Journal of Epidemiology

Causal Diagrams: Pitfalls and Tips

Etsuji Suzuki[1], Tomohiro Shinozaki[2], and Eiji Yamamoto[3]

[1]Department of Epidemiology, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, Okayama, Japan
[2]Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science, Tokyo, Japan
[3]Okayama University of Science, Okayama, Japan

ABSTRACT

Graphical models are useful tools in causal inference, and causal directed acyclic graphs (DAGs) are used extensively to determine the variables for which it is sufficient to control for confounding to estimate causal effects. We discuss the following ten pitfalls and tips that are easily overlooked when using DAGs: 1) Each node on DAGs corresponds to a random variable and not its realized values; 2) The presence or absence of arrows in DAGs corresponds to the presence or absence of individual causal effect in the population; 3) "Non-manipulable" variables and their arrows should be drawn with care; 4) It is preferable to draw DAGs for the total population, rather than for the exposed or unexposed groups; 5) DAGs are primarily useful to examine the presence of confounding in distribution in the notion of confounding in expectation; 6) Although DAGs provide qualitative differences of causal structures, they cannot describe details of how to adjust for confounding; 7) DAGs can be used to illustrate the consequences of matching and the appropriate handling of matched variables in cohort and case-control studies; 8) When explicitly accounting for temporal order in DAGs, it is necessary to use separate nodes for each timing; 9) In certain cases, DAGs with signed edges can be used in drawing conclusions about the direction of bias; and 10) DAGs can be (and should be) used to describe not only confounding bias but also other forms of bias. We also discuss recent developments of graphical models and their future directions.

Key words: bias; causal inference; causality; confounding; directed acyclic graphs

1. BACKGROUND ON THE TOPIC

Causal diagrams have been often used among epidemiologists as a tool to describe what is already known about relevant causal structures. In 1999, Greenland et al introduced formal theories of causal directed acyclic graphs (DAGs) within epidemiology,[1] and several comprehensive introductions to DAGs have been published,[2–7] including a review article in Japanese.[8] Consequently, the use of DAGs is now widespread among epidemiologists; when consulting *A Dictionary of Epidemiology*, although there was no entry for DAGs in its fourth edition that was published in 2001,[9] a definition of DAGs has been included in its later editions.[10,11]

Confounding is one of the primary concerns in epidemiologic research, and DAGs are used extensively to determine the variables for which it is sufficient to control for confounding to estimate causal effects. A confounder was traditionally identified based on the following three criteria[3–5]: a) it must be associated with the exposure; b) it must be associated with the outcome in the unexposed; and c) it must not lie on a causal pathway between exposure and outcome. Because these traditional criteria sometimes fail, however, the graphical criteria for identifying confounders in DAGs are especially useful. This point is often explained using an example of the so-called M bias.[12] In this case, even if one takes care to not adjust for variables affected by exposure or outcome in the traditional confounder-selection criteria, one may be led to adjust for a "collider" on the backdoor path from exposure to outcome and unnecessarily introduce bias.

Despite its widespread use, however, the lack of clear understanding of this tool could lead to inappropriate use and underappreciation of DAGs. The goal of this paper is not to provide a basic introduction of DAGs; rather, we aim to review some pitfalls that are easily overlooked when using DAGs and discuss tips, hoping that DAGs will be further used in epidemiology as well as in a wide range of other research disciplines. We also discuss recent developments of graphical models and their future directions.

2. PITFALLS AND TIPS

Simply speaking, a DAG is composed of nodes (or vertices) and arrows (or arcs/edges) connecting them such that the graph is acyclic. The first three pitfalls below address these basic components, followed by other pitfalls when drawing and interpreting DAGs.

2-1. Each node on DAGs corresponds to a random variable and not its realized values

When drawing DAGs, we are primarily concerned about causal relations among variables. DAGs are a simple way to encode our

---

Address for correspondence. Etsuji Suzuki, Department of Epidemiology, Graduate School of Medicine, Dentistry and Pharmaceutical Sciences, Okayama University, 2-5-1 Shikata-cho, Kita-ku, Okayama 700-8558, Japan (e-mail: etsuji-s@cc.okayama-u.ac.jp).
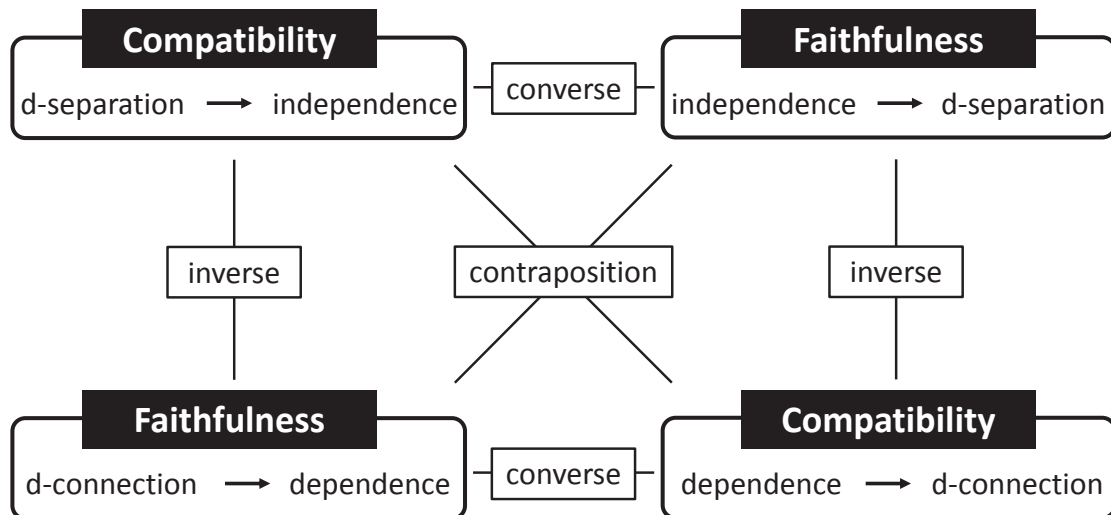
Figure 1. The relation between compatibility and faithfulness. See Section 2-2 and Box 1 for details.

subject-matter knowledge, or our assumptions, about the qualitative causal structure of interest. Consequently, each node on DAGs corresponds to a random variable and not its realized values. Incorrect use of nodes can lead to erroneous construction of the DAGs. For example, Steiner et al proposed that propensity score itself is described as a collider with respect to exposure and confounders, claiming that propensity score analysis removes the confounding bias by offsetting the relation from confounders to exposure via association due to conditioning on the collider, regardless of analytic method (eg, matching, stratification, or weighting).[13] Their approach, however, does not describe causal structures between random variables; rather, they are concerned about the realized values of propensity scores based on an empirical joint distribution of exposure A and confounders C in a sample. Also, the fact that (true) propensity score is the functional of the joint probability distribution $P(A, C)$ does not justify drawing arrows from A and C to the node representing propensity score. Arrows in DAGs should represent causal structure between each random variable but not the dependence forms defined by probability distributions; a related pitfall is presented in Section 2-2.

Note that statistical independencies implied by a DAG are only "in expectation", and they apply to the expected data distribution if the causal structure represented by the DAG is correct. Thus, the associations that may arise as a result of purely random events, such as those produced by randomization or random sampling, are not described in DAGs. A related issue is discussed in Section 2-5.

2-2. The presence or absence of arrows in DAGs corresponds to the presence or absence of individual causal effect in the population

We say that P affects Q in a population if and only if there is at least one individual for whom changing (intervening on) P would change Q.[4] In line with this, an arrow from P to Q is drawn when we suspect there is a direct causal effect (ie, an effect not mediated through any other variables in the DAG) for at least one individual in the population, or when we are unwilling to assume such individual causal effects do not exist.[4,5] Thus, the presence of an arrow from P to Q does not necessarily imply that these two variables are statistically dependent; if the quantity of positive

individual causal effect and that of negative individual causal effect are equivalent, they perfectly cancel out and the net effect of P on Q becomes null (this situation would be relatively rare in large samples). If this is the case, P and Q are statistically independent even if there is an arrow from P to Q. By contrast, if there are no (direct or indirect) paths from P to Q, it means that we assume that a sharp null hypothesis of no causal effect of P on any individual's Q holds in the population.[4,5] Thus, omitting arrows represents a strong assumption, though some may put more emphasis on drawing arrows. Recall that the mathematical theory underlying the graphical rules is known not as "d-connection" but as "d-separation".[2] To summarize, the presence or absence of arrows in DAGs corresponds to the presence or absence of individual causal effect in the population.

This point is sometimes explained using the rules of *compatibility* and *weak faithfulness* (see Figure 1 and Box 1 for their definitions).[3] Under these two rules, we can specify the relation between d-separation and statistical independence, and it is consistent with drawing arrows based on the presence or absence of individual causal effect. When we further assume *faithfulness* (see Figure 1 and Box 1 for its definition), we can also find the statistical dependencies implied by DAGs.[3] (Matching leads to unfaithfulness: see Section 2-7 and elsewhere

Box 1. Terminology in Section 2-2

- **d-separation/d-connection:** Two variables (or two sets of variables) are said to be d-separated if every path between them is blocked; otherwise they are d-connected.

- **Rule of *compatibility*:** Whenever two sets of variables are d-separated given a third set, the two sets are independent conditional on the third.

- **Assumption of *faithfulness*:** Whenever two sets of variables are d-connected given a third set, the two sets are associated conditional on the third. This is the converse property of compatibility.

- **Rule of *weak faithfulness*:** When two sets of variables are d-connected given a third set, the two sets are expected to be associated conditional on the third.

for details.[14,15]) We may often see apparent violations of faithfulness in small data sets when that would not be the case in a larger sample from the same population. However, because faithfulness is controversial as a routine assumption from practical as well as theoretical perspectives, Glymour and Greenland discuss only uses of graphical models that do not rely on its assumption in *Modern Epidemiology*,[3] whereas Hernán and Robins assume faithfulness throughout their book *Causal Inference: What If* unless mentioned otherwise.[5] The importance of faithfulness becomes clearer for "causal discovery". The details about these rules are beyond the scope of this paper, but it is notable that, under the rule of compatibility and the assumption of faithfulness (which are collectively referred to as *perfect compatibility*),[3] the presence or absence of arrows corresponds to the presence or absence of average causal effect in the target population, instead of individual causal effect. These may well illustrate the significance of distinguishing individual causal effect and average causal effect.

In the Appendix, a further discussion is provided following an example in Fine Point 6.2 of *Causal Inference: What If*.[5]

2-3. "Non-manipulable" variables and their arrows should be drawn with care

In the counterfactual framework, we think of the effect of a particular factor as a contrast between the potential outcomes when an exposure is set to a particular value, while holding all other components constant (unless they are assumed to mediate the causal path).[16,17] Therefore, if we can clearly address hypothetical interventions on manipulable variables, it is relatively straightforward to examine their causal effects. Because a counterfactual definition of causation requires hypothetical intervention(s), the question of whether non-manipulable variables can be considered "causal" has long been controversial.[4] For example, Holland famously argued that "causes are only those things that could, in principle, be treatments in experiments", putting forward the slogan "no causation without manipulation".[18] Although counterfactuals related to manipulable quantities are of great value, causation related to non-manipulable quantities can be of scientific interest and constitute a substantial portion of instances of causation in biomedical and social sciences.[19] For example, though not easily amenable to experimental interventions, we may be interested in examining "causal" effects of sex, race, or genetic ancestry. There have been similar arguments about age-period-cohort analyses; although this analytic method is used to measure the effect of time, hypothetical interventions to change "time" are beyond the realm of possibility, though may be conceivable.[20]

Accordingly, when drawing DAGs, each researcher should carefully think about whether it even makes sense to include non-manipulable variables at all with any arrows into or out of them. In this regard, it would be helpful to clearly define "hypothetical interventions" in the counterfactual framework.[21] If a "hypothetical intervention" on a non-manipulable variable is well-defined, the variable may be treated as a "hypothetically manipulable" variable. For particular research hypotheses of interest, we may need to represent "hypothetically manipulable" variables (either as exposures or as covariates) and show arrows emanating from such variables. In most cases, however, "hypothetically manipulable" variables will not have any causes (other than the hypothetical intervention) and therefore will not have arrows pointing into them.[4] Consequently, for example, the total causal effect of sex can be estimated simply by taking

the observed differences in the outcome of interest between male and female.[19] In other words, once we adopt the presence of "causal" effect of sex on a particular health outcome, there is no confounding because sex itself is randomized at conception. As one of a handful of exceptions, Glymour explained that, if sex ratios of newborns are thought to vary slightly in response to environmental conditions,[22] it may be useful to include such stressors as causes of sex.[4] To summarize, "non-manipulable" variables and their arrows should be drawn with care from a perspective of hypothetical interventions.

On a related issue, even if the exposure of interest is manipulable and hypothetical interventions on it are conceivable, there is a serious concern arising from ill-defined causal questions. If there are multiple versions of treatment, which is referred to as compound treatments,[5,23] estimating causal effects is challenging in randomized controlled trials (RCTs) as well as in observational studies. See Hernán and VanderWeele for DAGs with compound treatments.[23]

2-4. It is preferable to draw DAGs for the total population, rather than for the exposed or unexposed groups

The target population concept plays a key role in discussions of causal inference in epidemiology, and it has been well established that confounding depends on the population chosen as the target of inference.[24–26] The target can be the total population or the exposed, the unexposed, or conditional on the covariates. When drawing DAGs, however, it would be preferable to use the total population as the target population, which enables us to more readily identify the presence of bias.[26,27] This is because, as is further explained in Section 2-5, DAGs are primarily useful to examine the presence of confounding in distribution in the notion of confounding in expectation.[26] The exposed and unexposed groups are, by definition, determined by the specific pattern of exposure status. Thus, if these groups are used as target populations, the discussion is logically restricted to the notion of realized confounding for the two target populations in that particular pattern. Generally, drawing DAGs for the exposed or unexposed group should be avoided.[27,28]

If we intend to use the exposed and unexposed groups as target populations when drawing DAGs, we may need to draw a square around the node of exposure to explicitly indicate that we are conditioning on exposure status. On a related issue, other than unspecified component causes, when every sufficient cause either includes both the exposure and a certain covariate (ie, causal co-action), or includes the exposure alone without the covariate, confounding does not occur if the target population is the exposed group. This is because, irrespective of distributions of the covariate and the unspecified component causes, no outcome occurs in the actual unexposed group, which is used as a substitute of what would have occurred in the actual exposed group had they been unexposed. In this situation, Flanders et al proposed a simple modification of DAGs to use dotted arrows for all effects that are absent when exposure is absent.[29] Their approach, however, is also based on DAGs for the total population.

2-5. DAGs are primarily useful to examine the presence of confounding in distribution in the notion of confounding in expectation

Despite its significance, the different notions of confounding have not been fully appreciated in the literature, leading to confusion of

causal concepts in epidemiology. This lack of clear understanding could lead to inappropriate use and underappreciation of DAGs, which provide a simple algorithm for examining the presence of confounding in distribution in the notion of confounding in expectation.[26] We briefly explain these notions of confounding below.

First, the notion of confounding can be defined both with respect to marginal distributions of potential outcomes (ie, *confounding in distribution*) and with respect to a specific effect measure (ie, *confounding in measure*).[30,31] By definition, confounding in distribution is scale-independent, whereas confounding in measure is scale-dependent. No confounding in distribution is a sufficient condition for no confounding in measure. DAGs are primarily useful to examine the presence of confounding in distribution because they are completely nonparametric and provide qualitative assumptions behind a causal analysis.[1,26] As explained in Section 2-4, it is preferable to draw DAGs for total population as the target population. When the target is the total population, confounding depends on the notions of confounding in distribution and confounding in measure,[26,30–32] which highlights the significance of distinguish these notions.

Second, a further distinction can be drawn between *confounding in expectation* and *realized confounding*.[26,30,33] In an ideal RCT, the randomized groups will be comparable in their potential outcomes on average over repeated experiments. For any given experiment, however, the particular randomization may result in imbalances by chance because of the particular allocation or exposure assignment.[34] Such a scenario would be one in which there is no confounding in expectation but there is realized confounding for the particular experiment.[26,30] (This phenomenon has been also referred to as random confounding,[35] and is particularly of concern when the size of the population is small.) To grasp the profound distinction between these notions of confounding, we need to understand the mechanism that generates exposure events not the product of that mechanism. The lack of clear understanding of this mechanism could lead to inappropriate use of DAGs to examine the presence of realized confounding.[27,28] DAGs are, however, primarily useful to examine the presence of confounding in expectation,[26] which is a form of systematic error.[36] Therefore, one needs to draw DAGs based on the understanding of underlying causal structures among random variables, as mentioned in Section 2-1. This point is related to the fact that bias (or more strictly speaking, exact bias[5]) is defined by comparing the expected value of an estimator and the true value of the parameter.[37,38] If the size of population is large enough, DAGs are also practical tools to identify realized confounding.[26]

To summarize, DAGs are primarily useful to examine the presence of confounding in distribution in the notion of confounding in expectation.[26] See elsewhere for a detailed discussion about the notions of confounding.[26]

2-6. Although DAGs provide qualitative differences of causal structures, they cannot describe details of how to adjust for confounding

Analytic adjustment for confounders in observational studies has consequences quite different from those of physical control in RCTs. When randomization is implemented in ideal RCTs, there is no confounding in expectation, and all arrows pointing into exposure(s) can be erased or removed.[26] In observational studies, however, arrows pointing into exposure(s) cannot be erased, even

though one aims to adjust for confounding by using appropriate study designs as well as analytic methods. There are inherent distinctions in the underlying causal structures between observational studies and RCTs, which can be readily understood by considering theoretical data frequencies in these studies based on the counterfactual model.[39]

Although DAGs can provide qualitative differences of causal structures between observational studies and RCTs, they cannot describe details of how to adjust for confounding because DAGs are not quantitative models. For example, we often use a square box around a node to indicate that we are conditioning on it (or stratifying on its values). If we can identify a sufficient set of covariates to block all the backdoor paths from exposure to outcome, the statistical relation between exposure and outcome *within strata of the covariates* is due only to the effect of exposure on outcome *within the strata*. Indeed, blocking the flow of association between exposure and outcome through the common cause(s) is the graph-based justification to use stratification as a method to achieve exchangeability.[5] However, we cannot use DAGs to explain the fact that standardization can be used to obtain marginal causal effect in the target population; rather we need to analytically explain that marginal causal effect becomes equivalent to a weighted average of conditional causal effects. Thus, DAGs cannot be used to distinguish analytic methods that aim to obtain conditional causal effects (eg, stratified analyses and regression models) and those that aim to obtain marginal causal effects (eg, standardization). Further, DAGs cannot be used to explain whether specific estimators or estimation works well. Irrespective of the methods employed to adjust for confounding, however, DAGs can be used to determine the variables for which it is sufficient to control for confounding to estimate causal effects.

2-7. DAGs can be used to illustrate the consequences of matching and the appropriate handling of matched variables in cohort and case-control studies

The goal of matching differs by study design. In cohort studies, matching is used to prevent confounding by the matched variables, and then adjustment for these variables may be unnecessary to remove bias in point estimation. In case-control studies, matching is used to increase statistical efficiency. Unlike cohort matching, case-control matching does not and cannot remove confounding, but instead may induce selection bias that can itself be removed by adjustment for the matched variables. Although these differences were clarified long ago, misconceptions about the implications of matching remain common. DAGs can be used to illustrate the consequences of matching and the appropriate handling of matched variables in cohort and case-control studies.[14,15] Let A denote an exposure, Y denote an outcome, and C denote a confounder and matching variable. The variable S indicates whether an individual in the source population is selected for the matched study (1: selected, 0: not selected).

In cohort studies, matching process ensures that the distributions of matched variables are (nearly) identical across exposure groups. In a DAG, cohort matching is described by drawing arrows from C and A to S because selection into the sub-cohort depends on the values of C and A (Figure 2). The square around $S = 1$ indicates that the analysis is conditional on having been selected. Although the variables C and A are d-connected via two paths, $C \rightarrow A$ and $C \rightarrow \boxed{S = 1} \leftarrow A$, these variables are independent
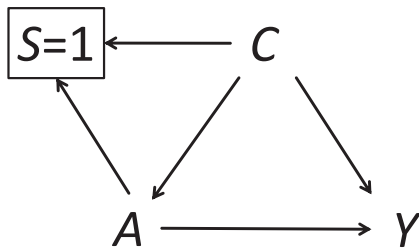
Figure 2. Cohort matching on a confounder. We let *A* denote an exposure, *Y* denote an outcome, and *C* denote a confounder and matching variable. The variable *S* indicates whether an individual in the source population is selected for the matched study (1: selected, 0: not selected). See Section 2-7 for details.

Figure 3. Case-control matching on a confounder. We let *A* denote an exposure, *Y* denote an outcome, and *C* denote a confounder and matching variable. The variable *S* indicates whether an individual in the source population is selected for the matched study (1: selected, 0: not selected). See Section 2-7 for details.
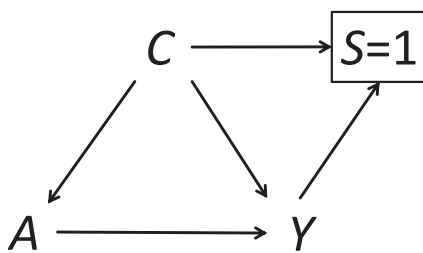
by design. In other words, the two paths "unfaithfully" cancel each other exactly. This graphical representation is useful to explain that, if there are unmatched confounders in matched cohort studies, ignoring the matching variables may leave bias even when additional adjustment is made for the unmatched confounders.[40] A similar issue occurs for propensity score balancing.[41]

By contrast, case-control matching ensures that the distributions of matched variables are (nearly) identical across outcome groups. In a DAG, case-control matching is described by drawing arrows from *C* and *Y* to *S* because, by definition, both *C* and *Y* affect *S* (Figure 3). (Here, we only discuss a situation in which there is an arrow from *A* to *Y*. See elsewhere for a situation without the arrow.[14,15]) The variables *C* and *Y* are d-connected via three paths, $C \rightarrow Y$, $C \rightarrow A \rightarrow Y$, and $C \rightarrow \boxed{S=1} \leftarrow Y$. The third path is induced by the matching and is of equal magnitude but opposite direction to the net association via the first two paths. Consequently, *C* and *Y* are independent unconditional on *A* in the matched sample (ie, unfaithfulness). This implies, however, that the net association over the two paths $C \rightarrow Y$ and $C \rightarrow \boxed{S=1} \leftarrow Y$ is not zero: *C* and *Y* are associated conditional on *A* in the matched sample. Therefore, the case-control matching does not break the biasing path $A \leftarrow C \rightarrow Y$, which graphically explains that case-control matching does not prevent the original confounding. Moreover, conditional on $S=1$, there is a biasing path for the effect of *A* on *Y*, $A \leftarrow C \rightarrow \boxed{S=1} \leftarrow Y$, which graphically explains that case-control matching induces selection bias. Adjustment for *C* is necessary to control both the intentional selection bias introduced by matching

and the original confounding. Furthermore, case-control matching on a non-confounder may lead to selection bias, which is illustrated in DAGs.[3,14,15] For details about DAGs for matching, see elsewhere.[14,15]

Finally, it is worth mentioning that, when using a square box around a node, there are two distinct ways. One is to simply add a square around a node, which is often used to indicate that all the strata of the variable are observed (eg, when conditioning is done analytically). The other is to use a square around a node by specifying the value on which conditioning was made (as in Figure 2 and Figure 3). This is also often employed to indicate that the strata except for the one indicated are missing (eg, when conditioning is done during selection or attrition, which could result in selection bias). Because the values on which conditioning was made could yield different associations, it is important to specify them when drawing DAGs.[42]

2-8. When explicitly accounting for temporal order in DAGs, it is necessary to use separate nodes for each timing

DAGs are, by definition, acyclic, and they should contain no feedback loops. Therefore, no variable is an ancestor or descendant of itself. If *A* causes *Y*, *Y* cannot also cause *A* at the same time. Though it may induce few problems in settings of a single-point exposure, we need to carefully draw DAGs for longitudinal analyses with time varying exposures by using separate nodes for each timing. If a prior value of *Y* affects *A*, and then *A* affects a subsequent value of *Y*, we must use separate nodes because these are separate variables (eg, $Y_0 \rightarrow A_1 \rightarrow Y_2$).

On a related issue, some researchers draw DAGs without a rigorous consideration of chronological order, putting more emphasis on their visibility. Others are more rigorously concerned about the chronological order, consistently drawing variables that occur temporarily earlier on the left side of DAGs.[5] Furthermore, some may draw DAGs in the order from top to bottom,[43] and others may draw in the order from right to left. These are merely differences of drawing styles of DAGs, and there is probably no single correct style. To avoid pitfalls in drawing DAGs, however, it may be helpful to prioritize clarity of communication. For example, although M bias is a classic example to illustrate benefits of using DAGs,[12] its form becomes quite different if chronological orders of variables are rigorously drawn. It is thus perhaps helpful to use different styles properly based on the aims of using DAGs.

2-9. In certain cases, DAGs with signed edges can be used in drawing conclusions about the direction of bias

Arrows in DAGs represent the presence of causal relations between the variables, and do not describe their directions or signs. Consequently, DAGs cannot describe direction of bias. This point is crucial because when interpreting epidemiologic findings, it is significant to consider not only the presence of but also the direction of bias—whether underestimation or overestimation occurs.

To address this issue within the context of DAGs, signs can sometimes be added to the edges of DAGs following the rigorous rules.[44–46] Briefly, a positive sign can be added to an edge from *A* to *Y* if a distributional positive monotonic effect of *A* on *Y* is met, which means that intervening on *A* will increase or leave unchanged the distribution of *Y* over the population, regardless of

the value of the parents of Y other than A.[45] A negative sign can be analogously added. If A has neither a distributional positive monotonic effect nor a distributional negative monotonic effect on Y, then the edge from A to Y is said to be without a sign. The sign of a path on a DAG is the product of the signs of the edges that constitute that path. If one of the edges on a path is without a sign, then the sign of the path is said to be undefined. Based on these rules, DAGs with signed edges, or signed DAGs, can be used to draw conclusions about the direction of the bias of unmeasured confounding under specific assumptions. Weaker assumptions can be drawn in certain situations.[44]

When using signed DAGs, the direction of nodes themselves should be also clearly indicated; for example, instead of a node for "sex", we need to use a node for "male".

2-10. DAGs can be (and should be) used to describe not only confounding bias but also other forms of bias

Because DAGs have been used extensively to determine the variables for which it is sufficient to control for confounding to estimate causal effects, some may misunderstand that DAGs can be used only for assessment of confounding bias. To obtain an accurate estimate of the causal effect of exposure on outcome, however, we need to deal with other important forms of bias.[36] If DAGs are used only as tools to assess confounding bias, we may overlook other important bias and fail to properly assess them.

DAGs can be used to illustrate not only confounding bias but also selection bias and measurement bias.[42,46,47] It is fairly well known that both confounding bias and selection bias imply a lack of exchangeability between the exposed and unexposed groups,[33,39,42,48] and they may be collectively referred to as nonexchangeability bias.[36] In some cases, confounding bias and selection bias cannot be clearly distinguished, and the same is true in DAGs. The blurred border between confounding bias and selection bias can be illustrated using M bias[42]; while some prefer to classify M bias as selection bias because it is induced by conditioning on a common effect of causes of exposure and outcome,[42] others may prefer to refer to it as confounding bias based on the position of the entire path in the DAG.[3] Even if the distinction between confounding bias and selection bias cannot be clearly made, DAGs can enhance communication among investigators to avoid unnecessary misunderstandings. Thus, even though analyses may be conditioned on participants in a study or survival to a certain age, it is preferable to include that selection process in the DAG, rather than omit it entirely. Unlike confounding bias and selection bias, measurement bias is explained by the presence of backdoor path(s) between the misclassified exposure and the misclassified outcome.[36] DAGs can be used to represent four types of measurement error: independent nondifferential, dependent nondifferential, independent differential, and dependent differential errors.[46,47]

To summarize, DAGs can be used to illustrate these three main source of bias (ie, confounding bias, selection bias, and measurement bias); a recent paper refers to these biases as structural error.[36] DAGs can be used at various steps during the process of conducting research, contributing to understanding of the overall picture of errors in causal inference. Even in RCTs, because randomization at baseline does not guarantee exchangeability during the post-baseline period (eg, due to differential loss to follow-up), it is significant to consider the whole process of obtaining data when drawing DAGs. For details of illustration, see elsewhere.[49,50]

3. FUTURE DIRECTIONS

Graphical models are useful tools in causal inference. In this paper, we have primarily focused on causal DAGs, which provide one basis of our clearer understanding of causality if appropriately used. In addition, there are other important causal models aside from graphical ones. Greenland and Brumback discussed four major causal model frameworks[51]; graphical models, potential outcome (counterfactual) models, sufficient cause models, and structural equation models. These frameworks are distinct but closely related, providing complementary perspectives of causality. For example, DAGs are often criticized for not allowing for the representation of interactions among variables. In this regard, VanderWeele and Robins proposed to incorporate sufficient causes into the DAG framework, in which an ellipse is put around the sufficient-cause nodes to indicate that the set is determinative.[52,53] Causal DAGs with sufficient causation structures thus allow for the graphical representation of interactions on DAGs. Furthermore, by formulating mediation in the sufficient cause framework,[54,55] Suzuki et al showed that operating mediation and mechanism can be identified from empirical data, illustrating it in causal DAGs with sufficient causation structures.[56] A similar approach was used to help distinguish and relate concepts of confounding and of covariate balance.[57] These approaches provide the link between the graphical models and the sufficient cause models among the four major causal model frameworks. Further, DAGs represent nonparametric structural equation models,[2,5] and the link between graphical models and potential outcome models becomes clearer using advanced types of causal diagrams, as described later. In these manners, attempts to integrate the four major causal model frameworks can lead to profound, universal perspectives of causality.[58–62] Further developments of causal theories would enhance versatility of causal diagrams in applied research. Meanwhile, the significance of study design itself should not be overlooked for valid causal inference, and it is important to have a clear research question before even beginning to draw a DAG.[8]

Although DAGs are useful tools to make decisions for confounder selection, one can rarely construct perfectly correct DAGs because complete knowledge of underlying causal structures is often unavailable. Uncertainties prevail in biomedical and social sciences.[63] Accordingly, VanderWeele proposed a practical approach to confounder selection decisions[64]: control for each covariate that is a cause of the exposure, or of the outcome, or of both; exclude from this set any variable known to be an instrumental variable; and include as a covariate any proxy for an unmeasured variable that is a common cause of both the exposure and the outcome. This approach is referred to as a "modified disjunctive cause criterion", which would be a useful guide for confounder selection, once we can properly use DAGs to avoid the pitfalls.

Finally, conventional causal diagrams do not include the underlying counterfactual variables on the graphs. In other words, the link between the graphical models and the potential outcome models has remained traditionally hidden.[5] In this regard, Richardson and Robins developed single world intervention graphs (SWIGs),[65] which explicitly connect the potential outcome framework with DAGs. Although DAGs are used to visually summarize hypothetical relations among observed variables, SWIGs allow us to show hypothetical relations between observed/unobserved factual random variables and potential (or

counterfactual) outcomes. This advanced type of causal diagram is introduced primarily in a chapter about confounding of *Causal Inference: What If*,[5] and its further use is expected. Although SWIGs have been used less in other settings, a recent letter discussed a practical example demonstrating the utility of SWIGs for selection bias.[66] As another approach to describe hidden causal structures, Suzuki et al proposed extended DAGs, which integrate response types and observed variables.[39] They show the conceptual link between unobservable response types and observed (or observable) data frequencies in the population. This is crucial because the causal effect of exposure on disease frequency in a population depends on the distribution of response types of individuals in that population. As an example of their usefulness, the principal stratification approach can be illustrated using extended DAGs to describe such contexts as truncation by death, mediation, and noncompliance.[67] Some other extensions of DAGs include selection diagrams[68] and missingness graphs.[69]

As addressed in this article, there are many pitfalls to avoid when using DAGs. Indeed, DAGs are not panacea, and there is no magic bullet for causal inference. Nevertheless, DAGs have proven to be a useful tool to clarify our causal thinking. As a visual aid, causal diagrams provide an incomparable help at the stages of study design, data collection and analysis, and interpretation of study findings. Furthermore, their value as a pedagogical tool cannot be overlooked in the causal inference literature. It is hoped that the utility of DAGs will be further advanced as an interdisciplinary tool, and thus contribute to elaborate scientific understanding of causality.

ACKNOWLEDGEMENTS

APPENDIX

As explained in Section 2-2 of the main text, an arrow from P to Q in a DAG indicates that P has a causal effect on Q of at least one individual in the population. In this Appendix, we illustrate this point following an example in Fine Point 6.2 of *Causal Inference: What If*.[5] We let A denote a binary exposure (1 = received heart transplant, 0 = did not receive heart transplant) and Y denote a binary outcome (1 = deceased, 0 = did not decease) and M denote a binary covariate (1 = men, 0 = women). A half of the target population was men, and the other half was women. We consider a simple RCT, letting p denote a probability of exposure in the population, irrespective of sex. The assumptions about the two random variables A and M can be respectively described as:

$$A \sim p_A(a) = \begin{cases} p & (a = 1) \\ 1 - p & (a = 0) \end{cases} \quad \text{(Eq. 1)}$$

and

$$M \sim p_M(m) = \begin{cases} 1/2 & (m = 1) \\ 1/2 & (m = 0) \end{cases} \quad \text{(Eq. 2)}$$
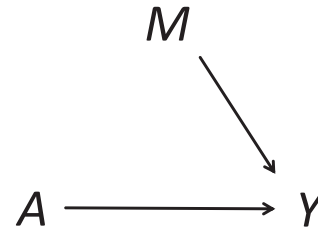

Figure 4. DAG that properly illustrate the underlying causal structures, before conditioning on M. We let A denote a binary exposure (1 = received heart transplant, 0 = did not receive heart transplant) and Y denote a binary outcome (1 = deceased, 0 = did not decease) and M denote a binary covariate (1 = men, 0 = women). DAG, directed acyclic graph.

Consider a situation in which heart transplant A has a causal effect to increase the risk of death Y in every woman and has a causal effect to decrease the risk of death in every man. In other words, the response type of all the women is "causal" whereas that of all the men is "preventive". Because the beneficial and harmful effects of A perfectly cancel out, the average causal effect in the target population is null. Despite the absence of average causal effect from A on Y, however, we need to draw an arrow from A to Y to make them d-connected because heart transplant A has a causal effect on the death Y of at least one (actually, all of the) individuals in the population (Figure 4). Below, we discuss this point by highlighting the difference between issues of causal structures and statistical (in)dependencies. Note that M is referred to as a direct effect modifier of A on Y.[70]

First, we discuss the relationship between A, M, and Y from the perspective of causal structures. Recall that a DAG represents a nonparametric structural equation model, and in Figure 4, we consider the following structural equations:

$$\begin{cases} A = f_A(\varepsilon_A) \\ M = f_M(\varepsilon_M) \\ Y = f_Y(A, M, \varepsilon_Y) \end{cases} \quad \text{(Eq. 3)}$$

where $f(\cdot)$ is an arbitrary function and ε represents all causes or random components of the corresponding variable not represented in the DAG. Note that these random components are independent of each other. Suppose that the nonparametric structural equation models in this example are parametrized as:

$$\begin{cases} A = f_A(\varepsilon_A) \\ M = f_M(\varepsilon_M) \\ Y = f_Y(A, M) = A(1 - M) + (1 - A)M \end{cases} \quad \text{(Eq. 4)}$$

Note that we dropped off ε_Y from Equation 4 because we consider an example in which Y is fully determined by A and M. Then, among men (ie, $M = 1$), the structural equation for Y in Equation 4 becomes

$$Y = f_Y(A, 1) = 1 - A \quad \text{(Eq. 5)}$$

which shows that the response type of all the men is "preventive". By contrast, among women (ie, $M = 0$), the structural equation for Y in Equation 4 becomes

$$Y = f_Y(A, 0) = A \quad \text{(Eq. 6)}$$

which shows that the response type of all the women is "causal". Note that Equations 5 and 6 show that Y is causally determined
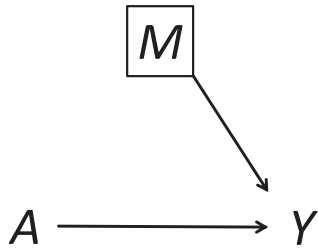
Figure 5. DAG that properly illustrate the underlying causal structures, after conditioning on M. We let A denote a binary exposure ($1 =$ received heart transplant, $0 =$ did not receive heart transplant) and Y denote a binary outcome ($1 =$ deceased, $0 =$ did not decease) and M denote a binary covariate ($1 =$ men, $0 =$ women). DAG, directed acyclic graph.
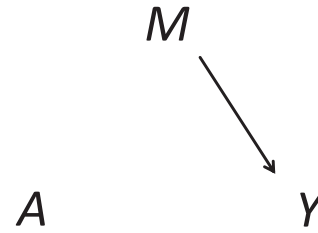

Figure 6. DAG that does not properly illustrate the underlying causal structures. We let A denote a binary exposure ($1 =$ received heart transplant, $0 =$ did not receive heart transplant) and Y denote a binary outcome ($1 =$ deceased, $0 =$ did not decease) and M denote a binary covariate ($1 =$ men, $0 =$ women). DAG, directed acyclic graph.

by A conditional on M. This is graphically illustrated as the arrow from A to Y in the DAG with a box around M (Figure 5). Thus, A and Y are conditionally d-connected.

Next, we discuss the relationship between A, M, and Y from the perspective of statistical (in)dependencies. From Equations 1, 2, and 4, the joint distribution of (A, M, Y) is described as:

$$(A, M, Y) \sim p_{AMY}(a, m, y)$$
$$= \begin{cases} 1/2 \cdot p & (a, m, y) = (1, 1, 0), (1, 0, 1) \\ 1/2 \cdot (1 - p) & (a, m, y) = (0, 1, 1), (0, 0, 0) \\ 0 & \text{otherwise} \end{cases} \quad \text{(Eq. 7)}$$

By marginalizing over M, the joint distribution of (A, Y) is given as:

$$(A, Y) \sim p_{AY}(a, y) = \begin{cases} 1/2 \cdot p & (a, y) = (1, 0), (1, 1) \\ 1/2 \cdot (1 - p) & (a, y) = (0, 1), (0, 0) \end{cases}$$
$$\text{(Eq. 8)}$$

Further, by marginalizing over A, the marginal distribution of Y is given as:

$$Y \sim p_Y(y) = \begin{cases} 1/2 & (y = 1) \\ 1/2 & (y = 0) \end{cases} \quad \text{(Eq. 9)}$$

which shows that the probability of death in the total population is 50% because the numbers of men and women are equal in this example. From Equations 1, 8, and 9, we obtain

$$\forall (a, y), \ p_{AY}(a, y) = p_A(a) p_Y(y) \quad \text{(Eq. 10)}$$

which clearly shows that A and Y are statistically independent when we do not condition on M. When we condition on M, the conditional distribution of $(A, Y)|_M$ is described as:

$$(A, Y)|_{M=1} \sim p_{AY|M}(a, y | m = 1) = \frac{p_{AMY}(a, m = 1, y)}{p_M(m = 1)}$$
$$= \begin{cases} p & (a, y) = (1, 0) \\ 1 - p & (a, y) = (0, 1) \\ 0 & \text{otherwise} \end{cases}$$

$$(A, Y)|_{M=0} \sim p_{AY|M}(a, y | m = 0) = \frac{p_{AMY}(a, m = 0, y)}{p_M(m = 0)}$$
$$= \begin{cases} p & (a, y) = (1, 1) \\ 1 - p & (a, y) = (0, 0) \\ 0 & \text{otherwise} \end{cases}$$
$$\text{(Eq. 11)}$$

Further, the conditional distributions of $A|_M$ and $Y|_M$ are respectively given as:

$$A|_{M=1} \sim p_{A|M}(a | m = 1) = \begin{cases} p & (a = 1) \\ 1 - p & (a = 0) \end{cases}$$
$$A|_{M=0} \sim p_{A|M}(a | m = 0) = \begin{cases} p & (a = 1) \\ 1 - p & (a = 0) \end{cases} \quad \text{(Eq. 12)}$$

and

$$Y|_{M=1} \sim p_{Y|M}(y | m = 1) = \begin{cases} 1 - p & (y = 1) \\ p & (y = 0) \end{cases}$$
$$Y|_{M=0} \sim p_{Y|M}(y | m = 0) = \begin{cases} p & (y = 1) \\ 1 - p & (y = 0) \end{cases} \quad \text{(Eq. 13)}$$

From Equations 11 to 13, we obtain

$$p_{AY|M}(a = 1, y = 0 | m = 1)$$
$$= p \neq p \cdot p = p_{A|M}(a = 1 | m = 1) \cdot p_{Y|M}(y = 0 | m = 1)$$
$$p_{AY|M}(a = 1, y = 1 | m = 0)$$
$$= p \neq p \cdot p = p_{A|M}(a = 1 | m = 0) \cdot p_{Y|M}(y = 1 | m = 0)$$
$$\text{(Eq. 14)}$$

which shows that A and Y are statistically dependent conditional on M.

Then, what if one draws an arrow from A to Y based on the statistical (in)dependencies between them? Because A and Y are statistically independent without conditioning on M, one would not draw an arrow from A to Y in the DAG without a box around M (ie, d-separation of A and Y) as shown in Figure 6. Note that one (implicitly) employ an assumption of *faithfulness* here. By contrast, because A and Y are statistically dependent conditional on M, one would draw an arrow from A to Y in the DAG with a box around M (ie, conditional d-connection of A and Y) as shown in Figure 5, which is based on the rule of *compatibility*. Therefore, if one draws an arrow from A to Y based on the statistical (in)dependencies, the presence of the arrow depends on the presence of a "box" around M. In other words, the presence of the arrow from A to Y becomes "unstable".

To summarize, although A and Y are statistically independent without conditioning on M, we need to draw an arrow from A to Y irrespective of whether we draw a box around M or not. This is because Y is causally determined by A. Thus, Figure 4 and Figure 5 properly illustrate the underlying causal structures, but Figure 6 does not. Under the rule of compatibility, d-connection is a necessary but not a sufficient condition for statistical dependence. The above discussion exemplifies that an arrow in

DAGs represents the presence of individual causal effect in the target population.

REFERENCES

1. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999;10:37–48.
2. Pearl J. *Causality: Models, Reasoning, and Inference*. 2nd ed. New York, NY. Cambridge University Press; 2009.
3. Glymour MM, Greenland S. Causal diagram. In: Rothman KJ, Greenland S, Lash TL, eds. *Modern Epidemiology*. 3rd ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2008:183–209.
4. Glymour MM. Using causal diagrams to understand common problems in social epidemiology. In: Oakes JM, Kaufman JS, eds. *Methods in Social Epidemiology*. 2nd ed. San Francisco, CA: Jossey-Bass; 2017:458–492.
5. Hernán MA, Robins JM. *Causal Inference: What If*. Boca Raton, FL. Chapman & Hall/CRC; 2020.
6. Fleischer NL, Diez Roux AV. Using directed acyclic graphs to guide analyses of neighbourhood health effects: an introduction. *J Epidemiol Community Health*. 2008;62:842–846.
7. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. BMC Med Res Methodol [Internet]. 2008 Oct [cited 2008 Oct 30];8:[70 p.]. Available from: www.biomedcentral.com/1471-2288/8/70.
8. Suzuki E, Komatsu H, Yorifuji T, Yamamoto E, Doi H, Tsuda T. Causal inference in medicine part II: directed acyclic graphs—a useful method for confounder selection, categorization of potential biases, and hypothesis specification. *Nihon Eiseigaku Zasshi*. 2009; 64:796–805.
9. Last JM. *A Dictionary of Epidemiology*. 4th ed. New York, NY. Oxford University Press; 2001.
10. Porta MS. *A Dictionary of Epidemiology*. 5th ed. New York, NY. Oxford University Press; 2008.
11. Porta MS. *A Dictionary of Epidemiology*. 6th ed. New York, NY. Oxford University Press; 2014.
12. Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiology*. 2003;14:300–306.
13. Steiner PM, Kim Y, Hall CE, Su D. Graphical models for quasi-experimental designs. *Sociol Methods Res*. 2017;46:155–188.
14. Mansournia MA, Hernán MA, Greenland S. Matched designs and causal diagrams. *Int J Epidemiol*. 2013;42:860–869.
15. Mansournia MA, Jewell NP, Greenland S. Case-control matching: effects, misconceptions, and recommendations. *Eur J Epidemiol*. 2018;33:5–14.
16. Little RJ, Rubin DB. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. *Annu Rev Public Health*. 2000;21:121–145.
17. Hernán MA. A definition of causal effect for epidemiological research. *J Epidemiol Community Health*. 2004;58:265–271.
18. Holland PW. Statistics and causal inference. *J Am Stat Assoc*. 1986; 81:945–960.
19. VanderWeele TJ, Hernán MA. Causal effects and natural laws: towards a conceptualization of causal counterfactuals for non-manipulable exposures, with application to the effects of race and sex. In: Berzuini C, Dawid P, Bernardinelli L, eds. *Causality: Statistical Perspectives and Applications*. Hoboken, NJ: Wiley; 2012:101–113.
20. Suzuki E. Time changes, so do people. *Soc Sci Med*. 2012;75:452–456; discussion 457–458.
21. VanderWeele TJ. On well-defined hypothetical interventions in the potential outcomes framework. *Epidemiology*. 2018;29:e24–e25.
22. Catalano R, Bruckner T, Smith KR. Ambient temperature predicts sex ratios and male longevity. *Proc Natl Acad Sci USA*. 2008;105: 2244–2247.
23. Hernán MA, VanderWeele TJ. Compound treatments and transportability of causal inference. *Epidemiology*. 2011;22:368–377.
24. Greenland S, Morgenstern H. Confounding in health research. *Annu Rev Public Health*. 2001;22:189–212.
25. Maldonado G, Greenland S. Estimating causal effects. *Int J Epidemiol*. 2002;31:422–429.
26. Suzuki E, Mitsuhashi T, Tsuda T, Yamamoto E. A typology of four notions of confounding in epidemiology. *J Epidemiol*. 2017;27:49–55.
27. Suzuki E, Mitsuhashi T, Tsuda T, Yamamoto E. A simple example as a pedagogical device? *Ann Epidemiol*. 2014;24:560–561.
28. Maldonado G. Toward a clearer understanding of causal concepts in epidemiology. *Ann Epidemiol*. 2013;23:743–749.
29. Flanders WD, Johnson CY, Howards PP, Greenland S. Dependence of confounding on the target population: a modification of causal graphs to account for co-action. *Ann Epidemiol*. 2011;21:698–705.
30. VanderWeele TJ. Confounding and effect modification: distribution and measure. Epidemiol Method [Internet]. 2012 Aug [cited 2012 Aug 29];1(1):[55–82 p.]. Available from: http://www.degruyter.com/view/j/em.2012.1.issue-1/2161-962X.1004/2161-962X.1004.xml?format=INT.
31. Suzuki E, Yamamoto E. Further refinements to the organizational schema for causal effects. *Epidemiology*. 2014;25:618–619.
32. Greenland S, Robins JM, Pearl J. Confounding and collapsibility in causal inference. *Stat Sci*. 1999;14:29–46.
33. Greenland S, Robins JM. Identifiability, exchangeability and confounding revisited. Epidemiol Perspect Innov [Internet]. 2009 Sep [cited 2009 Sep 4];6:[4 p.]. Available from: https://epi-perspectives.biomedcentral.com/articles/10.1186/1742-5573-6-4.
34. Greenland S. Randomization, statistics, and causal inference. *Epidemiology*. 1990;1:421–429.
35. Greenland S, Mansournia MA. Limitations of individual causal models, causal graphs, and ignorability assumptions, as illustrated by random confounding and design unfaithfulness. *Eur J Epidemiol*. 2015;30:1101–1110.
36. Suzuki E, Tsuda T, Mitsuhashi T, Mansournia MA, Yamamoto E. Errors in causal inference: an organizational schema for systematic error and random error. *Ann Epidemiol*. 2016;26:788–793.e1.
37. Everitt B, Skrondal A. *The Cambridge Dictionary of Statistics*. 4th ed. Cambridge, UK. Cambridge University Press; 2010.
38. Upton G, Cook I. *A Dictionary of Statistics*. 2nd ed. New York, NY. Oxford University Press; 2008.
39. Suzuki E, Mitsuhashi T, Tsuda T, Yamamoto E. A counterfactual approach to bias and effect modification in terms of response types. BMC Med Res Methodol [Internet]. 2013 July [cited 2013 July 31];13:[101 p.]. Available from: http://www.biomedcentral.com/1471-2288/13/101.
40. Sjölander A, Greenland S. Ignoring the matching variables in cohort studies - when is it valid and why? *Stat Med*. 2013;32:4696–4708.
41. Shinozaki T, Nojima M. Misuse of regression adjustment for additional confounders following insufficient propensity score balancing. *Epidemiology*. 2019;30:541–548.
42. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15:615–625.
43. Szklo M, Nieto FJ. *Epidemiology: Beyond the Basics*. 3rd ed. Burlington, MA. Jones & Bartlett Learning; 2012.
44. VanderWeele TJ, Hernán MA, Robins JM. Causal directed acyclic graphs and the direction of unmeasured confounding bias. *Epidemiology*. 2008;19:720–728.
45. VanderWeele TJ, Robins JM. Signed directed acyclic graphs for causal inference. *J R Stat Soc Series B Stat Methodol*. 2010;72:111–127.
46. VanderWeele TJ, Hernán MA. Results on differential and dependent measurement error of the exposure and the outcome using signed directed acyclic graphs. *Am J Epidemiol*. 2012;175:1303–1310.
47. Hernán MA, Cole SR. Invited commentary: causal diagrams and measurement bias. *Am J Epidemiol*. 2009;170:959–962; discussion 963–964.
48. Flanders WD, Eldridge RC. Summary of relationships between exchangeability, biasing paths and bias. *Eur J Epidemiol*. 2015;30: 1089–1099.
49. Mansournia MA, Higgins JPT, Sterne JAC, Hernán MA. Biases in randomized trials: a conversation between trialists and epidemiologists. *Epidemiology*. 2017;28:54–59.

50. Shinozaki T, Hagiwara Y, Matsuyama Y. Re: Biases in randomized trials: a conversation between trialists and epidemiologists. *Epidemiology*. 2017;28:e40.

51. Greenland S, Brumback B. An overview of relations among causal modelling methods. *Int J Epidemiol*. 2002;31:1030–1037.

52. VanderWeele TJ, Robins JM. Directed acyclic graphs, sufficient causes, and the properties of conditioning on a common effect. *Am J Epidemiol*. 2007;166:1096–1104.

53. VanderWeele TJ, Robins JM. Minimal sufficient causation and directed acyclic graphs. *Ann Stat*. 2009;37:1437–1465.

54. Hafeman DM. A sufficient cause based approach to the assessment of mediation. *Eur J Epidemiol*. 2008;23:711–721.

55. VanderWeele TJ. Mediation and mechanism. *Eur J Epidemiol*. 2009;24:217–224.

56. Suzuki E, Yamamoto E, Tsuda T. Identification of operating mediation and mechanism in the sufficient-component cause framework. *Eur J Epidemiol*. 2011;26:347–357.

57. Suzuki E, Tsuda T, Yamamoto E. Covariate balance for no confounding in the sufficient-cause model. *Ann Epidemiol*. 2018; 28:48–53.e2.

58. Greenland S, Poole C. Invariants and noninvariants in the concept of interdependent effects. *Scand J Work Environ Health*. 1988;14:125–129.

59. Flanders WD. On the relationship of sufficient component cause models with potential outcome (counterfactual) models. *Eur J Epidemiol*. 2006;21:847–853.

60. VanderWeele TJ, Hernán MA. From counterfactuals to sufficient component causes and vice versa. *Eur J Epidemiol*. 2006;21:855–858.

61. Suzuki E, Yamamoto E, Tsuda T. On the link between sufficient-cause model and potential-outcome model. *Epidemiology*. 2011;22:131–132.

62. Suzuki E, Yamamoto E, Tsuda T. On the relations between excess fraction, attributable fraction, and etiologic fraction. *Am J Epidemiol*. 2012;175:567–575.

63. Suzuki E, VanderWeele TJ. Mechanisms and uncertainty in randomized controlled trials: A commentary on Deaton and Cartwright. *Soc Sci Med*. 2018;210:83–85.

64. VanderWeele TJ. Principles of confounder selection. *Eur J Epidemiol*. 2019;34:211–219.

65. Richardson TS, Robins JM. *Single world intervention graphs (SWIGs): a unification of the counterfactual and graphical approaches to causality*. Center for Statistics and the Social Sciences, University of Washington, Working Paper. 2013;128.

66. Breskin A, Cole SR, Hudgens MG. A practical example demonstrating the utility of single-world intervention graphs. *Epidemiology*. 2018;29:e20–e21.

67. Suzuki E. Generalized causal measure: the beauty lies in its generality. *Epidemiology*. 2015;26:490–495.

68. Pearl J, Bareinboim E. External validity: from do-calculus to transportability across populations. *Stat Sci*. 2014;29:579–595.

69. 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;187:2705–2715.

70. VanderWeele TJ, Robins JM. Four types of effect modification: a classification based on directed acyclic graphs. *Epidemiology*. 2007;18:561–568.