How to use directed acyclic graphs: guide for clinical researchers

Timothy Feeney,[1] Fernando Pires Hartwig,[2,3] Neil M Davies,[4,5,6]

[1]Gillings School of Global Public Health, Department of Epidemiology, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA
[2]Postgraduate Program in Epidemiology, Federal University of Pelotas, Pelotas, Brazil
[3]Medical Research Council Integrative Epidemiology Unit at the University of Bristol, Bristol, UK
[4]Division of Psychiatry, University College London, London, UK
[5]Department of Statistical Sciences, University College London, London, UK
[6]K G Jebsen Centre for Genetic Epidemiology, Department of Public Health and Nursing, Norwegian University of Science and Technology, Trondheim, Norway

Correspondence to: T Feeney
feeney@unc.edu
(or @tfeend on X and @tfeend.bsky.social on bluesky;
ORCID 0000-0001-8139-5830)

Additional material is published online only. To view please visit the journal online.

Cite this as: *BMJ* 2025;388:e078226
http://dx.doi.org/10.1136/bmj-2023-078226

Accepted: 17 June 2024

Directed acyclic graphs are commonly used to illustrate and assess the hypothesised causal mechanisms in health and social research. These graphs can illuminate investigators' assumptions and help clearly describe each possible explanation for associations observed in data given researchers' assumptions, ranging from causal effects to confounding and selection bias, and thereby help identify variables that can be used to reduce or overcome bias. This article explains how to construct, interpret, and present directed acyclic graphs as part of clinical research studies and how they can help communicate a study's strengths or limitations.

Causal directed acyclic graphs (DAGs) are a type of graph that illustrates an assumed causal structure between variables of interest. These graphs can illustrate assumed links between possible causes (eg, a behaviour or a medical intervention; referred to in this article as exposure) to possible consequences (eg, presence or absence of disease; referred to in this article as outcome).[1] While causal graphs have long been used,[2] DAGs have a relatively short history in epidemiological research[3] but have become widespread as a way to think about the causal structure underlying an exposure-outcome association.[4 5] DAGs can be useful for many purposes, such as helping to identify confounders,[6 7] evaluating potential selection bias,[8 9] and understanding the roles that measurement error[10 11] and missing data[12] might have in effect estimation. Recent papers have highlighted how DAGs can improve epidemiological[13-15] and clinical studies.[16-19] However, they can also aid in understanding descriptive studies (eg, estimating the incidence of disease) and prediction studies (eg, modelling a patient's risk of disease). These graphs can also help communicate the assumptions necessary to interpret results to collaborators, researchers, reviewers, readers, and editors.

Despite their potential utility, wide variation in the use of DAGs can limit their effectiveness. In a review of 234 articles using DAGs, researchers found increasing use of these graphs but a wide variation in how they were used,[7] with relatively few studies reporting key information. Improper development and use could result in failure to reap the benefits of using these graphs. Here, we explain why DAGs are helpful for biomedical research, highlight some limitations of these graphs, and suggest how to construct and disseminate a DAG collaboratively when conducting clinical research. While these graphs can be useful in many scenarios, we focus here on observational research, where the goal is to estimate the total effect of an exposure on an outcome.

SUMMARY POINTS

- Directed acyclic graphs can help guide study design, data collection, and analyses assessing the causal effect of an exposure on an outcome (eg, a behaviour or a medical intervention) and outcomes (eg, presence or absence of disease)
- These graphs can illustrate potential sources of bias, determine key variables (observed or unobserved), and ascertain which variables should (and should not) be selected as covariates to control for these sources of bias
- Causal diagrams should be constructed based on expert clinical and subject matter knowledge
- This article provides a glossary of commonly used terms in directed acyclic graphs

Definition of directed acyclic graphs

DAGs are founded in graph theory and consist of a few basic elements, which can be illustrated with a simplified example investigating the effect of exposure to heavy alcohol use on the outcome of all cause mortality (fig 1). Figure 2 includes a glossary of terms and illustrations of the basic components of DAGs. Box 1 provides additional concepts and explanations on how these graphs can be used and interpreted.[20]

A DAG comprises nodes and edges. Nodes represent the variables in a study. In figure 1, the nodes are heavy alcohol use (the exposure), death (the outcome), and socioeconomic position (the confounder; fig 2 lists detailed definitions of the different aspects of DAGs). Nodes are connected by arrows, also known as directed edges, which are unidirectional and indicate the direction of a causal effect between two nodes. For clarity, we refer to edges as arrows. Specifically, an arrow indicates a direct causal effect not mediated by other variables in the DAG (however, such effects can also be, and likely are, mediated by other variables not included in the graph), with the direction of the arrow indicating causal ordering (and thus temporal ordering). The node at the tail of an arrow causes the node at the head of the arrow, which implies that, for at least one person in the studied population, an intervention that changes the variable at the tail would also result in a change in the variable at the head.

Arrows do not split or merge and should only originate from and end at a node; arrows should not start or end on the midpoint of another arrow. DAGs cannot indicate the magnitude of the effects, the form of

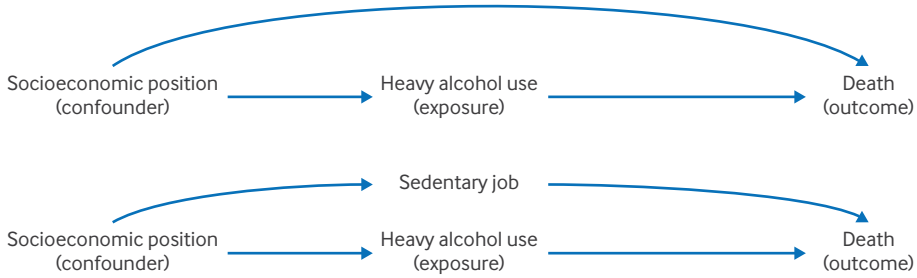


Fig 1 | Example of confounding shown in a directed acyclic graph with exposure to heavy alcohol use, outcome of death, and confounder of socioeconomic position. Top: socioeconomic position is shown as a confounder because it is a common cause of the exposure and outcome. Bottom: a sedentary job is shown as a confounder, owing to being on a confounding path between the exposure and outcome

the effect (eg, linear, quadratic) and whether the effect varies between individuals. The DAG in figure 1 has an arrow from the socioeconomic position to the heavy alcohol use node, which implies that socioeconomic position has a direct causal effect of unspecified direction, magnitude, and form on heavy alcohol use for at least one individual (and, if the effect exists for multiple individuals, its direction, magnitude, and/or form might differ between them).

Two nodes connected by a series of arrows oriented in the same direction imply an indirect causal effect. For instance, socioeconomic position has a direct causal effect on death in figure 1A, whereas it also has an indirect causal effect on death in figure 1B, mediated by a sedentary job. The total causal effect of socioeconomic position on death combines all these individual causal effects.

Finally, DAGs are acyclic, meaning that no loops of arrow-node combinations should lead out from one node and back to the initial node. The acyclic nature is because causal ordering implies temporal ordering—the variable at the tail of an arrow must come temporally before the variable at the head of that arrow. Thus, a graph with cyclical loops would imply that a variable in the future is affecting its past self. Accommodating time requires multiple nodes to represent different time points and avoid cycles; this concept is discussed in more detail below.

Whether a variable is an outcome, exposure, or confounder depends on the research question of interest (eg, socioeconomic position in figure 1 could be an exposure in another study). The validity of the conclusions obtained from a DAG depends on the plausibility of the assumptions encoded in the graph. For example, socioeconomic position is an assumed confounder of the exposure-outcome association in figure 1; however, this is not proof that it is a confounder in the study but merely a statement of the researchers' assumptions. Once the set of assumptions is given in the structure of the DAG, conclusions from the graph can be inferred. Therefore, the crucial step is drawing out the DAG, which is when subject matter knowledge is incorporated into the graph. Below, we provide general guidance on how to draw a DAG.

Use of directed acyclic graphs in clinical research

DAGs are useful because they force researchers to make explicit assumptions about the causal structure underlying the associations between variables. Therefore, DAGs can identify the variables that should be conditioned on and, equally importantly, those that should not be conditioned on according to the researchers' assumptions.[21 22] Below, we describe how researchers can use DAGs to identify and describe the role of different potential variables in analyses: the exposures, outcomes, confounders, minimum sufficient adjustment sets, mediators, selection, colliders, and effect modifiers.[23 24]

Identifying confounders and minimally sufficient adjustment sets

DAGs can identify confounders and minimally sufficient adjustment sets (box 1). Confounders are variables that lay on a confounding path, which is a path that includes common causes of the exposure and outcome nodes in the graph. This path might be as simple as a node with arrows going directly into the exposure and the outcome in the DAG (ie, socioeconomic position is a confounder and a direct common cause of exposure and outcome in fig 1A). A confounder might also be a node on a path that includes a common cause of the exposure and outcome. For example, in figure 1B, socioeconomic position is a confounder even though it indirectly causes death through the sedentary job node. Further, having a sedentary job is a confounder because it is on a common cause path that leads to the outcome and is connected to the exposure through its direct cause, socioeconomic position. These confounding paths create what are sometimes referred to as backdoor paths, even though not all backdoor paths are confounding paths (eg, they can arise by conditioning on a collider, a concept that we explain below).

Confounders induce an association between the nodes they cause, which could lead to bias when the goal is to estimate the causal effect of one node on another. For example, in figure 1, socioeconomic position is a confounder of the association between heavy alcohol use and death. This confounding means

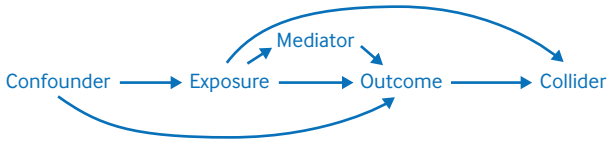
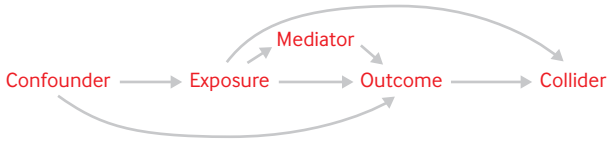
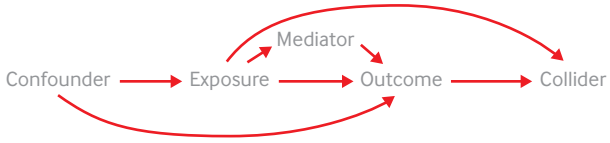
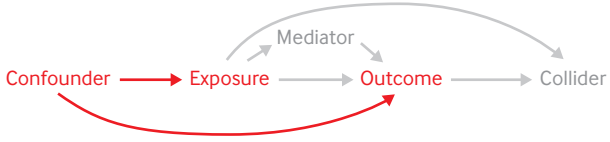
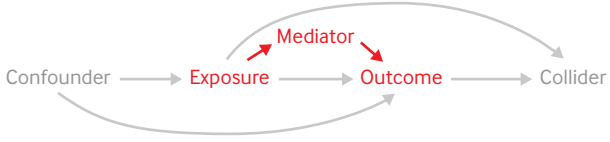
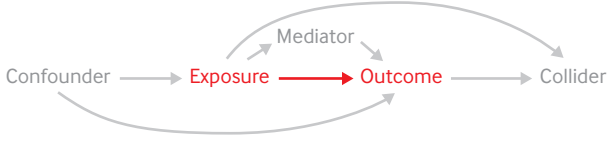
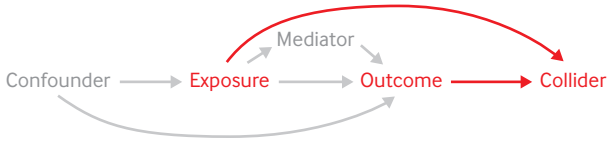
Term	Definition	Visual depiction on directed acyclic graph
Causal directed acyclic graph (DAG)	Graphical depiction of the assumed causal connection between the exposure or treatment of interest, outcome of interest, and variables that have an assumed direct causal effect on the exposure and outcome while also including connections between the variables included in the DAG	
Node(s)	A variable in a DAG. One node represents each variable at each time point	
Directed edge(s) or arrow	A connection from one node to another typically depicted as an arrow. The edge is unidirectional and indicates the direction of causation between two nodes	
Confounder	A variable on a confounding path that links the exposure and outcome. Can be a single node that is a common cause of both the outcome and the exposure or treatment. Every node on a confounding path is also a confounder. For example, a cause of the outcome that is affected by a cause of the exposure, or a cause of the exposure that is affected by a cause of the outcome	
Mediator	A variable on at least one causal path between the exposure or treatment and the outcome	
Indirect effects	Effects are mediated, indirectly, from exposure to outcome through a definable mediator variable	
Direct effects	Effects directly from exposure to outcome without mediation through an intermediate variable	
Collider	A variable that is a common consequence of two variables that, if conditioned on (eg, through regression adjustment, stratification, or matching), can open a non-causal path of association between the exposure or treatment and the outcome of interest	
Minimally sufficient adjustment set	The minimal set of variables needed to adjust for potential bias	
Selection bias	When the study population selected differs from the target population, which can be represented in a DAG by including a node denoting inclusion into the study. One type of selection bias is caused by conditioning on a collider (see figure for Collider above). See Lu et all for extensive discussion and graphical description[9]	

Fig 2 | Glossary of terms used in studies with directed acyclic graphs

that the crude (eg, not accounting for confounding by socioeconomic position) association of heavy alcohol consumption and death will be a biased estimate of the causal effect of heavy alcohol consumption on death.

A minimally sufficient adjustment set is a group of measured variables that is sufficient to remove bias due to confounding, leaving only the causal paths from the exposure to the outcome (ie, all paths formed by arrows oriented in the same direction from exposure into outcome).[4 15] This set is considered minimal because removing any variable from it would result in residual confounding. Minimally sufficient adjustment sets can be found even if relevant variables were not measured.[25] For instance, consider figure 3, which adds an unmeasured variable for dietary quality to figure 1. Dietary quality is a common cause of both heavy alcohol consumption and death, and the DAG indicates that dietary quality would need to be conditioned on; otherwise, there will be a confounding path between the exposure and outcome. Here, we used conditioning to refer to adjusting for a variable (eg, by including it in an outcome regression

Box 1: Guide to use and evaluate directed acyclic graphs

General considerations and rationale

- Were the directed acyclic graph(s) and analysis pre-registered? Pre-registration helps prevent p-hacking and avoids building a directed acyclic graph based on available data instead of the assumed data-generating mechanism, making the paper's inferences more reliable.
- Is the directed acyclic graph reasonable? Does it include all the relevant relations, and have confounding, selection and other sources of bias been considered? If the graph is not reasonable, or not all likely sources of bias have been considered, what nodes and/or edges should be added?
- Are there aspects of the graph that suggest additional analyses that can help evaluate whether a potential source of bias will likely affect the results (eg, quantitative bias analysis if the directed acyclic graph suggests unmeasured confounding) and corresponding inferences?

Suggested steps to building a directed acyclic graph and specific questions for checking

1. Identify the target population and define the research question, precisely defining the exposure and outcome of interest:
   - Is the definition of exposure and outcome precise enough to evaluate the question of interest?
   - If not, can the definitions be made more precise? Will this require adapting the original research question?
2. Identify all possible variables that cause at least one of the following variables: the exposure, the outcome, or selection into the study by reviewing the literature and discussing it with experts in the subject matter, including those familiar with constructing directed acyclic graphs:
   - Have variables been included even though they are likely to be unmeasured or unmeasurable?
   - Has a sufficient and replicable literature search been performed to understand the relevant connections between variables?
   - Have substantive experts been involved?
   - Have additional variables with direct effects on any pair of the variables identified as described above also been included?
   - Have all meaningful connections between any two pairs of variables been included?
   - Have omitted variables and connections been well motivated?
3. Iterate on the directed acyclic graph with additional experts until consensus is reached; include the consensus graph in any pre-registration:
   - Are all experts in agreement with the graph constructed?
   - If not, have areas of disagreement identified a set of candidate graphs that can be used to guide sensitivity analysis?
   - Has the graph been included in pre-registration documentation?
4. Identify variables for data collection or the appropriate dataset based on the consensus directed acyclic graph analysis methods based on the graph:
   - Have the necessary confounding variables been included in the graph?
   - Have potential biases from conditioning on a mediator or collider been considered?
   - If using data that have already been collected, have variables that lead to selection into that data been considered?
5. Choose analysis methods, measures of the outcome and exposure, and covariates based on those defined by the research question and the consensus directed acyclic graph.
6. Based on variables that are unavailable or prone to measurement error, identify sensitivity analyses that can evaluate the impact of unmeasured variables and measurement error:
   - Have all unmeasured and potentially unidentified variables been considered?
   - Have causal mechanisms been considered between unmeasured, potentially unidentified, and potentially mismeasured variables?
   - Has the mismeasurement of variables been considered?
   - Have sensitivity and quantitative bias analyses been considered to deal with unmeasured, unidentified, and mismeasured variables?
7. Include the directed acyclic graph in published works and refer to the graph when describing adjustment sets:
   - Has a complete graph been included in the main document or the supplement in the publication?
   - Has the analysis been explicitly guided by the graph?
   - Has the code to facilitate reproducibility been included?

model, matching, or stratification). Thus, because socioeconomic position and dietary quality are confounders and conditioning on them is sufficient to remove confounding bias, a minimally sufficient adjustment set for the DAG in figure 3 is socioeconomic position and dietary quality. The confounding path due to dietary quality can be blocked by conditioning on it. In this case, the confounder would either need to be measured, or if unmeasured it would need a sensitivity analysis performed using quantitative bias analysis to evaluate how much the bias from this confounder affects the results.

These examples illustrate how using DAGs assist in making causal assumptions explicit and aid in identifying the potential adjustment sets to mitigate bias. Importantly, there might be no minimally sufficient adjustment sets containing only measured variables. For instance, in figure 1B, there are two minimally sufficient adjustment sets, each comprising only one variable: (1) socioeconomic position and (2) sedentary job. Therefore, even if the actual common cause between exposure and outcome (here socioeconomic position) had not been measured, it would still be possible to remove confounding bias by adjusting for sedentary job. However, in figure 3 there is only one minimally sufficient adjustment set containing socioeconomic position and diet quality—there is no alternative set to get around the need to condition on these variables. Detecting these situations where requisite variables might be missing from the data or unmeasured aids researchers in assessing the likely bias in their results, designing sensitivity analysis,

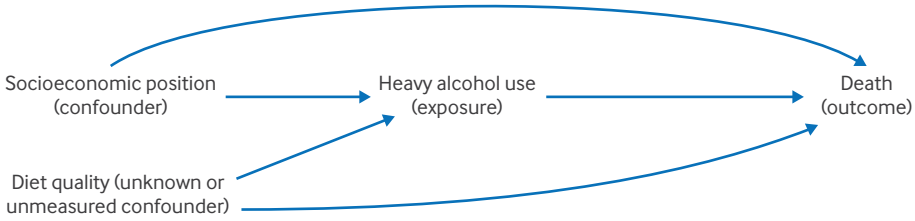


Fig 3 | Example of unmeasured confounding in a directed acyclic graph from figure 1 with an additional unmeasured confounder: diet quality. In this case, diet quality is unmeasured in the data but is a common cause of heavy alcohol use and death

and considering alternative estimation methods (eg, instrumental variables or negative controls).[26]

Identifying mediators

DAGs can help identify mediators (fig 1).[21 22] Conditioning on mediators can bias estimates of the total effect of an exposure on an outcome because it will block the mediated (indirect) effects.[27] For example, in figure 4, heavy alcohol use not only affects mortality but also affects fast food consumption, which in turn affects mortality. Thus, the total effect of heavy alcohol consumption is partly due to its direct effects (illustrated by the arrow from heavy alcohol use to death) and partly due to the indirect effects from heavy alcohol use to fast food consumption to death (the arrows from heavy alcohol use to smoking to death). From this, we can conclude that conditioning on fast food consumption will potentially lead to bias if the goal is to estimate the total effect of heavy alcohol consumption on mortality. Moreover, conditioning on a mediator can lead to collider bias (see below). Additional considerations are important when conditioning on a mediator (eg, when the goal is to estimate direct effects), which are beyond the scope of this manuscript.[28]

Identifying colliders and selection bias

DAGs can aid in the identification of variables impacting selection (fig 1). Selection bias might occur whenever the estimated causal effect in a population sample differs from the true causal effect in the population of interest as a result of selecting a sample from the population of interest.[9] Potential sources of selection bias include selection into the study (ie, eligibility criteria that differ across groups); missing data after selection into the study, which could happen because of loss of follow-up or non-response to questionnaire sections; intentional analytical stratification or restriction (eg, for estimation of subgroup specific effects). One specific form of selection bias relates to a node caused by two other nodes (ie, having two or more arrows pointing into it), which is known as a collider. If the exposure-outcome association is conditioned on such a variable, it can induce spurious associations, known as collider bias.[8 21] More recently, analyses of SARS-CoV-2 have been prone to this bias and are a poignant reminder of its importance.[29 30] Importantly, DAGs can help to identify colliders, which should generally not be conditioned on and can also help identify variables that can help ameliorate the bias from conditioning on a collider when conditioning on a collider is unavoidable (eg, when selecting a sample).

The example in figure 5 illustrates the effect of heavy alcohol use on all cause mortality, but is restricted to people aged 65 years and older. A selection node has been added, along with a node for age, which is a confounder. Many factors will influence the selection of this study; for example, to live to 65 years old, participants would have needed to live healthily and by definition not have died. Moreover, those that drink alcohol heavily are potentially less likely to end up in a research study. These variables will affect whether participants are selected, which suggests that there is an arrow leading from age, heavy alcohol use, and death into selection. Selection into the study is, therefore, a collider of the exposure-outcome association, leading to a biased estimate of the causal effect of the exposure on the outcome.

DAGs have many additional uses. For instance, these graphs can describe assumptions regarding measurement error and its causes, thus helping to understand information bias and identify mitigation strategies.[10] They can also be extended to conceptualise potential effect modifiers.[23] DAGs can also be applied to time-varying exposures, outcomes, and covariates and can describe their potentially complex mechanisms, such as past and current smoking.[31]

The use of DAGs to understand the impact of time-varying variables includes the identification of covariates to adjust for time-varying confounding when estimating the effect of exposure over multiple time points. For example, when investigating the influence of heavy alcohol use on death, using data on a friend's heavy alcohol use at baseline and at a later time point and smoking at two later time points can help understand the seemingly cyclical or apparent

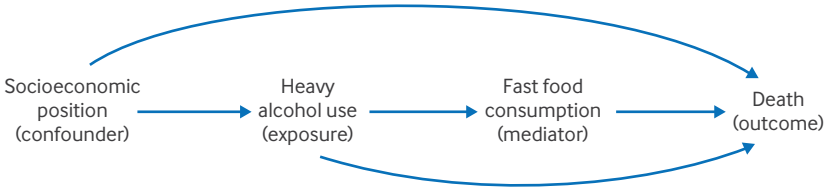


Fig 4 | Example of mediation shown in a directed acyclic graph. Fast food consumption is based on a mechanism by which the effect of heavy alcohol use on death is mediated

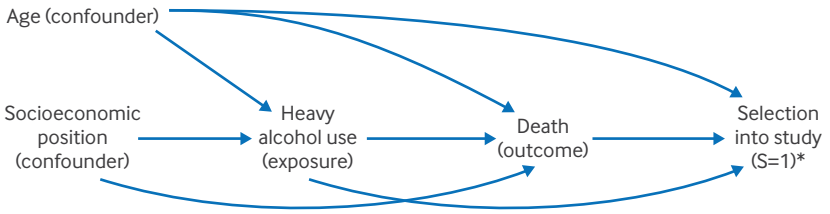


Fig 5 | Example of selection shown in a directed acyclic graph from figure 1 with the addition of selecting on age 65 years or above. Socioeconomic position and age are confounders, and mortality is the outcome of interest. This figure includes a node highlighting selection into the study and potential bias that can result from adjusting for this variable. *Selection (S) into study denotes restricting to S=1 (ie, to a single level of the variable selection)

bi-directional causal structure. We can use a DAG to illustrate the assumed causal mechanisms between these variables over time and, from these assumptions, select covariates for confounding adjustment when estimating the effect of sustained heavy alcohol use over time (supplemental fig 1).

Implementing directed acyclic graphs in clinical research

Based on previously published recommendations[7] (summarised in box 2) and our experience with DAGs, we now suggest a guideline of seven steps for drawing and using these graphs to improve study design and data analysis, and we provide a checklist to evaluate each step in box 1. These steps offer an introduction to DAGs and available algorithmic approaches.[15 32] Implementing these steps does not require software, but some tools can help construct DAGs to ease the process further [BMJ Medicine DAG TOOLs Primer.[33 34] For each suggested step, we describe building a DAG with examples. In our example, the effect of heavy alcohol use on death is investigated, and we will consider both a prospective and retrospective case when a dataset might be present or not.

While these steps represent an idealised scenario, the circumstances of how and when DAGs can be used vary greatly. Therefore, the following steps are intended to illustrate generally how these graphs can help researchers rather than proscribe how they should be used.

Box 2: Recommendations for using directed acyclic graphs in applied health research (Tennant et al 2021)[7]

- Associations of interest and estimands should be stated in the study's aims
- Studies should report directed acyclic graph for each association of interest
- Directed acyclic graphs should include all relevant measured and unmeasured variables
- Nodes should be arranged visually so that all edges flow in the same direction
- Edges should be assumed between all nodes unless there is a strong reason to assume otherwise
- Minimum sufficient adjustment sets should be clearly stated
- Estimates using the minimally sufficient adjustment set, or the nearest available approximation, should be reported
- Alternative adjustment sets should be justified and reported separately

Step 1: Identify the target population and define the question, with precise definition of the exposure or treatment and outcome of interest

This step is to agree on the target population of interest. Next, the exposure (ie, the potential cause—eg, a behaviour, medical intervention, or risk factor of interest) and outcome (ie, presence or absence of disease) should be defined, and the specific causal question of interest, such as the average treatment effect in the population, can be agreed on. The parameter implied by the casual question is often termed the "estimand," which is what we are interested in estimating.[35-38] These aspects of the study should be reported to readers.

Example

Our example considers the effects of alcohol use on death. However, that question can be more specific. In other words, the exposure needs to be defined precisely, for example, as a diagnosis of alcohol use disorder, which acts as a measurable proxy of heavy alcohol use, and the outcome defined as the risk of mortality at five years among individuals aged 18 years or older. Our target population is the whole population of those of drinking age in the UK. The estimand we are interested in estimating is the average treatment effect in those who are treated, which is the effect of comparing everyone with alcohol use disorder to the same group if, counter to fact, they had not had alcohol use disorder.

Step 2: Identify all variables involved in the causal effect of interest by reviewing the literature and discussing it with subject matter experts, including those familiar with constructing directed acyclic graphs

Once the question of interest has been agreed on, the construction of most DAGs will start with the exposure and outcome. We also recommend adding a selection node, which should correspond not only to remaining in a cohort study over time but also to the criteria used to define the target population (fig 5). The confounders, colliders, mediators, and variables related (directly or indirectly) to the exposure or the outcome can then be added based on either literature supported evidence or expert input from the subject matter. As a practical recommendation, one can include all known causes of exposure, outcome, and selection in the DAG.[6 39] Then, one should articulate the causal mechanisms among the selected variables by adding arrows between each included variable and consider whether there are any variables missing from the DAG that have direct causal effects in two or more variables already included in the DAG. If so, such missing variables should be included, and their causal mechanisms with the other variables in the DAG should be articulated. At this stage, there are likely to be variables that cannot be measured but should still be included in the DAG. Importantly, arrows that are omitted between variables encode strong assumptions that direct causal effects do not exist.

For prospective observational studies, DAGs can inform the minimal set of needed variables to condition on and, therefore, that should be collected. In a retrospective study, the graph is ideally constructed before identifying the dataset or determining which variables are available. In other words, building the DAG based on the assumed causal structure and then identifying a dataset that suffices for bias adjustment. However, in a practical sense, many researchers might want to answer a question with data that have already been collected. In this case, DAGs might also be useful for understanding the implications of available and omitted variables, thereby helping researchers to understand the limitations of this already available dataset. For instance, if a DAG suggests that a retrospective study using a particular dataset is unlikely to have data on a key source of confounding, researchers can consider whether the question can be reliably answered with available data or whether sufficient sensitivity analyses and quantitative bias analysis,[40-42] aided by the graph, can be designed to better answer the question with the variables at hand. In this case, we suggest building a graph as described above and noting in the DAG which variables are

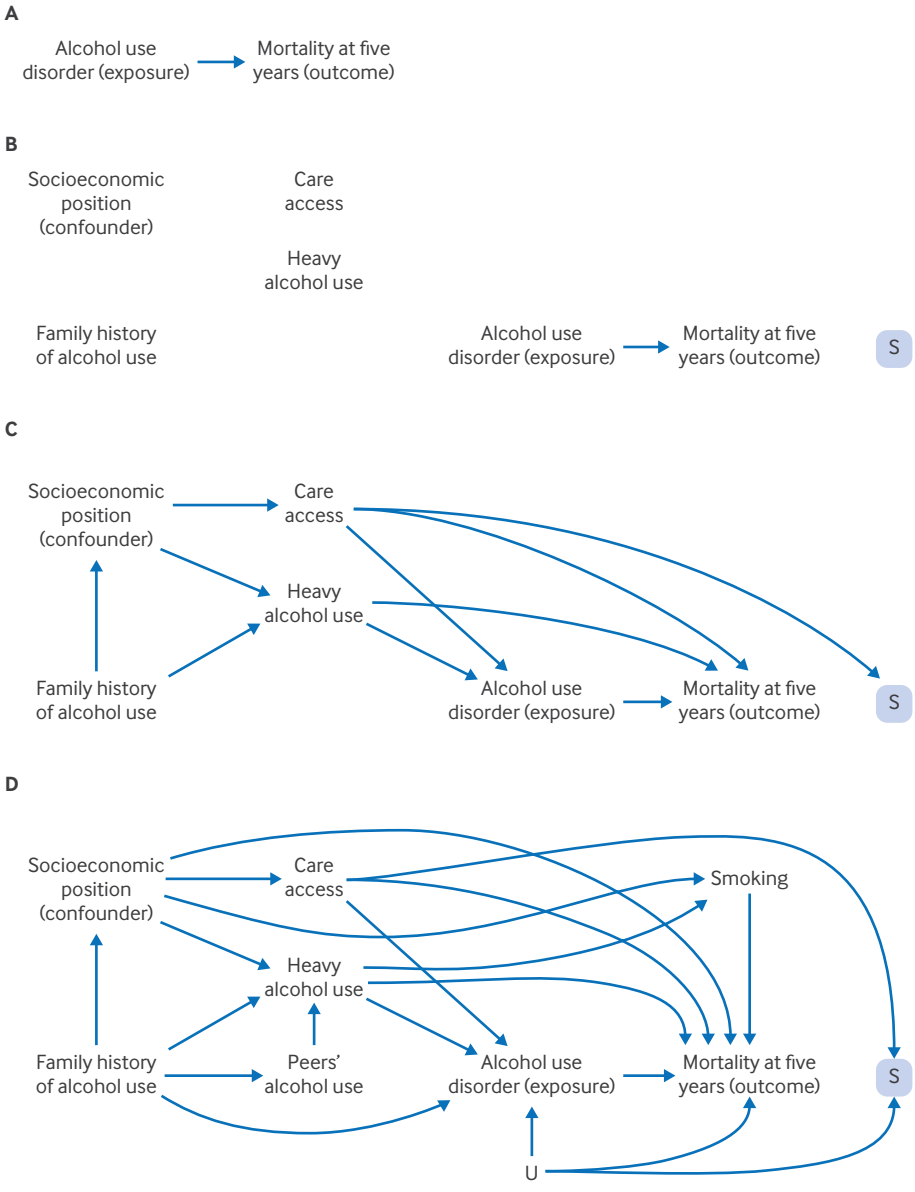


Fig 6 | Construction of a directed acyclic graph. (A) The process starts with defining the exposure (ie, a cause, such as a behaviour or medical intervention) and outcome (ie, presence or absence of disease) of interest. (B) Initial graphs can be constructed by using expert knowledge and previous studies to inform variables that are related to both the exposure and outcome, but also other variables. (C) These variables can then have causal effects indicated explicitly by adding arrows. (D) Refinement and consensus help complete the graph, including all relations between the variables, exposure, and outcome. U=unmeasured and possibly unknown confounders; S=selection (S=1 is selection into the study implied by the box around the S node)

missing so that adjustment sets can be searched within measured variables.

Example

Building the DAG begins by outlining the exposure-outcome effect of interest (fig 6A); at this stage, it is worth involving other experts, including other social scientists, substance use specialists, psychiatrists, and epidemiologists. Variables that have a relation with the exposure of interest and alcohol use disorder can then be added, including a selection node. In this example, expert knowledge and the literature suggest that four variables affect the diagnosis of alcohol use disorder: family history of alcohol use, socioeconomic position, care access, and alcohol use. Placing these four variables on the DAG, we can draw arrows to illustrate the causal assumptions (fig 6B and fig 6C):

- Family history affects socioeconomic position and alcohol abuse;
- Socioeconomic position affects care access and alcohol use;
- Care access affects alcohol use disorder diagnosis and mortality at five years;
- Alcohol use affects alcohol use disorder diagnosis and mortality at five years; and
- Care access affects selection into the study.

It is worth reiterating that not including a variable in the DAG is equivalent to assuming that such an omitted variable does not have a direct causal effect on any pair of variables already included in the DAG, and any omitted arrows imply that there is no direct causal effect between the pair of variables; these are strong assumptions.

Step 3: Develop the directed acyclic graph with additional experts until consensus is reached, and include consensus DAG in any pre-registration

The first draft of a DAG among a small group of researchers might not consider a wider array of variables that could be important. Critical connections between variables might also be missed. We suggest collaborating with a larger group of experts and potentially including at least one person with experience in building and using DAGs. It could take some iterations until the consensus DAG is reached. The consensus graph should be included in relevant pre-registration to ensure that the assumptions encoded in the graph are defined before analysis and are not dependent on data availability or results of statistical analyses.

Since our DAG started with the exposure, outcome, a selection node, and all known causes of at least one of them, the consensus DAG contains the confounding and selection structure according to the researchers' assumptions. This approach allows use of the consensus DAG to identify variables for both confounding and selection bias adjustment in accordance with the causal assumptions encoded in the DAG.

Example

The study team should review the initial DAG and assure that everyone involved is satisfied that all relevant variables are present (possibly including unmeasured variables) and all assumed causal effects are depicted. In this example, iteration identified a few additions to the graph (fig 6B). Firstly, the team chose not to ignore smoking, the impact of a person's peers' alcohol use, and the association between socioeconomic position and alcohol use. This decision, therefore, meant adding smoking and peers' alcohol use and arrows between socioeconomic position and alcohol to the graph (fig 6D).

Further, it was believed that omitting the direct arrow from socioeconomic position to mortality from the DAG was too strong an assumption because it is highly plausible that this effect is mediated by variables not included in the DAG. The team continued to consider variables likely to influence selection; however, they decided not to add any additional arrows to the selection node. Lastly, we explicitly acknowledged potential unidentified and unmeasured confounders (U) of our exposure-outcome association. For instance, familial support might be a variable that would reasonably cause the exposure, outcome, and selection given our study interest in those aged 18 years or older. These confounders were subsequently added to the updated consensus graph (fig 6D).

Based on the consensus DAG, the team determined a minimally sufficient adjustment set (ie, the set of variables necessary to remove bias from confounding, assuming that the graph is correct and the node U on fig 6D does not truly exist). This example had two sets: (1) socioeconomic position, care access, and heavy alcohol use; and (2) care access, heavy alcohol use, and family history of alcohol abuse. These sets can be used to identify a dataset containing some reasonable measurement of the variables that we identified as necessary for our analysis. At a minimum, this dataset would require variables for diagnosis of alcohol use disorder, mortality, socioeconomic position, access to care, alcohol abuse, and family history of alcohol abuse. If a dataset is already in hand, investigators can choose a preferable adjustment set—for instance, if there is a set with minimal missing data. Figure 6D has an added arrow from U, the potentially unmeasured variable (eg, familial support). If investigators believed that the presence of U was very likely, and that U is a variable that should be accounted for, then this would suggest that there is no minimally sufficient adjustment set based on measured variables. Either information on familial support would need to be obtained, or additional sensitivity analyses would be needed to evaluate the impact of the unmeasured variable on the outcome (see step 6).

Step 4: Identify variables for data collection or the appropriate dataset based on the consensus directed acyclic graph, and if already collected data are used, consider which variables affect selection

Ideally, DAGs should inform the dataset and variables used, not vice versa. Thus, data should be collected

only once the graph is constructed in a prospective study. However, this approach could have practical limitations. In the case of a retrospective study, a dataset might already be identified or in hand for analysis.

Once consensus is reached, if data are not collected prospectively, selection into the study should be revisited. Where a dataset is already in hand before the graph is created, it will be particularly important to draw assumptions about selection into the study to understand how the formation of the dataset has created potentially biased relations. For instance, if the dataset identified for use is only among those with low socioeconomic position, that should be accounted for by adding an arrow from socioeconomic position (if on the DAG) into the selection node.

Example

Using the consensus DAG from step 3, we identified two minimally sufficient adjustment sets (socioeconomic position, care access, and heavy alcohol use; or care access, heavy alcohol use, and family history of alcohol abuse). We also posited that an unmeasured variable, familial support, might exist. These variables should then be collected prospectively to adjust for in the analysis. If a dataset already exists, ideally it is chosen because it has the requisite variables already collected. In this case, investigators should also evaluate any additional variables, how they affect selection, and how they are causally related to the other variables in the DAG—particularly those in the minimally sufficient adjustment set. For instance, if it was decided that using US Medicaid data was useful it would be important to acknowledge that socioeconomic position would have an arrow into selection.

Step 5: Choose analysis methods, measures of the outcome and exposure, and covariates based on those defined by the research question (step 1) and the consensus directed acyclic graph (step 3)

The analytical model should be based on the consensus DAG, and any deviations should be explained. For example, if the minimum adjustment set is not fully measured, state which key variables were omitted and methods to overcome these omissions.

Step 6: Based on variables that are unavailable or prone to measurement error, identify sensitivity analyses that can quantify the likely influence of unmeasured variables or measurement error

When one or more covariates have been measured with error, a statistical adjustment might only partly eliminate bias, thus leading to residual bias in the estimates. Although measurement error can be incorporated in the DAG,[43 44] this is rarely done in practice because it would greatly complicate the graph, and the error structure is rarely known. Researchers typically draw the graph as if variables had been perfectly measured while acknowledging that this assumption is, at best, an approximation. A simple sensitivity analysis to explore this potentially erroneous assumption might use a wider covariate set, for example, for confounding adjustment, including the common cause and mediators of its effect on the exposure or the outcome (of course, while avoiding adjusting for colliders or mediators of the effect of the exposure on the outcome). Another simple sensitivity analysis would include measured proxies for unmeasured covariates in the adjustment set.[6]

Measurement errors in the exposure and outcome can also influence the results. Independent (ie, error sources in exposure and outcome are independent) and non-differential (ie, error sources in treatment are independent of the true outcome, and vice versa) error will typically, but not always, attenuate the effect estimate. This bias can be analytically corrected using modelling assumptions about how the sources of error affect the exposure and outcome.[45] Such approaches can also be applied to measurement errors in covariates. Other forms of measurement error (dependent or differential) are more difficult to adjust for and require more sophisticated methods and often strong assumptions.

Example

The penultimate step is to consider measurement error and sensitivity analyses. Since the DAG has helped to find a minimal set of needed variables, investigators can focus on analyses to explore the potential impact of measurement error.[46] For example, if the measure of heavy alcohol use comes from self-reported responses to questions asked during a clinical encounter, we might think that there is a level of mismeasurement in our exposure variable, and we can potentially correct this mismeasurement with validation data. Furthermore, for variables with that are potentially mismeasured or completely unmeasured, but are believed to be important, sensitivity analysis methods such as quantitative bias analysis[47-49] can be used to evaluate the alcohol-abuse-mortality association in the presence of multiple different values. In effect, this approach will allow us to query the impact of an effect under better measurement of variables or while also including unmeasured variables. Examples of this include, for instance, evaluating the exposure-outcome association conditioning on peers' heavy use of alcohol or family history, which might be difficult to obtain information prospectively or completely absent and impossible to retrieve in retrospective data analysis. Evaluating the impact of mismeasured of unmeasured variables using a bias analysis approach would allow us to gain information about the effect of interest, assuming the structure of the DAG, and illustrate how bias from these variables might affect our results.

Step 7: Include the directed acyclic graphs in future publications and refer to it when describing adjustment sets and sensitivity analyses

The DAG used to guide the study's development should be included in future publications in either the main document or supplementary material. The graph can

then explain analytical decisions to justify a particular dataset's use and analysis measures, motivate methods to account for measurement error and unmeasured variables and necessary sensitivity or quantitative bias analyses. In addition to the graph, code that allows recreating it can be included to simplify the recreation of the graph for readers. Examples from the DAGs in this paper are included in web appendices 2 and 3. These actions will both explain underlying assumptions to readers and serve as a starting point for an improved DAG that can be used for incorporating new knowledge.

Example

In our example, we would at the very least include the graph in supplemental files and refer to it when describing covariate selection. Using the graph will inform future researchers, help justify decisions, and clarify the underlying rationale. Moreover, it will help future researchers build their own DAGs that incorporate different assumptions or new knowledge around the impact of heavy alcohol use on mortality.

Drawbacks and limitations of directed acyclic graphs

As with any research method, the use of DAGs has limitations, the most important of which is that they merely reflect our assumptions. One main strong assumption is the variables we choose not to include on the graph. For example, omitting relevant confounding variables can result in residual confounding (eg, because a true confounding path has been missed from a graph) or misrepresent the causal mechanisms between confounding variables (eg, an arrow has been missed from a graph), leading to biased estimates. DAGs simplify real world mechanisms, so there will rarely be a perfect or completely correct graph, but it is important to consider the implications of excluded nodes and arrows. Furthermore, while DAGs are useful for identifying adjustment sets that might be sufficient to eliminate bias (more precisely, to achieve exchangeability), they cannot reduce or verify other key assumptions required to make causal inferences (eg, causal consistency and positivity).

Moreover, their application for bias adjustment is limited because they merely allow researchers to be explicit about the assumptions they are willing to make to guide data analysis and interpretation. In terms of assessing whether such assumptions are true, the most DAGs can do is imply expected patterns of statistical independence. These patterns can potentially be empirically verified and, if violated, would serve as evidence that the constructed graph is incorrect—a type of falsification. However, the lack of evidence against a DAG is, at best, partial because the statistical patterns implied by the graph are often compatible with several different causal structures. Showing that a specific DAG is inconsistent with the data might provide little guidance on improving it; furthermore, in finite samples, such conclusions depend on arbitrary significance cut-off thresholds and power to detect associations, as well as on correctly specifying the models describing the association patterns among the variables.

DAGs are non-parametric, meaning that they do not need to make assumptions about the data-generating function that links the variables. This strength is important because it means that conclusions from the graph depend only on its structure. However, using DAGs to describe the implications of different functional forms (eg, the difference between a threshold effect and a linear effect) is challenging and limits their use for conclusions beyond the dichotomy of "any bias versus no bias" to assess the likely magnitude of the bias. Finally, it can be challenging to correctly specify the temporal ordering of nodes in a DAG.

DAGs can help precisely define assumptions, but there are some instances where they can be inappropriately used or are unhelpful. Firstly, these graphs are not conceptual diagrams or mind maps meant to illustrate processes (eg, a diagram of intracellular processes or chemical reactions). Secondly, DAGs restricted in advance to measured variables are inadequate because such graphs would never, by construction, suggest that relevant variables were not measured. Thirdly, overly simplified graphs can misguide analysis; for example, if each covariate is drawn only with arrows into the exposure and the outcome, without including any arrows between the covariates, investigators would fail to identify colliders that should not be adjusted for. Furthermore, boxes around nodes on a DAG should generally indicate conditioning via statistical adjustment or stratification (box 1). Drawing boxes or shapes around nodes for other reasons can make the graphs ambiguous and difficult to read and interpret; thus, boxes or other shapes around variables are often avoided unless specifically drawn to indicate conditioning. Lastly, double headed arrows or arrows from a node to another edge, or vice versa, are often discouraged because double headed arrows can violate the acyclic property of the DAG and make relations ambiguous. Although in some contexts, double headed arrows denote unmeasured confounding between two nodes, explicitly including the unmeasured confounder in the graph is often clearer. Arrows leading into or out of an arrow, although potentially intuitively appealing in some cases, have no formal meaning within the theory underlying these graphs. Such arrows might also convey ambiguity about relations between variables and hamper the ability to determine causal and biasing paths from the graph.[20]

Furthermore, DAGs can become complicated and dense in real world applied studies if they have many nodes and arrows. Although software can help make such examples tractable, in some cases, the number of variables is very large, thus making it nearly impossible to draw a plausible graph. In such cases, alternative tools that do not require explicitly drawing a full graph might be useful. For example, for covariate selection for confounding adjustment, the disjunctive cause criterion might be used, which only requires knowing if the covariate is a cause of the exposure or outcome.[6]

However, this criterion assumes that the pool of variables from which the covariates will be selected is a subset sufficient for confounding adjustment. This assumption is difficult to justify without the full DAG. Therefore, this method can be complemented with tools that assess whether a candidate covariate set is adequate.[50] Alternatively, interactive procedures allow researchers to build only as little as needed of a DAG for selecting covariates (thus reducing the number of assumptions needed and automated covariate selection procedures).[51] These examples illustrate that, even when drawing a full graph is impractical, principled alternatives for covariate selection are available, which should be preferred over simply selecting pre-exposure variables that are available in the dataset.[52]

Reporting directed acyclic graphs

Reporting DAGs in empirical papers can be extremely helpful for readers, reviewers, and editors. They can also be reported in pre-registration that investigators undertake, which can assure readers that the causal structure was considered well before the data and analysis were agreed on. This reporting can help avoid bias and spurious statistical signals from cherry picking analyses or adjustment sets (eg, Andrew Gelman's garden forking paths[53]). When empirical papers are reported, DAGs can be presented as part of the analysis plan. They can either be included as a primary figure or, more typically, contained in the supplement to state the study's assumptions. Authors can include a list of people directly involved in building the graph and any literature supporting the relations presented along with the graph and reproducible code to create it (if software was used).

DAGs can help reviewers and editors understand a study, because they make explicit the assumptions underlying design, data collection, and analysis. Many empirical papers implicitly assume a causal structure. Including a DAG makes this underlying causal structure explicit and allows readers to assess the plausibility of assumptions. Thus, it can be helpful for reviewers and editors to consider requesting these graphs as part of peer review. Box 1 outlines questions that readers could consider when interpreting a DAG.

We thank Charlie Poole, F Thurston Drake, Alexandra Pitman, and Bernardo Horta for their useful and insightful comments on a draft of this manuscript; and David Richardson, Alexander Commanday, and Robert Platt for helpful comments on the revision.

Contributors: TF and NMD conceived the paper. TF wrote and revised the first draft, and NMD and FH revised it. All authors are guarantors. The corresponding author attests that all listed authors meet authorship criteria and that no others meeting the criteria have been omitted.

Funding: No specific funding was provided for this work.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/disclosure-of-interest/ and declare: no support from any organisation for the submitted work; for other work, funding was provided by the UK Medical Research Council (MRC) and the University of Bristol support the MRC Integrative Epidemiology Unit (MC_UU_00011/1) at the University of Bristol; NMD is supported via a Norwegian Research Council grant (No 295989) and the UCL Division of Psychiatry (https://www.ucl.ac.uk/psychiatry/division-psychiatry).

Code availability: Code for reproducing the directed acyclic graphs in Latex and Daggity is provided here https://github.com/tpfeeney/DAGs_for_clinicians.

Provenance and peer review: Not commissioned; externally peer reviewed.

1 Pearl J. Causal diagrams for empirical research. *Biometrika* 1995;82:669-88. doi:10.1093/biomet/82.4.669.
2 Wright S. Correlation and causation. *J Agricultural Res* 1921;557.
3 Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology* 1999;10:37-48. doi:10.1097/00001648-199901000-00008
4 Rothman KJ, Lash TL, VanderWeele TJ, et al. *Modern Epidemiology: Chapter 3* . Modern epidemiology. Wolters Kluwer, 2021.
5 Hernan MA, Robins JM. *What if: Chapter 6* . Causal Inference: What If. Chapman & Hall/CRC, 2020.
6 VanderWeele TJ. Principles of confounder selection. *Eur J Epidemiol* 2019;34:211-9. doi:10.1007/s10654-019-00494-6.
7 Tennant PWG, Murray EJ, Arnold KF, et al. Use of directed acyclic graphs (DAGs) to identify confounders in applied health research: review and recommendations. *Int J Epidemiol* 2021;50:620-32. doi:10.1093/ije/dyaa213
8 Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;15:615-25. doi:10.1097/01.ede.0000135174.63482.43
9 Lu H, Cole SR, Howe CJ, Westreich D. Toward a Clearer Definition of Selection Bias When Estimating Causal Effects. *Epidemiology* 2022;33:699-706. doi:10.1097/EDE.0000000000001516
10 Hernán MA, Cole SR. Invited Commentary: Causal diagrams and measurement bias. *Am J Epidemiol* 2009;170:959-62, discussion 963-4. doi:10.1093/aje/kwp293
11 Wardle MT, Reavis KM, Snowden JM. Measurement error and information bias in causal diagrams: mapping epidemiological concepts and graphical structures. *Int J Epidemiol* 2024;53:dyae141. doi:10.1093/ije/dyae141.
12 Moreno-Betancur M, Lee KJ, Leacy FP, White IR, Simpson JA, Carlin JB. Canonical Causal Diagrams to Guide the Treatment of Missing Data in Epidemiologic Studies. *Am J Epidemiol* 2018;187:2705-15. doi:10.1093/aje/kwy173
13 Etminan M, Collins GS, Mansournia MA. Using Causal Diagrams to Improve the Design and Interpretation of Medical Research. *Chest* 2020;158(1S):S21-8. doi:10.1016/j.chest.2020.03.011
14 Lipsky AM, Greenland S. Causal Directed Acyclic Graphs. *JAMA* 2022;327:1083-4. doi:10.1001/jama.2022.1816
15 Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Med Res Methodol* 2008;8:70. doi:10.1186/1471-2288-8-70
16 Byeon S, Lee W. Directed acyclic graphs for clinical research: a tutorial. *J Minim Invasive Surg* 2023;26:97-107. doi:10.7602/jmis.2023.26.3.97
17 Digitale JC, Martin JN, Glymour MM. Tutorial on directed acyclic graphs. *J Clin Epidemiol* 2022;142:264-7. doi:10.1016/j.jclinepi.2021.08.001
18 Piccininni M, Konigorski S, Rohmann JL, Kurth T. Directed acyclic graphs and causal thinking in clinical risk prediction modeling. *BMC Med Res Methodol* 2020;20:179. doi:10.1186/s12874-020-01058-z
19 Williams TC, Bach CC, Matthiesen NB, Henriksen TB, Gagliardi L. Directed acyclic graphs: a tool for causal studies in paediatrics. *Pediatr Res* 2018;84:487-93. doi:10.1038/s41390-018-0071-3
20 Hartwig F, Feeney T, Davies NM. D-separation for applied researchers: understanding how to interpret directed acyclic graphs. *arXiv* 2025;2502.13736.
21 Hernán MA, Monge S. Selection bias due to conditioning on a collider. *BMJ* 2023;381:p1135. doi:10.1136/bmj.p1135
22 Cashin AG, McAuley JH, VanderWeele TJ, Lee H. Understanding how health interventions or exposures produce their effects using mediation analysis. *BMJ* 2023;382:e071757. doi:10.1136/bmj-2022-071757
23 Webster-Clark M, Breskin A. Directed Acyclic Graphs, Effect Measure Modification, and Generalizability. *Am J Epidemiol* 2021;190:322-7. doi:10.1093/aje/kwaa185
24 Nilsson A, Bonander C, Strömberg U, Björk J. A directed acyclic graph for interactions. *Int J Epidemiol* 2021;50:613-9. doi:10.1093/ije/dyaa211
25 VanderWeele TJ, Robins JM. Minimal sufficient causation and directed acyclic graphs. *Ann Stat* 2009;37. doi:10.1214/08-AOS613.
26 Walker V, Sanderson E, Levin MG, Damraurer SM, Feeney T, Davies NM. Reading and conducting instrumental variable studies: guide, glossary, and checklist. *BMJ* 2024;387:e078093. doi:10.1136/bmj-2023-078093.
27 Westreich D, Greenland S. The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. *Am J Epidemiol* 2013;177:292-8. doi:10.1093/aje/kws412
28 Cole SR, Hernán MA. Fallibility in estimating direct effects. *Int J Epidemiol* 2002;31:163-5. doi:10.1093/ije/31.1.163
29 Griffith GJ, Morris TT, Tudball MJ, et al. Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nat Commun* 2020;11:5749. doi:10.1038/s41467-020-19478-2

30 Monge S, Pastor-Barriuso R, Hernán MA. The imprinting effect of covid-19 vaccines: an expected selection bias in observational studies. *BMJ* 2023;381:e074404. doi:10.1136/bmj-2022-074404
31 Hernán MA, Robins JM. *Causal Inference: What If* . 1st ed. Taylor and Francis, 2024.
32 Pearl J. *Causality: models, reasoning, and inference.* Cambridge University Press, 2022.
33 Textor J, van der Zander B, Gilthorpe MS, Liskiewicz M, Ellison GT. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. *Int J Epidemiol* 2016;45:1887-94. doi:10.1093/ije/dyw341. doi:10.1093/ije/dyw341
34 Feeney T, Davies NM, Hartwig F. Software aids for building directed acyclic graphs. 2025. https://osf.io/638nk/. doi:10.17605/OSF.IO/638NK.
35 European Medicines Agency. ICH E9 (R1) addendum on estimands and sensitivity analysis in clinical trials to the guideline on statistical principles for clinical trials. 2020. https://www.ema.europa.eu/en/documents/scientific-guideline/ich-e9-r1-addendum-estimands-sensitivity-analysis-clinical-trials-guideline-statistical-principles_en.pdf
36 Kahan BC, Hindley J, Edwards M, Cro S, Morris TP. The estimands framework: a primer on the ICH E9(R1) addendum. *BMJ* 2024;384:e076316. doi:10.1136/bmj-2023-076316
37 Lundberg I, Johnson R, Stewart BM. What Is Your Estimand? Defining the Target Quantity Connects Statistical Evidence to Theory. *Am Sociol Rev* 2021;86:532-65. doi:10.1177/00031224211004187
38 Little RJ, Lewis RJ. Estimands, Estimators, and Estimates. *JAMA* 2021;326:967-8. doi:10.1001/jama.2021.2886
39 VanderWeele TJ, Shpitser I. A new criterion for confounder selection. *Biometrics* 2011;67:1406-13. doi:10.1111/j.1541-0420.2011.01619.x.
40 VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann Intern Med* 2017;167:268-74. doi:10.7326/M16-2607
41 Jackson JW, Swanson SA. Toward a clearer portrayal of confounding bias in instrumental variable applications. *Epidemiology* 2015;26:498-504. doi:10.1097/EDE.0000000000000287
42 Davies NM, Thomas KH, Taylor AE, et al. How to compare instrumental variable and conventional regression analyses using negative controls and bias plots. *Int J Epidemiol* 2017;46:2067-77. doi:10.1093/ije/dyx014
43 Hernan MA, Robins JM. *What if: Chapter 9* . Causal Inference: What If. Chapman & Hall/CRC, 2020:119-27.
44 Wardle MT, Reavis KM, Snowden JM. Measurement error and information bias in causal diagrams: mapping epidemiological concepts and graphical structures. *Int J Epidemiol* 2024;53:dyae141. doi:10.1093/ije/dyae141.
45 Keogh RH, White IR. A toolkit for measurement error correction, with a focus on nutritional epidemiology. *Stat Med* 2014;33:2137-55. doi:10.1002/sim.6095
46 Batistatou E, McNamee R. Performance of bias-correction methods for exposure measurement error using repeated measurements with and without missing data. *Stat Med* 2012;31:3467-80. doi:10.1002/sim.5422
47 Brown JP, Hunnicutt JN, Ali MS, et al. Quantifying possible bias in clinical and epidemiological studies with quantitative bias analysis: common approaches and limitations. *BMJ* 2024;385:e076365. doi:10.1136/bmj-2023-076365
48 Fox MP, MacLehose RF, Lash TL. *Applying Quantitative Bias Analysis to Epidemiologic Data.* Springer International Publishing, 2021, doi:10.1007/978-3-030-82673-4.
49 Lash TL, Fox MP, MacLehose RF, Maldonado G, McCandless LC, Greenland S. Good practices for quantitative bias analysis. *Int J Epidemiol* 2014;43:1969-85. doi:10.1093/ije/dyu149
50 Hartwig FP, Tilling K, Smith GD. Empirically assessing the plausibility of unconfoundedness in observational studies. 2024.
51 Guo FR, Zhao Q. Confounder selection via iterative graph expansion.2023. http://arxiv.org/abs/2309.06053
52 Entner D, Hoyer P, Spirtes P. Data-driven covariate selection for nonparametric estimation of causal effects. Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics. *Proc Mach Learn Res* 2013;31:256-64. https://proceedings.mlr.press/v31/entner13a.html
53 Gelman A, Loken E. The garden of forking paths: Why multiple comparisons can be a problem, even when there is no "fishing expedition" or "p-hacking" and the research hypothesis was posited ahead of time. Department of Statistics, Columbia University. 2013;348:3.

Web appendix 1: Supplemental figure 1
Web appendix 2: Supplemental codes for DAGs
Web appendix 3: Code for reproducing the directed acyclic graphs in Latex