HHS Public Access
Author manuscript
Neurourol Urodyn. Author manuscript; available in PMC 2015 July 22.

Published in final edited form as:
Neurourol Urodyn. 2012 January ; 31(1): 115–120. doi:10.1002/nau.21183.

Reducing Bias in Pelvic Floor Disorders Research: Using Directed Acyclic Graphs as an Aid

Vivian W. Sung*

The Division of Urogynecology and Reconstructive Pelvic Surgery, Department of Obstetrics and Gynecology, Alpert Medical School of Brown University, Providence, Rhode Island

Abstract

Aims—The aim of most pelvic floor disorders (PFD) research is to obtain an unbiased effect estimate and to make causal inferences. New developments in epidemiologic research, including the use of causal directed acyclic graphs (DAGs), have shown that traditional analytical strategies for research can be inadequate, leading to unintended consequences such as introducing additional bias. Although DAGs have been proven to be useful in other medical fields, their use has been limited in PFD research. The aim of this paper is to introduce DAGs and then demonstrate their application in PFD research. This paper will also illustrate how relying purely on statistical techniques can lead to pitfalls in reducing bias in research studies.

Methods/Results—DAGs are a graphical epidemiologic tool that provide a method to select for potential confounders and minimize bias in the design and analysis of research studies. We start by providing an introduction to DAGs. We then describe six scenarios in PFD research in which DAGs can be helpful: (1) identifying appropriate confounding variables for adjustment; (2) identifying potential over-adjustment when conditioning on a mediator; (3) identifying unintended confounding due to inappropriate adjustment; (4) identifying unintended selection bias due to inappropriate adjustment; (5) planning analyses in cross-sectional studies; and (6) using DAGs as a framework to help plan data collection and analyses in PFD research.

Conclusions—We demonstrate how the application of DAGs as an aid to PFD research can help to decrease bias and discuss the insights and implications for study design and analytical approaches.

Keywords

bias (epidemiology); confounding; pelvic floor disorders; incontinence; research methods; women's health

*Correspondence to: Vivian W. Sung, M.D., M.P.H., Division of Urogynecology and Reconstructive Pelvic Surgery, Women and Infants' Hospital/Warren Alpert Medical School at Brown University, 695 Eddy Street, Providence, RI 02903. vsung@wihri.org.

Linda Brubaker led the review process.

Conflict of interest: none.

INTRODUCTION

Most biomedical research is designed to predict an outcome after a treatment is applied or after a harmful exposure is removed. In order to make causal inferences about an exposure, it is critical to be able to determine an unbiased estimate of effect for an exposure on the outcome of interest. This allows clinicians and policy makers to make valid conclusions and decisions influencing the health of individuals and populations.

Even under “ideal” study conditions (large sample size, randomization, 100% follow-up), there are many factors that may lead to inappropriate causal inferences. For example, one important step in avoiding bias is identification of potential confounders. It is not always clear in the design stage of a study which variables should be collected and which may be potential confounders. In the analysis phase, it is often also unclear which variables should be “adjusted for.” Inappropriate adjustment may result in over-adjustment, unnecessary adjustment, and can even introduce bias where none existed.[1]

There are several definitions of confounding bias.[2] One common approach is to consider a covariate to be a confounder if it is associated with the exposure, associated with the outcome, and changes the effect estimate compared to the crude estimate in a statistical model. This method assumes that the data will be what ultimately informs us about confounding. However, relying only on “blind statistical” approaches may be insufficient and may introduce unintended conditional associations.[3–5] This directly contrasts with the goal of eliminating bias.

Causal directed acyclic graphs (DAGs) are diagrams that provide a graphical representation of causal effects between variables. They have been used as an aid in causal inference in the fields of artificial intelligence and robotics[6] and their use in epidemiologic research was extended by Pearl and others.[7–10] DAGs provide a framework to help visualize confounding in a way that is grounded in substantive knowledge. DAGs can aid in variable selection and help complement traditional methods of controlling confounding.[11]

Although DAGs have been proven to be useful tools in epidemiologic research,[12] their application to specific research problems for female pelvic floor disorders (PFDs) is uncommon. The objective of this paper is to provide an introduction to DAGs and illustrate their application using PFD research scenarios. We will demonstrate that qualitative, a priori background knowledge is critical for accurately adjusting for confounders, consideration of biases, and designing appropriate analytic plans. Drawing from prior work by Hernan et al.[12] and others, six PFD scenarios are developed to illustrate the usefulness of DAGs for: (1) identifying appropriate confounding variables for adjustment; (2) identifying unintended consequences of conditioning on a mediator (over-adjustment); (3) identifying unintended confounding due to inappropriate adjustment; (4) identifying unintended selection bias due to inappropriate adjustment; (5) planning analyses in cross-sectional studies; and (6) using DAGs as a framework to identify measured and unmeasured confounders to help plan for data collection and analyses.

The examples provided in this paper are not intended to be exhaustive, and this paper is written from a conceptual and applied point of view targeted at clinical researchers.

Analytical “proof,” theoretical development, and algebraic solutions are summarized in previous literature and are beyond the scope of this paper.[13] Our PFD scenarios and causal diagrams are based on the work of Burgio et al.,[14] Delancey et al.,[15] and Bump and Norton.[16] Alternative DAGs for each scenario can be argued and the DAGs provided are meant only to be used as examples and are not intended to describe all possible causal mechanisms and relationships. When appropriate, additional references are provided for more detailed information regarding the applications of DAGs.

TERMINOLOGY AND BASIC RULES OF CAUSAL DAGs

DAGs include a set of arrows (directed edges) drawn along a timeline which help to characterize causal and temporal relationships between variables (nodes). They can never be a cycle (i.e., they are always acyclic) because we can never go back in time. Thus it is not possible to start at a node, follow the directed edges in the direction of the arrows, and end up back at the same node (self-causation cannot exist). An arrow connecting two variables indicates causation. Bi-directional arrows (the traditional representation of a confounder) are not allowed because temporal relationships are taken into consideration when constructing the DAG. Confounding variables that are “conditioned” on are denoted by a box around the variable (see Fig. 1). “Conditioning” on a variable means that one has used restriction, stratification, and/or adjustment through regression or other statistical methods to examine the association of exposure and outcome within levels of the conditioned variable.

Basic DAG Terminology

A *parent* is defined as a direct cause of a particular variable. An *ancestor* is a direct cause or indirect cause of a particular variable. A *child* is the direct effect of a particular variable. A *descendant* is a direct effect or indirect effect of a particular variable.

A path is a sequence of variables connected by arrows, regardless of the direction of arrowheads. A “back door” path is a path that can be connected “going backwards” against the direction of the arrows between variables. Paths can be “blocked” or “unblocked.” An unblocked path is a path that can be traced between two variables without traversing a pair of arrows that collide “head-to-head” with a third variable, again regardless of the direction of arrowheads (see Fig. 3). An unblocked path typically denotes the presence of a confounder, and can be blocked by conditioning on the confounder. A blocked path is a path between 2 variables that traverses a pair of arrows that collide at a third variable (the third variable is called a collider), or a path that has been blocked by conditioning on a confounder. A DAG with only blocked paths implies bias has been minimized.

SIMPLE SCENARIOS ILLUSTRATING THE USE OF DAGs IN PFD RESEARCH

Example 1: Identification of Confounders

Consider an example in which we want to assess if vaginal delivery causes fecal incontinence. In DAGs confounders must be a common cause (or be a marker for a cause) of both exposure and outcome. Figure 1a demonstrates that age is a direct cause of whether

someone has a vaginal delivery (or alternatively a cesarean delivery) and is also a direct cause of whether someone experiences fecal incontinence. In this model, age is predetermined (or temporally occurs) prior to the exposure (type of delivery) and outcome (fecal incontinence), and is a *common cause* of both the exposure and outcome. Thus age is a confounder in this model and there is an open back-door path from fecal incontinence to age to vaginal delivery (confounding). If we adjust for age, we close the backdoor path and can observe the true association between vaginal delivery and fecal incontinence. We could also condition on a descendant or a surrogate of a confounder. Figure 1b illustrates that menopause status is a descendant of age. Conditioning on menopause status instead of age would also close the backdoor path; however, because menopause is an imprecise measure of age, there would be a risk of residual confounding bias. Methods for conditioning typically include restriction, stratification, and regression.

Example 2: Consequences of Conditioning on a Mediator (Over-Adjustment Bias)

When a covariate is a descendant of the exposure and a cause of the outcome, it is a mediator. A mediator is not a confounder, as confounding only occurs when a covariate *causes* the exposure. A mediating variable lies along the causal pathway between exposure and outcome. We illustrate this in Figure 2a, where anal sphincter laceration is a mediator between vaginal delivery and fecal incontinence.

Controlling for a mediator would only allow one to determine how much of the causal relation between the exposure and outcome was accounted for by the mediator itself, but would not result in the *total* causal effect between exposure and outcome. This is sometimes desirable for researchers interested in decomposing total causal effects into direct and indirect effects. However, if one incorrectly assumes the mediator is a confounder, the interpretation will be inaccurate. This is also known as “over-adjustment bias” and will typically cause a null-biased estimate of the total causal effect.[1,17] Similarly, any adjustment for an ascending or descending proxy (or surrogate measure) of an unmeasured mediator will also cause over-adjustment bias. A “proxy” is a measure that approximates the real measure of interest. Consider Figure 2b, in which episiotomy could be considered an ascending proxy for anal sphincter laceration. Adjustment for episiotomy in this scenario would also result in over-adjustment bias. More information on over-adjustment bias and ascending and descending proxies can be found by Schisterman et al.[1]

Example 3: Consequences of Conditioning on a Collider-Inducing Confounding

When a covariate is a descendant of two other covariates, it is a *common effect*, or collider (two parent arrows “collide” at the node of the descendant). A path that contains a collider is NOT a confounding path and does not generate an association between the two parents. For example, suppose a dataset contains the three variables: street condition, brake condition, and motor vehicle accident. Poor street conditions and poor brake conditions can both cause motor vehicle accidents. Therefore, “accident” is a collider in this model. Street condition and brake condition are not associated with each other and we cannot use one to guess the condition of the other. However, if we stratify on the collider accident, we would erroneously induce an association between street and brake conditions. For example, if we only consider the stratum including those who had an accident, if we knew that street

conditions were good, we would assume that the brake conditions were poor. Stratifying on "accident" induces a false association between the two parent variables.

Figure 3a includes a simple DAG illustrating that pelvic floor muscle impairment is a common effect (collider) of vaginal delivery and menopausal status. Knowing someone's menopausal status does not provide any information about someone's history of vaginal delivery. The presence of a collider does not induce an association between its causes. However, conditioning on a collider can turn that path (erroneously) into a confounding path (denoted by a dotted line between the parent covariates). In the case of Figure 3a, there is no other variable to control that would remove the bias (close the path) caused by conditioning on pelvic floor muscle impairment. Therefore, adjustment for this collider is not only unnecessary but is harmful.[11]

An extension of this is shown in Figure 3b, where anal sphincter injury is the exposure and pelvic organ prolapse (POP) is the outcome. In this example, there is no true relationship between anal sphincter injury and POP and the crude association represents the true association. However, if we condition on pelvic floor muscle impairment, a collider, we open a backdoor path between anal sphincter injury and pelvic prolapse through the induced association between vaginal delivery and menopause status. But in reality, there is no true relationship between menopause status and vaginal delivery. If we condition, stratify or adjust for pelvic floor muscle impairment, we will induce confounding and the resulting effect estimate will be biased. In order to then get an unbiased estimate, we would need to condition on another variable to block the new path (either vaginal delivery or menopausal status). This highlights the point that unnecessary adjustment for variables may necessitate adjustment for even more variables.[9] A more intuitive example of the effects of conditioning on a collider can be found in a paper by Cole et al.[5]

Example 4: Consequences of Conditioning on a Collider-Inducing Selection Bias

Similar to collider-induced confounding, selection bias (also called collider-stratification bias) can also be caused by conditioning on a collider, but under some different circumstances.[5] In this case, the collider is a binary variable that indicates whether a person could have belonged to our sample or not. Conditioning on this variable results in computing a measure of effect for only one stratum of the collider and not both.[3]

Figure 4 illustrates the consequence of conditioning on a sampling variable (collider), "help-seeking for PFDs." In this example, our exposure is income and outcome is PFDs. There is no true relationship between income and PFDs and the crude estimate is unbiased. However, if we stratify our analyses based on women who did and did not seek help, we would induce bias by opening the backdoor path between education level and age, even though no causal relation exists between these two covariates. This would ultimately lead to a biased estimate between income and PFDs. In this specific example, conditioning on education level or age, in addition to our collider, would block the path and allow us to estimate the effect of income on PFDs in a sample of women seeking care for PFDs. A similar inadvertent mistake can be made if we restrict our analysis to only part of a sample based on a sampling collider.

Other common causes of selection bias occur when investigators condition on participants' willingness to participate, Berkson's bias (conditioning on whether someone is admitted to the hospital), and losses to follow up or missing data (conditioning on presence of data).[4] While beyond the scope of this article, some alternatives to adjusting by stratification that help to avoid creating additional bias include inverse weighting and g-analysis.[3]

Example 5: DAGs in Helping to Plan Analyses in Cross-Sectional Studies

In a cross-sectional study, information about exposure and outcome are collected at the same time and therefore, it is generally not possible to assess temporal sequences without strong a priori assumptions. In Figure 5a, we have an example of a cross-sectional study conducted to assess the effect of performing Kegel exercises on preventing urinary incontinence. However, it is possible that once a woman experienced symptoms or was diagnosed with urinary incontinence that she subsequently began to perform Kegel exercises as treatment. In a cross-sectional design, it may not be possible to distinguish between the "preventive" and "treatment" Kegel variables. If we had information on a variable that we were confident would only come after incontinence symptoms arose, such as referral to a specialist or treatment for symptoms, we could block the path and our effect estimate between Kegels and prevention of urinary incontinence would be closer to what was originally conceptualized (Fig. 5b).

Example 6: Using DAGs as a Framework to Identify Measured and Unmeasured Confounders and Plan Data Collection and Analyses

Suppose we are interested in whether vaginal delivery causes POP in a longitudinal study. Figure 6a represents a possible DAG framework for the effect of vaginal delivery on POP. In this DAG, genetic composition and pelvic floor injury are unmeasured. Because pelvic floor injury is a mediator between vaginal delivery and POP, there is no need to condition on this variable. In this DAG, there are four open backdoor paths between the outcome, POP and the exposure, vaginal delivery:

POP ← age → vaginal delivery.

POP ← obesity ← age → vaginal delivery.

POP ← menopause ← age → vaginal delivery.

POP ← race ← genetic composition → vaginal delivery.

Conditioning on the covariate age will close backdoor paths #1, #2, and #3. Because genetic composition is unmeasured, we cannot condition on this covariate to close path #4. However, conditioning on race, a descendant of genetic composition, would close backdoor path #4 assuming there are no alternative paths between genetic composition and POP. Therefore, in this model, a sufficient set of S variables for adjustment could include age and race. Additional reading on sufficient sets of variables for adjustment is available.[9,11] In theory, there would be no need to condition on obesity, menopause, or genetic composition in addition to age and race to get the causal effect of vaginal delivery on POP. One caveat the researcher should note is that whenever one adjusts for a confounder using a surrogate or proxy measure, residual confounding can still remain. Therefore, race in this model is only a surrogate for genetic composition and likely will not account for all relevant genetic factors,

resulting in some residual confounding. As such, DAGs are also useful for visualizing where residual confounding may still occur.

It is important to note in this model, if age or race were unmeasured we would not be able to obtain an unconfounded effect estimate between vaginal delivery and POP. Using the DAG allows us to identify measured and unmeasured confounders, which is helpful in both the planning and analytical phases. Details on a pragmatic six-step approach to using DAGs to help plan clinical research and minimize bias can be found by Shrier et al.[18]

DISCUSSION

The primary contribution of this paper is to illustrate how recent advances in epidemiologic research and the use of DAGs can help minimize bias in PFD research. We have provided six scenarios to illustrate why a priori knowledge and assumptions are critical for both study planning and analysis. We have also discussed scenarios in which adjusting for particular covariates can be harmful and can threaten the validity of a study. We have illustrated why relying purely on statistical techniques can lead to pitfalls and increase bias. A better approach is to use a combination of methods that can help to minimize bias: for example, using our a priori knowledge to develop a DAG and combining this with traditional statistical approaches. There are also additional extensions of DAGs that were not described in this paper including the use of DAGs in understanding information bias,[12] unnecessary adjustment of confounders,[1] and misclassification.[19]

There are limitations to DAGs. The DAGs presented in this paper are based on the assumptions that outcomes and exposures are binary variables, there is conditional independence, and that the exposures have monotonic effects. Also, in reality it is difficult to know the true underlying causal structure of a disease or health condition, including PFDs. We probably would not need to study it if we knew it. DAGs should be developed within the basic, epidemiologic, and clinical evidence available and it is obvious that collaboration between different disciplines and research areas within PFDs is needed for this to be successful. DAGs highlight the point that basic science and an understanding of disease mechanisms are necessary for appropriate epidemiologic studies. In addition, a DAG that is appropriate for one exposure and outcome may not be appropriate for a different exposure and the same outcome.

It is possible that even under the best circumstances, it will be unclear which DAG is most appropriate for a specific study. In this situation, some epidemiologists would recommend analytical approaches that utilize different DAGs (sensitivity analyses). DAGs cannot encode parametric assumptions. Finally, DAGs as we have discussed in this paper do not include effect modifiers or covariates that interact with other covariates. DAGs also do not eliminate or reduce other sources of bias, such as measurement bias and are not a statistical technique that will yield an effect estimate.

Despite these limitations, many research questions in PFDs can be usefully represented by DAGs, which can provide a useful framework. The use of DAGs will allow users of

traditional stratification and regression statistical techniques to reduce the magnitude of bias in their studies.

CONCLUSION

We have used simple examples in PFDs to illustrate the insights that DAGs can add to the investigation and analysis of research in this field. Decisions of adjusting for confounding should be based on sound knowledge and assumptions of underlying relationships between exposures, outcomes, and covariates. Although DAGs cannot address all possible sources of bias, their use would likely improve the scientific rigor of research in this field and reduce the magnitude of bias, especially in observational and epidemiologic research in female PFDs.

Acknowledgments

The author would like to thank Drs. Stephen L. Buka and Angela D. Paradis for reviewing the manuscript and providing valuable comments. Dr. Sung is supported by grant K23HD060665 from the *Eunice Kennedy Shriver* National Institute of Child Health and Human Development. The content is solely the responsibility of the author and does not necessarily represent the official views of the Eunice Kennedy Shriver National Institute of Child Health and Human Development or the National Institutes of Health.

References

1. Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiology. 2009; 20:488–95. [PubMed: 19525685]
2. Greenland S, Morgenstern H. Confounding in health research. Annu Rev Public Health. 2001; 22:189–212. [PubMed: 11274518]
3. Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. Epidemiology. 2004; 15:615–25. [PubMed: 15308962]
4. Greenland S. Quantifying biases in causal models: Classical confounding vs collider-stratification bias. Epidemiology. 2003; 14:300–6. [PubMed: 12859030]
5. Cole SR, et al. Illustrating bias due to conditioning on a collider. Int J Epidemiol. 39:417–20. [PubMed: 19926667]
6. Spiegelhalter DJ, et al. Bayesian analysis in expert systems (with discussion). Stat Sci. 1993; 8:219–83.
7. Robins J. A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods. J Chronic Dis. 1987; 40:139S–61S. [PubMed: 3667861]
8. Robins JM. Data, design, and background knowledge in etiologic inference. Epidemiology. 2001; 12:313–20. [PubMed: 11338312]
9. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. Epidemiology. 1999; 10:37–48. [PubMed: 9888278]
10. Glymour, MM.; Greenland, S. Causal diagrams. In: Rothman, KJ.; Greenland, S.; Lash, TL., editors. Modern epidemiology. Vol. 3. Philadelphia: Lippencott-Raven Publishers; 2008.
11. VanderWeele TJ, Robins JM. Directed acyclic graphs sufficient causes, and the properties of conditioning on a common effect. Am J Epidemiol. 2007; 166:1096–104. [PubMed: 17702973]
12. Hernan MA, et al. Causal knowledge as a prerequisite for confounding evaluation: An application to birth defects epidemiology. Am J Epidemiol. 2002; 155:176–84. [PubMed: 11790682]
13. Pearl, J. Causality: Models, reasoning and inference. Cambridge: University of Cambridge; 2000.
14. Burgio KL, et al. Risk factors for fecal and urinary incontinence after childbirth: The childbirth and pelvic symptoms study. Am J Gastroenterol. 2007; 102:1998–2004. [PubMed: 17573795]
15. Delancey JO, et al. Graphic integration of causal factors of pelvic floor disorders: An integrated life span model. Am J Obstet Gynecol. 2008; 199:610e1–5. [PubMed: 18533115]

16. Bump RC, Norton PA. Epidemiology and natural history of pelvic floor dysfunction. Obstet Gynecol Clin North Am. 1998; 25:723–46. [PubMed: 9921553]
17. VanderWeele TJ. On the relative nature of overadjustment and unnecessary adjustment. Epidemiology. 2009; 20:496–9. [PubMed: 19525686]
18. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. BMC Med Res Methodol. 2008; 8:70. [PubMed: 18973665]
19. VanderWeele TJ, Robins JM. Four types of effect modification: A classification based on directed acyclic graphs. Epidemiology. 2007; 18:561–8. [PubMed: 17700242]

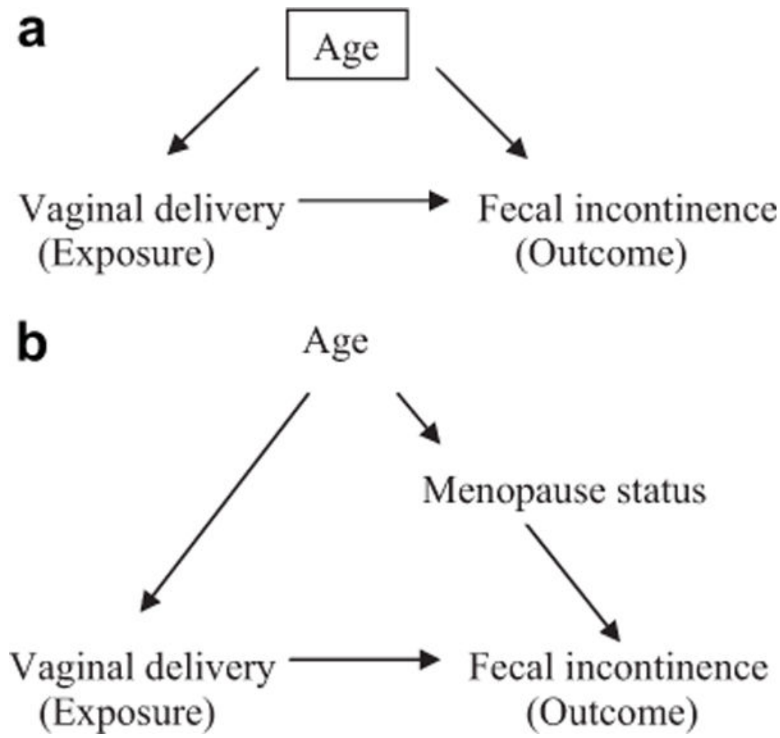

Fig. 1.

DAG illustrating age as a confounder for effect of vaginal delivery on fecal incontinence. **a**: DAG representation of confounder. Uni-directional arrows imply that covariate needs to be a common cause of exposure and outcome. To obtain an unbiased effect estimate of vaginal delivery on fecal incontinence, we would need to condition on the covariate age (conditioning denoted by box around covariate). **b**: DAG representation of descendant or surrogate of age as a confounder. To obtain an unbiased effect estimate, we would need to condition on either age or menopause status, but do not need to condition on both.

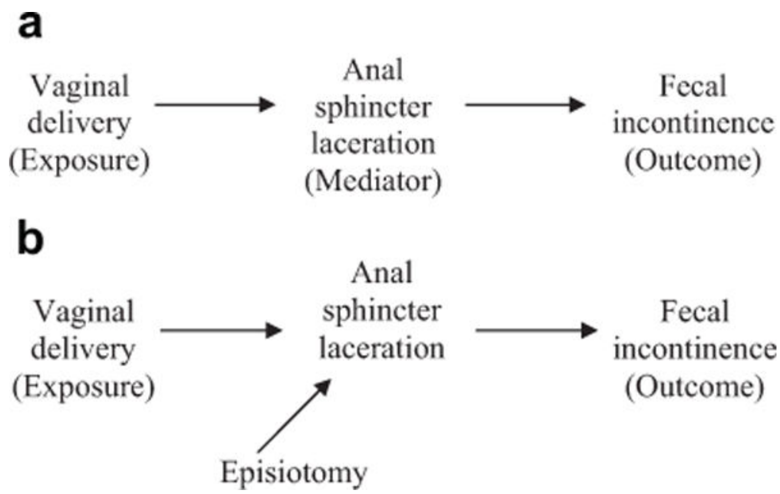

Fig. 2.

DAG illustrating over-adjustment bias. **a**: Conditioning on a mediator. Conditioning on anal sphincter laceration, a mediator, would result in an null-biased effect estimate that would only partially reflect the true association between vaginal delivery and fecal incontinence. **b**: Conditioning on an ascending proxy for mediator. Conditioning on episiotomy would also result in an effect estimate that would only partially reflect the true association between vaginal delivery and fecal incontinence.

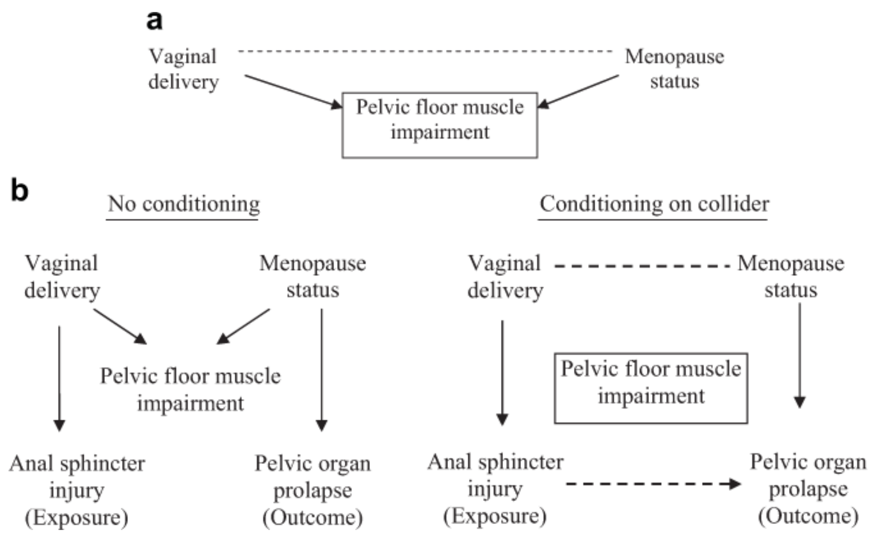

Fig. 3.

DAG illustrating collider (common effect). **a**: Simple DAG illustrating pelvic floor muscle impairment as a collider. Whether someone is menopausal is unrelated (not associated) to whether she had a vaginal delivery, but pelvic floor muscle impairment is a *common effect* of both parent covariates. Conditioning on a collider will induce an association between the two parent covariates (dotted line). **b**: DAG illustrating new confounding path due to conditioning on a collider.

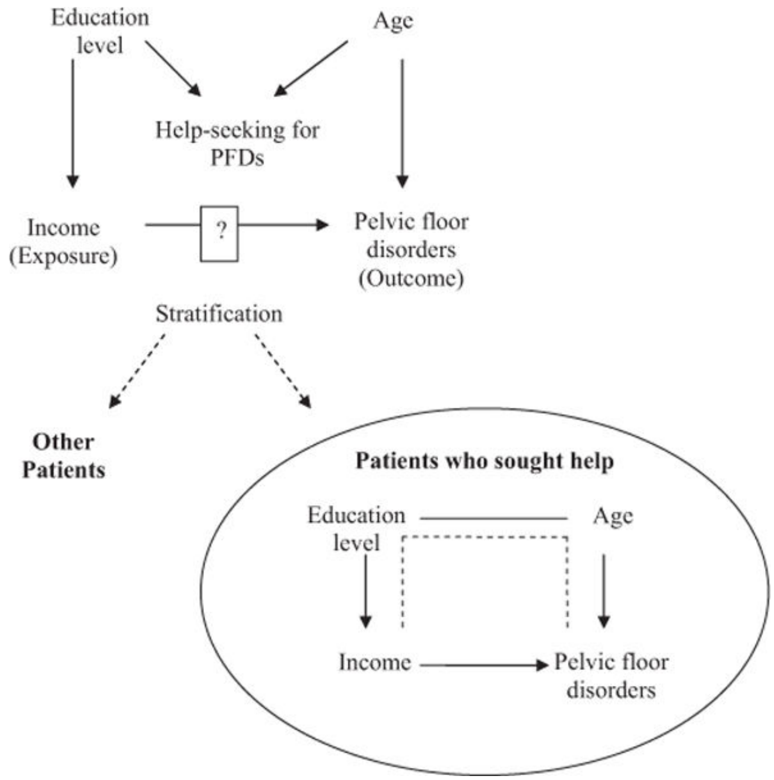

Fig. 4.

Selection bias due to conditioning on a collider. Education level and age are both causes of whether someone may seek care for PFDs, but they are not related to each other. Stratifying patients by help-seeking (a collider) will induce an association between education level and age, and open a confounding path between income and PFDs.

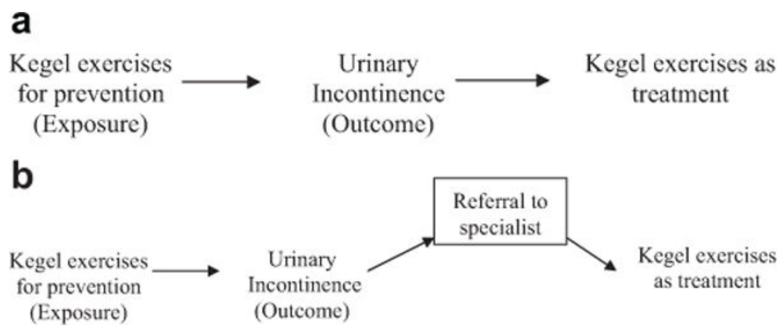

Fig. 5.

Using DAGs in analyzing cross-sectional studies. **a**: Cross-sectional study evaluating effect of Kegel exercises for prevention of urinary incontinence. Here, Kegel exercises can be a cause or effect of the variable urinary incontinence. **b**: By adjusting for referral to a specialist, we block the path from urinary incontinence to kegel exercises as treatment. The remaining path would be closer to our question of interest, evaluating Kegel exercises as a preventive measure for urinary incontinence.

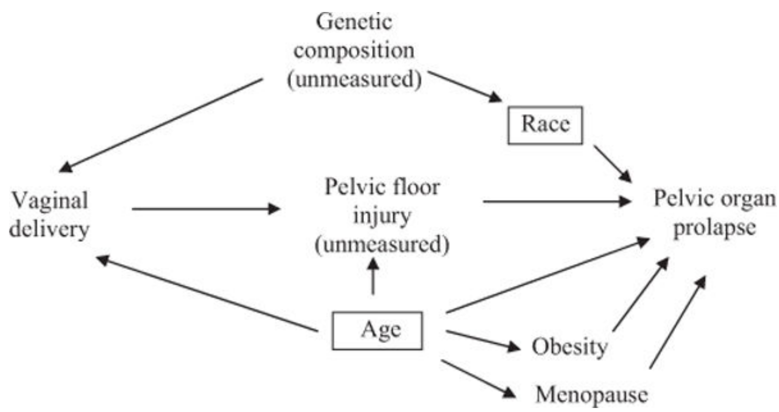

Fig. 6.

DAGs as an aid to help plan for data collection and analysis of a study. Based on this DAG, conditioning on Age and Race would be sufficient to close all backdoor paths and obtain an unconfounded effect estimate between vaginal delivery and pelvic organ prolapse.