Chapter 3. Confounding: a routine concern in the interpretation of epidemiological studies

CHAPTER 3.

Confounding: a routine concern in the interpretation of epidemiological studies

David B. Richardson, Sadie Costello, Jay S. Kaufman, Kaitlin Kelly-Reif, Sarah Lewis, Kyle Steenland, and Eric J. Tchetgen Tchetgen

3.1 Introduction

As noted in the Preface, confounding arises when the exposure and the outcome of interest share a common cause. Informally, confounding may be described as a condition in which the association of exposure with the outcome is, in part, due to differences in outcome risk between the exposed and the unexposed that are not due to exposure effects on the outcome. A confounder is then defined as a variable that is responsible for confounding; typically, such a variable is a cause of the outcome that is associated with exposure but not affected by exposure. More precise definitions can be provided within formal causal models, such as potential-outcome and graphical models (Greenland et al., 1999a; Hernán and Robins, 2023; see also Chapter 2); these models will not be discussed here, but the reader is warned that there can be various definitions of confounding and confounders in these more formal discussions.

At an *IARC Monographs* meeting, the epidemiological studies under review are typically observational, meaning that the investigators did not have control over the exposure of interest (or any other variables) and, importantly for this chapter, did not randomly assign study participants to exposure. In observational studies, it is seldom reasonable to assume that pre-exposure factors that affect the outcome are equally distributed across subgroups defined by exposure; rather, exposure is often influenced by other factors, some of which may be risk factors for the cancer outcome of interest. Consequently, confounding is a common concern for Working Group members. Thus, one of the primary questions posed to reviewers in an *IARC Monographs* Working Group is "Can we reasonably rule out confounding as an explanation for an observed exposure–cancer association?"

A standard approach to the problem of confounding is to measure the important factors (e.g. pre-exposure factors that are predictive of the outcome in a cohort study) that may differ between exposure groups and to match on them in the study design (to the extent possible) or adjust for them in the analysis. If all the important confounders were accurately measured, an investigator might be able to obtain a valid estimate of the causal effect of the exposure on the outcome. However, the choice of which variables to control for (a judgement informed by causal, in

addition to statistical, considerations) is crucial because bias in an estimate of a defined exposure–disease association can be induced, or increased, by inappropriate control for covariates (Greenland et al., 1999b; Cole et al., 2010). Occasionally, IARC reviewers may encounter a study that used an approach intended to control for unmeasured as well as measured potential confounders. A classic example of such an approach is a randomized controlled trial, but other examples encountered in observational studies include analyses that leverage a natural experiment or an instrumental variable (such as genetic variation in a Mendelian randomization analysis; see Side Box 3.1). However, many epidemiological studies of cancer cannot or do not use these approaches; hence, uncontrolled confounding is often an important consideration for reviewers.

In Section 3.2, the reader will gain an understanding of how to evaluate control for confounding in published studies. Directed acyclic graphs (DAGs) (Chapter 2) will be referenced to represent assumptions regarding causal relations between variables and to assist in identifying causal effects. In Section 3.3, the reader will gain an understanding of approaches to assess potential bias due to uncontrolled confounding.

Given the focus on cancer studies, throughout the chapter confounding is considered as it applies to analyses of a binary outcome variable and ratio measures of association (such as rate ratios, odds ratios, hazard ratios, or risk ratios, as typical of most cancer studies). It is assumed that reviewers are interested in the total effect of the exposure on an outcome; therefore, mediation analysis, which is covered in VanderWeele (2016) and Hernán and Robins (2023), is not addressed here.

Chapter 2 introduced the use of DAGs to frame the identification and control of confounding. The focus in this chapter is on the evaluation of confounding within the context of a review that aims at hazard identification. Consequently, the focus is on whether uncontrolled confounding of a particular study result is a major source of bias and could meaningfully change a conclusion regarding that study's contribution for (or against) evidence of an association between the agent under review and the cancer outcome of interest. Evaluation of control for confounders is also commonly included in systematic reviews and meta-analyses through the use of tools to assess study quality. Such approaches rely on methods to assess the risk of bias due to confounding. While tools to assess study quality can be useful for helping a reviewer to think systematically about sources of bias, they are best used by substantive experts who can also consider the direction and magnitude of potential confounding and consider a range of methods to assess it. An uncritical use of risk-of-bias tools can lead to unwarranted dismissal of some studies because of alleged but unimportant confounding (Steenland et al., 2020). Methods are described in this chapter for an assessment of potential confounding bias, which may be useful when reviewing studies that inform an *IARC Monographs* evaluation. As described in Chapter 1, IARC has published general guidelines regarding the assessment of bias, and the methods outlined here are consistent with this guidance (IARC, 2019).

While this chapter focuses solely on confounding, there may be factors that are modifiers of the association under study (as well, perhaps, as confounders of it). In addition to considering whether a factor is a confounder, a reviewer might consider whether that factor modifies the association under study, meaning that the association on the selected measurement scale (e.g. relative risk) varies across values of the factor. Given a published report, a reviewer may be limited in such considerations by the information reported. For example, if the authors of a publication only report a covariate-adjusted estimate, then a reviewer cannot distinguish confounding by that factor from effect measure modification. However, if results have been stratified on a factor, and if the association varies importantly across strata of that factor, then there is modification of the association on that effect scale. Conversely, if the association is the same across strata, then the factor is probably not a modifier of the association (but could be a confounder of the association in a crude analysis that collapses information across strata of the factor). A variable can be a confounder, an effect measure modifier, both, or neither.

3.2 Evaluating control for confounding

When evaluating control for confounding in a published study (Fig. 3.1), reviewers will typically consider the following four topics (these will be explained in Sections 3.2.1–3.2.4):

- the study design;
- the study setting and restriction of the study population;

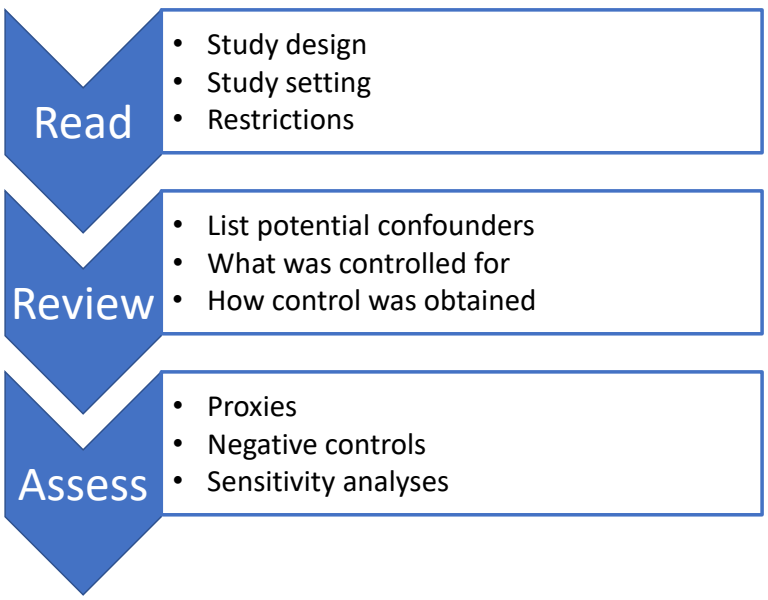
Side Box 3.1. Study designs or analyses and confounding considerations

Study design or analysis	Confounding consideration
Randomized controlled trial	Randomized trials are the most widely accepted method of addressing concern about confounding without necessarily measuring or adjusting for covariates. Nonetheless, randomized trials are rarely used in human subjects for evaluations of known or suspected carcinogens, because the administration of suspected carcinogens is unethical and because a long follow-up period is typically required in studies of cancer to observe the effect of exposure on cancer occurrence. Consequently, there are seldom many relevant randomized controlled trials to assess the carcinogenic potential of substances reviewed in *IARC Monographs*. Possible reasons for confounding in a randomized trial include imperfect allocation concealment, blinding, and adherence (i.e. compliance), as well as loss to follow-up; such considerations are important in studies of cancer outcomes because it is difficult to maintain adherence to a treatment protocol over many years.
Case-only (self-controlled)	Case-only designs, such as case-crossover and case-specular studies, are used to address concerns about potential confounding by characteristics that are constant over time. Only those participants who experience the outcome of interest are included, and study participants act as their own control. Because comparisons are made within individuals, confounding by characteristics that are constant over time is not possible. However, self-controlled designs do not typically lend themselves to investigations of cancer, which often feature long induction and latency periods. Possible reasons for bias in a self-controlled design include time-varying confounders and selection bias.
Matched fixed effects design (e.g. sibling or twin study)	Sibling and twin designs are used to address concerns about potential confounding by measured and unmeasured time-invariant factors. These studies involve pairs of participants who might be viewed as matched for a large number of potential confounders (e.g. genetics and childhood environment); these are shared or invariant characteristics within pairs and are often handled as fixed effects (intercepts) in a model. Possible reasons for bias in such designs include confounding by non-shared factors (i.e. those that may vary within a twin pair or sibling set) and selection bias (Frisell et al., 2012; Sjölander et al., 2022).
Instrumental variables, natural experiment, or Mendelian randomization analysis	Instrumental variables (IVs) are variables that are only associated with the outcome via their association with the exposure but are not affected by the exposure. A natural experiment is a type of IV analysis that involves settings in which the exposure variation for an individual (or a group of people) is due to an external factor, such as a natural disaster or an industrial accident, that is assumed to be related to the disease only through the exposure. A Mendelian randomization study is a type of IV analysis in which the investigators use genetic variation as the basis for a type of natural experiment. One important motivation for its use is that it offers the possibility to control for unmeasured confounders in an observational study (under certain strong identifying conditions). A Mendelian randomization analysis requires that the exposure under study has known genetic variants, which are strongly associated with it (an assumption termed *relevance*); it also requires that these genetic determinants are not associated with the outcome independently of the exposure (an assumption termed *exclusion restriction*), and it requires that the association between the instrument and the outcome is not confounded (an assumption termed *independence*) (Pierce et al., 2018). An increasing number of observational studies are using Mendelian randomization methods to study the effects of exposures, lifestyle factors, or biomarkers on cancer. If the necessary conditions do not hold, IV estimates may be biased (Hernán and Robins, 2006).
Cohort study	Cohort studies are commonly encountered in IARC evaluations of human carcinogens; these are observational in nature and are susceptible to confounding. This is particularly an issue when considering socioeconomic and lifestyle factors and cancer risk, because these factors tend to cluster and confounding within such studies can be substantial (Davey Smith et al., 2007). The collection of information on potential confounders in a retrospective cohort design may be limited by the available historical information, and that in a prospective cohort design may be limited by the knowledge at the time of study enrolment. Nested case–control studies (nested within cohorts) may be conducted for efficiency or to collect information on important confounders that may not have been collected in the original cohort study.

Side Box 3.1. Study designs or analyses and confounding considerations (continued)

Study design or analysis	Confounding consideration
Case–control study	Case–control studies often involve the collection of detailed information on at least a few potential confounders that are of primary focus; they typically focus on one outcome and are usually smaller than cohort studies, allowing for a richer collection of data on risk factors for the single outcome of interest and other covariates than a cohort study. Matching on potential confounders, more commonly observed in case–control studies than in cohort studies, can improve efficiency when adjusting for confounders (and may permit control for confounding that is otherwise difficult to achieve, as in neighbourhood matching). Within case–control studies, there may be issues with the control participants not being representative of the population from which the case participants arose, which could introduce confounding; also, confounders may not be measured well if individuals are asked to recall lifestyle factors that occurred before cancer symptoms were observed. In addition, confounders may be recalled differently by case and control participants (selection bias and information bias; see Chapters 4 and 5). From a confounding perspective, nested case–control studies have similar issues to cohort studies.
Ecological study	These are studies in which the exposure is studied at a population level rather than an individual level, and variation in outcome is examined in relation to variation in population prevalence of exposure. Ecological studies often have limited or no information on individual-level confounders and are consequently susceptible to confounding (including a particular form of bias that may arise in ecological study analyses because of confounding or effect measure modification between groups under comparison).
Cross-sectional study	Such studies typically play a minor role in cancer evaluations. Confounding is an issue; cross-sectional studies often have the additional complexity of temporal ambiguity. It may be unclear whether the exposure preceded the disease; it may also be unclear whether a covariate preceded the exposure and thus whether it is a confounder. (text continues on page 65)

Fig. 3.1. Steps to take when assessing confounding in individual studies.



- the set of covariates that were adjusted for in the analysis (and how those covariates were measured and modelled); and
- important confounders that were not controlled for.

3.2.1 Study design

The study design is an important starting point for evaluating control for confounding; it is possible to control for confounding in the study design phase. The choice of study design may direct a reviewer's attention to certain key areas for consideration, such as the appropriateness of an external comparison group for the analysis of standardized mortality ratios in an occupational cohort study. It may even obviate the need to focus attention on the adequacy of control for certain types of confounders. For example, matched designs involving siblings born of the same mother are sometimes used to control for maternal factors that remain constant between pregnancies, such as maternal genetics and some aspects of lifestyle and socioeconomic status (see Side Box 3.2).

3.2.2 Study setting and restrictions

A careful decision regarding study setting can help to minimize confounding, for example by finding populations that lack an association between a confounder and the exposure of concern. For instance, a large cohort of Seventh-Day Adventists offers a setting with little or no confounding by alcohol consumption or smoking, because these behaviours are largely absent in that population (Butler et al., 2008). Similarly, restriction of the study population (e.g. by sex, geography) can help to control for confounding. Sometimes restriction on a confounder can provide control over factors that would otherwise be difficult to measure and control for in an analysis. For example, restriction to a single continental population, such as Europeans (Auton et al., 2015), to minimize population stratification (confounding by ancestry) is common in genome-wide association studies (although many contemporary genome-wide analyses also adjust for finer population structure). As another example, occupational cohort studies are often conducted in a setting in which the workers involved share similarities in terms of education, income, access to medical care, geography, and lifestyle factors (e.g. diet). Consequently, in occupational studies with internal comparisons, such factors are usually of less concern as confounders than they are in environmental studies, because these lifestyle factors should have limited associations with occupational exposure.

However, inappropriate restriction can lead to bias (e.g. if restriction is on an intermediate or mediating variable or collider; see Chapters 2 and 5). Moreover, restriction necessarily affects the generalizability of results (and reduces sample size), so it should be carefully assessed.

3.2.3 Covariates that were (and were not) adjusted for in a published analysis

A standard approach to addressing the problem of confounding is to measure important factors that may differ between exposure groups and adjust for them in the analysis. Here, the focus is on analyses where the aim is to control for confounding by adjustment for measured variables (e.g. adjusting for the variable in a regression model for the outcome).

It is important to consider both the confounder–outcome association and the confounder–exposure association. Those involved in an expert review, such as an *IARC Monographs* evaluation, will often come to a consensus on the important potential confounders of an association under evaluation. One source of information about such potential confounders is the study publications under review; authors often provide useful guidance in their publications about measured and unmeasured potential confounders, as well as omitted potential confounders. However, regardless of the authors' description of important potential confounders, reviewers may have a different view. Authors often describe their approach to the final selection of their covariate adjustment set; again, regardless of how the adjustment variables were selected in a given publication, the reviewers' responsibility at this stage is to assess whether the important potential confounders have been sufficiently controlled for.

Reviewers may wish to start by considering the confounder–outcome association, focusing on those factors that are established causes of the cancer outcome under study. Useful sources of such information are the *IARC Monographs* and the IARC list of classifications of agents for which there is *sufficient* and *limited* evidence of carcinogenicity in humans by cancer site; similarly, the *IARC Handbooks of Cancer Prevention* can provide information on potential

Side Box 3.2. Some approaches to control for confounding

Matching in the design of a study can sometimes allow for control for factors that would otherwise be difficult to adjust for efficiently in the analysis (e.g. in the absence of a matched design, because of sparse data).

Matching may be used in cohort or case–control studies. In a matched cohort study, an investigator might enumerate an unexposed group of study participants who match the exposed study participants in terms of some characteristics (such as age and sex) that are of concern as potential confounders; a comparison of the occurrence of cancer between the exposed and unexposed groups will not be confounded by those factors that were matched on in the design. Matching is often used in case–control studies of cancer outcomes, with the aim of improving efficiency in a case–control analysis when it would otherwise be necessary to adjust for a matching factor, such as attained age. Similarly, in a population-based case–control study of a rare cancer, neighbourhood matching of case and control participants may allow for adjustment for characteristics that are shared by neighbours, such as socioeconomic, diet, or lifestyle factors, but that may be difficult to adjust for in the analysis in the absence of such matching, because of sparse data or difficulty in obtaining sufficient or accurate data to control for such hard-to-quantify variables. In certain settings, self-matching can be used (e.g. the case–control status is determined by the location of the tumour in relation to the exposure within the body, as in Example 3.1). However, as noted in Chapter 2, an important difference from matching in cohort studies is that case–control matching is a form of selection bias that distorts associations and trends (Mansournia et al., 2018). To control this bias, the analysis must include adjustment for the matching variables in a form at least as detailed as the form used for matching; this means, for example, that if age matching is done in 5-year categories, then the adjustment must use age as a categorical variable with categories at least as narrow as 5 years.

Example 3.1. Self-matching to control for confounding

A case-only study (Maclure, 1998) was conducted of mobile phone use and glioma (Larjavaara et al., 2011). The location of the actual tumour site (i.e. the case site) was compared with a control site, defined as the mirror image site obtained across the midpoint of the axial and coronal planes of the patient's brain (i.e. within the same person). The control sites were effectively matched to the case sites on each pair being within the same patient's brain. The case and control sites were then compared with respect to estimated mobile phone exposure, to determine whether the phone was used on the side of the brain where the tumour occurred. In this design, participants with cancer each served as their own control; therefore, confounding by personal characteristics (such as age, sex, income, or diet) was judged to be unlikely in these analyses.

Side Box 3.2. Some approaches to control for confounding (continued)

Other approaches to study design that are sometimes used in cancer research to address potential confounding involve leveraging situations in which exposure was determined by factors beyond the control of the investigator but that arguably mimic random exposure assignment. Such studies are sometimes called natural experiments or quasi-experimental designs, as explained in Example 3.2.

In a natural experiment, the assignment mechanism is a form of instrumental variable (IV), because it influences exposure but only influences the outcome through its effect on the exposure. Quasi-experimental designs have been used in evaluations of interventions on tobacco, air pollutants, and petrochemical exposures. One version of IV analysis that is sometimes encountered in epidemiological studies of cancer outcomes is Mendelian randomization, in which the IV is the random inheritance of genetic variants that are known to predict exposure, under the classic assumption that genetic factors are inherited independently of each other (note that this assumption may not hold for genetic variants that are located near one another on the same chromosome). Genetic variants are usually not subject to confounding by lifestyle and environmental factors (Smith and Ebrahim, 2003). In Mendelian randomization analyses, populations are grouped according to the presence of genetic variants (alleles) that are associated with the exposure of interest. Comparison of cancer risk between genetic groups that are associated with the exposure can provide an unconfounded estimate of the effect of the exposure on cancer (Yarmolinsky et al., 2018). (text continues on page 68)

Example 3.2. An example of a natural experiment to control for confounding: a military conscription lottery

A situation that has been used in a natural experiment is a military conscription lottery (where one compares those drafted with those not drafted for cancer outcomes, such as in studies that have examined effects of service in the Viet Nam era on cancer occurrence).

confounders that are cancer-preventive factors (see Section 6.3.1 for more examples). For a study of a given cancer outcome, a reviewer can readily refer to such lists of known or suspected causes of that cancer to inform consideration of potential confounders. Note that because confounder–outcome associations are rarely homogeneous from one cancer site to another, the list of potential confounders of concern will also vary by cancer site. As noted in Chapter 1, the evaluation of human evidence regarding carcinogenicity is also specific to each cancer site; therefore, concern about a potential confounding factor (e.g. smoking) might be reasonably excluded for certain cancer types (e.g. melanoma) but not others (e.g. lung cancer).

In addition to the confounder–outcome association, it is also necessary to consider the confounder–exposure association. An important consideration is whether potential confounders precede the exposure of interest. Therefore, reviewers may often rely on information on the distribution and determinants of exposure. A reviewer may encounter situations in which adjustment was made for a covariate that was measured after the exposure of interest occurred. In such situations, careful consideration should be given to whether exposure influenced that covariate; however, there are settings for which an investigator may reasonably assume that the measured value of such a covariate is a good approximation of its pre-exposure value and is unaffected by the exposure of interest (e.g. educational attainment, assessed after exposure, in a study of the effect of an exposure in a population of middle-aged adult patients). The factors that influence exposure to an agent may vary over time and between populations and may depend on economic and social factors, laws and regulations, and social and behavioural factors. Consequently, in assessment of confounding, information should be obtained and used on how the association of a potential confounding factor with the exposure and the disease may vary across different study populations.

A reviewer of a published article will consider whether any important potential confounders were not accounted for (e.g. not controlling for smoking in a study of a given exposure in relation to lung cancer). However, the fact that a variable that a reviewer posited as a potential confounder was not adjusted for in a published study does not necessarily mean that it was a strong confounder (or even a confounder at all). Often the authors of a publication will describe the rationale for exclusion of a variable from the adjustment set and may report results that were obtained with different sets of adjustments for covariates. As shown in Example 3.3, a factor could be an established cause of cancer but might not confound the association of interest in the population under study.

Another consideration in a review of a published article is whether any of the variables adjusted for in the published analysis were not potential confounders but rather could induce or exacerbate confounding through inappropriate control. The term *overadjustment* is sometimes used to refer to bias induced by adjustment for intermediate variables or variables downstream from exposure – to use the language described in Chapter 2, to disrupt a chain from exposure to outcome. Adjusting for a variable that is on the causal pathway is an example of overadjustment (Schisterman et al., 2009). Overadjustment can also sometimes refer to a different problem: the bias (or loss of precision) that can occur in an analysis that controls for a strong predictor of exposure that is not associated with the outcome. In some settings, adjustment for a strong predictor of the outcome that is not associated with exposure also can induce a form of overadjustment bias, because such adjustment may push an estimate of the log odds ratio away from the null (Greenland et al., 2016). To help fully understand and discuss potential confounders, a diagram, such as a DAG, showing presumed causal relations among variables (and their measurements) can represent the assumed underlying causal associations and any confounding pathways implied (see Chapter 2), as shown in Example 3.4.

3.2.4 How confounders were measured and modelled

Consideration of how the confounders included in an adjustment set were measured and modelled is important because it relates to concerns about residual confounding by the factor after adjustment. Imperfect measurement of a confounding variable will usually lead to incomplete control of confounding (i.e. residual confounding) that is proportional to the amount of confounding originally present (Greenland, 1980; Greenland and Robins, 1985; Savitz and Barón, 1989; Ogburn and VanderWeele, 2012). For example, smoking may be imperfectly controlled in an analysis that classifies whether a person has ever smoked but does not account for whether the person is a current smoker or a former smoker or for the amount and duration of smoking. If the amount of original confounding was substantial, then – regardless of the fraction that was controlled – the amount that was not controlled may

Example 3.3. Adjustment for body mass index in studies on red meat consumption and colorectal cancer

In a meta-analysis (IARC, 2018), it was noted that many studies did not adjust for body mass index (BMI) because estimates of the association between red meat consumption and colorectal cancer (CRC) did not change after adjustment for BMI, although it is considered a potential confounder in the literature (Chan et al., 2011). Some may consider BMI to be a mediator on the pathway between red meat consumption and CRC (e.g. Example 2.1a), but in much of the literature BMI is considered to be a confounder that can affect both red meat consumption (those with higher BMI are likely to eat more red meat) and risk of colorectal cancer. For the Working Group's deliberations regarding the association between red meat consumption and colon cancer, the observation that inclusion of BMI in a regression model does not change the estimate of the association between red meat consumption and colon cancer suggests that BMI is neither an important mediator nor a confounder. (text continues above)

Example 3.4. Overadjustment as a concern in studies on shift work and cancer

In the *IARC Monographs* evaluation of the literature on night shift work in relation to breast cancer, the reviewers considered confounding and adjustment for other covariates (IARC, 2020). These considerations were particularly important because day workers are usually taken as the reference group and there may be many important differences between day workers and night workers with respect to risk factors for breast cancer. The Working Group consulted the literature to determine the degree to which lifestyle factors of day workers and night workers differ, to help in deciding whether a particular covariate was a potential confounder. For example, the reviewers noted that reproductive factors (parity, age at first birth, and menopause) are considered risk factors for breast cancer. They then cited studies that found differences between day workers and night workers with regard to reproductive factors; however, they noted that these associations were not strong. One could conclude that reproductive factors are potential confounders for the association between night shift work and breast cancer, but they are not likely to be strong confounders. Thus, a study that did not include reproductive factors may not suffer from much confounding bias. The reviewers also cited references indicating that several risk factors for breast cancer may be affected by night shift work, including disrupted sleep, physical activity, eating behaviours, and consumption of alcohol. Note that the total effect of night shift work on breast cancer includes the effect mediated by other factors. For example, perhaps night shift work increases the risk of breast cancer because people who work at night experience work-induced changes to exercise and diet that, in turn, lead to breast cancer. This does not imply that one would need to adjust for diet or exercise to obtain an unbiased result. In fact, the opposite is true: to obtain an unbiased result for the total effect of night work on breast cancer, one should not disrupt the causal chain by adjusting for behavioural factors that are affected by night work. (text continues on page 71)

still be important in absolute terms. Conversely, if a covariate is a weak confounder, residual confounding will have only a minor influence on the estimate of association. For example, smoking might be imperfectly controlled through next-of-kin reporting about whether a patient with breast cancer had ever smoked, but the residual confounding might be minor, given the weak smoking–breast cancer associations. Theoretically, in the extreme case of a very poorly measured confounder that suffers from systematic misclassification, adjustment for such an error-prone variable can make confounding worse (Ogburn and VanderWeele, 2012); however, such a scenario is typically implausible, and in most applications adjustment for an error-prone measure of a confounder will not make confounding worse (Greenland, 2012).

Importantly, some types of confounding are more difficult to control for than others. For example, specific exposures, such as tobacco smoking, lend themselves to careful measurement, whereas other factors that might confound an association of interest, such as socioeconomic conditions or health behavioural factors that influence exposure and cancer detection, are often almost impossible to measure well and fully control for in an analysis. If there are major differences at the outset (e.g. in a between-country comparison of breast cancer incidence rates), an investigator may have adjusted for a large set of covariates, yet the reviewers may remain sceptical that important confounding factors were adequately controlled. Another example of confounding that may be difficult to control arises in occupational studies when co-exposure occurs in the workplace to multiple correlated agents that could be carcinogenic (see Example 3.5).

(a) Time-varying confounders and time-varying confounders affected by prior exposure

So far, the discussion has been limited to confounding at one point in time, implying that study authors are interested in estimating the effect of an exposure that occurred at one time point on cancer. Many cohort studies involve the analysis of data

Example 3.5. Examining confounding by co-exposures in the workplace

In the *IARC Monographs* evaluation of the literature on night shift work in relation to breast cancer, the reviewers considered the association between occupational circadian rhythm disruption and breast cancer incidence among female flight attendants (IARC, 2020). Metrics of circadian rhythm disruption included employment duration, hours flying in the standard sleep interval, and number of time zones crossed. A potential confounder of concern was occupational exposure to cosmic radiation, which was highly correlated with employment duration. In the context of the IARC review, the concern was primarily with respect to positive confounding of the association between circadian rhythm disruption and breast cancer incidence; given the lack of observed association between circadian rhythm disruption and breast cancer incidence among female flight attendants (Pinkerton et al., 2016), the concern was not substantiated. However, in many situations when the primary exposure of interest is highly correlated with a potential confounder, reviewers may express concern about the ability to estimate the effect of the exposure of interest with adequate control for the confounding factor. (text continues on page 72)

on exposure at multiple different time points, and authors are frequently interested in the effect of lifetime or cumulative exposure (or possibly a lagged metric of cumulative exposure) on cancer. When exposures vary over time, so can confounding in an analysis that allows for time-dependent exposures. In the most straightforward scenario, a predictor of the outcome might also predict exposure at each time point. For example, in studies of occupational exposure on cancer, age could be a time-dependent confounder. In this situation, time-varying confounders can usually be treated in a standard manner; for example, in a regression model, a term for the confounder might be included at each time point.

However, some time-varying confounders can be affected by prior exposure. Reviewers of papers can evaluate the plausibility of confounders at a given point in time being influenced by prior exposure. Consideration of whether time-varying confounders are affected by prior exposure is important because in such situations standard outcome modelling of the associations is not guaranteed to yield unbiased results, as shown in Example 3.6 (Cook et al., 2002; Hernán and Robins, 2023).

3.3 Tools for assessing bias due to confounding

Control for confounding is rarely, if ever, sufficient to remove bias entirely. Rather, control is a matter of degree and often warrants a critical assessment of whether the control achieved in a published paper may be adequate to make a reasonable judgement regarding the effect of the exposure on cancer. After having reviewed the control for confounding in a study, a reviewer might suspect that the published analysis suffers from substantial confounding by uncontrolled covariates or suffers from residual confounding due to inadequately controlled covariates (e.g. confounders that were poorly measured, inadequately modelled, or poorly specified).

Given concern about possible confounding of an exposure–cancer association that was reported in an individual study, the next step is to assess the direction and magnitude of the confounding bias. This can help to understand the impact of uncontrolled confounding on the evidence under review.

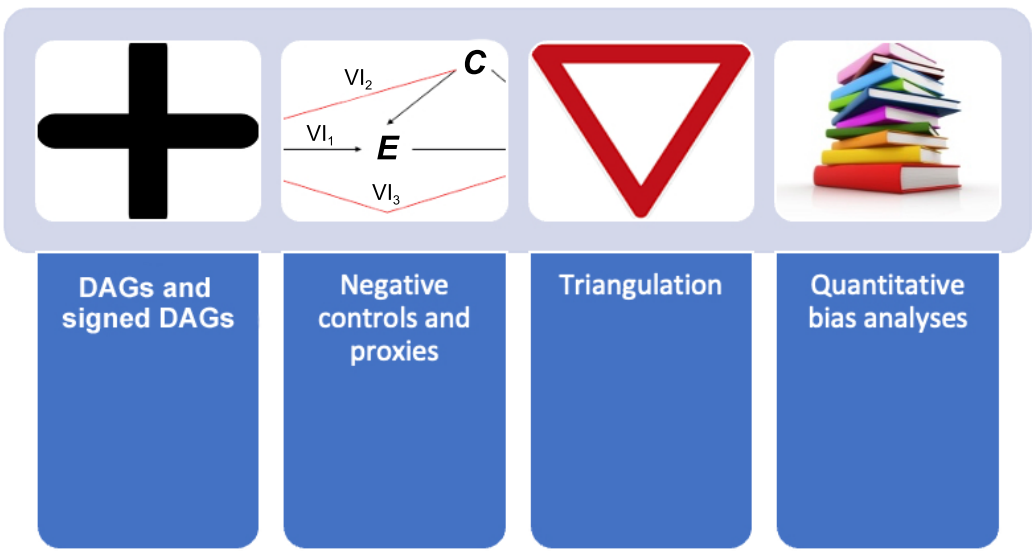
Various approaches (tools for assessing confounding, numbered C-# below) are available to inform evaluations of confounding by unmeasured variables of an observed association between cancer and exposure to an agent under evaluation in an IARC review. Investigators should consider the following, which will be developed further in subsequent subsections (Fig. 3.2):

- Tool C-1: DAGs and signed DAGs (i.e. causal relations between variables based on substantive knowledge);
- Tool C-2: negative control outcomes (or exposures) and proxies (evidence of confounding within a study);
- Tool C-3: triangulation (evidence of confounding between studies that differ meaningfully); and

Example 3.6. Healthy worker survivor bias

In occupational studies of carcinogens, a common concern is a form of confounding that is often referred to as healthy worker survivor bias, whereby workers who are less susceptible to the health effects of the exposure survive longer in the workplace and therefore may accrue more cumulative exposure. In this instance, leaving work is a time-varying confounder that is affected by prior exposure because (i) leaving work may predict cancer diagnosis (e.g. if a person left work for cancer-related reasons), (ii) leaving work will affect accrual of occupational exposure, and (iii) prior exposure to hazardous material may affect a worker's current employment status. Such a bias may occur if people who work nights quit when they cannot tolerate the lifestyle anymore. If years of night work have already taken a health toll on the worker in ways that are on the pathway to cancer, there may be time-varying confounding affected by prior exposure. Including time-varying explanatory variables in an outcome regression model for factors such as whether individuals are currently employed or the duration of each person's employment will not remove the bias. Special methods, known collectively as g-methods (generalized methods), are needed to address this issue and produce unbiased estimates (Robins, 1986; Hernán and Robins, 2006, 2023; Buckley et al., 2015). These g-methods enable researchers to model long-term exposure in a different way from ordinary outcome regression modelling. For example, rather than estimating risk from cumulative exposure over many years, g-computation, one type of g-method, estimates the risk of cancer from exposure in each year and then sums up the risks of cancer over time. This approach allows the researcher to adjust only for confounders that precede exposure in each year, thus addressing the bias from the healthy worker survivor effect. Unless g-methods are used, estimates of occupational exposure–cancer associations from studies that are affected by healthy worker survivor bias will typically be attenuated (i.e. biased downwards). (text continues on page 73)

Fig. 3.2. Tools to consider when evaluating the impact of probable confounding. Each of the approaches proposed requires substantive expertise, which may include expert judgement, information derived from internal substudies, or findings from external studies.



- Tool C-4: quantitative bias analyses.

For simplicity, the focus in this chapter is on assessment of the impact of a single primary confounder of concern. Often, for clarity in a review, it is useful to focus on assessment of the impact of one key potential confounder of concern at a time. Section 6.4 discusses some approaches for multiple-bias analysis, where more than one confounder (or other source of bias) is of concern.

Most of the approaches described here are premised on the ability to explicitly name a factor of concern as a confounder. This requires hypothesizing why that factor is associated with both the exposure and the outcome. Like substantive hypotheses, hypotheses about why a confounding factor is associated with exposure, and with disease, should be specific, should describe substantively important associations, and should make quantitative predictions of the confounding effect (Hertz-Picciotto, 2000). Given well-specified hypotheses about confounding of observed associations, a reviewer may be able to assess the degree to which results from observed data are likely to be substantially affected by the hypothesized confounding. A review is strengthened by explaining which factors were considered as potential confounders and why, as well as their likely effects (see Example 3.7).

For known confounders that have not been measured, it may be feasible to perform a bias analysis to suggest the possible effect of the unmeasured confounder. Of course, it might be the case that a reviewer does not wish to posit (or name) a specific confounder but rather wishes only to express a general concern that an observed association between exposure and disease might be confounded by a factor as yet unknown (at least to the investigator). A general concern about uncontrolled confounding might arise if a reviewer were to conclude that the important risk factors for a given cancer outcome have simply not yet been identified. In general, vague statements regarding entirely unknown confounders are less amenable to evaluation using most of the approaches described here. The less that is understood about disease etiology or exposure assignment, the greater the potential for unknown factors to be important confounders. The latter threat to validity can be minimized by focusing a hazard identification on a well-defined exposure (e.g. benzene) rather than a vague exposure or contextual factors (e.g. green space) (Hernán, 2016). Again, a Working Group's discussion of the role of confounding when evaluating evidence regarding the carcinogenicity of an agent will be most informative when the confounding factor is explicitly named, and when hypotheses regarding why that factor is associated with exposure and disease can be discussed and evaluated. Quantitative bias analysis can be used to assess whether the study results are sufficiently robust to render uncontrolled confounding unlikely (see Section 3.3.4).

3.3.1 Tool C-1: DAGs

As noted in Chapter 2, a simple DAG can serve as a starting point for the analysis of uncontrolled confounding. The drawing of a DAG requires substantive knowledge about covariates and their causal relations to the exposure and outcome of interest. Without such substantive knowledge, a DAG is largely speculative. Although a DAG is not an oracle that can provide infallible identification of confounding in a particular study, given substantive expertise (which often exists in expert Working Groups), it can be useful for reasoning about systematic bias and making the causal assumptions of Working Group members involved in an *IARC Monographs* evaluation explicit and clear.

A signed DAG (i.e. one in which the direction of the effect of a confounder is specified) can aid Working Group members in assessing the probable direction of bias due to confounding (see Section 2.6). Also, DAGs can inform the assessment of time-varying confounders. For

Example 3.7. Relative importance of confounders

In its examination of the carcinogenicity of red meat (IARC, 2018), the Working Group specified which confounders were thought to be important (physical activity, BMI, caloric intake) and gave more weight to studies that controlled for these confounders (or that demonstrated that adjustment for the covariate of concern did not have meaningful impact on the estimate of interest). (text continues above)

example, in an occupational cohort mortality study to investigate a suspected carcinogen, healthy worker survivor bias is a common concern. A signed DAG can help to judge whether this form of confounding is likely to be present; relevant considerations include the need for associations between (i) prior exposure and employment status and (ii) employment status and mortality (Naimi et al., 2013). A DAG can help to answer these questions and guide a reviewer's assessment of the likelihood of such bias.

3.3.2 Tool C-2: negative control outcomes (or exposures) and proxies

Sometimes a reviewer is able to indirectly assess confounding by an unmeasured factor using evidence available from within the published study, based on approaches that involve negative controls and proxies. These methods all share similar assumed causal structures between variables (Fig. 3.3). However, as discussed next (Sections 3.3.2(a) to 3.3.2(c)), Working Group members may find useful conceptual distinctions between negative control outcomes, negative control exposures, and proxies for an unmeasured confounder.

(a) Negative control outcomes

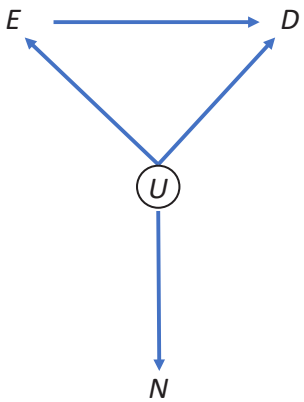
Suppose that a reviewer is concerned about potential confounding in a cohort study of the association between a suspected carcinogen and a site-specific cancer, but that the potential confounder was unmeasured in the study under review. A negative control outcome approach proceeds by examining the association between the suspected carcinogen and another outcome that (i) is caused by the hypothesized confounding factor and (ii) is not caused by the suspected carcinogen of interest.

Fig. 3.3 illustrates the causal associations described: *E* denotes the exposure of interest, *D* the outcome, *U* the unmeasured confounder, and *N* the negative control outcome. Note that *U* has a causal effect on *N* but *E* does not.

Under these conditions, an observed association between *E* and *N* would be entirely due to confounding by *U*. Therefore, the absence of an association between *E* and *N* would argue against the hypothesis that the *E*–*D* association is confounded by *U* (Example 3.8).

This approach is well suited to the evaluation of cohort studies where information on many outcomes (e.g. cause-specific mortality) has been collected; this may enable an investigator to examine not only the association between the exposure of interest and the outcome of primary interest but also the association between that exposure and an outcome that a reviewer posits as a useful negative control outcome. Absence of evidence of an association between *E* and *N* would help to nullify claims of confounding by *U*. This can be thought of as an example of internal (i.e. within-study) triangulation of evidence, where the examination of associations between exposure and outcomes with different presumed causal structures can be compared to indirectly assess bias (Pearce et al., 1986).

Fig. 3.3. Diagram for analyses involving a negative control or proxy, *N*. *E*, exposure; *D*, outcome; *U*, confounder.



Example 3.8. Negative control outcomes

In an investigation of the effect of red meat consumption (*E*) on cancer (*D*), where tobacco smoking (*U*) is not measured but is considered a potential confounder, an investigator might posit emphysema as a valid negative control outcome (*N*). If that assumption were correct, the absence of an association between red meat consumption and emphysema would be evidence that tobacco is not a confounder of a red meat consumption–cancer association. (text continues above)

A related approach is sometimes used in the interpretation of standardized mortality ratio (SMR) analyses to, as it were, correct SMRs for bias. The SMR for an outcome that is presumed to be susceptible to the same confounding factors as the outcome of primary interest, but is presumed not to be strongly associated with the exposures of interest, serves as a measure of the bias due to confounding. This approach has been used both qualitatively, to indirectly assess confounding when interpreting cause-specific SMRs, and quantitatively, to derive an adjusted SMR for the outcome of interest (and associated confidence interval) by taking a ratio of the measures.

As shown in Example 3.9, expert groups can quantitatively evaluate uncontrolled confounding by calculating an adjusted SMR using published results if appropriate negative control outcomes can be identified (Side Box 3.3) and are reported.

(b) Negative control exposures

Suppose that a reviewer is concerned about potential confounding of the association between a suspected carcinogen and a site-specific cancer, but that the potential confounder was unmeasured in the study under review. A negative control exposure approach proceeds by examining the association of the site-specific cancer outcome of interest with another exposure variable that (i) is associated with the hypothesized confounding factor and (ii) is not a cause of the site-specific cancer outcome of interest.

Fig. 3.3 can also illustrate the causal associations required for a valid negative control exposure if *N* is now taken to denote the negative control exposure: *N* shares common cause *U* with *E*, but *N* does not cause *D*.

Under these conditions, an observed association between *N* and *D*, adjusted for *E* (or within a stratum of *E*), would be entirely due to confounding by *U*, whereas the absence of such an association between *N* and *D* would be evidence against confounding of the *E*–*D* association by *U* (Example 3.10 and Side Box 3.4).

(c) Proxies for a confounder

Proxies are indirect measures of unavailable variables of interest; this chapter focuses on proxies that are used as surrogates for potential

Example 3.9. Indirect adjustment of SMRs to reduce healthy worker biases in aluminium smelting work

In a study of bladder cancer among workers in an aluminium smelting plant, confounding through healthy worker biases was a concern (McClure et al., 2020). The investigators quantitatively evaluated healthy worker effects through negative control outcomes and derived an adjusted SMR. They did this by selecting a group of diseases (e.g. non-malignant blood disorders, diabetes, psychological disorders) that satisfied the conditions of a negative control outcome because they were thought to be unaffected by smelting work exposure but would be affected by healthy worker effects in a fashion similar to bladder cancer. The unadjusted SMR for bladder cancer was 2.27, and the unadjusted SMR for the negative control group was 0.65. The adjusted SMR, derived by taking the ratio of the two SMRs, was 3.47; this indicated that the confounding from healthy worker effects downwardly biased the SMR for bladder cancer. (text continues above)

Side Box 3.3. Information needed to facilitate use of negative control outcomes to evaluate confounding

Several elements are required to use negative control outcomes to evaluate confounding. The first requirement is for a suitable negative control outcome, i.e. an outcome that is related to the confounder but is not caused by exposure to the agent under evaluation. Notably, the negative control outcome may be identified by the expert reviewer but not by the original researchers. Required results include the association between the agent of interest and the negative control outcome, as well as the primary association between the agent and the outcome of interest. (text continues above)

Example 3.10. Negative control exposures

In studies that assess exposure information by questionnaire, investigators will often include questions about exposure to agents that are thought to be unrelated to the outcome of interest; these may serve as negative control exposures. (text continues on page 77)

Side Box 3.4. Information needed to facilitate use of negative control exposures to evaluate confounding

Several elements are required to use negative control exposures to evaluate confounding. The first requirement is for a suitable negative control exposure, i.e. an exposure that is related to the confounder but is not a cause of the disease outcome under evaluation. As with the negative control outcome, the negative control exposure may be identified by only the expert reviewers. Required results include the negative control exposure–disease association, adjusted for exposure 1 (or negative control exposure–disease association within a stratum of exposure 1) between the negative control exposure and the outcome of interest, adjusted for the exposure of interest (or the negative control exposure association with the disease of interest within a stratum of the main exposure of interest), as well as the primary association between the agent and the outcome of interest. (text continues on page 77)

confounders. Here, a proxy is taken to be a variable associated with an uncontrolled confounder *U* that would be irrelevant for confounding adjustment had *U* been measured and controlled for (Example 3.11).

A valid proxy for a confounding variable should (i) be associated with the hypothesized confounding factor *U* after controlling for exposure and (ii) not be associated with the outcome of interest except via *U*. Fig. 3.3 illustrates an example of causal associations required for a valid proxy, where *N* is now the proxy for *U* (Lipsitch et al., 2010).

Sometimes results are reported with stratification or restriction on a proxy variable in the form of subgroup analyses, in which strata were defined by a measured proxy variable. In other situations, results are reported with regression model adjustment for a proxy variable (sometimes results are reported with and without adjustment for a covariate that is a proxy for the confounder).

Example 3.12 illustrates the point that an analysis restricted to one level of a valid proxy variable (e.g. in which there is presumed to be little variation in the confounder *U*) might be viewed as less susceptible to confounding by *U*. However, as noted previously regarding residual confounding, the degree to which the proxy variable is a good surrogate for the unmeasured confounder will affect the degree by which confounding by *U* is minimized (Ogburn and VanderWeele, 2012; Ogburn et al., 2021). Moreover, the degree of residual bias that remains is typically proportional to the amount of confounding originally present (Greenland and Robins, 1985; Savitz and Barón, 1989).

3.3.3 Tool C-3: triangulation across studies

As described in Chapter 1, an *IARC Monographs* evaluation of an agent typically involves comparing findings across studies; this permits consideration of results across a set of studies that may differ in control for a confounder of concern within the wider context of the strengths and limitations of the available studies. The term *triangulation* is used to describe a variety of approaches in which analysts use different types of evidence from different study designs or types that have different identifying conditions; these approaches leverage variation between studies, focusing on settings in which biases vary across study types. Triangulation involves comparing results for a common effect from two or more studies that are

Example 3.11. Using a proxy variable to evaluate confounding in a cohort of Seventh Day Adventist adherents

A Working Group can evaluate concern about confounding by smoking if the reported results include analyses restricted to one level of a variable that is a proxy for smoking (the unmeasured potential confounder). An example is the study of chronic disease in the Adventist Health Study cohort, in which recruitment is restricted to a religious group who mostly do not smoke, to serve as a proxy for not smoking (Butler et al., 2008). (text continues on page 78)

Example 3.12. Restriction to one level of a proxy variable to examine residual confounding

Sheikh et al. (2020) examined the association between opium use (*E*) and oesophageal cancer (*D*) in the Islamic Republic of Iran; the Working Group discussed concerns about potential residual confounding by tobacco use (*U*). Sex was a measured variable in the study; it is presumed to be associated with tobacco smoking, because tobacco use is very rare among women in this population. A Working Group could consider sex as a proxy variable to indirectly assess residual confounding of the association between opium use and oesophageal cancer by smoking. In an analysis restricted to women, a positive association between opium use and oesophageal cancer was observed, and the association observed among women was similar in magnitude to that observed among men. Results conditioned on sex, if sex is considered a valid proxy for smoking, should be less susceptible to confounding by smoking. In this example, results suggested that the (sex- and smoking-adjusted) association between opium use and oesophageal cancer was unlikely to be substantially biased by residual confounding by tobacco smoking. (text continues on page 78)

thought to differ in susceptibility to confounding, or where the presumed confounder is thought to act in opposing directions; deliberate use could be made of studies conducted in contexts with differing confounding structures (Lawlor et al., 2016). Triangulation between covariate-adjusted analyses and instrumental variable analyses (such as Mendelian randomization studies) can offer some insight into whether the covariate-adjusted studies are likely to be confounded, because of the different identifying conditions required for covariate-adjusted analyses and Mendelian randomization studies (Example 3.13). Notably, there are also more advanced methods, such as multivariable Mendelian randomization, that adjust for known confounders to test the independence assumption in the Mendelian randomization studies (Brookhart et al., 2010; Burgess and Thompson, 2015). For further discussion of the use of triangulation in evidence synthesis, see Chapter 6.

Insight into possible bias can also be obtained by comparing results from two or more studies that are thought to differ in susceptibility to confounding. For example, a reviewer may raise a concern about a confounder that is uncontrolled in one or more studies (e.g. no control for smoking in studies among workers exposed to diesel fumes where lung cancer is the outcome) but observe that other studies of the same association reported similar results after adjustment for smoking (e.g. Bhatia et al., 1998). This offers another possible method to assess confounding; however, such simple comparisons across studies may not be valid. Rather, it would be surprising if the bias in one study applied perfectly to other studies (or even to other study samples drawn from the same source population). Confounding is seldom, if ever, the only bias of concern. When multiple biases are present, comparison between studies becomes more difficult (see Section 6.3 for further discussion and examples).

Example 3.13. Evidence triangulation to evaluate confounding

The *IARC Monographs* Volume 124 on night shift work found *limited* evidence that night shift work causes cancer in humans, with convincing evidence that it disrupts circadian rhythms (IARC, 2020). In that review, an example of triangulation between covariate-adjusted analyses and an instrumental variable analysis was discussed. A multivariable regression analysis demonstrated that, when examining chronotype (morning or evening preference) as a measure of circadian rhythm, morning preference was inversely associated with breast cancer incidence among participants in the UK Biobank study. The investigators identified genetic variants related to chronotype and undertook a Mendelian randomization study of chronotype and breast cancer incidence; they found a protective effect of morning preference on breast cancer risk (Richmond et al., 2019). This lends indirect support to the hypothesis that shift work is related to cancer risk because it disrupts this biological pathway. (text continues on page 79)

3.3.4 Tool C-4: bias adjustment

Investigators may be concerned about confounding by an unmeasured variable (the total or residual confounding) or the confounding produced by specific unmeasured variables. In the latter case, suppose that a reviewer has drawn a simple signed DAG for posited confounder–exposure and confounder–disease associations, implying potential bias in the study under review.

It is then necessary to assess how large this bias is likely to be, relative to the observed exposure–disease association. In assessing the potential impact of an unmeasured or incompletely adjusted confounder, reviewers may be able to estimate the size of the bias induced and decide whether it is indeed relevant. A variety of methods are available to quantitatively assess confounding under specified scenarios (or to identify bounds on bias due to an unmeasured confounder). Not all proposed methods are reviewed here; only a few approaches that are well suited to the *IARC Monographs* process are highlighted. Although subject matter knowledge is necessary, it need not be certain or complete; a range of values can be examined to assess plausible scenarios.

In the following subsections, many of the quantitative bias analyses are framed to guide judgement regarding whether a published estimate of association could plausibly be attributed entirely to an unmeasured confounder. As noted in Chapter 1, this reflects one of the primary questions posed to experts involved in an *IARC Monographs* review: can confounding reasonably be ruled out as an explanation for all of an observed exposure–cancer association? Simple expressions (and spreadsheet calculators) are also provided to facilitate the assessment of a range of bias. The focus throughout is on a single unmeasured confounder of primary concern; in Chapter 6, methods are extended to address multiple-bias analysis.

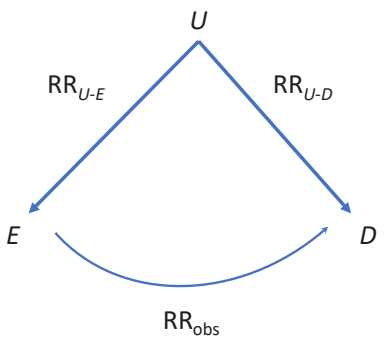
(a) Bounding

Concern about potential unmeasured confounders is often focused first on established cancer risk factors that have a strong independent association with the cancer of interest. This is because the understanding of strong risk factors for cancer outcomes is often better than that of the determinants of exposure. For simplicity, let us focus on settings where the hypothesized confounder increases the risk of cancer (i.e. $RR_{U\text{-}D} \geq 1$), where $RR_{U\text{-}D}$ denotes the magnitude of the confounder–outcome relative risk (this magnitude is typically estimated from prior information; Fig. 3.4).

Key message

If $RR_{U\text{-}D}$ is less than RR_{obs}, the reported relative risk between the exposure and the outcome in the study under review, then confounding by *U* cannot entirely explain an observed association.

Fig. 3.4. Diagram for analyses involving bounding and correction for the effect of a confounder.



Example 3.14. Use of bounding to examine confounding scenarios

Suppose that, in a study under review, it was reported that the observed association between opium use and laryngeal cancer (unadjusted for tobacco use) was RR_{obs} = 2.0. Suppose that confounding by tobacco smoking is of concern but had not been assessed in the study. On the basis of prior literature (Bakhshaee et al., 2017; Alizadeh et al., 2020), it can be hypothesized that the smoking–laryngeal cancer association in the study population was no larger than $RR_{U\text{-}D}$ = 5.5. In that situation, bounds on the smoking-adjusted association between opium use and laryngeal cancer are [0.36, 2.00]. With these assumptions, a reviewer could conclude that the observed association between opium use and laryngeal cancer could be due to confounding by smoking. (text continues below)

As shown in Example 3.14, if we know just the magnitude of the confounder–outcome association, $RR_{U\text{-}D}$, then, given a reported association,

$$RR_{obs} \times (1/RR_{U\text{-}D}) = \text{lower bound} \quad (3.1)$$

$$RR_{obs} \times 1 = \text{upper bound} \quad (3.2)$$

Key message

From these expressions, it follows that if the association between the confounder and outcome is small (i.e. $RR_{U\text{-}D}$ is close to 1) then the amount of uncontrolled bias from this confounder is also likely to be small.

RR_{obs}, it is possible to identify bounds (under a worst-case scenario, in which all the exposed have the confounder but none of the unexposed has the confounder) on the association of interest after adjustment for *U* (Flanders and Khoury, 1990):

Alternatively, if the magnitude of the confounder–exposure relative risk ($RR_{U\text{-}E}$) is less than RR_{obs}, then confounding by *U* cannot entirely explain an observed association. In other words, for confounding to entirely explain the observed association, both of the underlying associations ($RR_{U\text{-}E}$ and $RR_{U\text{-}D}$), not just one of them, must be larger than the published relative risk estimate, RR_{obs} (Cornfield et al., 1959). More informative bounds can be obtained using these two pieces of information ($RR_{U\text{-}E}$ and $RR_{U\text{-}D}$) (Flanders and Khoury, 1990; VanderWeele and Ding, 2017).

(b) Bias adjustment

As shown in Example 3.15, a simple bias-adjusted (Bross, 1966; Axelson, 1978; Schlesselman, 1978) estimate of the association can be derived, based on posited values for the strength of the confounder–outcome ($RR_{U\text{-}D}$) association and the prevalence of the confounder among the unexposed ($p_0 = \Pr[U = 1 \mid E = 0]$) and the exposed ($p_1 = \Pr[U = 1 \mid E = 1]$):

$$RR_{adj} = RR_{obs} \frac{RR_{U\text{-}D}\, p_0 + (1 - p_0)}{RR_{U\text{-}D}\, p_1 + (1 - p_1)} \quad (3.3)$$

Either unique values for p_0 and p_1 can be posited, along with the confounder–disease association ($RR_{U\text{-}D}$), or a range of plausible values for each can be posited and a distribution developed of the probable effects of bias due to an unmeasured confounder, using either Monte Carlo simulations or Bayesian priors (Steenland and Greenland, 2004). If the prevalences of the confounder among the unexposed and the exposed are not known, a Working Group member might take the latter approach to investigate what prevalence of smoking would be needed to entirely explain the observed association, and then consider the plausibility of such a pattern in the study population. Implementation of such calculations in a spreadsheet facilitates exploration (Fox et al., 2021).

(c) Unknown uncontrolled confounders and E*-values*

If a concern is expressed about an unknown confounder, a reviewer might undertake a quantitative bounding analysis, following the principles outlined previously in Section 3.3.4(a). Such an evaluation could be considered when doubts remain about causality, despite the lack of an identified confounder. For example, if it is arbitrarily assumed that the magnitudes of the associations of the confounder with exposure and outcome are equal on a risk-ratio scale (i.e. $RR_{U\text{-}E} = RR_{U\text{-}D}$) then, for an observed positive exposure–outcome association to be entirely due to a confounder *U*, $RR_{U\text{-}E}$

Example 3.15. Bias adjustment to evaluate confounding

Consider the possibility of unmeasured smoking as a potential confounder in a study of opium use and lung cancer. Suppose that the prevalence of smoking in the unexposed is 20%, the prevalence in the exposed is 30% ($RR_{U\text{-}E}$ = 1.5), smoking has a hypothesized $RR_{U\text{-}D}$ of 10, and the observed relative risk for opium and lung cancer is 2.0 (exposed versus unexposed). Let p_0 be the proportion of smokers among the unexposed and p_1 be the proportion of smokers among the exposed. The risk of lung cancer among those unexposed due solely to smoking will be a weighted average of the risks of lung cancer in non-smokers and smokers, i.e. $RR_{U\text{-}D}\ p_0 + (1 - p_0)$, and the risk of lung cancer among the exposed, due to smoking alone, is $RR_{U\text{-}D}\ p_1 + (1 - p_1)$. The relative risk of exposed versus unexposed, due to smoking alone, is $[RR_{U\text{-}D}\ p_1 + (1 - p_1)]/[RR_{U\text{-}D}\ p_0 + (1 - p_0)]$, and we can adjust the observed relative risk due to opium by this factor to indirectly adjust for the estimated confounding by smoking (Flanders and Khoury, 1990). (text continues on page 81)

$$\mathrm{RR_{adj}} = \mathrm{RR_{obs}}\frac{\mathrm{RR}_{U-D}p_0 + (1-p_0)}{\mathrm{RR}_{U-D}p_1 + (1-p_1)} = \mathrm{RR_{obs}}\frac{10(0.2) + (1-0.2)}{10(0.3) + (1-0.3)} = \mathrm{RR_{obs}}\frac{2.8}{3.7} \quad \text{(E3.1)}$$

For a simple bias adjustment, as given by this equation, one can correct the observed risk ratio for the potential confounding; if the observed risk ratio were 2.00, the adjusted risk ratio would be

$$\frac{2.8}{3.7} \times 2.0 = 1.5 \quad \text{(E3.2)}$$

and $RR_{U\text{-}D}$ must equal, or exceed, RR_{obs} + sqrt[RR_{obs} × (RR_{obs} − 1)], a quantity that has been termed the *E*-value (VanderWeele and Ding, 2017). Note that this value is derived using just the observed (potentially confounded) association between agent and outcome, RR_{obs}, without specification of the confounder–outcome or confounder–exposure association (other than assuming that they are equal). It also unrealistically assumes that the prevalence of the uncontrolled confounder among the exposed is 100% or, equivalently, that the prevalence of the exposure among those without the confounder is 0%, and hence can be misleadingly small compared with what is needed for an actual confounder to fully explain the magnitude of RR_{obs} (MacLehose et al., 2021), as shown in Example 3.16.

Bias analyses (Flanders and Khoury, 1990; Lash et al., 2009; Fox et al., 2021; MacLehose et al., 2021) allow one to relax the assumptions used by the *E*-value that $RR_{U\text{-}E}$ equals $RR_{U\text{-}D}$ and that the prevalence of the confounder is 100% among the exposed.

3.4 Summary

Confounding is typically of concern in observational studies. Expert reviewers can assess the impact of confounding on the observed exposure–cancer association in several ways. Some study designs can minimize confounding, for example by matching on probable confounders ahead of time. In other studies, the investigators will have measured potential confounders and controlled for them in the design or analysis.

Key message

When there is concern about unknown confounders, a quantitative bounding analysis, as discussed previously in Section 3.3.4(a), can clarify what magnitudes of confounder–disease association, and what prevalences of confounder among exposed and unexposed, would be needed to entirely explain an observed exposure–disease association (see Side Box 3.5).

Example 3.16. The *E*-value to evaluate confounding

The reported association between opium use and oesophageal cancer (unadjusted for some unknown confounder *U*) was RR_{obs} = 2.0 (Example 3.15). Suppose that confounding by the unknown confounder *U* is suspected. The resultant *E*-value would take a value of 2 + sqrt[2 × (2 − 1)] = 3.4, meaning that if a reviewer posited that $RR_{U\text{-}E}$ and $RR_{U\text{-}D}$ were positive, equal, and both less than 3.4, it could be concluded that confounding by *U* could not entirely explain the observed positive exposure–disease association.

However, a reviewer might assume the confounder–opium use association, $RR_{U\text{-}E}$, to be larger than 3.4. This illustrates one important caution concerning interpretation of the *E*-value: although it might be tempting to say that both associations need to be at least as large as the *E*-value, that is incorrect. In fact, $RR_{U\text{-}D}$ could be less than the *E*-value, while $RR_{U\text{-}E}$ could be substantially larger than the *E*-value, allowing for confounding to completely explain the association. Conversely, both $RR_{U\text{-}D}$ and $RR_{U\text{-}E}$ could be substantially larger than 3.4 and still not completely explain the association, for the simple reason that the unknown confounder *U* could have a prevalence substantially less than 100% among opium users. (text continues on page 82)

Side Box 3.5. Information needed to facilitate use of bias assessment to evaluate confounding

For bounding approaches, the original (or associated) studies should report the value of the probable magnitude of association of the confounder with the outcome of interest in the population under study, and the association of the confounder with the exposure of interest.

For quantitative bias assessment, the original or associated studies should also report, more specifically, the prevalence of the confounder among those unexposed (p_0) and exposed (p_1) to the agent of interest. (text continues on page 82)

In both settings, reviewers will want to consider whether the confounder was well measured and controlled (i.e. whether residual confounding is likely to remain). Reviewers may also consider whether, based on the literature, there are likely to be important unmeasured confounders.

If potential confounders were not measured or were inadequately controlled in a study, then reviewers need to make informed judgements about the direction of residual confounding and its probable magnitude, and, in particular, the extent to which residual confounding could explain the observed exposure–disease association. The reliability of such judgements will be greatly improved to the extent that they make use of background information about the relations of uncontrolled potential confounders to the exposure and disease under study, and the results of other studies that did control for those potential confounders.

References

Alizadeh H, Naghibzadeh Tahami A, Khanjani N, Yazdi-Feyzabadi V, Eslami H, Borhaninejad V, et al. (2020). Opium use and head and neck cancers: a matched case-control study in Iran. Asian Pac J Cancer Prev. 21(3):783–90. doi:10.31557/APJCP.2020.21.3.783 PMID:32212808

Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, et al.; 1000 Genomes Project Consortium (2015). A global reference for human genetic variation. Nature. 526(7571):68–74. doi:10.1038/nature15393 PMID:26432245

Axelson O (1978). Aspects on confounding in occupational health epidemiology. Scand J Work Environ Health. 4(1):98–102. doi:10.5271/sjweh.2720 PMID:644270

Bakhshaee M, Raziee HR, Afshari R, Amali A, Roopoosh M, Lotfizadeh A (2017). Opium addiction and risk of laryngeal and esophageal carcinoma. Iran J Otorhinolaryngol. 29(90):19–22. PMID:28229058

Bhatia R, Lopipero P, Smith AH (1998). Diesel exhaust exposure and lung cancer. Epidemiology. 9(1):84–91. doi:10.1097/00001648-199801000-00017 PMID:9430274

Brookhart MA, Rassen JA, Schneeweiss S (2010). Instrumental variable methods in comparative safety and effectiveness research. Pharmacoepidemiol Drug Saf. 19(6):537–54. doi:10.1002/pds.1908 PMID:20354968

Bross ID (1966). Spurious effects from an extraneous variable. J Chronic Dis. 19(6):637–47. doi:10.1016/0021-9681(66)90062-2 PMID:5966011

Buckley JP, Keil AP, McGrath LJ, Edwards JK (2015). Evolving methods for inference in the presence of healthy worker survivor bias. Epidemiology. 26(2):204–12. doi:10.1097/EDE.0000000000000217 PMID:25536456

Burgess S, Thompson SG (2015). Multivariable Mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects. Am J Epidemiol. 181(4):251–60. doi:10.1093/aje/kwu283 PMID:25632051

Butler TL, Fraser GE, Beeson WL, Knutsen SF, Herring RP, Chan J, et al. (2008). Cohort profile: the Adventist Health Study-2 (AHS-2). Int J Epidemiol. 37(2):260–5. doi:10.1093/ije/dym165 PMID:17726038

Chan DS, Lau R, Aune D, Vieira R, Greenwood DC, Kampman E, et al. (2011). Red and processed meat and colorectal cancer incidence: meta-analysis of prospective studies. PLoS One. 6(6):e20456. doi:10.1371/journal.pone.0020456 PMID:21674008

Cole SR, Platt RW, Schisterman EF, Chu H, Westreich D, Richardson D, et al. (2010). Illustrating bias due to conditioning on a collider. Int J Epidemiol. 39(2):417–20. doi:10.1093/ije/dyp334 PMID:19926667

Cook NR, Cole SR, Hennekens CH (2002). Use of a marginal structural model to determine the effect of aspirin on cardiovascular mortality in the Physicians' Health Study. Am J Epidemiol. 155(11):1045–53. doi:10.1093/aje/155.11.1045 PMID:12034583

Cornfield J, Haenszel W, Hammond EC, Lilienfeld AM, Shimkin MB, Wynder EL (1959). Smoking and lung cancer: recent evidence and a discussion of some questions. J Natl Cancer Inst. 22(1):173–203. PMID:13621204

Davey Smith GD, Lawlor DA, Harbord R, Timpson N, Day I, Ebrahim S (2007). Clustered environments and randomized genes: a fundamental distinction between conventional and genetic epidemiology. PLoS Med. 4(12):e352. doi:10.1371/journal.pmed.0040352 PMID:18076282

Flanders WD, Khoury MJ (1990). Indirect assessment of confounding: graphic description and limits on effect of adjusting for covariates. Epidemiology. 1(3):239–46. doi:10.1097/00001648-199005000-00010 PMID:2081259

Fox MP, MacLehose RF, Lash TL (2021). Misclassification. In: Applying quantitative bias analysis to epidemiologic data. Statistics for biology and health. 2nd ed. Cham, Switzerland: Springer; pp. 141–195. doi:10.1007/978-3-030-82673-4_6

Frisell T, Öberg S, Kuja-Halkola R, Sjölander A (2012). Sibling comparison designs: bias from non-shared confounders and measurement error. Epidemiology. 23(5):713–20. doi:10.1097/EDE.0b013e31825fa230 PMID:22781362

Greenland S (1980). The effect of misclassification in the presence of covariates. Am J Epidemiol. 112(4):564–9. doi:10.1093/oxfordjournals.aje.a113025 PMID:7424903

Greenland S (2012). Intuitions, simulations, theorems: the role and limits of methodology. Epidemiology. 23(3):440–2. doi:10.1097/EDE.0b013e31824e278d PMID:22475828

Greenland S, Mansournia MA, Altman DG (2016). Sparse data bias: a problem hiding in plain sight. BMJ. 352:i1981. doi:10.1136/bmj.i1981 PMID:27121591

Greenland S, Pearl J, Robins JM (1999b). Causal diagrams for epidemiologic research. Epidemiology. 10(1):37–48. doi:10.1097/00001648-199901000-00008 PMID:9888278

Greenland S, Robins JM (1985). Confounding and misclassification. Am J Epidemiol. 122(3):495–506. doi:10.1093/oxfordjournals.aje.a114131 PMID:4025298

Greenland S, Robins JM, Pearl J (1999a). Confounding and collapsibility in causal inference. Stat Sci. 14(1):29–46. doi:10.1214/ss/1009211805

Hernán MA (2016). Does water kill? A call for less casual causal inferences. Ann Epidemiol. 26(10):674–80. doi:10.1016/j.annepidem.2016.08.016 PMID:27641316

Hernán MA, Robins JM (2006). Estimating causal effects from epidemiological data. J Epidemiol Community Health. 60(7):578–86. doi:10.1136/jech.2004.029496 PMID:16790829

Hernán MS, Robins JM (2023). Causal inference: what if. Boca Raton (FL), USA: Chapman and Hall/CRC. Available from: https://www.hsph.harvard.edu/miguel-hernan/causal-inference-book/.

Hertz-Picciotto I (2000). Invited commentary: shifting the burden of proof regarding biases and low-magnitude associations. Am J Epidemiol. 151(10):946–8, discussion 949–50. doi:10.1093/oxfordjournals.aje.a010136 PMID:10853632

IARC (2018). Red meat and processed meat. IARC Monogr Eval Carcinog Risks Hum. 114:1–506. Available from: https://publications.iarc.who.int/564 PMID:29949327

IARC (2019). Preamble to the *IARC Monographs* (amended January 2019). Lyon, France: International Agency for Research on Cancer. Available from: https://monographs.iarc.who.int/iarc-monographs-preamble-preamble-to-the-iarc-monographs/.

IARC (2020). Night shift work. IARC Monogr Identif Carcinog Hazard Hum. 124:1–371. Available from: https://publications.iarc.who.int/593 PMID:33656825

Larjavaara S, Schüz J, Swerdlow A, Feychting M, Johansen C, Lagorio S, et al. (2011). Location of gliomas in relation to mobile telephone use: a case-case and case-specular analysis. Am J Epidemiol. 174(1):2–11. doi:10.1093/aje/kwr071 PMID:21610117

Lash TL, Fink AK, Fox MP (2009). Multiple bias modeling. In: Applying quantitative bias analysis to epidemiologic data. Statistics for biology and health. New York (NY), USA: Springer; pp. 151–73. doi:10.1007/978-0-387-87959-8_9

Lawlor DA, Tilling K, Davey Smith G (2016). Triangulation in aetiological epidemiology. Int J Epidemiol. 45(6):1866–86. doi:10.1093/ije/dyw314 PMID:28108528

Lipsitch M, Tchetgen Tchetgen E, Cohen T (2010). Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology. 21(3):383–8. doi:10.1097/EDE.0b013e3181d61eeb PMID:20335814

MacLehose RF, Ahern TP, Lash TL, Poole C, Greenland S (2021). The importance of making assumptions in bias analysis. Epidemiology. 32(5):617–24. doi:10.1097/EDE.0000000000001381 PMID:34224472

Maclure M (1998). The case-specular study design and counterfactual controls. Epidemiology. 9(1):6–7. doi:10.1097/00001648-199801000-00003 PMID:9430260

Mansournia MA, Jewell NP, Greenland S (2018). Case–control matching: effects, misconceptions, and recommendations. Eur J Epidemiol. 33(1):5–14. doi:10.1007/s10654-017-0325-0 PMID:29101596

McClure ES, Vasudevan P, DeBono N, Robinson WR, Marshall SW, Richardson D (2020). Cancer and noncancer mortality among aluminum smelting workers in Badin, North Carolina. Am J Ind Med. 63(9):755–65. doi:10.1002/ajim.23150 PMID:32649003

Naimi AI, Cole SR, Hudgens MG, Brookhart MA, Richardson DB (2013). Assessing the component associations of the healthy worker survivor bias: occupational asbestos exposure and lung cancer mortality. Ann Epidemiol. 23(6):334–41. doi:10.1016/j.annepidem.2013.03.013 PMID:23683709

Ogburn EL, Rudolph KE, Morello-Frosch R, Khan A, Casey JA (2021). A warning about using predicted values from regression models for epidemiologic inquiry. Am J Epidemiol. 190(6):1142–7. doi:10.1093/aje/kwaa282 PMID:33350434

Ogburn EL, VanderWeele TJ (2012). On the nondifferential misclassification of a binary confounder. Epidemiology. 23(3):433–9. doi:10.1097/EDE.0b013e31824d1f63 PMID:22450692

Pearce NE, Smith AH, Howard JK, Sheppard RA, Giles HJ, Teague CA (1986). Non-Hodgkin's lymphoma and exposure to phenoxyherbicides, chlorophenols, fencing work, and meat works employment: a case-control study. Br J Ind Med. 43(2):75–83. doi:10.1136/oem.43.2.75 PMID:3753879

Pierce BL, Kraft P, Zhang C (2018). Mendelian randomization studies of cancer risk: a literature review. Curr Epidemiol Rep. 5(2):184–96. doi:10.1007/s40471-018-0144-1 PMID:30034993

Pinkerton LE, Hein MJ, Anderson JL, Little MP, Sigurdson AJ, Schubauer-Berigan MK (2016). Breast cancer incidence among female flight attendants: exposure–response analyses. Scand J Work Environ Health. 42(6):538–46. doi:10.5271/sjweh.3586 PMID:27551752

Richmond RC, Anderson EL, Dashti HS, Jones SE, Lane JM, Strand LB, et al. (2019). Investigating causal relations between sleep traits and risk of breast cancer in women: mendelian randomisation study. BMJ. 365:l2327. doi:10.1136/bmj.l2327 PMID:31243001

Robins J (1986). A new approach to casual inference in mortality studies with a sustained exposure period – application to control of the healthy worker survivor effect. Math Model. 7(9–12):1393–512. doi:10.1016/0270-0255(86)90088-6

Savitz DA, Barón AE (1989). Estimating and correcting for confounder misclassification. Am J Epidemiol. 129(5):1062–71. doi:10.1093/oxfordjournals.aje.a115210 PMID:2705426

Schisterman EF, Cole SR, Platt RW (2009). Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiology. 20(4):488–95. doi:10.1097/EDE.0b013e3181a819a1 PMID:19525685

Schlesselman JJ (1978). Assessing effects of confounding variables. Am J Epidemiol. 108(1):3–8. PMID:685974

Sheikh M, Shakeri R, Poustchi H, Pourshams A, Etemadi A, Islami F, et al. (2020). Opium use and subsequent incidence of cancer: results from the Golestan Cohort Study. Lancet Glob Health. 8(5):e649–60. doi:10.1016/S2214-109X(20)30059-0 PMID:32353313

Sjölander A, Öberg S, Frisell T (2022). Generalizability and effect measure modification in sibling comparison studies. Eur J Epidemiol. 37(5):461–76. doi:10.1007/s10654-022-00844-x PMID:35312926

Smith GD, Ebrahim S (2003). 'Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 32(1):1–22. doi:10.1093/ije/dyg070 PMID:12689998

Steenland K, Greenland S (2004). Monte Carlo sensitivity analysis and Bayesian analysis of smoking as an unmeasured confounder in a study of silica and lung cancer. Am J Epidemiol. 160(4):384–92. doi:10.1093/aje/kwh211 PMID:15286024

Steenland K, Schubauer-Berigan MK, Vermeulen R, Lunn RM, Straif K, Zahm S, et al. (2020). Risk of bias assessments and evidence syntheses for observational epidemiologic studies of environmental and occupational exposures: strengths and limitations. Environ Health Perspect. 128(9):95002. doi:10.1289/EHP6980 PMID:32924579

VanderWeele TJ (2016). Mediation analysis: a practitioner's guide. Annu Rev Public Health. 37(1):17–32. doi:10.1146/annurev-publhealth-032315-021402 PMID:26653405

VanderWeele TJ, Ding P (2017). Sensitivity analysis in observational research: introducing the E-value. Ann Intern Med. 167(4):268–74. doi:10.7326/M16-2607 PMID:28693043

Wade KH, Yarmolinsky J, Giovannucci E, Lewis SJ, Millwood IY, Munafò MR, et al.; with the MR in Nutrition, Cancer working group (2022). Applying Mendelian randomization to appraise causality in relationships between nutrition and cancer. Cancer Causes Control. 33(5):631–52. doi:10.1007/s10552-022-01562-1 PMID:35274198

Yarmolinsky J, Wade KH, Richmond RC, Langdon RJ, Bull CJ, Tilling KM, et al. (2018). Causal inference in cancer epidemiology: what is the role of Mendelian randomization? Cancer Epidemiol Biomarkers Prev. 27(9):995–1010. doi:10.1158/1055-9965.EPI-17-1177 PMID:29941659