AOGS
Acta Obstetricia et Gynecologica Scandinavica

AOGS REVIEW

An overview of confounding. Part 2: how to identify it and special situations

PENELOPE P. HOWARDS

Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA, USA

Key words
Bias, causality, confounding factors (epidemiology), data analysis, epidemiologic methods, epidemiologic research design

Correspondence
Penelope Howards, Department of Epidemiology, 1518 Clifton Rd. NE, CNR 3029, Atlanta, GA 30306, USA.
E-mail: penelope.howards@emory.edu

Conflict of interest
The author has stated explicitly that she has no conflicts of interest with this article.

Please cite this article as: Howards PP. An overview of confounding. Part 2: how to identify it and special situations. Acta Obstet Gynecol Scand. 2018; 97:400–406.

Received: 21 December 2017
Accepted: 28 December 2017

DOI: 10.1111/aogs.13293

Abstract

Confounding biases study results when the effect of the exposure on the outcome mixes with the effects of other risk and protective factors for the outcome that are present differentially by exposure status. However, not all differences between the exposed and unexposed group cause confounding. Thus, sources of confounding must be identified before they can be addressed. Confounding is absent in an ideal study where all of the population of interest is exposed in one universe and is unexposed in a parallel universe. In an actual study, an observed unexposed population represents the unobserved parallel universe. Thinking about differences between this substitute population and the unexposed parallel universe helps identify sources of confounding. These differences can then be represented in a diagram that shows how risk and protective factors for the outcome are related to the exposure. Sources of confounding identified in the diagram should be addressed analytically and through study design. However, treating all factors that differ by exposure status as confounders without considering the structure of their relation to the exposure can introduce bias. For example, conditions affected by the exposure are not confounders. There are also special types of confounding, such as time-varying confounding and unfixable confounding. It is important to evaluate carefully whether factors of interest contribute to confounding because bias can be introduced both by ignoring potential confounders and by adjusting for factors that are not confounders. The resulting bias can result in misleading conclusions about the effect of the exposure of interest on the outcome.

Abbreviations: DAG, directed acyclic graph.

Introduction

Confounding is one of three types of bias that can distort the results of epidemiologic studies and potentially lead to erroneous conclusions. In the companion paper in this journal (1), we discuss how confounding occurs and how to address it. In short, confounding can be considered the confusion of the effect of the exposure on the outcome with the effects of other risk and protective factors for the outcome that are present differentially by exposure status (2). The ideal study design would avoid confounding through the use of parallel universes where all study participants are exposed in one universe and unexposed in the other universe. Only exposure status and consequences of being exposed would differ across the universes. Under these conditions, if the outcome were more common in one universe than another, this difference would be caused by the exposure and would not be the result of other risk or protective factors for the

Key message

Diagramming the relations among the exposure, outcome, and other factors affecting the outcome or exposure can help identify sources of confounding. Confounding can be addressed, but bias can occur from treating a non-confounder as a confounder.

outcome because those factors would be equally present or absent in each universe.

In a real study, the presence of risk and protective factors for the outcome may differ between the exposed and unexposed study participants, and this imbalance can result in confounding. However, not all differences between the exposed and unexposed populations cause confounding. To identify potential confounders, we need to think about risk and protective factors for the outcome and how they relate to the exposure of interest. In this paper, we describe how to identify potential sources of confounding. We also consider variables that should not be treated as confounders, and we introduce special confounding conditions.

Identifying confounders

Many people learned to identify confounders by considering the following three criteria: (1) the exposure is associated with the confounder, (2) the outcome is associated with the confounder among the unexposed (or sometimes more specifically, the confounder is a risk factor or a proxy for a risk factor for the outcome), and (3) the confounder is not an intermediate (2). However, more recent developments in the field of epidemiology have found that this approach is inadequate and may introduce bias into a study if implemented (3–5). Directed acyclic graphs (DAGs) are tools that allow for a more comprehensive examination of potential sources of confounding. They provide a graphical representation of hypothesized relations among the exposure of interest (E), the outcome of interest (O), and other variables related to the exposure and the outcome. DAGs have formal rules regarding representation and interpretation, which have been described in more detail elsewhere (4,6,7).

Typically, DAGs represent hypothesized causal relations in an infinitely large source population where chance relations are removed. DAGs are composed of factors of interest, including unmeasured factors, and arrows connecting these factors. The arrows indicate causal relations with the factor at the tail of the arrow causing or preventing the condition at the head of the arrow. However, the arrows are not deterministic in the sense that not all exposed women get the outcome, for example, but exposure causes the outcome in at least some of the women. Two-headed arrows are not allowed because a condition cannot be the result of something it caused. When tempted to draw a two-headed arrow, it is often helpful instead to consider multiple time points (for example, weight at time 1 may cause dieting, which may affect weight at time 2).

In Figure 1A, magnesium supplementation (E) prevents preterm birth (O), and older maternal age (C) causes women to use magnesium supplements and causes preterm birth. There is confounding in this case because there is a path (a series of variables and arrows) from the exposure to the outcome through maternal age that is not caused by magnesium supplementation. This path leads to an association between magnesium supplement use and preterm birth, but it is not part of the causal effect of magnesium on preterm birth. Thus, if this DAG is correct, the unadjusted estimate of the effect of magnesium on preterm birth is a confusion of the causal effect and the non-causal association represented by these two paths. To estimate the effect of magnesium on preterm birth alone, we must block the path through maternal age using one of the methods described in the companion paper (1). Essentially, the arrow between maternal age and magnesium supplementation exists in our source population, but we can remove that arrow in our study (through study design or analytically) and block the confounding path.

DAGs can represent different sources of bias introduced by non-causal paths (8,9). Confounding paths

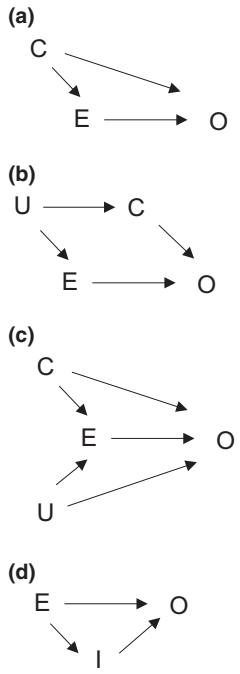

Figure 1. (a) The exposure (E) causes or prevents the outcome (O). However, there is confounding by a variable (C) that causes or prevents both the exposure and the outcome. (b) The exposure (E) causes or prevents the outcome (O). However, there is confounding by an unmeasured variable (U) that causes or prevents the exposure and another condition (C), which in turn affects the outcome. (c) The exposure (E) causes or prevents the outcome (O). However, there is confounding by a variable (C) that causes or prevents both the exposure and the outcome as well as confounding by an unmeasured factor (U). (d) The exposure (E) causes or prevents the outcome (O) directly and by affecting another condition (I), which in turn affects the outcome.

include a subset of non-causal paths from the exposure to the outcome that start with an arrow pointing towards the exposure. Thus, confounders are often thought of as factors that affect the exposure and the outcome, as in Figure 1A (9,10). However, Figure 1B illustrates a situation where the factor that causes both the exposure and the outcome is unmeasured (U). Nevertheless, we could block the confounding path by adjusting for C, even though C does not cause the exposure. For example, suppose we are interested in whether drinking alcohol (E) increases the risk of having a low birthweight birth (O). Smoking (C) causes low birthweight but does not cause alcohol consumption. Nevertheless, there may be an unmeasured factor (U) that causes women to drink alcohol and to smoke. This factor creates a non-causal path from alcohol to low birthweight. If this DAG is correct, the unadjusted estimate of the effect of drinking alcohol on low birthweight would be confounded, but adjusting for smoking would block the confounding path even though smoking does not cause alcohol consumption. Of note, smoking would be associated with drinking alcohol even though it did not cause alcohol consumption because the unmeasured factor causes both smoking and drinking.

In Figure 1C, there is a non-causal path through an unmeasured confounder with no other factors on the path. The unmeasured confounder could be a hypothesized but unidentified factor that affects the exposure and the outcome, or it could be a known confounder that was not measured in the study. Suppose income (U) is unmeasured but affects a woman's ability to afford magnesium supplements and low income also increases her risk of preterm birth. If this DAG is correct, the unadjusted estimate of the effect of magnesium supplements on preterm birth is a mixture of the true effect and confounding by maternal age and income. If we adjust for maternal age, we block the path through maternal age and remove confounding by maternal age, but our estimate is still confounded by income because we did not have data to adjust for it [see Supporting Information Appendix S2 of the companion paper (1) for a mathematical example]. If we measured both age (C) and income (U), we could address confounding by both factors analytically.

Factors caused by exposure are not confounders

Confounding paths start with an arrow pointing towards the exposure. In contrast, paths with variables affected by the exposure are not confounding paths. This statement is more general but consistent with the traditional third criterion described above that a confounder not be an intermediate. Even in the ideal study, factors affected by the exposure, change between the two universes. In other words, in addition to the outcome, other conditions can be affected by exposure; these factors are not confounders and should not be treated as such (3,5,10,11). Specifically, factors on the causal path between the exposure and the outcome are often referred to as intermediates or mediators, and their effect on the outcome is part of the total effect of the exposure on the outcome. For example, a researcher interested in the effect of gestational diabetes (E) on macrosomia (O) might be tempted to adjust for insulin treatment (I) (Figure 1D). However, insulin treatment is a result of gestational diabetes and should not be considered a confounder. Gestational diabetes is what caused the woman to be treated with insulin in one universe but not the other. Any effect of insulin on macrosomia contributes to the total effect of gestational diabetes on macrosomia. Thus, even though the exposed women (those with gestational diabetes) have a different distribution of insulin use than the unexposed women (those without gestational diabetes), insulin use is not a confounder. Figure 1D illustrates the absence of confounding (in this unrealistically simple scenario) by the absence of a non-causal path from the exposure to the outcome. If the researcher were interested in insulin use, this would be a separate research question. For example, one could compare the effect of treatment with insulin on macrosomia with not being treated with insulin among women with gestational diabetes.

Similarly, maternal smoking (E) causes low birthweight (I) and infant mortality (O). In addition, low birthweight causes infant mortality (Figure 1D). The total effect of maternal smoking on infant mortality includes both the path directly to infant mortality and the path through low birthweight. In this case, our unadjusted estimate of the effect of maternal smoking on infant mortality combines the effect of these two paths, which is appropriate because both paths are ways that maternal smoking affects infant mortality. However, some researchers might be interested in the direct effect of maternal smoking on infant mortality, which corresponds to the path that does not include low birthweight. Investigators might mistakenly opt to apply a technique used to address confounding, such as adjusting for low birthweight in a model or stratifying on low birthweight, but such analyses can produce misleading results under some conditions (10,12–14). Interest in direct effects requires mediation analysis (15,16), which is distinct from confounding and beyond the scope of this paper.

Another example involves pre-pregnancy obesity (E), which is strongly associated with postpartum iron deficiency (O), compared with not being obese (17). Pre-pregnancy obesity is also associated with a higher incidence of postpartum hemorrhage (I), which is

associated with postpartum iron deficiency (Figure 1D). It might be tempting to adjust for postpartum hemorrhage. However, postpartum hemorrhage is likely an intermediate between pre-pregnancy obesity and postpartum iron deficiency. Therefore, the total effect of pre-pregnancy obesity on postpartum iron deficiency includes the direct effect and the effect through postpartum hemorrhage. Mediation analysis would be required if there were an interest in separating these effects (15,16).

Even factors caused by the exposure but not on the causal path are not confounders; adjusting for such factors can introduce bias even when they are not intermediates (3,5). To understand why, we need to learn about colliders.

Conditioning on colliders creates spurious associations

In the simplest cases, the traditional confounding criteria are consistent with DAGs. However, these criteria are not equipped to address colliders. Colliders are variables where two arrowheads meet (Y in Figure 2). Thus far, we have identified two ways that factors may be associated. The first is that one factor causes the other, such as magnesium supplementation preventing preterm birth (E→O in Figure 1A). The second is that both factors are caused by a third factor. For example, maternal age affects magnesium supplementation and it affects preterm birth, so magnesium supplementation is also associated with preterm birth because they are both caused by maternal age (for example, confounding: E←C→0 in Figure 1A). While adjusting for a non-collider, such as maternal age, (C) blocks a path, adjusting for a collider opens a path and may introduce bias instead of removing it (4,6).

Thus, conditioning on a collider is a third way two variables can appear to be associated although this association is a statistical artifact. To understand how this bias can occur, it is helpful first to consider the relation among three factors outside of a study context. For example, in a population of reproductive-aged women, one would not expect being sexually active (X) to be associated with Turner's syndrome (Z) (Figure 2). Some

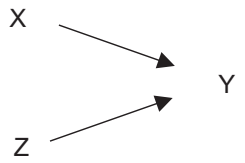

Figure 2. Factors X and Z both affect Y. Factor Y is a collider. X and Z are not associated in the total population, but stratifying on Y would introduce a spurious association between X and Z in at least one stratum of Y.

women who are sexually active have Turner's syndrome and some do not. Similarly, some women who are not sexually active have Turner's syndrome and some do not. However, being sexually active increases the probability of spontaneous pregnancy (Y), and Turner's syndrome decreases the probability of spontaneous pregnancy. If we stratified on spontaneous pregnancy, the pregnant women would be sexually active and they would not have Turner' syndrome. Thus, among pregnant women, there would be an inverse association between being sexually active and Turner's syndrome. However, this association would not be because sexual activity causes Turner's syndrome or vice versa, and it would not be because another factor causes both sexual activity and Turner's syndrome. Instead, this association would be an artifact of stratifying on pregnancy. This is important because without drawing a DAG, we might mistakenly treat a collider as a confounder and introduce bias into our study (4,6).

In fact, one of the reasons adjusting for a factor caused by the exposure may introduce bias is that the factor may be a collider. For example, women who have had an adverse pregnancy outcome (for example, spontaneous abortion, preterm birth, etc.) are at greater risk of having a second pregnancy with an adverse pregnancy outcome than are women whose first pregnancy was healthy. Although in most cases it is unlikely that the first pregnancy outcome causes the second pregnancy outcome directly, having a prior adverse pregnancy outcome is often treated as a confounder. The hypothesis is likely that the prior adverse pregnancy outcome can stand in for an unmeasured confounder. However, treating pregnancy history as a confounder may introduce bias depending on the underlying relations among the variables. Consider a study of whether women exposed to diethylstilbestrol (DES) in utero (E) have an increased risk of having a pregnancy ending in a spontaneous abortion (O) compared with women who were not exposed to DES in utero (18). One might be tempted to adjust for having a history of a previous spontaneous abortion (P) in order to adjust for an unmeasured risk factor for spontaneous abortion (U). If in utero exposure to DES causes women to be at greater risk for spontaneous abortion, perhaps through adverse effects on the cervix which affect both pregnancies, then having a history of a spontaneous abortion is a result of being exposed to DES in utero (Figure 3). Thus, having a history of spontaneous abortion is a collider that is affected by the exposure and the unmeasured factor. It is not on a confounding path, and adjusting for it would introduce bias because adjusting for a collider on a non-causal path from the exposure to the outcome induces a spurious association between the exposure and the outcome (E→P←U→O). If the DAG is correct, bias would be introduced even though

pregnancy history is not an intermediate between DES and spontaneous abortion. Of note, the unmeasured risk factor for spontaneous abortion is not a confounder in Figure 3 because it is not on a non-causal path that starts with an arrow pointing at the exposure. Other structures could be hypothesized and evaluated for potential confounding, which might give the same conclusions or might give different conclusions (5,19).

It is important to consider colliders in other contexts as well. Suppose you are interested in whether diabetes (E) increases the risk of fetal death (O). Obesity (C) increases the risk of diabetes and of having hypertension (H). Hypertension may increase the risk of fetal death, and smoking (S) may increase the risk of hypertension and fetal death (Figure 4). In this DAG, there is a path from diabetes to obesity to hypertension to fetal death (E←C→H→O), which confounds the unadjusted estimate of the effect of diabetes on fetal death. In contrast, the path from diabetes to obesity to hypertension to smoking to fetal death (E←C→H←S→O) does not affect the unadjusted estimate because hypertension (H) is a collider and blocks this path. If this DAG were correct, we could adjust for obesity (C), which would provide an unconfounded estimate of the effect of diabetes on fetal death.

The traditional confounding criteria would suggest hypertension is a confounder because of the following: (1) it is associated with diabetes (because they are both caused by obesity), (2) it is associated with fetal death (because it causes fetal death), and (3) it is not an intermediate between diabetes and fetal death. However, the DAG reveals that adjusting for hypertension alone would introduce bias. Although hypertension is on the confounding path, it is also a collider on another path. Therefore, adjusting for hypertension would induce a spurious association corresponding to the path from diabetes to obesity to hypertension to smoking to fetal death (E←C→H←S→O). If obesity (C) were unmeasured, it would be necessary to adjust for hypertension (H) to block the confounding path, but it would then also be necessary to adjust for smoking (S) to block the path from diabetes to obesity to hypertension to smoking to the outcome (E←C→H←S→O).

In this paper, the DAGs represent simple situations for clarity. In an actual study, there would likely be numerous non-causal paths from the exposure to the outcome. To minimize bias, all non-causal paths would need to be blocked while leaving all causal paths opens. There may be more than one way to block a confounding path, as in Figure 4 where confounding could be prevented by adjusting for obesity (C) or by adjusting for both hypertension and smoking (H and S). When developing an analytic plan to address confounding, it is important to consider how adjusting for factors on one non-causal path might affect other paths.

Special confounding conditions

Time-varying confounding

For clarity, this paper focused on confounding in simple scenarios. However, there are more complex situations that require more complex methods to address confounding. One example of note is time-varying confounding. This can occur when there is a time-varying exposure and there is a factor that is an intermediate (F) between early exposure (E_1) and the outcome (O) that is also a confounder of the exposure-outcome relation at the later time point (E_2) (Figure 5). For example, if the research question is whether a higher cumulative dose of antiepileptic medication during pregnancy increases the risk of cognitive developmental delay (O). Antiepileptic medication dose might be measured early in pregnancy (E_1) and later in pregnancy (E_2). However, the medication dose at time 1 might affect whether the woman has a seizure (F), which in turn might affect the dose of her medication at time 2. Further having a seizure during pregnancy might affect the child's cognitive development. Thus, seizure would be an intermediate between early exposure and the outcome ($E_1 \rightarrow F \rightarrow E_2 \rightarrow O$) but would also be a confounder for the later exposure-outcome relation ($E_2 \leftarrow F \rightarrow O$). Although beyond the scope of this paper, this situation can be addressed through analytic methods such as marginal structural models (20).

Unfixable confounding

In contrast, unfixable confounding cannot be addressed analytically and can occur even in randomized studies. Unfixable confounding occurs under two circumstances. First, if everyone in the exposed group also experiences the confounder and no one in the unexposed group experiences the confounder, then it is impossible to disentangle the effect of the exposure from the effect of the confounder. Secondly, when no one in the exposed group has the confounder and everyone in the unexposed group has the confounder it is also impossible to separate the effect of the exposure of interest from the confounder. For example, it was once hypothesized that an unidentified factor present in potatoes caused spina bifida, and studies were performed comparing women with a potato-free diet to women who ate potatoes during pregnancy (21). Suppose investigators interested in this hypothesis enrolled women in a study where they were either randomized to a potato-free diet (the intervention) or a diet that allowed participants to eat potatoes. Further suppose

that the investigators replaced potatoes with beans for the potato-free diet. Beans are rich in folate, which was not known to protect against spina bifida at the time. For the sake of the example, ignore the fact that other foods have folate and assume the women eating potatoes never ate beans. In this fictitious study, all the women exposed to the intervention (those on the potato-free diet) are also exposed to folate, while all the women not receiving the intervention (those eating potatoes) are not. Therefore, there is unfixable confounding, which would make it appear that avoiding potatoes protects against spina bifida when in fact avoiding potatoes has no effect on spina bifida. The lack of an effect of potatoes on spina bifida is masked because of confounding by folate consumption by the women on the potato-free diet. Even if it were recognized that eating beans might prevent NTDs, the effect of beans could not be separated from the effect of avoiding potatoes analytically because everyone avoiding potatoes ate beans and no one eating potatoes ate beans.

Summary

Using DAGs to evaluate confounding has several advantages. They allow us to clarify our hypotheses to ourselves and others. Some researchers are uncomfortable with DAGs because they have doubts about the underlying relations among the variables. However, all research makes assumptions about underlying causal relations; DAGs make these assumptions explicit. Further, more than one DAG may be drawn to assess whether different possible scenarios would require adjustment for different factors or not. DAGs allow us to consider confounding by multiple factors at one time and can accommodate unmeasured factors. They help us to identify factors that might intuitively seem like confounders that are not confounders (such as intermediates and some colliders). They also help us to identify more complex situations, such as the examples illustrated in Figures 3–5, and provide insight into how to limit confounding in these situations.

Although DAGs help to identify sources of confounding, they do not provide insight into the magnitude of confounding. Some researchers interested in the magnitude of confounding use a change in estimate approach. This approach compares the unadjusted estimate of effect for an exposure on an outcome with the estimate adjusted for the potential confounder. If the estimate changes by a prespecified amount (for instance, 10%), the factor is judged to be a confounder. Although this approach has the appeal of assessing the magnitude of confounding, it may be misleading. The change in estimate may be the result of introducing rather than removing bias by adjusting for a factor that is not a confounder. For example, adjusting for an intermediate between the exposure and the outcome is not appropriate, but it could change the effect estimated, leading to the false conclusion that the intermediate is a confounder. Therefore, the change in estimate approach should not be used without also considering a DAG. Similarly, it is not appropriate to use the *p*-value to identify potential confounders because the *p*-value also does not consider the relations among the variables and may introduce bias rather than removing it.

DAGs can clarify the likelihood that the unexposed population represents the unobserved experience of the exposed group in the parallel universe where they were unexposed, by identifying non-causal ways that exposure

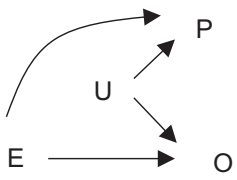

Figure 3. Exposure (E) causes the outcome (O) and causes the same outcome during an earlier pregnancy (P). An unmeasured factor causes the outcome and the outcome of the previous pregnancy. There is no confounding; the path E→P←U→O is blocked because P is a collider. Adjusting for P would introduce bias into the study.

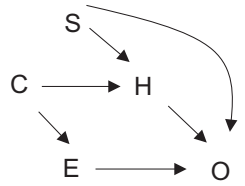

Figure 4. The exposure (E) causes or prevents the outcome (O). A confounding factor (C) affects the exposure and affects the condition H. The condition H affects the outcome. Condition S causes or prevents both H and the outcome. There is a confounding because of the path E←C→H→O. However, the path E←C→H←S→O does not introduce confounding. Adjusting for C prevents confounding. Although adjusting for H would block the confounding path E←C→H→O, it would induce a spurious association between the exposure and the outcome through the path E←C→H←S→O. Adjusting for both H and S would prevent confounding because adjusting for H would block the path E←C→H→O and adjusting for S would block the path opened by adjusting for H (E←C→H←S→O).

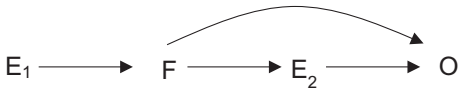

Figure 5. Exposures (E_1) causes condition F, which affects exposure at time 2 (E_2) which affects the outcome (O). Condition F also affects the outcome. F is an intermediate between exposure at time 1 and the outcome, but a confounder of the exposure at time 2 and the outcome.

status could be associated with the outcome. Further, they help identify how to address confounding when it occurs. Avoiding bias from adjusting for variables that are not confounders is as important as identifying and minimizing sources of confounding. DAGs facilitate the identification of both conditions.

Acknowledgments

I thank Olga Basso, Carey Drews-Botsch, Melanie Jacobson, Candice Johnson, and Lisa Shandley for providing feedback on an earlier version of this manuscript. I thank the Whiteley Center for providing a peaceful place to write.

Funding

This study has not been funded by any external sources.

References

1. Howards PP. An overview of confounding. Part 1: the concept and how to address it. Acta Obstet Gynecol Scand. 2018;97:394–399.
2. Rothman KJ, Greenland S, Lash TL. Modern Epidemiology, 3rd edn. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins, 2008.
3. Weinberg CR. Toward a clearer definition of confounding. Am J Epidemiol. 1993;137:1–8.
4. Hernan MA, Hernandez-Diaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. Am J Epidemiol. 2002;155:176–84.
5. Howards PP, Schisterman EF, Poole C, Kaufman JS, Weinberg CR. "Toward a clearer definition of confounding" revisited with directed acyclic graphs. Am J Epidemiol. 2012;176:506–11.
6. Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. Epidemiology. 2003;14:300–6.
7. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology. 1999;10:37–48.
8. Hernan MA, Cole SR. Invited Commentary: causal diagrams and measurement bias. Am J Epidemiol. 2009;170:959–62; discussion 63-4.
9. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. Epidemiology. 2004;15:615–25.
10. Ananth CV, Schisterman EF. Confounding, causality, and confusion: the role of intermediate variables in interpreting observational studies in obstetrics. Am J Obstet Gynecol. 2017;217:167–75.
11. Wilcox AJ, Weinberg CR, Basso O. On the pitfalls of adjusting for gestational age at birth. Am J Epidemiol. 2011;174:1062–8.
12. Cole SR, Hernan MA. Fallibility in estimating direct effects. Int J Epidemiol. 2002;31:163–5.
13. Hernandez-Diaz S, Schisterman EF, Hernan MA. The birth weight "paradox" uncovered? Am J Epidemiol. 2006;164:1115–20.
14. Hernandez-Diaz S, Wilcox AJ, Schisterman EF, Hernan MA. From causal diagrams to birth weight-specific curves of infant mortality. Eur J Epidemiol. 2008;23:163–6.
15. Valeri L, Vanderweele TJ. Mediation analysis allowing for exposure-mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. Psychol Methods. 2013;18:137–50.
16. VanderWeele TJ. Explanation in Causal Inference: methods for Mediation and Interaction. New York: Oxford University Press, 2015.
17. Bodnar LM, Cogswell ME, McDonald T. Have we forgotten the significance of postpartum iron deficiency? Am J Obstet Gynecol. 2005;193:36–44.
18. Kaufman RH, Adam E, Hatch EE, Noller K, Herbst AL, Palmer JR, et al. Continued follow-up of pregnancy outcomes in diethylstilbestrol-exposed offspring. Obstet Gynecol. 2000;96:483–9.
19. Howards PP, Schisterman EF, Heagerty PJ. Potential confounding by exposure history and prior outcomes: an example from perinatal epidemiology. Epidemiology. 2007;18:544–51.
20. Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology. 2000;11:550–60.
21. Nevin NC, Merrett JD. Potato avoidance during pregnancy in women with a previous infant with either anencephaly and/or spina bifida. Br J Prev Soc Med. 1975;29:111–5.