

# READERS' TOOLBOX

## Understanding Research Methods

### An Introduction to Causal Diagrams for Anesthesiology Research

Amy L. Gaskell, M.B.Ch.B., F.A.N.Z.C.A.,  
Jamie W. Sleight, M.D., F.A.N.Z.C.A.


(ANESTHESIOLOGY 2020; 132:951–67)

#### An Introduction to Causal Diagrams for Anesthesiology Research

Review of basic causal inference principles and use of causal diagrams


Directed Acyclic Graphs provide a framework to

- Address sources of bias
- Discover causal effects




Causal Diagram Essentials

- Exposure and outcome variables
- Arrow represents causal relationship, directed from exposure to outcome variable ( $A \rightarrow B$ )



Examples:

- Causal pathway:  $X \rightarrow Y$
- Confounding pathway:  $X \leftarrow C \rightarrow Y$
- Collider pathway:  $X \rightarrow Z \leftarrow Y$



Causal diagrams provide a rigorous approach to bias and can inform better study design and analysis.

Gaskell AL, et al. ANESTHESIOLOGY. February 2020.

**ANESTHESIOLOGY**  
Trusted Evidence: Discovery to Practice®

Copyright © 2020 Wolters Kluwer Health. All rights reserved. Published by Lippincott, Williams & Wilkins.

#### SUMMARY

Making good decisions in the era of Big Data requires a sophisticated approach to causality. We are acutely aware that association  $\neq$  causation, yet untangling the two remains one of our greatest challenges. This realization has stimulated a Causal Revolution in epidemiology, and the lessons learned are highly relevant to anesthesia research. This article introduces readers to directed acyclic graphs; a cornerstone of modern causal inference techniques. These diagrams provide a robust framework to address sources of bias and discover causal effects. We use the topical question of whether anesthetic technique (total intravenous anesthesia vs. volatile) affects outcome after cancer surgery as a basis for a series of example directed acyclic graphs, which demonstrate how variables can be chosen to statistically control confounding and other sources of bias. We also illustrate how controlling for the wrong variables can introduce, rather than eliminate, bias; and how directed acyclic graphs can help us diagnose this problem.

This is a rapidly evolving field, and we cover only the most basic elements. The true promise of these techniques is that it may become possible to make robust statements about causation from observational studies—without the expense and artificiality of randomized controlled trials.

(ANESTHESIOLOGY 2020; 132:951–67)

The causal relationship between smoking and lung cancer is firmly established; however, this was not always the case. In fact, Ronald Fisher, the eminent statistician, fervently rejected the notion that smoking caused lung cancer<sup>1</sup>—a denial perhaps fuelled by his own significant tobacco habit. He instead asserted that the strong association could be due to a smoking gene responsible for both smoking behavior and lung cancer (*i.e.*, a case of simple confounding). It took many years for clinicians and scientists to finally resolve the issue; a delay that cost many lives.

Many of the questions we strive to answer with anesthetic research are also of a causal nature; for example, *does hypotension cause perioperative stroke?* or *does total intravenous anesthesia improve survival after cancer surgery?* To improve clinical outcomes, it is imperative to determine cause–effect relationships, rather than simply describe associations. If we fail to do so, we risk subjecting patients to futile or harmful interventions or missing beneficial treatments.

Systematic error—comprising confounding and other bias—is a key barrier to the valid estimation of causal effects.

Image: Jorge A. Galvez, M.D., M.B.I., and Allan F. Simpao, M.D., M.B.I.

Submitted for publication April 14, 2019. Accepted for publication January 17, 2020. From the Department of Anaesthesia and Pain Medicine, Waikato District Health Board, Hamilton, New Zealand; and the Department of Anesthesiology, Faculty of Medical and Health Sciences, University of Auckland, Auckland, New Zealand.

Copyright © 2020, the American Society of Anesthesiologists, Inc. All Rights Reserved. Anesthesiology 2020; 132:951–67. DOI: 10.1097/ALN.0000000000003193

## Box 1. What to Look for in Research Using This Method

Is the primary causal question clearly stated?

Is the causal diagram consistent with existing knowledge?

Are there any important variables missing from the causal diagram?

The diagram must include:

- All relevant variables (not just those easily available or measured)
- Any variables which influence any two variables in the directed acyclic graph
- Any variables used for statistical control

Is the causal diagram consistent with the data-generating process?

- Consider study design and inclusion and exclusion criteria.

Have any arrows been omitted? Absent arrows represent strong assumptions.

Check that no feedback loops exist.

Are the statistical methods used consistent with the directed acyclic graph?

Random error can also mislead us but is generally overcome by increasing sample size. Systematic error, however, is not eliminated by increasing sample size (in fact a biased estimate might misleadingly appear to be more certain due to the ensuing lower *P* values) and so alternative strategies are required. We heavily rely on randomization to minimize selection bias and confounding and to establish causation, with blinding to minimize measurement bias. However, randomized, controlled trials are not immune from bias, and they are a limited resource due to the expense, effort, and large numbers of participants required. Thus, there is a patent need to obtain better estimates of causal effects from observational studies, particularly with the advent of Big Data.<sup>2</sup> Recent years have seen a flurry of interest in causality and refinements of causal inference techniques—particularly in the fields of epidemiology and the social sciences—which could address this need in perioperative research.

The main focus of this article is directed acyclic graphs, which have been popularized by Judea Pearl<sup>3</sup> and others.<sup>4,5</sup> These causal diagrams visually capture the complex interrelationships between important variables, and offer a robust framework to comprehensively address bias. Readers are increasingly likely to encounter these diagrams, as some journals have already recommended their inclusion to authors reporting observational studies.<sup>6,7</sup> We introduce the components and taxonomy of directed acyclic graphs in general terms and follow with specific examples relevant to anesthesiology, to illustrate how these diagrams might help researchers with study design and analysis and clinicians with interpretation of analyses reported in the literature.

The underlying ideas are fairly simple but do require the reader to devote some time to diligently think about the sometimes-subtle nuances within the relationships of *all* the factors that influence the outcome of any study under consideration. Box 1 is a brief summary of important points for the reader to consider when encountering causal diagrams in scientific reports. Because some of the terminology is quite technical and will be unfamiliar to most readers, a glossary is included at the end of this article.

## Causal Diagram Essentials

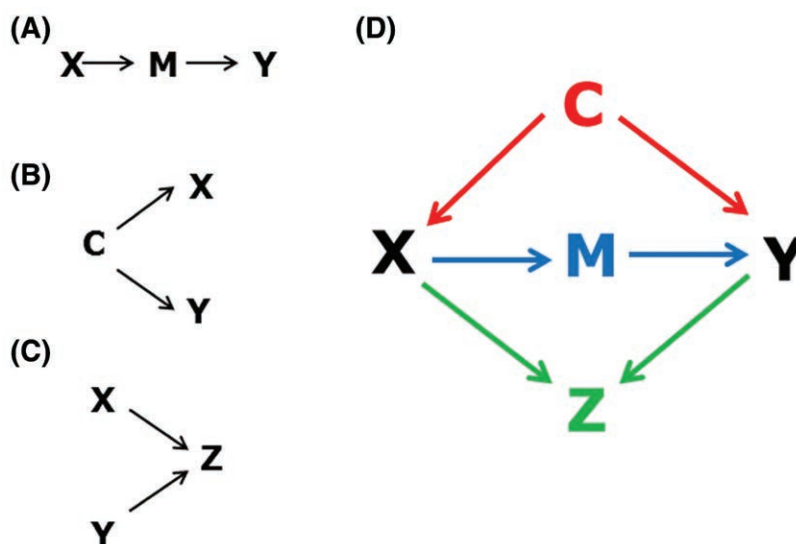
The primary research question defines the causal relationship of interest. This relationship typically takes place within a wider system of interrelated variables. The directed acyclic graph is a visual representation of this causal structure. The directed acyclic graph is constructed around the *exposure* (variable which exerts influence) and the *outcome* (affected variable). Arrows represent causal relationships between variables, which (with sufficient sample size) generate statistical associations. The arrows are always directed from the variable exerting the influence toward the affected variable. Figure 1 is a simple example which illustrates the three basic elements (*A–C*) of directed acyclic graphs, and how these might come together to form a complete directed acyclic graph (*D*).

A *pathway* refers to any series of arrows that connects two variables, regardless of the direction of the arrows. Pathways may be open or closed. Open pathways generate statistical associations, but closed pathways do not. The exposure and outcome may be linked by causal pathways and noncausal pathways. The research aim is to quantify the strength of the causal pathway, which can only be done if any noncausal pathways are closed or (equivalently) *blocked*.

## Causal Pathways (e.g., $X \rightarrow M \rightarrow Y$ )

The exposure (*X*) might cause the outcome (*Y*) directly or through an intermediate process or variable called a *mediator* (*M*). The causal relationship (fig. 1A) can therefore be represented by a single arrow ( $X \rightarrow Y$ ), or by a chain of arrows containing one or more mediators (e.g.,  $X \rightarrow M \rightarrow Y$ ). The single arrow indicates that *X* influences *Y* directly, whereas the chain indicates that *X* influences *Y* first by influencing an intermediate variable *M*, which in turn influences variable *Y*. For example, smoking (*X*) might cause lung cancer (*Y*) directly, or indirectly by first causing inflammation and chronic obstructive pulmonary disease (*M*), which is also thought to play a causal role in the development of lung cancer.

A causal pathway is defined as one in which all arrows point in the same direction from the exposure toward the outcome. There may be more than one causal pathway between the exposure and outcome if there is more than one potential mechanism of action, and the choice of how many mediators to include will depend on the question at hand. The total effect of the exposure on the outcome is the sum or net effect of all direct and indirect causal pathways.



**Fig. 1.** Elements of directed acyclic graphs. (A) A causal pathway between the exposure (X) and the outcome (Y). X influences Y through its effect on the mediator (M). (B) A noncausal pathway between X and Y due to a shared cause (the confounder-C). (C) A closed pathway between X and Y due to a shared effect (the collider-Z). (D) A complete causal graph of these pathways. The causal relationship of interest (blue pathway) cannot be quantified because of confounding (red pathway). The green pathway does not bias the relationship because it is closed by the collider (Z).

### Confounding Pathways (e.g., $X \leftarrow C \rightarrow Y$ )

Confounding pathways occur when the exposure and the outcome have shared causes or parents. These *noncausal* pathways are naturally open and represent associations already present in the populations we study. In figure 1B, X and Y will vary together in response to the confounder (C) and the association generated by this pathway will bias the estimated causal effect, unless it is addressed somehow. For example, Fisher postulated a smoking gene (C) might both influence smoking behavior (X) and cause lung cancer (Y). This could not be confirmed or refuted in his time—but such a gene has since been found to exist; however, it accounts for only a small fraction of the overall association between smoking and lung cancer.<sup>8</sup>

Confounding pathways begin with an arrow directed toward the exposure, and the arrows don't all point in the same direction. As with causal pathways, there may be multiple confounding pathways between the exposure and the outcome, and the confounding bias is the sum or net effect of all these pathways.

### Collider Pathways (e.g., $X \rightarrow Z \leftarrow Y$ )

A variable which is caused or influenced by (a *descendant* of) two other variables is known as a collider. Equivalently, a collider is any variable in a directed acyclic graph that has two arrows colliding into it, for example variable Z in figure 1C. Pathways are blocked by colliders, meaning that pathways containing colliders do not naturally generate statistical

associations between the variables they link. Extending our example, smoking (X) and lung cancer (Y) could both decrease weight (Z); smoking because of appetite suppressant effects of nicotine, and lung cancer because of cachexia. As expected intuitively, weight doesn't affect the observed association between smoking and lung cancer. In directed acyclic graph terms, this is because weight is a collider, so the pathway is closed and does not generate a statistical association.

### Path Rules and Bias

An *open* pathway is any pathway which generates a statistical association between the connected variables; the simplest special case of this being a causal pathway. A *closed* pathway is a pathway which does not generate a statistical association; the simplest example being when two variables are linked by a collider. Any noncausal pathways that do not contain collider variables are open and represent the phenomenon known as confounding. Conversely, noncausal pathways that contain colliders are closed and do not generate bias.

Correct specification of the direction of arrows in a directed acyclic graph is critical due to the distinct properties of collider variables. For example, in figure 1D the pathways  $X \rightarrow M \rightarrow Y$  and  $X \leftarrow C \rightarrow Y$  are open, and the pathway  $X \rightarrow Z \leftarrow Y$  is closed. The observed association is the net effect of all *open* pathways, so the causal  $X \rightarrow Y$  association is confounded by the open  $X \leftarrow C \rightarrow Y$  pathway, but not by the closed  $X \rightarrow Z \leftarrow Y$  pathway.

**Table 1.** Forms of Conditioning: Essentially Any Action That Renders an Association Conditional on a Particular Variable by Limiting That Variable to a Certain Level or Range

**Common Forms of Conditioning**

Restriction/selection  
Standardization/weighting  
Stratification/subgroups  
Matching  
Inclusion in multivariable regression models

## Conditioning

The status of pathways can be changed by *conditioning*. Conditioning refers to any action which renders an association conditional on other variables; essentially by restricting the variable or variables to a particular level or range of values. Some common forms of conditioning are listed in table 1. Conditioning includes most statistical control (and we use these terms interchangeably); for example, when variables are included in multivariable models or used to classify participants into subgroups for analysis. Conditioning is indicated graphically by drawing a box around the variable.

## Closing Pathways

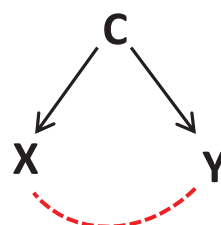
Conditioning on a variable on an open pathway will block that pathway. The aim is usually to remove bias arising from confounding pathways. If we consider figure 1D, controlling for C closes the  $X \leftarrow C \rightarrow Y$  pathway and removes the confounding. For example, when examining the association between smoking (X) and lung cancer (Y), performing the comparison after first dividing participants into groups according to their genotype—or alternatively including the genotype in a multivariable statistical model—would remove the confounding by smoking gene (C). Conditioning on a mediator is generally detrimental, as this blocks a causal pathway and will usually bias the estimation of the causal effect toward the null. For instance conditioning on chronic obstructive pulmonary disease (M) would remove part of the causal association between smoking (X) and lung cancer (Y) by blocking the  $X \rightarrow M \rightarrow Y$  pathway.

## Opening Pathways

Conditioning on a collider variable has a quite different effect (see box 2 for a comparison of confounder and collider variables). Conditioning on a collider variable *opens* the pathway where it was previously closed, and the opened pathway will introduce bias as it is noncausal. For instance, if we were to control for variable Z in figure 1D, we would open the previously closed pathway ( $X \rightarrow Z \leftarrow Y$ ). Returning to our smoking example, if we were to control for the collider weight (Z), we would introduce bias by

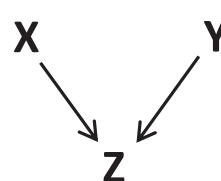
## Box 2. Important Differences between Confounder and Collider Variables

Confounder Variable (C)

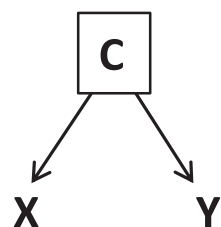


When left unadjusted *creates* a noncausal association between X and Y—*confounding*.

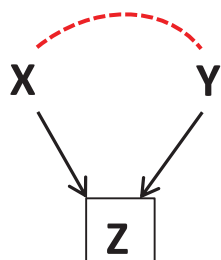
Collider Variable (Z)



When left unadjusted does *not* create any noncausal association between X and Y.



Adjustment for the confounder *removes* the noncausal association between X and Y.



Adjustment for the collider *creates* a noncausal association between X and Y—*selection bias*.

erroneously opening a *noncausal pathway* between smoking (X) and lung cancer (Y).

With the causal diagram approach, *selection bias* is defined broadly as any associations arising from pathways opened by conditioning on collider variables.<sup>9,10</sup> Selection bias is often introduced before analysis, for example when participant inclusion is based on specific inclusion or exclusion criteria, or restricted or influenced in other ways. For instance, if we examined the association between smoking and lung cancer in patients attending a hospital respiratory clinic, then clinic attendance would act as an alternative collider variable (Z) for our example, because patients might attend owing to lung cancer or lung cancer symptoms or to other smoking relating diseases. The association between smoking (X) and lung (Y) cancer in our sample would be biased because we conditioned on clinic attendance (Z) as we collected our data, and in doing so created a noncausal pathway. Box 3 provides a more detailed explanation and example of this phenomenon, which is somewhat less intuitive than confounding.

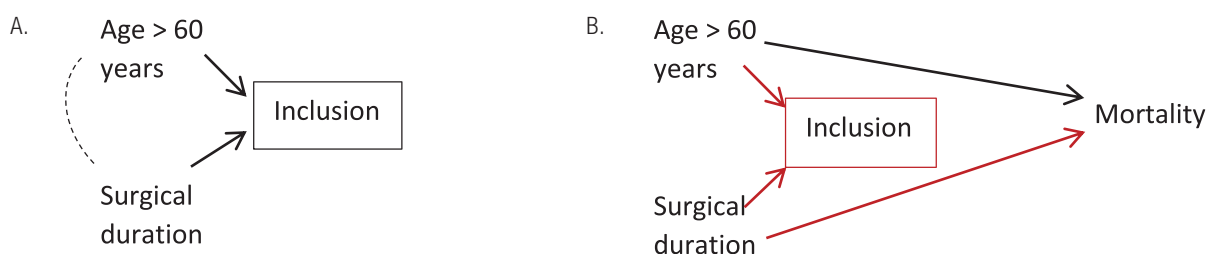


### Box 3. Colliders – Why Conditioning on a Common Effect Can Introduce Bias

Imagine a prospective observational study investigating factors affecting postoperative mortality. To reduce the burden of data collection and ensure a reasonable event rate, participants are selected according to certain inclusion criteria; namely at least one of (1) age over 60 yr or (2) surgical duration over 2 h. We assume here for simplicity that there is no baseline association between age and surgical duration (they are completely independent). However, when we come to analyze our data, we are likely to observe a negative association between these two variables due to our participant selection process.

Why does this happen? The data-generating process is depicted in the first directed acyclic graph *A*; and we can see that inclusion (in the study) is a collider between age and surgical duration that has already been conditioned on when selecting participants. Our inclusion criteria mean that knowing some information about one variable confers information about the other variable. If the patient is not aged over 60 yr, then they must have surgical duration of more than 2 hr to be included. Conversely those patients with surgical duration under 2 hr must be aged over 60 yr to be included. Having one of the inclusion criteria reduces your likelihood of having the other relative to those who don't have the first criterion. Hence, the variables *age* and *surgical duration* will be negatively correlated within our sample; in the language of directed acyclic graphs this is due to the pathway we opened by conditioning on inclusion criteria. These associations are sometimes represented by a dashed line on a directed acyclic graph.

If we examine the association between age and mortality (directed acyclic graph *B*), we will find ourselves with a biased estimate of the causal effect. This is because the crude (unadjusted) association consists of not only the causal pathway of interest (age→mortality) but also the spurious pathway (shown in *red*) that was induced by conditioning on inclusion criteria (age→inclusion←surgical duration→mortality), which is shown in *red*. To recover the causal effect of age on mortality, we would need to also condition on surgical duration to block the noncausal pathway.



This phenomenon can also occur whenever we statistically control for a variable. If the variable is caused or influenced by two parent variables (*i.e.*, is a collider variable), adjustment will introduce a spurious association between the parent variables. This is because statistical adjustment has a similar effect as selection by essentially limiting the adjustment variable to a certain level; this links the parent variables because the collider is statistically dependent on both.

### Box 4. How to Build a Directed Acyclic Graph in Seven Steps

1. Define the primary causal relationship of interest and begin the causal diagram with the exposure and the outcome.
2. Insert important mediators; there may be more than one causal path between the exposure and the outcome.
3. Consider important causes of the exposure and important causes of the outcome (both measured and unmeasured).
4. Consider whether any two variables already on the directed acyclic graph share a common influence; if so this variable should be included.
5. Review variables in a pairwise manner—should any arrows be added? Absent arrows represent strong assumptions.
6. Ensure that any selection procedures are adequately captured.
7. Ensure that there are no feedback loops present; it may be necessary to include a variable at different time points (*e.g.*, baseline, t1, t2) to maintain causality.

### Using Directed Acyclic Graphs to Make Causal Inferences

The directed acyclic graph must be complete to allow valid estimation of causal effects. The steps involved in constructing a complete directed acyclic graph are outlined in box 4. It is important to include all shared causes of the exposure and the outcome, and any shared causes of any two variables already in the DAG. All relevant variables must be included, even those which cannot be measured. The directed acyclic graph should accurately depict the structure of confounding pathways and any pathways opened by conditioning during data collection or analysis. The term *acyclic* refers to the rule that feedback loops are not permitted; you cannot follow a pathway forward from a variable back to itself. The directed acyclic graph encodes the assumptions which underpin the analysis, and also the implied statistical relationships between variables. Of note, omission of an arrow between a pair of variables represents a strong assumption

of no effect, and needs to be just as carefully considered as those arrows that are present.

The directed acyclic graph identifies which pathways must be closed for valid estimation of the causal effect. The aim is to isolate the causal pathway of interest by:

1. Blocking all open noncausal paths between the exposure and the outcome;
2. Leaving the causal paths between the exposure and the outcome unperturbed;
3. Avoiding methods which might open spurious non-causal pathways between the exposure and the outcome.

Directed acyclic graphs can be used to streamline data collection by identifying the most efficient sets of variables to accomplish this. Pathways need only be blocked in one place if the controlled variable is measured accurately, and it may be possible to close more than one pathway by adjusting for a single variable. A pathway opened by conditioning on a collider can be closed by conditioning on another variable in the pathway. The status of a variable within a particular system is not fixed; rather, it depends upon the causal relationship of interest. A variable might act as a confounder or mediator on one pathway and a collider on another, and so it is quite possible to simultaneously open one pathway and close another with adjustment, as we demonstrate in some later examples.

It should also be noted that conditioning on a descendant of a variable has a similar effect to conditioning on the variable itself—but the effect will be reduced. This includes imperfect markers or measures of the variable. Conditioning on a descendant of a variable on an open pathway (*i.e.*, mediators or confounders) will partially close the pathway, and conditioning on a descendant of a collider can partially open a pathway. The extent to which the pathway is closed or opened depends upon how closely the variable and the descendant are associated; the stronger the association, the greater the effect. Our ability to close noncausal pathways therefore relies on the measurement accuracy of the variables we use for statistical control. If measurement error is suspected, it may be prudent to close the pathway by conditioning on more than one variable along the pathway.

## Interpretation of Reported Statistical Analyses

Anyone reading papers using multivariable analysis is often struck by uncomfortable and nonsensical results. Why are

some predictive variables clearly counterintuitive? Why do different studies obtain correlations in opposite directions? Why do some variables lose (or gain) significant effects depending on which other variables are included in the model? Some common misconceptions about multivariable adjustment are summarized in table 2. Although adjustment for variables on confounding pathways will generally remove bias, adjustment for mediators and colliders will generally introduce bias. Typically multivariable analyses implicitly assume a causal structure resembling figure 2A, when in reality the included variables are often an assortment of independent causes, confounders, mediators, and colliders. The true causal association can be better assessed if adjustment is guided by causal structures such as that shown in figure 2B, which account for the relationships between variables as well as the direct relationships to the outcome.

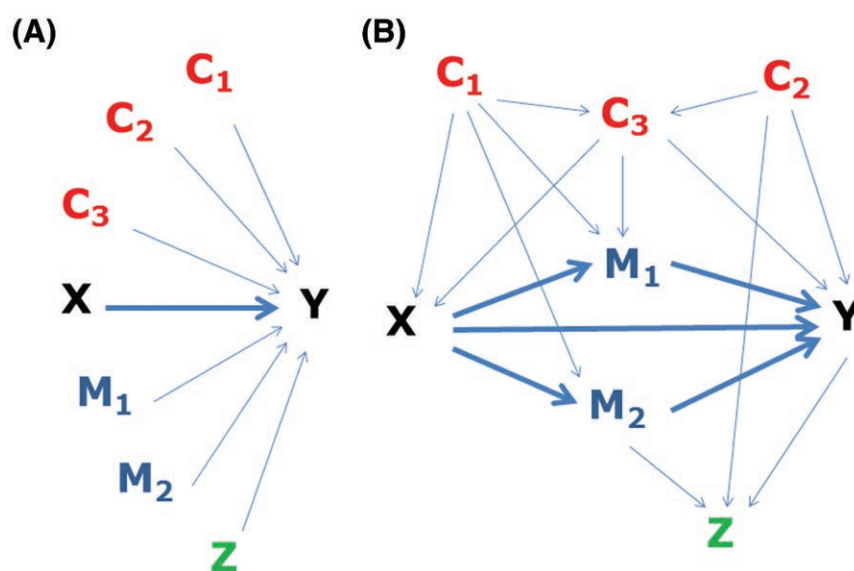
Often it is assumed that controlling for more variables is better, but injudicious inclusion of variables can make the adjusted associations awkward to interpret. This problem has been dubbed the *table 2 fallacy*,<sup>11</sup> because the second results table in observational studies customarily contains the adjusted associations between all of the variables included in the statistical model and the outcome. The adjusted associations are often interpreted as the independent effect of each explanatory variable on the outcome; however, the type of effect represented is entirely dependent on the other variables included in the model. The adjusted association will be the net effect of any pathways between the explanatory variable and the outcome remaining open after conditioning on the other included variables. The adjusted association might therefore represent either a total or partial (direct) causal effect, which could be distorted by residual confounding or opened collider pathways. Causal diagrams are immensely useful in diagnosing which of these is the case and can facilitate constructive critique and discussions such as those held in a journal club or at a scientific conference.

## Worked Examples: Anesthetic Technique and Mortality following Colon Cancer Surgery

We have created a set of example directed acyclic graphs, which are inspired by a recent retrospective observational study reported in *ANESTHESIOLOGY* by Wu *et al.*<sup>12</sup> This study

**Table 2.** Three Common Misconceptions about Multivariable Statistical Adjustment with Explanation

Misconception	Explanation
Increasing sample size will reduce bias	Increased sample size decreases <i>random error</i> and may result in narrower CIs but will not correct structural bias. You will simply achieve a more precise estimate of the wrong effect.
Controlling for more variables will reduce confounding	Controlling for variables on confounding pathways will generally remove bias and improve validity, but controlling for mediators and colliders can introduce bias.
The model with the best predictive ability should be used	The estimate for the causal association between a particular variable and the outcome will only be unbiased if appropriate covariates have been carefully selected for that particular relationship. A directed acyclic graph is needed to ensure this.



**Fig. 2.** Multivariable regression models typically assume a causal structure such as A, whereas a more realistic causal structure is depicted in B. X, exposure; Y, outcome; C, confounders; M, mediators; Z, colliders.

examined the association between anesthetic technique and mortality in patients having surgery for colon cancer. The main finding was that the use of propofol-based total intravenous anesthesia was associated with an approximately threefold reduction in long-term mortality compared to desflurane-based anesthesia. We use this topical example to illustrate the concepts around the applications of causal diagrams.

We first show how a directed acyclic graph might be developed around the causal question, and discuss how variables might be selected to control confounding, beginning with simplified diagrams and then adding complexity. We follow with some examples from the study analysis, which illustrate how controlling for some variables might be harmful rather than helpful. Our last examples use causal diagrams to explain some striking differences in the reported univariable and multivariable associations between explanatory variables and mortality. Note that we do not mean to conduct a full critique of the analyses conducted, which would require access to the raw data; rather, we hope to highlight the utility of causal diagrams in guiding and interpreting analyses.

### Building a Directed Acyclic Graph: Defining the Causal Question

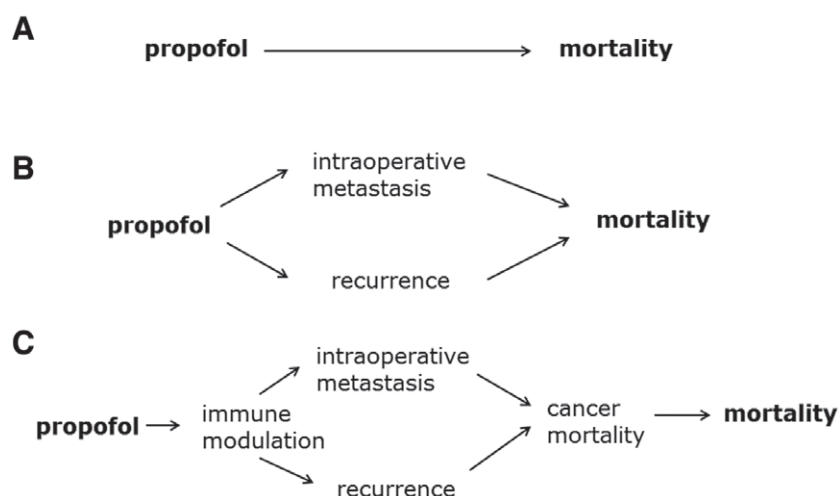
The causal question is *Does propofol-based total intravenous anesthesia reduce postoperative mortality in patients having colon cancer surgery?* The directed acyclic graph therefore begins with an arrow from propofol to mortality (fig. 3A). The authors hypothesized that propofol might reduce mortality by two mechanisms: (1) reducing intraoperative metastasis and (2) reducing postoperative cancer recurrence. Intraoperative metastasis and postoperative recurrence are therefore added

as possible mediators (fig. 3B). The decision of how many mediators to include in the diagram involves a careful balance of the needs for parsimony, *versus* full elucidation (fig. 3C).

### Closing Confounding Pathways by Statistical Control

The unadjusted association should always be regarded as potentially confounded. In this case, the exposure (propofol *vs.* desflurane) was chosen by the treating clinician. Although this may seem a somewhat arbitrary form of treatment allocation, we see that there were significant imbalances in important baseline variables between the propofol and desflurane groups. Anesthetic technique is often influenced by patient and surgical factors of prognostic importance. Indeed, many factors associated with a better prognosis (younger average age; lower comorbidity score; lower tumor-node-metastasis stage; greater functional capacity [metabolic equivalents]) were associated with propofol use. Therefore, the crude association would be expected to erroneously show an apparent mortality benefit of propofol, even if there were no causal effect.

Figure 4 shows how the problem of confounding might be approached for this example. Some pathways are relatively simple, involving only single confounders. For example, the date of surgery (relative to the first included surgical case) was included as a variable in the multivariable analysis conducted in the study (fig. 4A). Propofol use became more prevalent over the time-course of the study, but it is also reasonable to expect that medical and surgical management of cancer improved over this time and reduced the mortality rate. Date of surgery—though not technically a cause—acts as a good proxy for medical advances across the time course of the



**Fig. 3.** The causal pathways between the exposure (propofol) and the outcome (mortality) can be expressed simply (A) or in greater detail including mediators (B and C).

study that might confound the propofol and mortality relationship, and was appropriately controlled for in the study.

Often confounding has a more complex structure and variables may be interrelated, such as age and comorbidities (fig. 4B). As more variables are included, the directed acyclic graph can start to resemble a web-like structure (fig. 4C). Variables are selected for the statistical model in an attempt to close the confounding pathways and reveal the causal association of interest (fig. 4C). The emphasis is on closing pathways rather than needing to consider whether each individual variable should be classified as a confounder. This differs from the classical statistical approach, as is elaborated in box 4. Provided we have carefully considered the causal structure, it may be possible to close pathways using fewer covariates than we regularly see used in multivariable analyses, which often include all available variables. For example, in figure 4C, the variable *noncancer mortality* need not be included because those confounding pathways it lies on are already blocked.

Note that our example directed acyclic graphs are not intended to comprehensively account for confounding of the propofol-mortality association; instead, we use them to illustrate how the usual goal of statistical adjustment (elimination of confounding) might be achieved. A further two studies (Williamson *et al.*<sup>13</sup> and Staplin *et al.*<sup>14</sup>) are included in the references as examples of the application of causal diagrams to select confounders for multivariable regression models in different settings.

## When Statistical Adjustment Can Be Harmful

The goal of statistical adjustment is usually to reduce confounding, as illustrated in the previous examples. It is not widely recognized that adjustment is not always benign. If putative confounders are in fact mediators or colliders, such

control will instead introduce bias. The next few illustrative examples are based around variables included in the analysis by Wu *et al.*<sup>12</sup> to control for confounding of the association between propofol total intravenous anesthesia and mortality.

## Conditioning on a Known Mediator: Overcontrol Bias

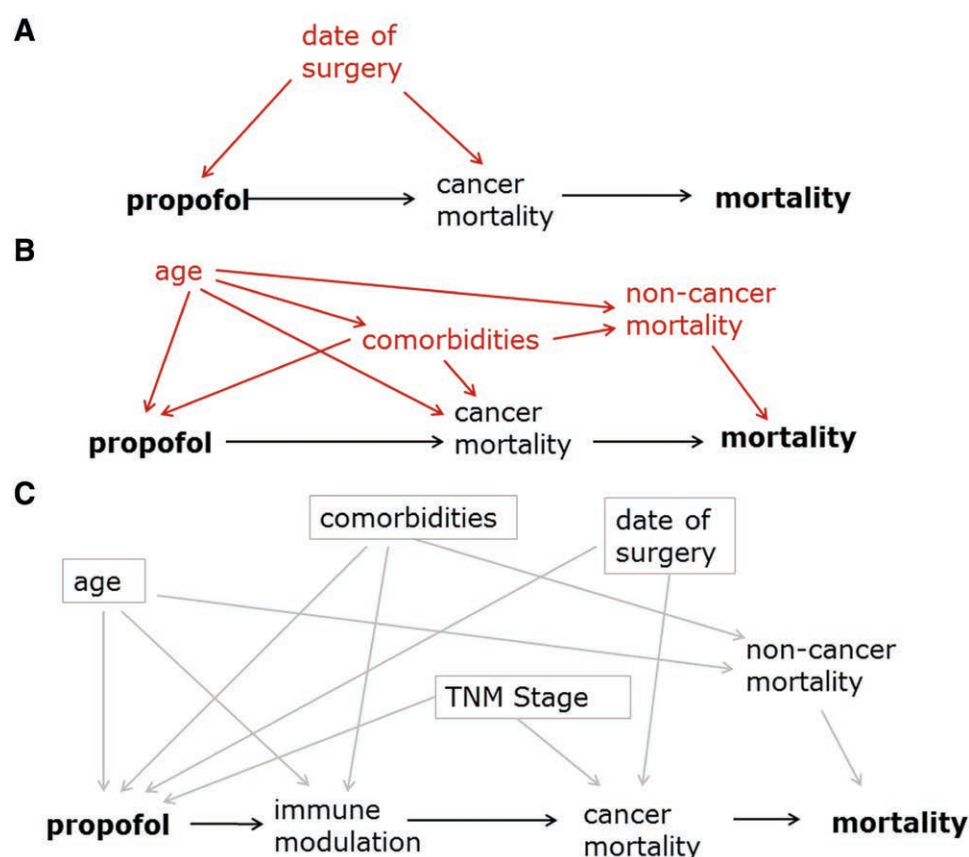
One of the variables included in the multivariable model was *postoperative recurrence*, which is a mediator on one of the causal pathways between propofol and mortality. Conditioning on postoperative recurrence blocks this causal pathway (fig. 5A), and essentially removes the contribution of cancer recurrence from the estimated total causal effect. We might therefore expect to see a diminished apparent causal relationship. This can be referred to as overcontrol, overmatching, or overadjustment bias.<sup>15</sup> It is therefore not advisable to include mediators in the covariate set used for statistical analysis.

## Conditioning on a Known Mediator: Collider Bias

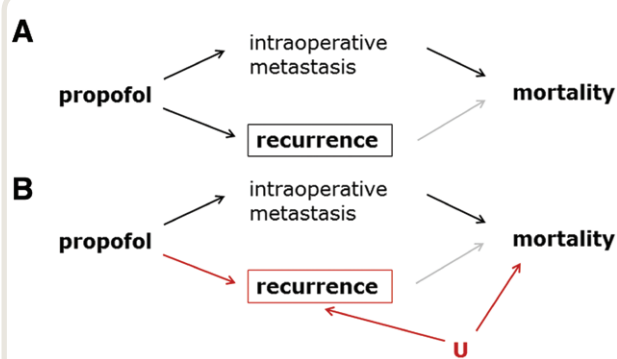
Controlling for mediators can have additional undesirable consequences. If we extend this example and consider the presence of unmeasured factors (*U*) that might influence *both* cancer recurrence and mortality, we realize that not only has one of the causal pathways of interest been closed, but another spurious noncausal pathway (the red pathway in fig. 5B) has been opened by controlling for postoperative recurrence.

The unmeasured causes of the outcome do not normally bias the propofol-mortality relationship, because *recurrence* is a collider on the noncausal pathway (propofol→recurrence←*U*→mortality) and so the path is closed; propofol and the unmeasured factors are statistically independent. However adjusting for recurrence opens the pathway because recurrence is caused by (and is therefore statistically





**Fig. 4.** The causal structure of confounding. A simple example of a classic confounding pathway (A) involving date of surgery. A more complex example (B) might involve interrelated confounders such as age and comorbidities, which may affect outcome *via* mediators along the pathway of interest (cancer mortality) or by alternative pathways to the outcome (noncancer mortality). As more confounders are included, the directed acyclic graph begins to resemble a web (C). All noncausal pathways must be blocked to derive the causal association of interest. Red arrows show open pathways; grey arrows show pathways which have been closed by adjustment (indicated by box around the variable). TNM, tumor-node-metastasis.



**Fig. 5.** Controlling for a mediator (recurrence) blocks one of the causal pathways of interest (A). If there are any unmeasured factors (U) which influence both cancer recurrence and noncancer mortality, a spurious pathway (red) is opened between propofol and mortality (B).

dependent upon) both propofol and the unmeasured factors. Adjustment for recurrence therefore creates a new statistical association or link between propofol and the unmeasured factors. This association extends to mortality because of the causal relationship between the unmeasured factors and mortality, resulting in a misleading estimate of the propofol–mortality association.

We included this example to highlight how adjusting for mediators can be particularly harmful, because it can close causal pathways *and* open spurious pathways. With these dual effects, the adjusted association becomes unintelligible.

### Conditioning on a Postexposure Variable: Unrecognized Mediator

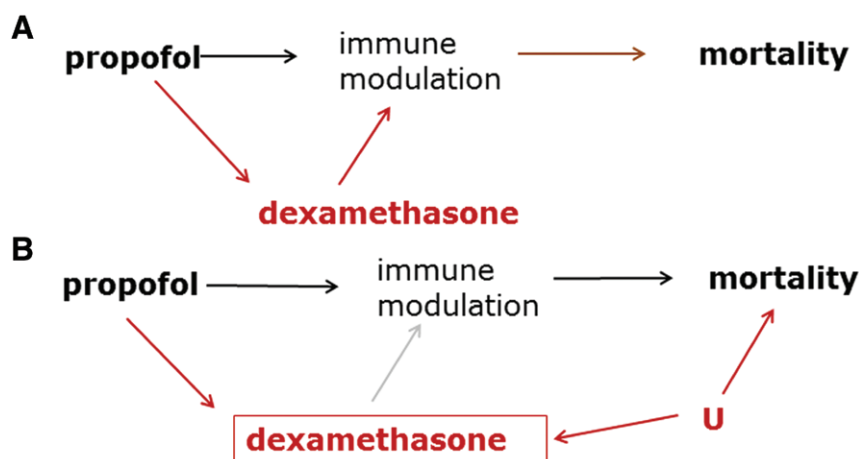
Any descendants of the exposure, not just known mediators, should be treated with caution. Typically, several agents are administered during the course of anesthesia. These adjuncts are usually treated as simple confounders, but this can be a harmful oversimplification of their role. Dexamethasone

is one such adjunct, which (quite appropriately) wasn't included in the analysis by Wu *et al.*<sup>12</sup> Dexamethasone is not a simple confounder because propofol influences the use of dexamethasone by reducing the need for antiemetic therapy, so the direction of the arrow in the directed acyclic graph is from propofol toward dexamethasone (fig. 6A). If we speculate that dexamethasone might also have causal or preventative actions on cancer metastasis, then dexamethasone becomes another mediator of the causal effect of propofol. This effect is indirect and may not be of primary interest, but it forms a covert causal pathway from the exposure toward the outcome. If we adjust for dexamethasone in an attempt to isolate the propofol→immune modulation→mortality pathway, we risk opening a false noncausal pathway (fig. 6B) containing unmeasured factors

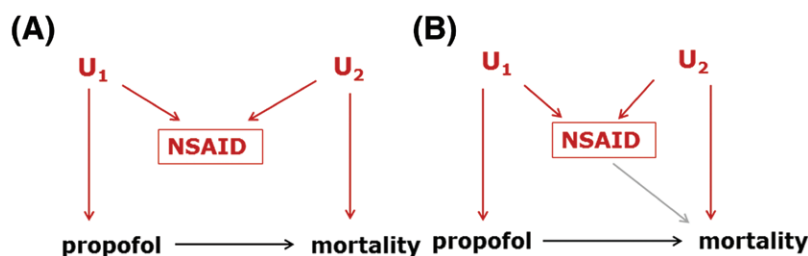
(U). This is why the advice not to include mediators might be extended to caution against including any postexposure variables—particularly known descendants of the exposure.

### Conditioning on a Collider: M-bias

Even when not influenced directly by the exposure of interest, the administration of adjuncts may be influenced by some of the same unmeasured factors that influence selection of the primary exposure; for example, individual clinician preferences. Such preferences may lead to associations between certain combinations of anesthetic agents. Wu *et al.*<sup>12</sup> included postoperative nonsteroidal antiinflammatory drugs (NSAIDs) in their multivariable model; however, propofol and NSAIDs are associated because of the shared influence of clinician preference. Furthermore, it is plausible that another



**Fig. 6.** Choice of anesthesia adjuncts may be affected by the exposure. The adjuncts are often regarded as confounders; however, they may in fact be mediators on an alternative causal pathway (A) between the exposure (propofol) and the outcome (mortality), which may not be of primary research intent. Adjustment for dexamethasone (B) in an attempt to isolate the causal pathway of interest may open a spurious pathway (red) involving unmeasured factors (U) that influence dexamethasone use and mortality.



**Fig. 7.** Adjusting for a variable (the collider – nonsteroidal antiinflammatory drug [NSAID]) that is influenced by an unmeasured cause of the exposure and an unmeasured cause of the outcome, creates a spurious noncausal pathway (A). The spurious pathway created (red) is M-shaped if variables are arranged in this pattern, and so this is sometimes referred to as M-bias. The unmeasured cause of propofol exposure might be clinician preferences (U<sub>1</sub>) and the unmeasured cause of mortality might be frailty or renal impairment (U<sub>2</sub>) which relatively contraindicates NSAID. If NSAID were thought to be a cause of mortality, then controlling for NSAID will simultaneously block a confounding pathway (B) and may be of net benefit.

unmeasured patient factor that influences mortality might also specifically influence NSAIDs use, such as frailty or renal impairment. Conditioning on postoperative NSAIDs could therefore open up a spurious noncausal pathway (fig. 7A) involving the unmeasured variables because NSAIDs is a collider. The bias arising from this sort of pathway is sometimes referred to as M-bias because the pathway can be drawn in the directed acyclic graph as a distinctive M-shape (fig. 7). Any factor that is influenced by both unmeasured causes of the exposure and unmeasured causes of the outcome has the potential to act in this way. The degree of bias induced by these pathways is unpredictable, because the collider variable may also (at least partially) block confounding pathways at the same time. This would be the case if we believed postoperative NSAIDs affect mortality; conditioning on NSAIDs would close a confounding pathway as well as opening the M-bias pathway (fig. 7B) and so could reduce bias overall.<sup>16,17</sup>

### Explaining Contrasting Univariable and Multivariable Associations

When multivariable models are selected, variables are often added or removed depending on the predictive ability of the model. If the goal is unbiased estimation of true causal effects, rather than an exercise in prediction, then this strategy is likely to fail. The specific combination of variables included determines the estimated association between each variable and the outcome, and how these effects should be interpreted. There can be quite striking differences between unadjusted (univariable) and adjusted (multivariable) associations, and it may be unclear which associations, if any, to trust. We now take some interesting examples of such changes from table 2 of Wu *et al.*<sup>12</sup> (summarized in our table 3) and show how directed acyclic graphs may help us explain them.

### Pre-exposure Collinearity: Loss of Causal Association

Our first example examines the associations between the tumor-node-metastasis stage of cancer and mortality. The staging criteria were developed to show a gradient in hazard and as such reflect features of tumors which are causally related to mortality. The expected relationship is evident in the unadjusted associations (table 3), as the hazard ratio increases steadily

with progressing cancer stage. However, the previously strong association between tumor-node-metastasis Stage 4 (which indicates metastatic cancer) and mortality essentially disappears in the multivariable model (table 3). So, are we really to believe that the crude relationship between metastatic disease and mortality was all attributable to confounding, and that metastatic disease doesn't affect cancer survival? Clearly, the adjusted association is implausible. This is an extreme example, but it forcibly demonstrates why adjusted associations are not necessarily more valid than unadjusted ones.

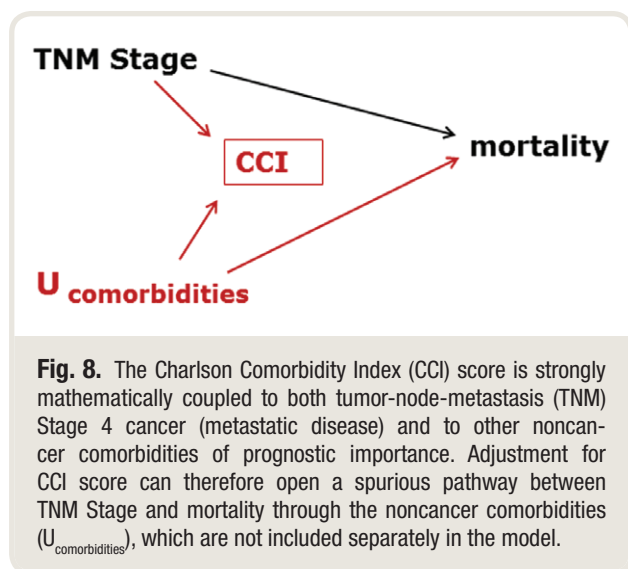
We suspect this dramatic change was attributable to conditioning on a collider in the presence of collinearity (high correlation) of certain baseline variables; the directed acyclic graph allows us to diagnose this problem. Composite scores of pre-existing illness are often used as covariates for statistical control in perioperative research. Usually these scores are regarded as simple confounder variables; however, they are also liable to act as colliders because they are descendants of multiple factors. The Charlson Comorbidity Index score,<sup>18</sup> a composite score used to predict 10 yr survival, was included in the multivariable model by Wu *et al.*<sup>12</sup> Cancer metastasis contributes strongly to the score, so tumor-node-metastasis Stage 4 and the Charlson Comorbidity Index score will be highly correlated because of mathematical coupling. Noncancer comorbidities contribute most of the remainder of the score and will also be highly correlated with the Charlson Comorbidity Index score, again due to mathematical coupling. The noncancer comorbidities are not included as a separate variable in the multivariable model, so controlling for the Charlson Comorbidity Index score creates a spurious pathway between tumor-node-metastasis 4 and mortality *via* the comorbidities (the *red pathway* in fig. 8). This problem could have been avoided by including noncancer comorbidities separately in the model, which would close the spurious pathway. The difference between univariable and multivariable hazard ratios was so pronounced because all of the correlations between the variables involved in this pathway were strong.

Baseline variables are often considered safe for adjustment purposes; however, this example shows why this is not necessarily the case. This is why it is traditionally advised to avoid including highly correlated variables (multicollinearity)<sup>19</sup> in statistical adjustment models and reinforces the need for caution with any descendants of the exposure as they may act as colliders.

**Table 3.** Variables with Extreme Differences between Univariable (Unadjusted) and Multivariable (Adjusted) Hazard Ratios for Mortality: Selected Findings from Table 2 by Wu *et al.*<sup>12</sup>

	Univariable HR (95% CI)	P Value	Adjusted HR (95% CI)	P Value
TNM 1	1	ref	1	ref
TNM 2	8.9 (5.3–15.1)	< 0.001	4.6 (2.6–7.9)	< 0.001
TNM 3	10.2 (6.1–17.3)	< 0.001	5.9 (3.4–10.3)	< 0.001
TNM 4	46.9 (27.8–78.8)	< 0.001	<b>1.12 (0.54–2.33)</b>	<b>0.754</b>
Age	1.03 (1.02–1.04) per yr	< 0.001	<b>0.95 (0.94–0.97) per yr</b>	<b>&lt; 0.001</b>

Bold text indicates major change in hazard ratio following multivariable adjustment which might result in misinterpretation of causal effects. HR, hazard ratio; TNM, tumor-node-metastasis stage.

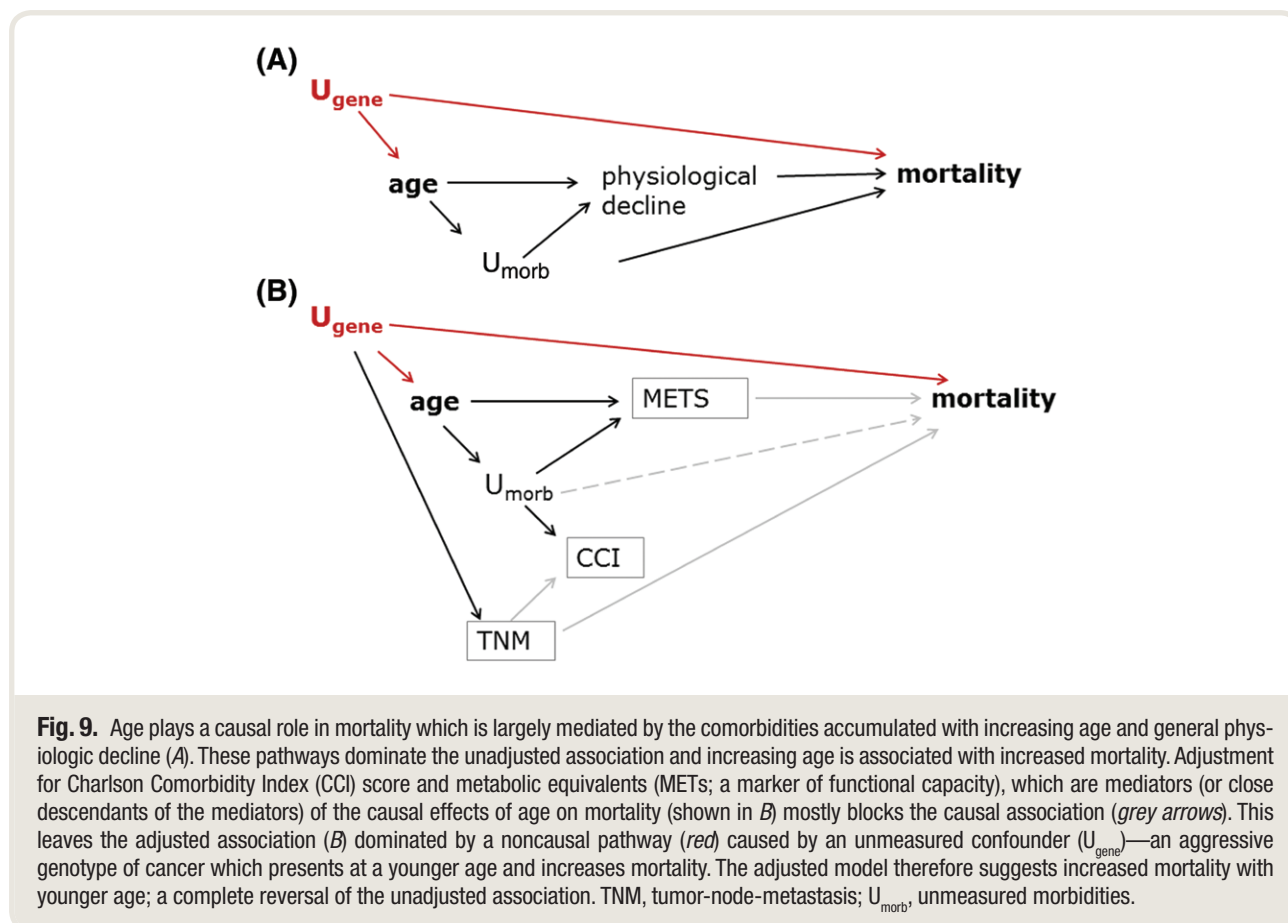


**Fig. 8.** The Charlson Comorbidity Index (CCI) score is strongly mathematically coupled to both tumor-node-metastasis (TNM) Stage 4 cancer (metastatic disease) and to other noncancer comorbidities of prognostic importance. Adjustment for CCI score can therefore open a spurious pathway between TNM Stage and mortality through the noncancer comorbidities ( $U_{\text{comorbidities}}$ ), which are not included separately in the model.

### Inclusion of Mediators: Reversal in Direction of Association

A second example from table 2 of Wu *et al.*<sup>12</sup> is the relationship between age and mortality. In this case there was a

complete reversal of the *direction* of association (table 3). The unadjusted association showed increased mortality with advancing age, which largely reflects a causal relationship mediated predominantly by physiologic decline and accumulated comorbidities. However, for the adjusted association, younger age was associated with increased mortality (table 3). A directed acyclic graph can help us to explain why this might be the case. Aggressive cancer genotypes, which cause greater mortality, often present when patients are younger; so there is a confounding pathway (fig. 9A), because age of cancer presentation and age at time of surgery will closely correspond. The unadjusted association is likely dominated by the expected causal relationship, but when the mediators of this causal relationship—comorbidities (the Charlson Comorbidity Index score) and functional reserve (metabolic equivalents)—are included in the model, the causal pathways are largely blocked (fig. 9B). The confounding pathway remains open because cancer genotype was not measured, and the adjusted association is likely largely attributable to this residual confounding effect. Unlike our previous example, the Charlson Comorbidity Index score doesn't cause a problem as a collider variable in this example, because tumor-node-metastasis stage is also included in the model (fig. 9B).



**Fig. 9.** Age plays a causal role in mortality which is largely mediated by the comorbidities accumulated with increasing age and general physiologic decline (A). These pathways dominate the unadjusted association and increasing age is associated with increased mortality. Adjustment for Charlson Comorbidity Index (CCI) score and metabolic equivalents (METS; a marker of functional capacity), which are mediators (or close descendants of the mediators) of the causal effects of age on mortality (shown in B) mostly blocks the causal association (grey arrows). This leaves the adjusted association (B) dominated by a noncausal pathway (red) caused by an unmeasured confounder ( $U_{\text{gene}}$ )—an aggressive genotype of cancer which presents at a younger age and increases mortality. The adjusted model therefore suggests increased mortality with younger age; a complete reversal of the unadjusted association. TNM, tumor-node-metastasis;  $U_{\text{morb}}$ , unmeasured morbidities.



This example highlights the importance of interpreting *adjusted associations in the context of the other variables in the model*. If the causal effects of several variables are of interest, a number of different statistical models may be required, depending on the directed acyclic graph for each effect.

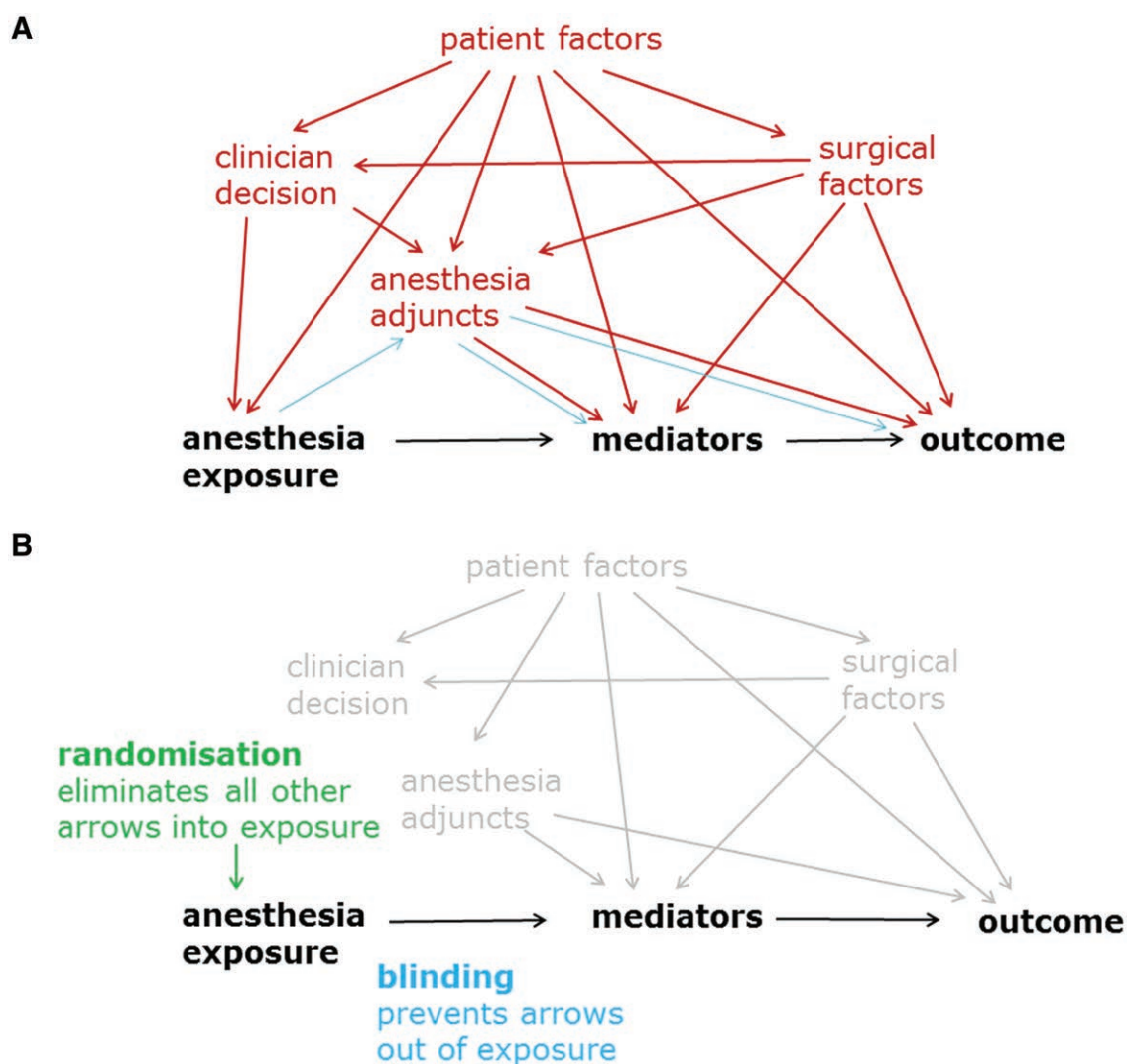
### Suggested Approaches to Confounding in Anesthesia Research

The structure of confounding can be highly complex with variables interacting to form a web of many interlinking pathways. A simplified causal framework for anesthesia-related

research is presented in figure 10A. Important potential confounding factors could broadly be categorized as:

1. Patient factors: demographic and illness related variables (*e.g.*, age, American Society of Anesthesiologists Physical Status)
2. Anesthesia factors (*e.g.*, technique, dose, *etc.*)
3. Surgical factors: intra- and perioperative variable (*e.g.*, duration, type of surgery)

Each group might be considered a nested structure with variables linking with each other and to variables from other



**Fig. 10.** A suggested framework for considering confounding (*red*) pathways for anesthesia interventions (A). Each group might include several factors. Intraoperative factors such as anesthesia adjuncts and surgical factors are particularly likely to act as colliders because they are influenced by many other factors. The anesthesia exposure may induce alternative causal pathways (*blue*), which are not of primary interest. Randomized controlled trials can minimize both these problems (B) because (1) exposure is random (removes all other arrows into the exposure) and (2) blinding will prevent the exposure having effects other than those of primary interest (removes arrows arising from the exposure).

groups and interacting with mediators along the causal pathway. Of particular note, anesthetic and surgical factors have clear potential to act as colliders, because they are subject to multiple influences. It is also important to consider downstream consequences of the exposure, which may result in additional bias (e.g., the blue pathway in fig. 10A).

Careful choice of primary outcome might reduce the number of confounding pathways. Shorter pathways containing fewer mediators will generally provide less opportunity for confounding. In our worked example (fig. 4), if cancer mortality is specified as the primary outcome rather than all-cause mortality, we do not have to worry about the confounding pathways containing noncancer mortality. Similarly, composite outcomes will have greater potential for confounding than more specific outcomes, because there will be more pathways between the exposure and outcome.

Attempting to close all noncausal pathways by statistical adjustment can be a formidable task; we have summarized some guidance in table 4. The ability to remove confounding is only as good as the measurement of the variables used for adjustment,<sup>20</sup> and it may be prudent to condition on more than one variable to block a pathway if significant measurement error is suspected. Categorization of continuous variables can have the same effect as measurement error and allow residual confounding and so should be avoided where possible.<sup>21</sup> Different pathways may bias the association of interest in different directions and will sometimes cancel out. The magnitude of bias arising from each pathway may be difficult to predict—it will depend on the strengths of relationships and the number of variables within the pathway. Longer pathways tend to generate less bias because the overall association will reduce with each additional step in the pathway, and the overall association cannot exceed the weakest association within the pathway. This is only general guidance, however, and in practice

difficult informed subjective choices will need to be made around the relative importance of the different pathways.

Often only a limited selection of measured variables is available, particularly with retrospective observational studies. It is extremely important to consider any relevant unmeasured factors which might give rise to bias. The clinical decision-making process is one such factor which may be very hard to measure. Further examples might include psychologic, lifestyle, socioeconomic, and genetic factors, intraoperative events, and surgical intensity and duration; these may affect the exposure or outcome or both. The unmeasured factors might result in classic residual confounding or bias from spurious noncausal pathways created if we mistakenly adjust for collider variables that link them. Adjustment for mediators (or other descendants of the exposure) is particularly likely to introduce collider bias and should be carefully avoided as the spurious pathways generated tend to be quite short and contain strong associations.

Sometimes the value of the causal diagram may lie in recognition that the causal effect cannot be identified using statistical adjustment, and that an alternative approach is required. Figure 10B demonstrates the elegance of randomized, controlled trials using a directed acyclic graph. Because the exposure becomes solely dependent on randomization, all other arrows into the exposure can be removed so there are no confounding pathways to worry about. This is why randomized, controlled trials are usually considered the gold standard in causal inference; however, they can suffer other forms of bias, which can be illustrated using directed acyclic graphs. The randomized exposure might influence the measurement of the outcome, such as with observer bias, or patient ascertainment or recall bias. Blinding is commonly used to address these problems (fig. 10B); however, it is often impractical or impossible to blind the clinicians caring for the patient to the exposure (or consequences of the exposure), which may influence subsequent patient management and bias the outcome.

**Table 4.** Guidance for Selection of Variables for Statistical Adjustment Using Causal Diagrams

**Selection of Covariates for Statistical Adjustment**

1. Define the primary causal relationship of interest.
  - Construct a directed acyclic graph around the exposure and outcome.
  - Ensure the directed acyclic graph is complete (box 2).
2. Identify any open noncausal pathways between the exposure and outcome.
  - Select the minimum set of variables necessary to close the open noncausal paths, without closing the causal pathway.
  - If measurement error is likely you may elect to block a pathway at additional points using further variables.
  - Avoid collinear explanatory variables.
3. Caution when including any postexposure variables:
  - Do not include mediators or colliders.
  - Carefully consider the implications of including any descendants of the exposure.
4. Consider causes (particularly unmeasured) of any variables you are including.
  - Any collider pathways that are opened should be closed.

## Advantages and Limitations of Causal Diagrams

Causal diagrams require us to clearly define our primary research question, and they intuitively and efficiently communicate our existing understanding of relationships between important variables. Directed acyclic graphs can identify confounding and selection bias, and enable researchers to find strategies to overcome these. Furthermore, directed acyclic graphs can facilitate constructive critique of study design and analyses by making any underlying assumptions explicit—and therefore open to debate; they can also help us to make sense of adjusted associations we see reported in studies.

Composing directed acyclic graphs is not always easy; even with seemingly simple causal questions the diagram can soon evolve into a large complex network. The accuracy of any causal inferences made will be limited by the fidelity of the causal diagram on which they are based, which in turn relies upon existing subject matter knowledge. The

direction of arrows or even the presence or absence of relationships may be uncertain. This may necessitate sensitivity analyses, adjusting for different variables for each putative causal structure. Directed acyclic graphs are qualitative rather than quantitative, showing the presence or absence, and the direction of relationships; they do not reveal the nature or strength of relationships and may fail to capture interactions. Causal diagrams cannot protect us from unknown unmeasured confounding, and it is unclear whether the assumption of no residual or unmeasured confounding ever holds in real-world situations. Nevertheless, the directed acyclic graph at least makes these, and other important assumptions, explicit.

Directed acyclic graphs have well understood mathematical properties that are useful and readily exploited in more complex causal contexts that presented here. Examples include parametric and nonparametric structural equation modeling,<sup>22</sup> instrumental variable analysis,<sup>23</sup> quantitative bias analysis,<sup>24</sup> and mediation techniques.<sup>25</sup> Causal diagrams are testable, in that the causal structure implies certain statistical associations and independences that can be confirmed or refuted with data; however, this is not a simple endeavor as data may be consistent with multiple causal diagrams. Also, we cannot ignore random error, and because the first iteration of causal diagrams will rarely contain definitive information about magnitudes of the associations represented, the information required to power a study will be even more subjective than conventional single variable power analyses. An iterative process of using data to test models, and using the models to determine causal effects, is likely to be the way forward.

## The Future of Causal Diagrams in Anesthesiology Research

We are in the midst of a reproducibility crisis in the scientific and clinical literature. Where causal effects truly exist, the evidence from different methodologies should converge; but we often see unexpected and conflicting results from randomized, controlled trials and observational studies, which might contradict our understanding of basic science principles. More rigorous handling of random error is typically seen as the solution to this problem; however, this usually only addresses error arising from sampling variation. Currently much emphasis is placed on rigid prespecification of analysis plans,<sup>26</sup> with relatively little attention paid to rigorous assessment of potential sources of bias and how this will be managed. In contrast, we believe that much of the problem lies in poor study design and analysis of data, and that the use of causal diagrams could raise standards considerably.

Anesthesiology and perioperative medicine is particularly fraught with confounding, given the wide variety of patients, surgeries, and anesthesia techniques encountered. We are often looking to identify relatively small effects within complex causal systems. Indications and selection criteria for surgery itself might produce unexpected baseline associations between important variables within our surgical populations, and we are susceptible to the introduction

### Box 5. Confounding – Traditional *versus* Causal Approach

Confounding has historically been defined in terms of individual confounder variables rather than conceptually as confounding pathways. Statistically, confounders have often been defined as variables *associated* with both the exposure and the outcome. Unfortunately this definition encourages the inclusion of mediators and colliders as covariates in statistical models, and makes interpretation of the adjusted associations near impossible.

The causal approach incorporates subject matter knowledge about the nature of associations between variables and focuses on blocking noncausal *pathways*, rather than considering each variable individually. One advantage of this approach is that data collection can be streamlined by avoiding unnecessary and misleading variables. Collider variables need not always be avoided if the pathway can be blocked using another variable, but generally this needs very careful consideration.

of odd collider biases when variables are included in statistical models without careful consideration. Observational anesthesia research is always potentially compromised by selection bias, measurement error and confounding; but with careful study design and analysis this need not be the case. The era of Big Data will bring us more information than ever before and great opportunity, but needs a consistent framework or we will be drawn to wrong conclusions. Causal diagrams can provide this framework and give us more confidence in the validity of our results.

Directed acyclic graphs have been largely absent from our literature thus far, and they may be unfamiliar to most clinicians and researchers. Box 1 contains a checklist readers

### Box 6. Where to Find More Information on This Topic

1. *The Book of Why: The New Science of Cause and Effect*, by J. Pearl and D. MacKenzie. An accessible, plain-language introduction to the Causal Revolution that includes further examples of causal diagrams.
2. [www.DAGitty.net](http://www.DAGitty.net) is a website where you can create your own directed acyclic graphs, visualize causal and noncausal pathways, and identify variable sets for statistical adjustment. It also provides the statistical independences implied by your directed acyclic graph that can be tested with data to support or refute your causal model. It also contains links to further articles and freely available R-based statistical packages.
3. <https://online-learning.harvard.edu/course/causal-diagrams-draw-your-assumptions-your-conclusions> (accessed August 2019). Miguel Hernan presents a series of introductory videos in this course from Harvard University (basic content free of charge).

may wish to refer to should they encounter causal diagrams in the literature, and box 6 contains some links to further introductory material. Causal diagrams are not featured in the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) recommendations for the reporting of observational research<sup>27</sup>; however, a consortium of clinical journals has recently recommended their inclusion in reports of observational studies,<sup>6</sup> and they are well-established in the epidemiology literature. Many anesthesia reports would benefit from inclusion of causal diagrams.

## Conclusions

Modern causal inference techniques promise to take us beyond the paralyzing *association ≠ causation* mantra, which has limited the advancement of clinical knowledge from observational research in recent times. By providing a rigorous approach to bias, causal diagrams can inform better study design and analyses and potentially allow us to make robust statements about causation from observational data. When we properly understand causal structures, the clinician will be clearer about what—and what not—to worry.

## Acknowledgments

The authors thank Prof. Tony Blakely, M.B.Ch.B., M.P.H., Ph.D., F.N.Z.C.P.H.M., University of Melbourne, Victoria, Australia and University of Otago, Dunedin, New Zealand, for carefully reviewing the initial manuscript and providing some guidance. The authors also thank colleagues from Waikato Hospital and the reviewers for their useful comments and suggestions which have helped to further improve the manuscript.

## Research Support

Supported by Project Grant No. 17/009 from the Australian and New Zealand College of Anaesthetists (Melbourne, Australia; to Dr. Gaskell). This work was also generously supported by the Department of Anaesthesia and Pain Medicine, Waikato Hospital, Hamilton, New Zealand and the Department of Anaesthesiology, University of Auckland, Auckland, New Zealand.

## Competing Interests

The authors declare no competing interests.

## Correspondence

Address correspondence to Dr. Gaskell: Department of Anaesthesia and Pain Medicine, Waikato Hospital, Corner Pembroke Street, Hamilton 3240, New Zealand. amy.gaskell@waikatodhb.health.nz. Information on purchasing reprints may be found at [www.anesthesiology.org](http://www.anesthesiology.org) or on the masthead page at the beginning of this issue. ANESTHESIOLOGY's articles are made freely accessible to all readers, for personal use only, 6 months from the cover date of the issue.

## References

1. Stolley PD: When genius errs: R. A. Fisher and the lung cancer controversy. *Am J Epidemiol* 1991; 133:416–25; discussion 426–8
2. Sessler DJA: Big Data—and its contributions to peri-operative medicine. *Anaesthesia* 2014; 69: 100–105
3. Pearl J: Causal diagrams for empirical research. *Biometrika* 1995; 82: 669–688
4. Greenland S, Pearl J, Robins JM: Causal diagrams for epidemiologic research. *Epidemiology* 1999; 10:37–48
5. Shrier I, Platt RW: Reducing bias through directed acyclic graphs. *BMC Med Res Methodol* 2008; 8:70
6. Maslove DM, Leisman DE: Causal inference from observational data: New guidance from pulmonary, critical care, and sleep journals. *Crit Care Med* 2019; 47:1–2
7. Sargeant J, O'Connor AM, Dohoo I, Erb H, Cevallos M, Egger M, Ersbøll AK, Martin S, Nielsen LR, Pearl D: Methods and processes of developing the Strengthening the Reporting of Observational Studies in Epidemiology–Veterinary (STROBE–Vet) statement. *J Vet Int Med* 2016; 30: 1887–95
8. VanderWeele TJ, Asomaning K, Tchetgen Tchetgen EJ, Han Y, Spitz MR, Shete S, Wu X, Gaborieau V, Wang Y, McLaughlin J, Hung RJ, Brennan P, Amos CI, Christiani DC, Lin X: Genetic variants on 15q25.1, smoking, and lung cancer: An assessment of mediation and interaction. *Am J Epidemiol* 2012; 175:1013–20
9. Elwert F, Winship C: Endogenous selection bias: The problem of conditioning on a collider variable. *Annu Rev Sociol* 2014; 40:31–53
10. Hernán MA, Hernández-Díaz S, Robins JM: A structural approach to selection bias. *Epidemiology* 2004; 15:615–25
11. Westreich D, Greenland S: The table 2 fallacy: Presenting and interpreting confounder and modifier coefficients. *Am J Epidemiol* 2013; 177:292–8
12. Wu ZF, Lee MS, Wong CS, Lu CH, Huang YS, Lin KT, Lou YS, Lin C, Chang YC, Lai HC: Propofol-based total intravenous anesthesia is associated with better survival than desflurane anesthesia in colon cancer surgery. *ANESTHESIOLOGY* 2018; 129:932–41
13. Williamson EJ, Aitken Z, Lawrie J, Dharmage SC, Burgess JA, Forbes AB: Introduction to causal diagrams for confounder selection. *Respirology* 2014; 19:303–11
14. Staplin N, Herrington WG, Judge PK, Reith CA, Haynes R, Landray MJ, Baigent C, Emberson J: Use of causal diagrams to inform the design and interpretation of observational studies: An example from the Study of Heart and Renal Protection (SHARP). *Clin J Am Soc Nephrol* 2017; 12:546–52
15. Schisterman EF, Cole SR, Platt RW: Overadjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiology* 2009; 20:488–95



16. Greenland S: Quantifying biases in causal models: Classical confounding vs collider-stratification bias. *Epidemiology* 2003; 14:300–6
17. Liu W, Brookhart MA, Schneeweiss S, Mi X, Setoguchi S: Implications of M bias in epidemiologic studies: A simulation study. *Am J Epidemiol* 2012; 176:938–48
18. Charlson M, Szatrowski TP, Peterson J, Gold J: Validation of a combined comorbidity index. *J Clin Epidemiol* 1994; 47:1245–51
19. Schisterman EF, Perkins NJ, Mumford SL, Ahrens KA, Mitchell EM: Collinearity and causal diagrams—A lesson on the importance of model specification. *Epidemiol* 2017; 28: 47
20. Greenland S: The effect of misclassification in the presence of covariates. *Am J Epidemiol* 1980; 112:564–9
21. Brenner H, Blettner M: Controlling for continuous confounders in epidemiologic research. *Epidemiology* 1997; 8:429–34
22. Lomax RG, Schumacker RE: A beginner's guide to structural equation modeling, 2nd edition, Psychology Press, 2004
23. Greenland S: An introduction to instrumental variables for epidemiologists. *Int J Epidemiol* 2000; 29:722–9
24. Lash TL, Fox MP, Fink AK: Applying quantitative bias analysis to epidemiologic data, Springer Science & Business Media, 2011
25. VanderWeele TJ: Mediation analysis: A practitioner's guide. *Annu Rev Public Health* 2016; 37:17–32
26. Sleight J: Is prespecification taking the search out of research? *Paediatr Anaesth* 2019; 29:216–7
27. von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative: The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Ann Intern Med* 2007; 147:573–7

## GLOSSARY OF TERMS

Term (Alternative)	Meaning
Adjusted association	The association observed after statistical analysis. This will consist of causal pathways, any residual confounding pathways that exist, and any pathways that have been opened due to conditioning.
Ancestor (cause)	A factor which influences another specific variable, no matter how distantly along a pathway. A direct cause of another variable can be referred to as a parent.
Arrow (directed edge)	In graphs represents a directed causal relationship between two variables
Bias	Systematic deviation from the true causal relationship. Nonrandom error.
Collider	A variable which is a common effect of two other variables. It is the variable at an inverted fork in a directed acyclic graph ( <i>e.g.</i> , variable Z in the pathway $X \rightarrow Z \leftarrow Y$ ). The collider blocks the pathway.
Conditioning	Refers to a variety of actions which essentially limit a variable to a certain level or range ( <i>e.g.</i> , selection criteria, stratification, matching, inclusion in multivariable regression model). Conditioning is denoted in a directed acyclic graph by a box around the variable.
Confounder	A shared cause of two variables or indeed any variable on a confounding pathway which is not a collider on that pathway.
Confounding	Noncausal association which occurs because the exposure and the outcome share common causes. Evident in directed acyclic graphs as any open noncausal pathway between the exposure and the outcome, which begins with an arrow directed into the exposure.
Crude (unadjusted) association	The association observed in the data before any statistical adjustment. This is the net association generated by all open pathways between the exposure and outcome of interest ( <i>i.e.</i> , the causal pathways, any confounding pathways that exist, and any pathways that have been opened as a result of selection factors).
Descendant (effect)	A factor that is influenced by another specific factor. In the path $C \rightarrow X$ , X would be regarded as a descendant of C. A direct descendant is sometimes referred to as a child.
Directed Acyclic Graph	A diagram consisting of variables linked by arrows which indicate directed causal relationships. No feedback loops are present.
Exposure (treatment, cause)	The causal variable of interest.
Mediator (intermediary variable)	Any intermediate on the causal pathway between the exposure and the outcome; this factor is responsible for part of the causal effect of the exposure.
Outcome (effect)	The effect of interest.
Pathways	Any series of arrows and variables linking two variables.
Open pathway	A pathway which generates a statistical association.
Closed pathway	A pathway which does not generate a statistical association.
Causal pathway	A pathway in which all arrows are directed from the exposure to the outcome.
Noncausal pathway	Any pathway between variables which is not causal.
Selection bias	Statistical association generated by a noncausal pathway opened by conditioning on a collider variable.
Variable (node, covariate)	A factor or construct that can take different values and is considered to be of importance in the causal system.