurrence of acute MI. For example, when the authors compared self-reported and objectively measured BMI using a digital scale and wall stadiometer, on average, they reported that cases more accurately reported measures of height and weight compared with control participants, which resulted in differential misclassification of BMI category according to outcome status. As a result, the association between BMI and acute MI was overestimated in certain age- and sex-specific subgroups when self-reported data were used compared with when objectively measured data were used, whereas in other subgroups, the association was underestimated.

Although the measurement error of continuous BMI appeared relatively small and potentially negligible when compared between cases and control participants (eg, self-reported BMI was overestimated by 0.23 for cases and 0.66 for control participants among men aged 45 years and younger), it still introduced substantial bias into the association of interest. Notably, the categorization of an inherently

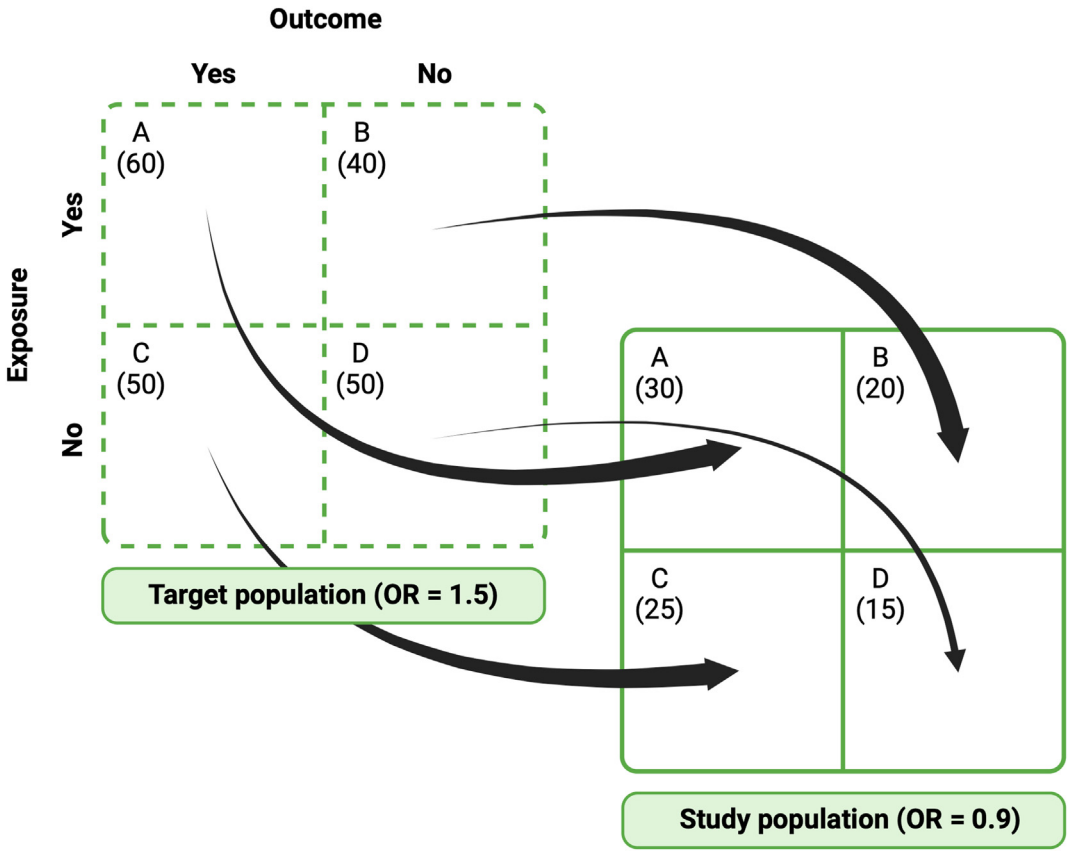

Figure 2. Selection bias due to the under-representation of individuals from a single exposure-outcome group. In this scenario, unexposed individuals without the outcome have a lower probability of being selected from the target population and into the study population (D = 30%) compared with individuals from the other 3 exposure-outcome groups (A, B, and C = 50%). When the odds ratio (OR; ie, [A × D]/[B × C]) calculated from the target population (OR = 1.5) is compared with that from the study population (OR = 0.9), this under-representation biases the measure of association toward and beyond the null value of 1. This scenario often occurs in case-control studies, in which unexposed "healthy" controls might be less likely to participate, especially in studies focused on the effect of unhealthy behaviours. This reluctance to participate might be attributed to a lack of incentives or benefits, perceived relevance, and scheduling conflicts or time constraints. This scenario can also occur in longitudinal cohort studies, where unexposed individuals might lose interest and drop out, which can affect the study's internal validity.

continuous variable—a methodological *faux pas* that has plagued the medical literature for decades—might have also amplified this bias.[48] Although common in everyday life (eg, labelling temperatures as "hot" or "cold") and clinical practice (eg, classifying cholesterol levels as low, borderline, or abnormal), translating this practice into analysis might lead to spurious associations between exposures and outcomes, turn nondifferential measurement errors into differential misclassification, and/or result in loss of information and reduced statistical power.[49-51] This "dichotomania" degrades rich continuous measures to overly simplistic, binary conclusions.[52] Such dichotomous thinking has also promoted the use of a $P < 0.05$ threshold to define statistically significant and "important" associations, which can distort the interpretation of meaningful results.[52]

Confounding

Confounding occurs when the effect between an exposure and outcome is distorted because of the influence of a third variable (or group of variables), referred to as a confounder. Of note, confounding is distinct from effect measure modification (ie, the magnitude of the effect of an exposure on an outcome differs according to strata of a third variable) and interaction (ie, joint effects of 2 or more exposures on an outcome), both of which should be reported if present and not considered as biases.[53]

Confounding can be separated into 3 types according to the direction of the distortion effect. Positive confounding is defined as an overestimation of the main effect of interest due to the influence of a confounder (ie, biased away from the null). Negative confounding is defined as an underestimation of the main effect of interest due to the influence of a confounder (ie, biased toward the null). In its most extreme form, qualitative confounding can occur if the confounding effect results in an inversion of the direction of the main effect of interest.[54] Even if all apparently important confounders are measured and controlled for in a study, confounding might persist (residual confounding).[11] Sources of residual confounding might include the presence of unmeasured confounders or analytical issues, such as the improper classification or definition of measured confounders (eg, adjusting for age using 2 categories, younger than 65 years and 65 years and older), confounder misclassification, or the use of

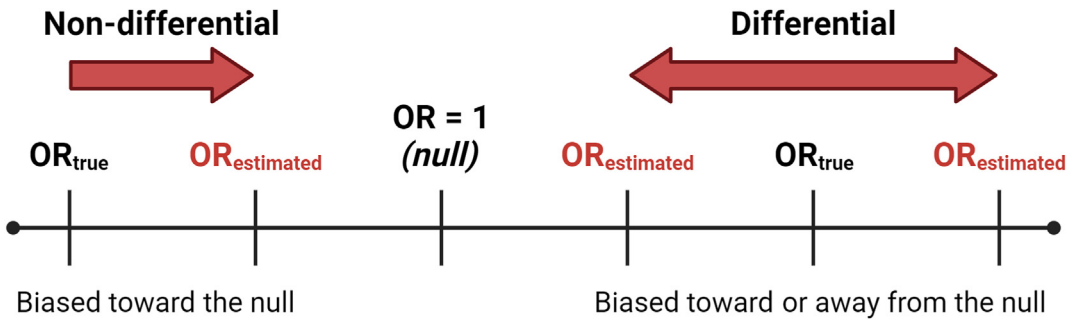

Figure 3. Nondifferential and differential misclassification of an exposure or outcome. Under nondifferential misclassification, the true odds ratio (OR; or any other measure of association) between the exposure and outcome might be biased toward the null value of 1, although this simple rule of thumb is not true in all cases. With differential misclassification, the true OR between the exposure and outcome might be biased either toward or away from the null value of 1.

imperfect surrogate measures (eg, resting heart rate used as a proxy for cardiovascular fitness).[55] Supplemental Table S1 presents a reference tool that can be used to predict the direction of confounding bias according to known associations between an exposure, outcome, and an unmeasured confounder.[56]

The identification and control of confounders are fundamental components of designing a study that intends to evaluate a potential causal association. The following 3 conceptual criteria are commonly used to determine if a variable is a confounder of an exposure-outcome relationship:[54]

1. The confounder is causally associated with the outcome.
2. The confounder is noncausally or causally associated with the exposure.
3. The confounder is not an intermediary variable along the causal pathway between the exposure and the outcome (ie, this intermediary variable is referred to as a mediator).

Regarding criterion 3, erroneous adjustment of a mediator or a descending proxy of this intermediate variable can lead to overadjustment bias.[57]

In reference to the previous example, which described a cardioprotective effect of HRT among postmenopausal women as reported in observational studies, the lack of control for potential confounding variables was also noted as a possible reason for this misleading finding.[58-60] A strong confounder, which has well documented associations with frequent HRT use and lower CAD risk, is socioeconomic status.[61-63] Notably, in their systematic review, Humphrey et al.[64] observed that studies that adjusted for socioeconomic status did not show a significant benefit of current HRT use on CAD incidence (RR = 0.97; 95% confidence interval, 0.82-1.16). In contrast, those that did not adjust for socioeconomic status reported effect estimates suggestive of a beneficial effect of HRT on CAD incidence (RR = 0.71; 95% confidence interval, 0.64-0.78). As illustrated in Figure 4, confounding by socioeconomic status would be classified as positive confounding on the basis of the direction of associations between HRT, CVD, and socioeconomic status.

Various approaches can be used to minimize the effect of confounding on a measured association of interest, which can be used at either the study design or analysis stage. At the study design phase, these can include restriction (ie, limiting the analytic sample to participants within a single category of a confounder), matching (ie, matching participants from comparison groups according to key confounders, such as age and sex), or randomization (ie, randomly assigning participants to exposure or treatment groups in a clinical trial).[65] Confounding control can also occur in the analysis phase through stratification, standardization, multivariable regression adjustment, or other sophisticated analytical approaches, such as propensity score methods.[66] Each of these approaches has distinct advantages and disadvantages that must be carefully evaluated to ensure adequate control of confounding, while also taking into account the time, cost, and resource constraints involved in conducting the study.[66,67]

Specialized Biases Commonly Encountered in Cardiovascular Research

Specialized biases, such as competing risks, immortal time bias, and confounding by indication, are frequently encountered in cardiovascular research. Although these biases are not specific to the cardiovascular field,[68,69] they are becoming increasingly prevalent, especially in studies that draw on retrospective health administrative data for observational analyses.

Competing risks

Time-to-event or "survival" analyses are common in clinical research and unique in that they not only assess if an event of interest occurred but also when it occurred—specifically, the length of time from recruitment or beginning of follow-up to event occurrence.[70] However, some participants might not experience the event before the end of the observation period (ie, they remain event-free or drop out), thus, their actual time-to-event is unknown.[70] To properly account for the time they contribute to the study, these individuals are treated as censored in the analysis.

Conventional approaches to time-to-event analysis (eg, Cox proportional hazards models) are carried out under the assumption of noninformative censoring, meaning that the likelihood of experiencing the outcome of interest for censored participants is the same as for uncensored participants.[71] Yet, in the presence of ≥ 1 competing risk, this assumption is violated. A competing risk is an event that precludes the occurrence of the primary outcome.[72,73] For example, in a time-to-event study on the association between preeclampsia and acute MI,

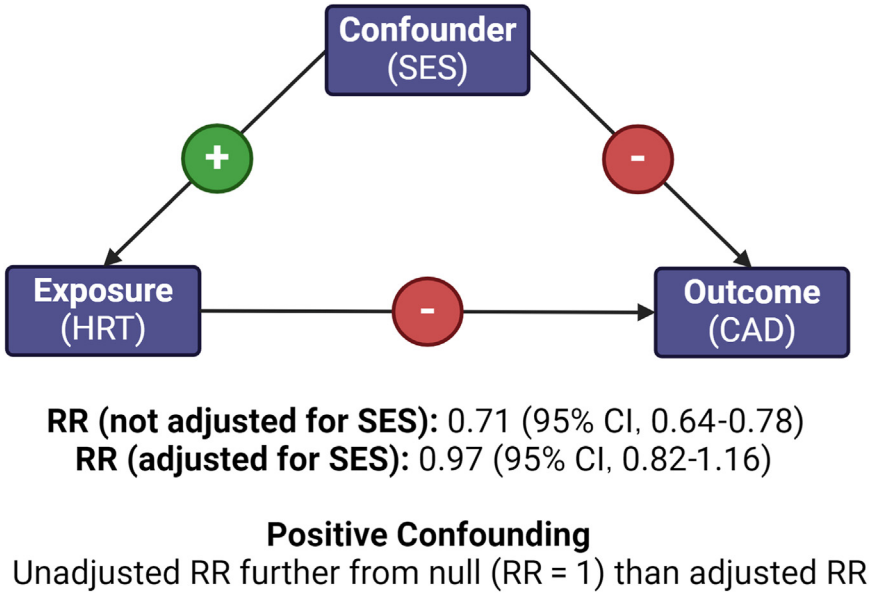

Figure 4. Causal diagram representing positive confounding due to an uncontrolled confounder (socioeconomic status; SES) in the association between hormone replacement therapy (HRT) and coronary artery disease (CAD). Higher SES is positively associated with HRT use and negatively associated with incident CAD. Without adjustment for SES, HRT use is negatively associated with incident CAD. In the presence of positive confounding, the unadjusted measure of association (RR = 0.71) is biased away from the null value of 1 (ie, overestimated) compared with the measure of association adjusted for the punitive confounder (RR = 0.97). CI, confidence interval.

all-cause mortality before an MI diagnosis would be a competing risk, because it precludes the occurrence of an MI in these individuals. Although common in cardiovascular research,[73,74] competing risks are often ignored in epidemiological studies, which can result in inflated estimates of incidence and upward-biased effect estimates.[71,73,75]

Fortunately, competing risks can be easily handled using simple extensions of conventional time-to-event models, such as the Cox proportional hazards model.[72] The Fine-Gray subdistribution hazards model is one of the most widely used methods for handling competing risks data.[72,74] However, it is best suited for questions related to incidence and prediction.[71,76,77] When assessing etiologic relationships in the presence of competing risks, the cause-specific hazards model is recommended.[75] Unlike differential losses to follow-up, which mainly stem from data limitations and can introduce selection bias if not properly addressed, competing risks do not inherently cause selection bias; they reflect real-world alternative events that prevent the occurrence of the event of interest.[78] As such, distinct design strategies and statistical methods are required to address bias from differential loss-to-follow-up (eg, participant retention strategies, inverse probability of censoring weights) and manage competing risks (eg, multistate models, Cox proportional hazards model extensions).

Immortal time bias

Epidemiologic studies are particularly vulnerable to time-dependent biases that threaten internal validity. Immortal time bias, a common time-dependent error, might arise when follow-up time and exposure (or treatment) status are not properly accounted for during design and analysis.[79] This bias can compromise the validity of effect estimates,[80] often making an exposure seem more favourable or effective than it truly is.[79] Although often discussed in pharmacoepidemiologic research,[81] immortal time bias can occur in any observational study that has a window of time between cohort entry and initiation of the exposure of interest (eg, cardiac surgery after heart failure diagnosis, filling a prescription after hospital discharge).[82] Bias is introduced when this "immortal" period is either misclassified (ie, the immortal period is misclassified as exposed time when it should be classified as unexposed time; information bias) or excluded (ie, the immortal period is excluded from the analysis; selection bias).[83,84]

Figure 5 illustrates how immortal time bias can arise in a hypothetical cohort study on the effect of statins (exposure) on reducing cardiovascular events (outcome) in individuals with high cholesterol levels. The period after a diagnosis of high cholesterol but before statin therapy initiation is deemed as "immortal" for participants in the exposed group because they are required to survive and be event-free until they fill their first statin prescription.[79] If follow-up begins at the time of high cholesterol diagnosis for the exposed and unexposed groups, immortal time bias occurs, because the "immortal" period is mistakenly counted (ie, misclassified) as time receiving statin therapy for a participant in the exposed group. Immortal time bias can also be introduced if follow-up begins at the time of statin therapy initiation for the exposed group and at high cholesterol diagnosis for the unexposed group. By design, participants in the exposed group are only those who survived the period between their high cholesterol diagnosis and first statin prescription (ie, the "immortal" period is excluded). As a result, healthier participants (and presumably those who are less likely to suffer a cardiovascular event) are included in the exposed group, which can lead to selection bias.

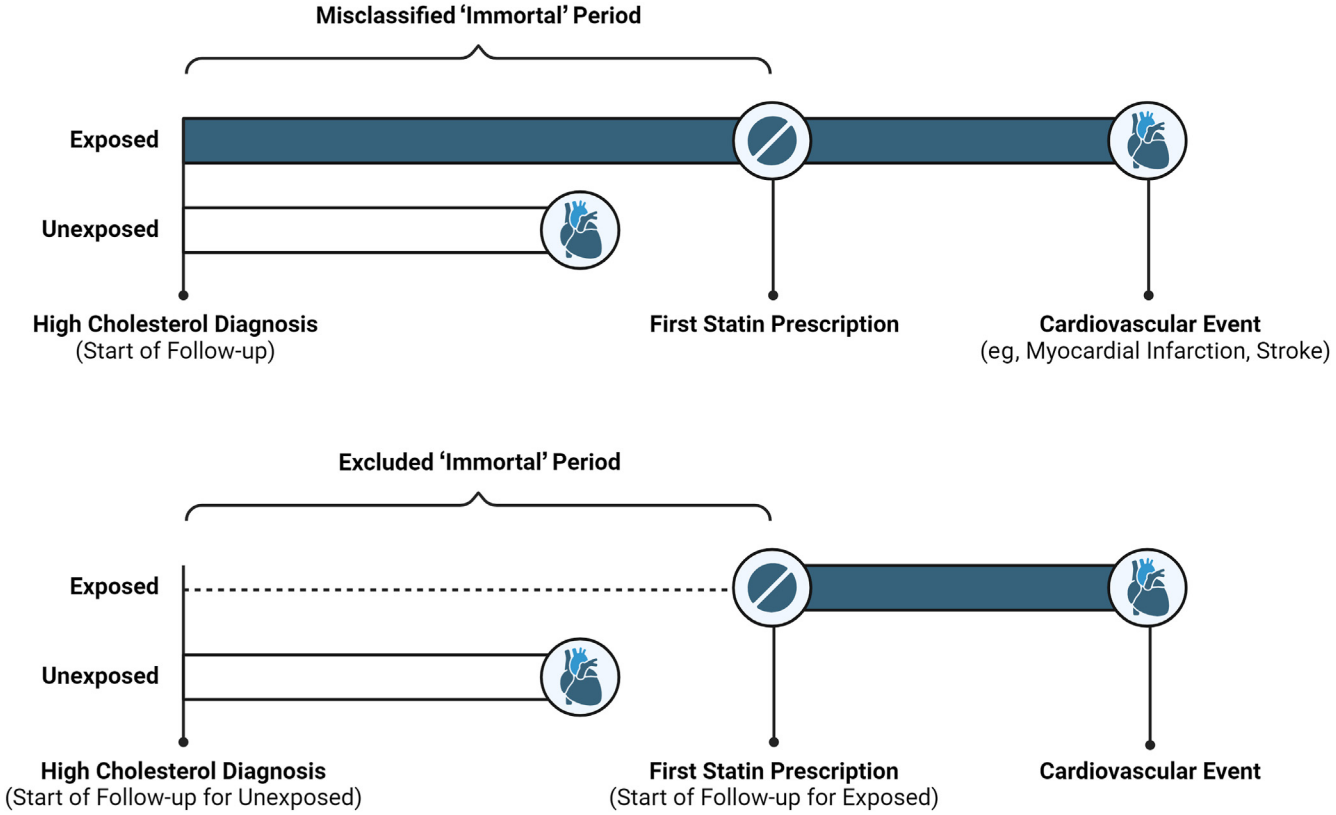

Figure 5. Immortal time bias in a hypothetical cohort study evaluating the association of statin initiation (exposure) and cardiovascular events (outcome) in individuals with high cholesterol. Immortal time bias can occur when the period between diagnosis of high cholesterol and initiation of statin therapy is either misclassified (top) or excluded (bottom). In the top scenario, the 'immortal' period, where participants must be event-free before initiating statin therapy, is incorrectly counted as time being exposed, leading to information bias. In the bottom scenario, follow-up for the exposed group begins at the time of statin initiation, excluding the 'immortal' period and introducing selection bias, as only participants who survived and remained event-free during this period are included in the exposed group. Modified from Lévesque et al.[79] with permission from BMJ Publishing Group Ltd.

Immortal time bias is best handled during the design phase, where careful planning (eg, making use of the target trial framework, as described later in this review, and prevalent new-user study designs) can help prevent it.[79,81,85] However, if it is not avoided during the design stage, such as in a situation in which exposure status might change throughout the follow-up period (ie, time-varying exposures), its effects can be minimized by using time-dependent analyses wherein the period before exposure is classified as unexposed, and the period after exposure is classified as exposed for each participant.[85] Landmark analyses (ie, determining exposure status for all participants at a predefined point in time, when follow-up begins) or matching on time (ie, exposed participants are matched to nonexposed participants who were followed-up for the same amount of time) are also reasonable alternatives to avoid immortal time bias in observational studies.[86]

Confounding by indication

Confounding by indication is a type of bias that occurs when a clinical indication for a treatment (eg, disease severity, frailty, health insurance in settings without universal health care) can also affect the outcome.[87] Since the clinical indication is associated with the treatment and the outcome, and does not lie on the causal treatment-outcome path, it fulfils the 3 criteria for confounding and must be managed accordingly. This form of confounding frequently arises in clinical research and observational pharmacoepidemiologic studies, because health care professionals tailor medication prescriptions and perform recommended procedures for patients expected to gain the greatest benefit.[88] At times, the presence of confounding by indication can erroneously suggest that certain treatments or interventions cause the very diseases or events that they aim to prevent.[89]

Observational studies that compared tracheal intubation and bag-valve-mask ventilation with survival after in-hospital cardiac arrest have highlighted the influence of confounding by indication.[90,91] Clinical indications, such as frailty, respiratory comorbidities (eg, asthma, cystic fibrosis), or preexisting airway obstructions (eg, tumours, swelling), often affect a health care professional's decision to perform or not perform an invasive tracheal intubation procedure. These factors must be accounted for because they are also associated with poor survival outcomes in cardiac arrest patients. In studies on this association in pediatric and adult in-hospital cardiac arrests, unadjusted analyses noted that tracheal intubation was strongly associated with poorer survival outcomes (RR in pediatric patients = 0.64; RR in adult patients = 0.58).

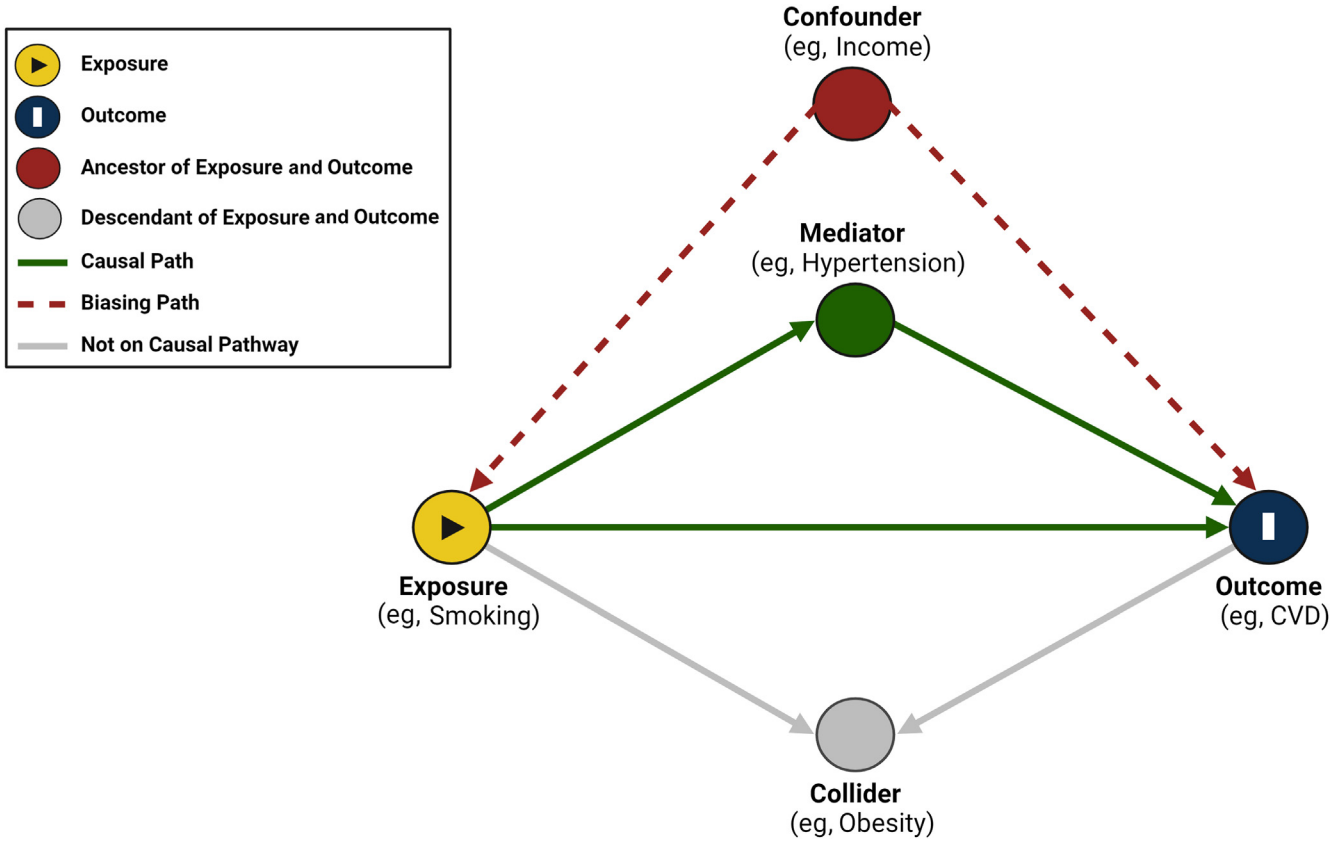

Figure 6. Directed acyclic graph depicting a hypothesized causal relationship between an exposure and an outcome. Each study variable is denoted by a **circle** (node), whereas **arrows** (edges) represent the direction of the hypothesized relationship between each node. In this simple directed acyclic graph of the causal relationship between smoking and cardiovascular disease (CVD), the minimally sufficient adjustment set would be income, because adjustments should not be made for obesity (ie, a collider) or hypertension (ie, a mediator).

However, after adjustment for various clinical indications using propensity score matching, this association weakened (RR in pediatric patients = 0.89; RR in adult patients = 0.84).

Strategies to Evaluate and Mitigate Bias in Observational Studies

In addition to adopting best practices in observational study design—such as reducing barriers to participation, using validated measures of key study variables, and following a meticulously planned protocol—new strategies have emerged over the past 2 decades to help researchers evaluate and mitigate bias. In this section, 4 valuable tools are presented, which can be readily applied to critically assess bias during the study design, analysis, and interpretation stages of the research process.

Target trial framework

Ideally, causal associations between treatments and outcomes would be examined through rigourously conducted randomized controlled trials. However, feasibility and ethical concerns might preclude this.[92] Hernán and Robins[92] initially introduced the concept of "target trials" as a strategy to reduce bias in observational studies, with the general aim of preventing critical mistakes that could lead to incorrect causal conclusions.[93] The application of this framework is becoming increasingly prevalent in CVD research,[94-97] underscoring its relevance and applicability to this field.

Briefly, target trials adapt key features of randomized controlled trial design for use on observational data using a 2-step process.[93] The first step requires that researchers clearly articulate and/or frame their causal question as if they were designing an ideal trial to obtain the desired answer. The second step requires that researchers develop a research protocol for the trial they aim to emulate in which 7 key components are predefined: eligibility criteria, treatment strategies, treatment assignment, outcomes, follow-up, causal contrasts, and statistical analysis.[92] Specific consideration of these target trial design elements is an effective way to prevent systematic error commonly observed in observational studies (eg, prevent prevalent user and immortal time bias).[81] Further, findings from observational studies planned using a target trial framework are comparable with those of randomized controlled trials,[92] which makes this framework an appealing option for cardiovascular researchers who wish to address questions that cannot or will not be explored through formal experiments. At present, reporting guidelines for observational studies emulating a target trial are being developed.[98]

The directed acyclic graph

When undertaking etiological research, a conceptual approach to study design is recommended. This approach underscores the importance of avoiding statistically driven decisions about covariate selection, such as adjusting for covariates found to be unevenly distributed with P values < 0.05 in a descriptive table or following the 10% change-in-estimate rule of thumb (ie, defining confounding as a situation where the inclusion of a potential confounder in a regression model changes the exposure-outcome effect estimate by 10% or more).[99] Indeed, variable types (eg, confounders, mediators, colliders) cannot be determined or distinguished using such approaches.[100] Moreover, model selection methods (eg, backward elimination, forward stepwise) cannot ensure that effect estimates are unconfounded or that statistical models have been correctly specified.

Covariate selection should be highly tailored to each research question. As such, it is important to develop a conceptual framework that clearly defines the research question(s), identifies key study variables, and thoughtfully characterizes their hypothesized interrelationships.[45] For instance, a variable that serves as a confounder in one study might function as a mediator in another, both of which must be managed differently when estimating a causal association.

The creation of causal diagrams, such as directed acyclic graphs (DAGs), can facilitate careful variable selection and is a systematic way for researchers to visually map out their causal assumptions (Fig. 6).[101] The use of DAGs is grounded in the principles of causal epidemiology informed by Judea Pearl's pioneering work on structural causal models and the development of the do-calculus.[102] In brief, DAGs summarize all potential or known causal or noncausal pathways between an exposure and outcome by connecting variables (known as nodes) with unidirectional arrows (known as edges). A key feature of DAGs that distinguishes them from other conceptual diagrams is their acyclic nature, meaning that no directed path from any node to itself (ie, a feedback loop) is permitted.[103]

DAGs are useful for identifying variables that should be controlled for (eg, confounders) and those that should not (eg, colliders), and can help researchers identify variables that might warrant further investigation as part of sensitivity or subgroup analyses (eg, mediators and effect measure modifiers).[45,101] These causal diagrams allow for the identification of a minimally sufficient adjustment set—that is, the minimum set of covariates that, if conditioned on, allow for an unbiased causal estimate of the association (assuming no other forms of bias, such as selection and information bias, are present). As such, they can help researchers achieve model parsimony (which is particularly relevant for small sample sizes or data collection constraints), avoid overadjustment, and improve the precision of effect estimates. Several resources are available to orient readers to the development of DAGs and the steps required to determine a minimally sufficient adjustment set, commonly referred to as "solving a DAG."[103-106]

Quantitative bias analysis

In any study, bias is always probable and can never be completely diminished through study design or statistical analysis. This might raise questions surrounding the extent to which residual bias might affect the results. Quantitative bias analysis (QBA) is a type of sensitivity analysis that can be used to estimate the direction, magnitude, and uncertainty arising from residual bias on study results.[107] In general, QBA can be used to estimate what an observed association would have been in the absence of systematic error due to selection bias, information bias, and/or uncontrolled confounding.[108] QBA relies on the input of prespecified bias parameters (eg, the strength of association between an unmeasured confounder and outcome, sensitivity and specificity of exposure classification within subgroups of the outcome), which are often obtained from internal validation data, external validation data, expert opinion, or a combination of these sources.[1]

Various QBA methods have been described, ranging from simple sensitivity analyses that only consider a single bias and assign 1 fixed value to each bias parameter, to multiple bias models, which deal with > 1 bias simultaneously and assign probability distributions to bias parameters.[107] Publicly available Web-based applications, such as Apisensr,[109] as well as downloadable Excel (Microsoft Corp, Redmond, WA) spreadsheets and statistical software code (SAS [SAS Institute, Cary, NC], R [R Foundation for Statistical Computing, Vienna, Austria], Stata [StataCorp LLC, College Station, TX])[110] provide a broad array of options for researchers interested in using QBA procedures in their studies.

A simple bias analysis that has garnered attention in the clinical research community is the E-value. Introduced by VanderWeele and Ding[111] in 2017, the E-value provides a way to quantify how susceptible an observed association is to unmeasured confounding. The E-value for relative measures of association can be calculated using the following formula, with extensions for other measures of associations detailed elsewhere:[111]

$$E-value = RR + \sqrt{RR \times (RR - 1)}$$

It should be noted that if a protective association is observed, 1/RR should be used instead of RR.

This measure can be interpreted as the minimum strength of association, on the RR scale, that an unmeasured confounder (or a summary of multiple unmeasured confounders) must have with the exposure and the outcome, conditional on the measured covariates, to completely explain away the observed association.[111] While useful, E-values rely on several simplifying assumptions that might be unrealistic, such as equivalent magnitudes of confounder-exposure and confounder-outcome associations, and the independence of an unmeasured confounder from controlled confounders.[112-114] Moreover, E-values are monotonically related to the effect estimate and are prone to misuse and misinterpretation.[112,113] For these reasons, critics have argued that other QBA approaches that incorporate multiple sources of bias, as opposed to placing sole emphasis on unmeasured confounding, are preferred to contextualize the potential effect of bias on study results.[115]

Risk of bias assessment tools

Various tools exist to critically appraise methodological study quality for observational studies.[116] Although these tools

were originally designed for critical appraisal, they are particularly useful for identifying potential biases and guiding study design. Many tools assess the risk of bias by evaluating several key domains, including the selection of study participants, comparability of study groups, exposure measurement, outcome ascertainment, and attrition or selective reporting of study participants. Supplemental Table S2 highlights some commonly used risk of bias assessment tools for observational study designs, including cohort, case-control, and cross-sectional studies. The Newcastle-Ottawa Scale (NOS) is the most commonly used tool for cohort and case-control studies.[116] However, the Risk Of Bias In Non-randomized Studies - of Interventions (ROBINS-I) tool is the most comprehensive for bias assessment in these study designs and is the preferred tool for use in Cochrane reviews.[2,117] For analytical cross-sectional studies, the Joanna Briggs Institute checklist has been recommended, whereas several other tools are suitable to evaluate the risk of bias in descriptive cross-sectional studies.[116]

Risk of bias assessment tools have also been tailored for comparative effectiveness studies that compare real-world outcomes for different treatment strategies, such as the validated Good Research for Comparative Effectiveness (GRACE) Checklist.[118] Most recently, the Risk Of Bias In Non-randomized Studies - of Exposures (ROBINS-E) instrument was introduced in 2024 to provide a standardized framework for examining the risk of bias in cohort studies that evaluate the causal association of an exposure on an outcome.[119] This tool addresses bias within 7 domains using a series of signalling questions, after which overall judgements are made for the risk of bias, predicted direction of bias, and threat to conclusions.[119] Other tools used for case series or case reports, diagnostic studies, prediction models, randomized controlled trials, and other study designs (eg, qualitative studies, systematic reviews and/or meta-analyses, etc) are also available.[116]

Concluding Remarks

In the absence of bias—a challenging assumption to uphold—a causal interpretation might be tenable from well conducted observational studies specifically designed to evaluate causal questions.[9,10] However, the potential for selection bias, information bias, and confounding, is always probable and should never be dismissed as unimportant. This presents an inherent challenge that faces investigators in all observational studies. As outlined in this review, there are several threats to validity that might affect observational studies, making it difficult to assume that a study is entirely free from bias.

High-quality evidence does not necessarily mean "free of error." Indeed, this can never be guaranteed in clinical research—only minimized with rigourous study design and analysis. The higher the quality of evidence, the greater the ability to make causal claims about an association of interest. Importantly, researchers must be aware of what qualifies as high-quality evidence, which depends on factors such as the design of a study, the scientific rigour used by those who conducted the study, as well as the risk of bias.[120]

In the pursuit of scientific truth in cardiovascular research, absolute certainty remains elusive. However, minimizing bias through meticulous design and analysis elevates the quality of evidence and strengthens one's ability to make meaningful causal claims. The use of strategies to critically assess the risk of bias—from design through analysis—and systematically quantify the extent to which bias might overturn conclusions enhances academic discourse and strengthens the foundation to advance clinical recommendations, treatment strategies, and public health policies.[121]

Acknowledgements

The authors thank the individuals who supported the content development for this article, which was originally presented as a workshop at Vascular 2023 in Montreal, Quebec. These contributors included Robert Avram, Katherine Allan, Hailey Banack, Shaun Goodman, Amer Johri, Ryan Perry, and Shelley Zieroth. The authors also thank Brice Batomen for insightful discussions that helped clarify components of this article. All illustrations were created using BioRender (BioRender.com).

Ethics Statement

The research reported in this article did not require ethics approval.

Patient Consent

The authors confirm that patient consent does not apply to this article. This review involved the synthesis of aggregated and deidentified data from previously published studies.

Funding Sources

Nicholas Grubic is supported by the Canadian Institutes of Health Research Vanier Canada Graduate Scholarship.

Disclosures

The authors have no conflicts of interest to disclose.

Editorial Disclaimer

Given her role as Associate Editor, Varinder Randhawa had no involvement in the peer review of this article and has no access to information regarding its peer review.

References

1. Lash TL, Fox MP, Maclehose RF, et al. Good practices for quantitative bias analysis. Int J Epidemiol 2014;43:1969-85.

2. 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.

3. van Diepen M, Ramspek CL, Jager KJ, Zoccali C, Dekker FW. Prediction versus aetiology: common pitfalls and how to avoid them. Nephrol Dial Transplant 2017;32(suppl 2):ii1-5.

4. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med 2019;170:51.

5. Moons KGM, Wolff RF, Riley RD, et al. PROBAST: a tool to assess risk of bias and applicability of prediction model studies: explanation and elaboration. Ann Intern Med 2019;170:W1.

6. Wang HE, Landers M, Adams R, et al. A bias evaluation checklist for predictive models and its pilot application for 30-day hospital readmission models. J Am Med Informatics Assoc 2022;29:1323-33.

7. Cochrane Training, Higgins J, Savović J, Page M, Elbers R, Sterne J. Chapter 8: Assessing Risk of Bias in a Randomized Trial. Available at: https://training.cochrane.org/handbook/current/chapter-08.

8. Mansournia MA, Higgins JPT, Sterne JAC, Hernán MA. Biases in randomized trials: a conversation between trialists and epidemiologists. Epidemiology 2017;28:54-9.

9. Dahabreh IJ, Bibbins-Domingo K. Causal inference about the effects of interventions from observational studies in medical journals. JAMA 2024;331:1845-53.

10. Hernán MA. The C-word: scientific euphemisms do not improve causal inference from observational data. Am J Public Health 2018;108:616-9.

11. Porta M, ed. A Dictionary of Epidemiology. 6th Ed. Oxford University Press, 2014.

12. Lash TL, Fink AK. Semi-automated sensitivity analysis to assess systematic errors in observational data. Epidemiology 2003;14:451-8.

13. Chatburn RL. Evaluation of instrument error and method agreement. J Am Assoc Nurse Anesth 1996;64:261-8.

14. Ishigami J, Charleston J, Miller ER, et al. Effects of cuff size on the accuracy of blood pressure readings: the Cuff(SZ) randomized crossover trial. JAMA Intern Med 2023;183:1061-8.

15. Catalogue of Bias. Available at: https://catalogofbias.org. Accessed August 2, 2024.

16. Sackett DL. Bias in analytic research. J Chronic Dis 1979;32:51-63.

17. Choi BCK, Pak AWP. A catalog of biases in questionnaires. Prev Chronic Dis 2005;2:1-13.

18. Delgado-Rodríguez M, Llorca J. Bias. J Epidemiol Community Health 2004;58:635-41.

19. Infante-Rivard C, Cusson A. Reflection on modern methods: selection bias—a review of recent developments. Int J Epidemiol 2018;47:1714-22.

20. Lu H, Cole SR, Howe CJ, Westreich D. Toward a clearer definition of selection bias when estimating causal effects. Epidemiology 2022;33:699-706.

21. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. Epidemiology 2004;15:615-25.

22. Kaufman JS. Causal inference challenges in the relationship between social determinants and cardiovascular outcomes. Can J Cardiol 2024;40:976-88.

23. Lesko CR, Buchanan AL, Westreich D, et al. Generalizing study results: a potential outcomes perspective. Epidemiology 2017;28:553-61.

24. Degtiar I, Rose S. A review of generalizability and transportability. Annu Rev Stat Appl 2023;10:501-24.

25. Szklo M, Nieto J. Understanding lack of validity: bias. In: Epidemiology: Beyond the Basics. 4th ed. John & Bartlett Learning, 2019:127-74.

26. Howe CJ, Cole SR, Lau B, Napravnik S, Eron JJ. Selection bias due to loss to follow up in cohort studies. Epidemiology 2016;27:91-7.

27. Wacholder S, McLaughlin JK, Silverman DT, Mandel JS. Selection of controls in case-control studies: I. Principles. Am J Epidemiol 1992;135:1019-28.

28. Wacholder S, Silverman DT, McLaughlin JK, Mandel JS. Selection of controls in case-control studies: II. Types of controls. Am J Epidemiol 1992;135:1029-41.

29. Lash TL, Rothman KJ. Case-control studies. In: Lash TL, Vanderweele TJ, Haneuse S, Rothman KJ, eds. Modern Epidemiology. 4th ed. Wolters Kluwer, 2021:161-84.

30. Hernán MA. Invited commentary: selection bias without colliders. Am J Epidemiol 2017;185:1048-50.

31. Grubic N, Khattar J, De Rubeis V, et al. The weight of trauma: navigating collider stratification bias in the association between childhood maltreatment and adult body mass index. CJC Pediatr Congenit Heart Dis 2024;3:98-101.

32. Admon AJ, Bohnert ASB, Cooke CR, Taylor SP. Beyond confounding: identifying selection bias in observational pulmonary and critical care research. Ann Am Thorac Soc 2022;19:1084-9.

33. Danaei G, Tavakkoli M, Hernán MA. Bias in observational studies of prevalent users: lessons for comparative effectiveness research from a meta-analysis of statins. Am J Epidemiol 2012;175:250-62.

34. Vandenbroucke JP. The HRT controversy: observational studies and RCTs fall in line. Lancet 2009;373:1233-5.

35. Hernán MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments. Epidemiology 2008;19:766-79.

36. Canto JG, Kiefe CI, Rogers WJ, et al. Number of coronary heart disease risk factors and mortality in patients with first myocardial infarction. JAMA 2011;306:53-8.

37. Banack HR, Harper S, Kaufman JS. Coronary heart disease risk factors and mortality. JAMA 2012;307:1137-8.

38. Banack HR, Harper S, Kaufman JS. Accounting for selection bias in studies of acute cardiac events. Can J Cardiol 2018;34:709-16.

39. Banack HR, Kaufman JS. The obesity paradox: understanding the effect of obesity on mortality among individuals with cardiovascular disease. Prev Med 2014;62:96-102.

40. Hernandez-Diaz S, Schisterman EF, Hernan MA. The birth weight "paradox" uncovered? Am J Epidemiol 2006;164:1115-20.

41. Levy PS. Measurement error and misclassification in statistics and epidemiology: impacts and Bayesian adjustments. Am J Epidemiol 2004;159:911-2.

42. Lash TL, Vanderweele TJ, Rothman KJ. Measurement and measurement error. In: Lash TL, Vanderweele TJ, Haneuse S, Rothman KJ, eds. Modern Epidemiology. 4th ed. Wolters Kluwer, 2021:287-314.

43. Alfred Frost S, Alexandrou E. Misclassification and measurement error - planning a study and interpreting results. Nurse Res 2021;29:21-5.

44. Jurek AM, Greenland S, Maldonado G, Church TR. Proper interpretation of non-differential misclassification effects: expectations vs observations. Int J Epidemiol 2005;34:680-7.

45. Johnston A, Smith GN, Tanuseputro P, Coutinho T, Edwards JD. Assessing cardiovascular disease risk in women with a history of hypertensive disorders of pregnancy: a guidance paper for studies using administrative data. Paediatr Perinat Epidemiol 2024;38:254-67.

46. Yland JJ, Wesselink AK, Lash TL, Fox MP. Misconceptions about the direction of bias from nondifferential misclassification. Am J Epidemiol 2022;191:1485-95.

47. Oliveira A, Ramos E, Lopes C, Barros H. Self-reporting weight and height: misclassification effect on the risk estimates for acute myocardial infarction. Eur J Public Health 2009;19:548-53.

48. Altman DG, Royston P. The cost of dichotomising continuous variables. BMJ 2006;332:1080.1.

49. van Walraven C, Hart RG. Leave 'em alone: why continuous variables should be analyzed as such. Neuroepidemiology 2008;30:138-9.

50. Flegal KM, Keyl PM, Nieto FJ. Differential misclassification arising from nondifferential errors in exposure measurement. Am J Epidemiol 1991;134:1233-46.

51. Ragland DR. Dichotomizing continuous outcome variables: dependence of the magnitude of association and statistical power on the cutpoint. Epidemiology 1992;3:434-40.

52. Greenland S. Invited commentary: the need for cognitive science in methodology. Am J Epidemiol 2017;186:639-45.

53. Corraini P, Olsen M, Pedersen L, Dekkers O, Vandenbroucke J. Effect modification, interaction and mediation: an overview of theoretical insights for clinical investigators. Clin Epidemiol 2017;9:331-8.

54. Szklo M, Nieto J. Identifying noncausal associations: confounding. In: Epidemiology: Beyond the Basics. 4th ed. John & Bartlett Learning, 2019:175-207.

55. Szklo M, Nieto J. Stratification and adjustment: multivariate analysis in epidemiology. In: Epidemiology: Beyond the Basics. 4th ed. John & Bartlett Learning, 2019:259-338.

56. Mehio-Sibai A, Feinleib M, Sibai TA, Armenian HK. A positive or a negative confounding variable? A simple teaching aid for clinicians and students. Ann Epidemiol 2005;15:421-3.

57. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiology 2009;20:488-95.

58. Petitti D. Commentary: hormone replacement therapy and coronary heart disease: four lessons. Int J Epidemiol 2004;33:461-3.

59. Newton J. Rapid response: confounding is mistakenly called selection bias. BMJ 1998;316:1842.

60. von Elm E, Egger M. The scandal of poor epidemiological research. BMJ 2004;329:868-9.

61. Wang T, Li Y, Zheng X. Association of socioeconomic status with cardiovascular disease and cardiovascular risk factors: a systematic review and meta-analysis [e-pub ahead of print]. Z Gesundh Wiss https://doi.org/10.1007/s10389-023-01825-4.

62. Schultz WM, Kelli HM, Lisko JC, et al. Socioeconomic status and cardiovascular outcomes. Circulation 2018;137:2166-78.

63. Lawlor DA, Davey Smith G, Ebrahim S. Socioeconomic position and hormone replacement therapy use: explaining the discrepancy in evidence from observational and randomized controlled trials. Am J Public Health 2004;94:2149-54.

64. Humphrey LL, Chan BKS, Sox HC. Postmenopausal hormone replacement therapy and the primary prevention of cardiovascular disease. Ann Intern Med 2002;137:273-84.

65. Howards PP. An overview of confounding. Part 1: the concept and how to address it. Acta Obstet Gynecol Scand 2018;97:394-9.

66. Kahlert J, Gribsholt SB, Gammelager H, Dekkers OM, Luta G. Control of confounding in the analysis phase - an overview for clinicians. Clin Epidemiol 2017;9:195-204.

67. Braga L, Farrokhyar F, Bhandari M. Confounding: what is it and how do we deal with it? Can J Surg 2012;55:132-8.

68. Martinuka O, von Cube M, Wolkewitz M. Methodological evaluation of bias in observational coronavirus disease 2019 studies on drug effectiveness. Clin Microbiol Infect 2021;27:949-57.

69. Jamme M, Geri G. Time-dependent effect, immortal bias, and competing risk: 3 components that should be handled to assess the impact of covariates on occurrence of acute kidney injury. Kidney Int 2020;98:1348.

70. Schober P, Vetter TR. Survival analysis and interpretation of time-to-event data: the tortoise and the hare. Anesth Analg 2018;127:792-8.

71. Schuster NA, Hoogendijk EO, Kok AAL, Twisk JWR, Heymans MW. Ignoring competing events in the analysis of survival data may lead to biased results: a nonmathematical illustration of competing risk analysis. J Clin Epidemiol 2020;122:42-8.

72. Wolbers M, Koller MT, Stel VS, et al. Competing risks analyses: objectives and approaches. Eur Heart J 2014;35:2936-41.

73. Austin PC, Lee DS, Fine JP. Introduction to the analysis of survival data in the presence of competing risks. Circulation 2016;133:601-9.

74. Andersen PK, Geskus RB, de Witte T, Putter H. Competing risks in epidemiology: possibilities and pitfalls. Int J Epidemiol 2012;41:861-70.

75. Abdel-Qadir H, Fang J, Lee DS, et al. Importance of considering competing risks in time-to-event analyses. Circ Cardiovasc Qual Outcomes 2018;11:e004580.

76. Austin PC, Fine JP. Practical recommendations for reporting Fine-Gray model analyses for competing risk data. Stat Med 2017;36:4391-400.

77. Lau B, Cole SR, Gange SJ. Competing risk regression models for epidemiologic data. Am J Epidemiol 2009;170:244-56.

78. Austin PC, Fine JP. Accounting for competing risks in randomized controlled trials: a review and recommendations for improvement. Stat Med 2017;36:1203-9.

79. Levesque 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.

80. van Walraven C, Austin P. Administrative database research has unique characteristics that can risk biased results. J Clin Epidemiol 2012;65:126-31.

81. Tyrer F, Bhaskaran K, Rutherford MJ. Immortal time bias for life-long conditions in retrospective observational studies using electronic health records. BMC Med Res Methodol 2022;22:86.

82. Zheng Q, Otahal P, Cox IA, et al. The influence of immortal time bias in observational studies examining associations of antifibrotic therapy with survival in idiopathic pulmonary fibrosis: a simulation study. Front Med 2023;10:1157706.

83. Ukah UV, Aibibula W, Platt RW, et al. Time-related biases in perinatal pharmacoepidemiology: a systematic review of observational studies. Pharmacoepidemiol Drug Saf 2022;31:1228-41.

84. Yadav K, Lewis RJ. Immortal time bias in observational studies. JAMA 2021;325:686-7.

85. Abrahami D, Hudson M, Suissa S. Statins and lower mortality in rheumatic diseases: an effect of immortal time bias? Semin Arthritis Rheum 2021;51:211-8.

86. Dekkers OM, Groenwold RHH. When observational studies can give wrong answers: the potential of immortal time bias. Eur J Endocrinol 2021;184:E1-4.

87. Kyriacou DN, Lewis RJ. Confounding by indication in clinical research. JAMA 2016;316:1818-9.

88. Brookhart MA, Wyss R, Layton JB, Stürmer T. Propensity score methods for confounding control in nonexperimental research, 6. Circ Cardiovasc Qual Outcomes, 2013:604-11.

89. Walker AM. Confounding by indication. Epidemiology 1996;7:335-6.

90. Andersen LW, Raymond TT, Berg RA, et al. Association between tracheal intubation during pediatric in-hospital cardiac arrest and survival. JAMA 2016;316:1786-97.

91. Andersen LW, Granfeldt A, Callaway CW, et al. Association between tracheal intubation during adult in-hospital cardiac arrest and survival. JAMA 2017;317:494-506.

92. 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.

93. Hernán MA, Wang W, Leaf DE. Target trial emulation: a framework for causal inference from observational data. JAMA 2022;328:2446-7.

94. Mésidor M, Sirois C, Guertin JR, et al. Effect of statin use for the primary prevention of cardiovascular disease among older adults: a cautionary tale concerning target trials emulation. J Clin Epidemiol 2024;168:111284.

95. Russo RG, Danaei G. Can observational data and target trial emulation inform cardiovascular disease prevention and management guidelines? Circ Cardiovasc Qual Outcomes 2024;17:e010979.

96. Chen A, Ju C, Mackenzie IS, et al. Impact of beta-blockers on mortality and cardiovascular disease outcomes in patients with obstructive sleep apnoea: a population-based cohort study in target trial emulation framework. Lancet Reg Health Eur 2023;33:100715.

97. Schmidt M, Arendt-Nielsen L, Hauge EM, Sørensen HT, Pedersen L. High- vs. low-dose diclofenac and cardiovascular risks: a target trial emulation. Eur Heart J Cardiovasc Pharmacother 2023;9:453-61.

98. 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.

99. VanderWeele TJ. Principles of confounder selection. Eur J Epidemiol 2019;34:211-9.

100. van Zwieten A, Tennant PWG, Kelly-Irving M, et al. Avoiding over-adjustment bias in social epidemiology through appropriate covariate selection: a primer. J Clin Epidemiol 2022;149:127-36.

101. Shimonovich M, Pearce A, Thomson H, Keyes K, Katikireddi SV. Assessing causality in epidemiology: revisiting Bradford Hill to incorporate developments in causal thinking. Eur J Epidemiol 2021;36:873-87.

102. Pearl J. Causality. 2nd ed. Cambridge University Press, 2009.

103. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology 1999;10:37-48.

104. Glymour M. Using causal diagrams to understand common problems in social epidemiology. In: Oakes MJ, Kaufman JS, eds. Methods in Social Epidemiology. 2nd Ed. Jossey-Bass/Wiley, 2006:393-428.

105. Digitale JC, Martin JN, Glymour MM. Tutorial on directed acyclic graphs. J Clin Epidemiol 2022;142:264-7.

106. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. BMC Med Res Methodol 2008;8:70.

107. Fox MP, MacLehose RF, Lash TL. Applying Quantitative Bias Analysis to Epidemiologic Data. 2nd ed. Springer International Publishing, 2021.

108. Petersen JM, Ranker LR, Barnard-Mayers R, Maclehose RF, Fox MP. A systematic review of quantitative bias analysis applied to epidemiological research. Int J Epidemiol 2021;50:1708-30.

109. apisensr: Quantitative Bias Analysis With apisensr. Available at: https://dhaine.shinyapps.io/apisensr/. Accessed August 20, 2024.

110. bias.analysis: Applying Quantitative Bias Analysis to Epidemiologic Data. Available at: https://sites.google.com/site/biasanalysis/Home?authuser=0. Accessed August 20, 2024.

111. VanderWeele TJ, Ding P. Sensitivity analysis in observational research: introducing the E-value. Ann Intern Med 2017;167:268-74.

112. Blum MR, Tan YJ, Ioannidis JPA. Use of E-values for addressing confounding in observational studies - an empirical assessment of the literature. Int J Epidemiol 2020;49:1482-94.

113. Ioannidis JPA, Tan YJ, Blum MR. Limitations and misinterpretations of E-values for sensitivity analyses of observational studies. Ann Intern Med 2019;170:108-11.

114. Barberio J, Ahern TP, MacLehose RF, et al. Assessing Techniques for Quantifying the Impact of Bias Due to an Unmeasured Confounder: An Applied Example. Clin Epidemiol 2021;13:627-35.

115. Fox MP, Arah OA, Stuart EA. Commentary: the value of E-values and why they are not enough. Int J Epidemiol 2020;49:1505-6.

116. Ma LL, Wang YY, Yang ZH, et al. Methodological quality (risk of bias) assessment tools for primary and secondary medical studies: what are they and which is better? Mil Med Res 2020;7:7.

117. Cochrane Training. Sterne J, Hernan M, McAleenan A, Reeves B, Higgins J. Chapter 25: assessing risk of bias in a non-randomized study. In: Higgins J, Thomas J, Chandler J, et al., eds. Cochrane Handbook for Systematic Reviews of Interventions. Available at: https://training.cochrane.org/handbook/current/chapter-25.

118. Dreyer NA, Bryant A, Velentgas P. The GRACE Checklist: a validated assessment tool for high quality observational studies of comparative effectiveness. J Manag Care Spec Pharm 2016;22:1107-13.

119. Higgins JPT, Morgan RL, Rooney AA, et al. A tool to assess risk of bias in non-randomized follow-up studies of exposure effects (ROBINS-E). Environ Int 2024;186:108602.

120. Sargeant JM, Brennan ML, O'Connor AM. Levels of evidence, quality assessment, and risk of bias: evaluating the internal validity of primary research. Front Vet Sci 2022;9:960957.

121. Miller M, Swanson SA, Azrael D. Are we missing something pertinent? A bias analysis of unmeasured confounding in the firearm-suicide literature. Epidemiol Rev 2016;38:62-9.

Supplementary Material

To access the supplementary material accompanying this article, visit the online version of the *Canadian Journal of Cardiology* at www.onlinecjc.ca and at https://doi.org/10.1016/j.cjca.2024.12.022.