Published in final edited form as:

Pharmacoepidemiol Drug Saf. 2013 November ; 22(11): 1139–1145. doi:10.1002/pds.3506.

A Review of Covariate Selection for Nonexperimental Comparative Effectiveness Research

Brian C. Sauer, PhD,

Salt Lake City Veteran's Affairs Career Development Awardee, Assistant Research Professor,
Division of Epidemiology, Department of Internal Medicine, University of Utah

Alan Brookhart, PhD, Jason Roy, PhD, and

Address: Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, 618
Blockley Hall, 423 Guardian Drive, Philadelphia, PA 19104, Tele: 215-746-4225, Fax:
215-573-1050, jaroy@upenn.edu

Tyler Vanderweele, PhD

Abstract

This paper addresses strategies for selecting variables for adjustment in non-experimental comparative effectiveness research (CER), and uses causal graphs to illustrate the causal network that relates treatment to outcome. Variables in the causal network take on multiple structural forms. Adjustment for on a common cause pathway between treatment and outcome can remove confounding, while adjustment for other structural types may increase bias. For this reason variable selection would ideally be based on an understanding of the causal network; however, the true causal network is rarely known. Therefore, we describe more practical variable selection approaches based on background knowledge when the causal structure is only partially known. These approaches include adjustment for all observed pretreatment variables thought to have some connection to the outcome, all known risk factors for the outcome, and all direct causes of the treatment or the outcome. Empirical approaches, such as forward and backward selection and automatic high-dimensional proxy adjustment, are also discussed. As there is a continuum between knowing and not knowing the causal, structural relations of variables, we recommend addressing variable selection in a practical way that involves a combination of background knowledge and empirical selection and that uses the high-dimensional approaches. This empirical approach can be used to select from a set of *a priori* variables based on the researcher's knowledge to be included in the final analysis or to identify additional variables for consideration. This more limited use of empirically-derived variables may reduce confounding while simultaneously reducing the risk of including variables that may increase bias.

Keywords

covariate selection; confounding; comparative effectiveness research; propensity scores;
backwards selection; stepwise selection; bias; nonexperimental methods

INTRODUCTION

Non-experimental studies that compare the effectiveness of treatments are often strongly affected by confounding. For example, consider two drugs used to treat hypertension – calcium channel blockers (CCB) and diuretics. Since CCBs are perceived by many clinicians as being particularly useful in treating high-risk patients with hypertension, patients with a higher risk for experiencing cardiovascular events are more likely to be channeled into the CCB group, thus confounding the relation between antihypertensive treatment and the clinical outcomes of cardiovascular events.¹ The difference in treatment groups is a result of the differing baseline risk for the outcome and the treatment effects (if any). Any attempt to compare the causal effects of CCBs and diuretics on cardiovascular events would require taking patients' underlying risk for cardiovascular events into account through some form of covariate adjustment. The use of statistical methods to make the two treatment groups similar with respect to measured confounders is sometimes called statistical adjustment, control, or conditioning.

The purpose of this paper is to address the complex issue of selecting variables for adjustment in order to compare the causative effects of treatments. Recommendations for variable selection in this paper focus primarily on fixed treatment comparisons when employing the so-called “incident user design”. We present variable selection approaches based on full and partial knowledge of the data generating process as represented in causal graphs. We also discuss approaches to selecting covariates from a high-dimensional set of variables based on statistical association, and how these approaches may be used to complement variable selection based on background knowledge.

CAUSAL MODELS AND VARIABLE TYPES

Treatment effects

The goal of comparative effectiveness research (CER) is to determine if a treatment is more effective or safer than another. Treatments should be “well-defined” in that they should characterize and operationalize exposure in observational CER, and represent manipulable units, e.g., drug treatments. With that as background, we focus in this paper on which types of variables should be controlled for and how to identify them. We begin with a brief description of different types of variables (defined by their relationship with treatment, outcome, and other variables).

Risk factors

Covariates that are predictive of the outcome but having no influence on treatment status are often referred to as pure risk factors. Conditioning on such risk factors is unnecessary to remove bias but can result in efficiency gains in estimation^{2,3} and does not induce bias in regression or propensity score models.⁴ Researchers need to avoid including variables affected by the outcome, as adjustment for such variables can increase the risk of bias.² We recommend including risk factors in statistical models to increase the efficiency/precision of an estimated treatment effect without increasing bias.⁴

Confounding

The central threat to the validity of non-experimental CER is confounding. Due to the ways in which providers and patients choose treatments, the treatment groups may not have similar underlying risk for the outcome, and confounding is often illustrated as a common cause pathway between the treatment and outcome. Measured variables that (1) influence treatment assignment (2) are predictive of the outcome, and (3) remove confounding when adjusted for, are often called confounders. Unmeasured variables on a common cause pathway between treatment and outcome are referred to as unmeasured confounders. For example, in Figure 1 unmeasured variables $U1$ and $U2$ are causes of treatment assignment (A_0) and outcome (Y_1). In general, sources of confounding in observational comparative effectiveness studies include provider actions, patient actions, and social and environmental factors. Here, unmeasured variable $U1$ has a measured confounder C_0 that is a proxy for $U1$, such that conditioning on C_0 removes confounding by $U1$, while the unmeasured variable $U2$ does not.

Intermediate variables

An intermediate variable is generally thought of as a post-treatment variable influenced by treatment that may or may not lie on the causal pathway between the treatment and the outcome. When the goal of CER is to estimate the total causal effect of the treatment on the outcome, adjustment for variables on the causal pathway between treatment and outcome is unnecessary and likely to induce bias² toward a relative risk of 1.0, though the direction can sometimes be in the opposite direction. The ability to adequately adjust for baseline confounders and not intermediate variables is one reason the incident user design is so highly valued in pharmacoepidemiology.

Time-varying confounding

The intention-to-treat analogue of a randomized trial, where subjects are assigned to the treatment they are first exposed to regardless of discontinuation or switching treatments, may not be the optimal design for all non-experimental CER. For example, if we were interested in comparing cardiovascular events between subjects who were completely adherent to CCBs versus completely adherent to diuretics, then we may consider a time-varying treatment design where subjects are censored when they discontinue the treatment they were first assigned. If joint predictors of compliance and the outcome are present, then some sort of adjustment for the time-varying predictors must be made. Standard adjustment methods may not produce unbiased effects when the predictors of adherence and the outcome are affected by prior adherence, and a newer class of causal effect estimators, such as inverse-probability of treatment weights or g-estimation, may be warranted.⁵

Collider variables

Colliders are the result of two independent causes having a common effect. When we include a common effect of two independent causes in our statistical model, the previously independent causes become associated, thus opening a backdoor path between the treatment and outcome (see Figure 2).

Bias resulting from conditioning on a collider when attempting to remove confounding by covariate adjustment is referred to as *M*-collider bias.⁶ Pure pretreatment *M*-type structures that statistically behave like confounders may be rare; nevertheless, anytime we condition on a variable that is not a direct cause of either the treatment or outcome but is merely associated with the two, we have the potential to introduce *M*-bias.⁷ However, we currently lack many compelling examples of pure *M*-bias for pretreatment covariates.

Instrumental variables

An instrumental variable is a pretreatment variable that is a cause of treatment but has no causal association with the outcome other than through its effect on treatment, such as Z_0 in Figure 3. It has been established that inclusion of variables strongly associated with treatment (A_0) but not independently associated with the outcome (Y_1) in statistical models will increase the standard error and decrease the precision of the treatment effect.^{2,4,8,9} It is less well known, however, that the inclusion of such instrumental variables into statistical models intended to remove confounding can increase the bias of an estimated treatment effect. The bias produced by the inclusion of such variables has been termed “*Z*-bias” as *Z* is often used to denote an instrumental variable.

Z-bias arises when the variable set is insufficient to remove all confounding, and for this reason *Z*-bias has been described as bias-amplification.^{10,11} Figure 3 illustrates a data generating process where unmeasured confounding exists along with an instrumental variable. In this situation, the variation in treatment (A_0) can be partitioned into 3 components: the variation explained by the instrument (Z_0), the variation explained by U_1 and the unexplained variation. The magnitude of unmeasured confounding is determined by the proportion of variation explained by U_1 , along with the association between U_1 and Y_1 . When Z_0 is statistically adjusted, one source of variation in A_0 is removed, making the variation explained by U_1 a larger proportion of the remaining variation. This is what amplifies the residual confounding bias.¹²

We have discussed bias amplification due to adjusting for instrumental variables. The use of instrumental variables, however, can be employed as an alternative strategy to deal with unmeasured confounding.¹³

SELECTION OF VARIABLES TO CONTROL CONFOUNDING

We have presented multiple types of variable structures, with a focus on variables that either remove or increase bias when adjusted. The dilemma is that many of these variable types statistically behave like confounders, which are the only structural type needing adjustment to estimate the average causal effect of treatment.^{14,15}

We present two general approaches to selecting variables to control confounding in non-experimental CER. The first approach selects variables based on background knowledge about the relationship of the variable to treatment and outcome, and the second approach relies primarily on statistical associations to select variables for control of confounding and can be described as employing high dimensional automatic variable selection techniques.

Variable selection based on background knowledge

Causal Graph Theory—Assuming a well-defined fixed treatment employing an intention-to-treat paradigm and no set of covariates predicts treatment assignment with 100% accuracy, control of confounding is all that is needed to estimate causal effects with non-experimental data.^{14,15} The problem, as described above, is that colliders, intermediate variables, and instruments can all statistically behave like confounders. This dilemma has led many influential epidemiologists to take a strong position for selecting variables for control based on background knowledge of the causal structure connecting treatment to outcome.^{16–20}

When sufficient knowledge is available to construct a causal graph, analyzing the structural basis for evaluating confounding is the most robust approach to selecting variables for adjustment. The goal is to use the graph to identify a sufficient set of variables to achieve unconfoundedness, sometimes also called conditional exchangeability.²¹ If the graph is correct, then the “backdoor criterion” can be used to identify a sufficient set of covariates (C) for estimating an effect of treatment (A_0) on the outcome (Y_1).

Although it is quite technical, causal graph theory has formalized the theoretical justification for variable selection, added precision to our understanding of bias due to under and over adjustment, and unveiled problems with historical notions of statistical confounding. The main limitation of causal graph theory is that it presumes that the causal network is known. In most observational pharmacoepidemiologic studies, such complete knowledge of causal networks is unlikely.^{22,23}

Since we rarely know the true causal network, investigators have proposed more practical variable selection approaches based on background knowledge when the causal structure is only partially known. These strategies include adjusting for all observed pretreatment variables thought to have some connection to the outcome²⁴, all known risk factors for the outcome^{4,12,25}, and all direct causes of the treatment or the outcome.²³

Adjustment for all observed pretreatment covariates—Emphasis is often placed on the treatment assignment mechanism and on trying to reconstruct the hypothetical broken randomized experiment that led to the observational data.²⁴ Propensity score methods are often employed for this purpose, and can be used in pharmacoepidemiology to statistically control large numbers of variables when outcomes are infrequent.^{26,27} The probability of treatment is estimated conditional on a set of covariates and the predicted probability is then used as a balancing score or matching variable across treatment groups to estimate the treatment effect. Importance is often placed on balancing all pretreatment covariates. However, when such attempts are made, regardless of their structural form, biases--e.g., from including strong instruments and colliders--can result.^{7,23,28}

Adjustment for all possible risk factors for the outcome—Confounding pathways require common cause structures between the outcome and treatment. A common strategy for removing confounding without incidentally including strong instruments and colliders is to only include variables thought to be direct causes of the outcome, i.e., risk factors, in propensity score models.^{4,25,29} This approach only requires background knowledge of

causes of the outcome and does not require an understanding of the treatment assignment mechanism or how variables that influence treatment are related to risk factors for the outcome. This strategy, however, may fail to include measured variables that predict treatment assignment but have an unmeasured ancestor that is an outcome risk factor ($A_0 \leftarrow C_0 \leftarrow UI \rightarrow Y_I$), as illustrated in Figure 1.²³

Disjunctive cause criterion—The main practical use of causal graphs is to avoid inconsistencies between beliefs and data analysis by not adjusting for known instruments and colliders.¹⁷ Thus, in practice, one only needs to partly know the causal structure of variables relating treatment to the outcome. The disjunctive cause criterion is a formal statement of the conditions in which variable selection based on partial knowledge of the causal structure can remove confounding.²³ It states that all observed variables that are a cause of treatment, or a cause of outcome, or a cause of both should be included for statistical adjustment. When any subset of observed variables is sufficient to control confounding, the set obtained by applying the disjunctive cause criteria will also constitute a sufficient set. This approach requires more knowledge of the variables' relationship to the treatment and outcome than the other approaches based on background knowledge. The approach performs well when a sufficient set of variables is measured, but presents problems when unmeasured confounding remains.

When unmeasured confounding remains, strong arguments exist for error on the side of over-adjustment (adjusting for instruments and colliders) rather than failing to adjust for measured confounders (under-adjustment).^{6,12} Nevertheless, adjustments for instrumental variables have been found to amplify bias in practice.

Empirical variable selection approaches—Historically, data collection for non-experimental studies was primarily collected prospectively, and thoughtful planning was needed to ensure complete measurement of all important study variables. We now live in an era where every interaction between the patient and the healthcare system produces hundreds, if not thousands, of data points that are recorded for clinical and administrative purposes.³⁰ These large multi-use data sources are highly dimensional in that every disease, medication, laboratory result and procedure code, along with any electronically accessible narrative statements, can be treated as variables.

The new challenge to the researcher is to select a set of variables from this high-dimensional space that characterizes the patient's baseline status at the time of treatment selection to enable identification of causal effects or, at least, to produce the least biased estimates.

Forward and backward selection procedures—When using traditional regression it is not uncommon to use, for the purposes of covariate selection, what are sometimes called forward and backward selection procedures. Forward selection procedures begin with an empty set of covariates and then consider whether, for each covariate, the covariate is associated with the outcome conditional on treatment (usually using a p-value cut-off in a regression model of 0.05 or 0.10). The variable that is most strongly associated with outcome (based on having the smallest p-value below the cut-off) is then added to the collection of variables for which control will be made. The process repeats until all

remaining covariates are independent of the outcome conditional on the treatment and the covariates that have been previously selected for control. Backward selection is similar, but instead begins with all covariates in the model and removes one at a time based on the p-value.

Provided that the original set of covariates with which one begins suffices for unconfoundedness of treatment effects estimates, then if the backward selection process correctly discards variables that are independent of the outcome conditional on the treatment and other covariates, the final set of covariates will also yield a set of covariates that suffices for conditional exchangeability.²³ Likewise, under an additional assumption of “faithfulness,”²³ the forward selection procedure will identify a set of covariates that suffices for unconfoundedness, provided that the original set of covariates with which one begins suffices to achieve unconfoundedness and that the process correctly identifies the variables that are and are not independent of the outcome conditional on the treatment and other covariates. The forward and backward procedures can thus be useful for covariate reduction, but both of them suffer from the need to specify a set of covariates to begin with that suffice for unconfoundedness. Thus, even if an investigator intends to employ forward or backward selection procedures for covariate reduction, other approaches will be needed to decide on what set of covariates these forward and backward procedures should begin with. Variable selection procedures also suffer from the fact that estimates about treatment effects are made after having already used the data to decide on covariates.

Similar but more sophisticated approaches using machine learning algorithms such as boosting, random forest, and other ensemble methods have become increasingly common, as have sparsity based methods such as LASSO, in dealing with high-dimensional data. All of these empirically driven methods are, however, limited in that they are in general unable to distinguish between instruments, colliders, and intermediates on the one hand and genuine confounders on the other.

Automatic high-dimensional “proxy” adjustment—In an attempt to capture important proxies for unmeasured confounders, Schneeweiss et al. proposed an algorithm that creates a very large set of empirically-defined variables from healthcare utilization data.²² The created variables capture the frequency of codes for procedures, diagnoses, and medication fills during a pre-exposure period. The variables created by the algorithm are required to have a minimum prevalence in the source population and to have some marginal association with both treatment and outcome. After they are defined, the variables can be entered into a propensity score model. In several example studies where the true effect of a treatment was approximately known from randomized controlled trials, the algorithm appeared to perform as well as, or better, than approaches based on simply adjusting for an a priori set of variables.³¹ By defining variables prior to treatment, propensity score methods will not “over-adjust” by including causal intermediates. Using statistical associations to select potential confounders can result in selection and adjustment of colliders and instruments. Therefore, the analyst should attempt to remove such variables from the set of identified variables.

A practical approach combining causal analysis with empirical selection—

There is a continuum between knowing and not knowing the causal, structural relations of variables. We suggest that a practical approach to variable selection may involve a combination of (1) *a priori* variable selection based on the researcher's knowledge of causal relationships together with (2) empirical selection using the high-dimensional approach described above. The empirical approach could be used to select from a set of *a priori* variables based on the researcher's knowledge and will ultimately select those to be included in the analysis. This more limited use of empirically-derived variables may reduce confounding while simultaneously reducing the risk of including variables that could increase bias.

CONCLUSION

In practice, the particular approach that one adopts for observational research will depend on the researcher's knowledge, the data quality, and the number of covariates. A deep understanding of the specific clinical and public health risks and opportunities that lie behind the research question often drives these decisions.

Regardless of the strategy employed, researchers should clearly describe how variables are measured and provide a rationale for *a priori* selection of potential confounders, ideally in the form of a causal graph. If the researchers decide to augment the model by using an empiric variable selection technique, then they should present both models and describe how the additional variables were measured and selected. Researchers should consider whether or not they believe adequate measurement is available in the dataset when employing a specific variable selection strategy. In addition, all variables included for adjustment should be listed in the manuscript or final report. When empiric selection procedures are newly developed or modified, researchers are encouraged to make the protocol and code publicly available to improve transparency and reproducibility.

However, even when using the methods we describe in this paper, confounding can persist. In such case sensitivity analysis techniques are useful for assessing residual confounding resulting from unmeasured and imperfectly measured variables.^{32–40} Sensitivity analysis techniques assess the extent to which an unmeasured variable would have to be related to the treatment and outcome of interest in order to substantially change the conclusions drawn about causal effects.

Reference

1. Huse DM, Roht LH, Hartz SC. Selective use of calcium channel blockers to treat high-risk hypertensive patients. *Pharmacoepidemiol Drug Saf.* 2000; 9(1):1–9. [PubMed: 19025797]
2. Schisterman EF, Cole SR, Platt RW. Over adjustment bias and unnecessary adjustment in epidemiologic studies. *Epidemiol.* 2009; 20(4):488–495.
3. Robins JM, Greenland S. The role of model selection in causal inference from nonexperimental data. *Am J Epidemiol.* 1986; 123(3):392–402. [PubMed: 3946386]
4. Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Stürmer T. Variable selection for propensity score models. *Am J Epidemiol.* 2006; 163(12):1149–1156. [PubMed: 16624967]

5. Cain LE, Cole SR. Inverse probability-of-censoring weights for the correction of time-varying noncompliance in the effect of randomized highly active antiretroviral therapy on incident AIDS or death. *Stat Med*. 2009; 28(12):1725–1738. [PubMed: 19347843]
6. Greenland S. Quantifying biases in causal models: classical confounding vs collider-stratification bias. *Epidemiol*. 2003; 14(3):300–306.
7. Pearl, J. Myth, confusion, and science of causal analysis [Unpublished Manuscript]. Los Angeles, CA: University of California; 2009. Available at: http://ftp.cs.ucla.edu/pub/stat_ser/r348-warning.pdf. [Accessed March 29, 2012]
8. Robinson LD, Jewell NP. Covariate adjustment. *Biometrics*. 1991; 47(1):342–343. [PubMed: 2049511]
9. Austin PC. The performance of different propensity score methods for estimating marginal odds ratios. *Stat Med*. 2007; 26(16):3078–3094. [PubMed: 17187347]
10. Pearl J. Invited commentary: understanding bias amplification. *Am J Epidemiol*. 2011; 174(11):1223–1227. [PubMed: 22034488]
11. Wooldridge, J. Should instrumental variables be used as matching variables? [Unpublished Manuscript]. East Lansing, MI: Michigan State University; 2009. Available at: <https://www.msu.edu/~ec/faculty/wooldridge/current%20research/treat1r6.pdf>. [Accessed March 29, 2012]
12. Myers JA, Rassen JA, Gagne JJ, et al. Effects of adjusting for instrumental variables on bias and precision of effect estimates. *Am J Epidemiol*. 2011; 174(11):1213–1222. [PubMed: 22025356]
13. Angrist JG, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *J Am Stat Assoc*. 1996; 91:28.
14. Rubin DB. Causal inference using potential outcomes: design, modeling, decisions. *J Am Stat Assoc*. 2005; 100(469):10.
15. Hernán MA, Alonso A, Logan R, et al. Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiol*. 2008; 19(6):766–779.
16. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiol*. 1999; 10(1):37–48.
17. Hernán MA, Hernández-Díaz S, Werler MM, Mitchell AA. Causal knowledge as a prerequisite for confounding evaluation: an application to birth defects epidemiology. *Am J Epidemiol*. 2002; 155(2):176–184. [PubMed: 11790682]
18. Robins JM. Data, design, and background knowledge in etiologic inference. *Epidemiol*. 2001; 12(3):313–320.
19. Pearl J. Causal diagrams for empirical research. *Biometrika*. 1995; 82:669–688.
20. Glymour MM, Weuve J, Chen JT. Methodological challenges in causal research on racial and ethnic patterns of cognitive trajectories: measurement, selection, and bias. *Neuropsychology Rev*. 2008; 18(3):194–213.
21. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Med Res Methodol*. 2008; 8:70. [PubMed: 18973665]
22. Schneeweiss S, Rassen JA, Glynn RJ, Avorn J, Mogun H, Brookhart MA. High-dimensional propensity score adjustment in studies of treatment effects using health care claims data. *Epidemiol*. 2009; 20(4):512–522.
23. Vanderweele TJ, Shpitser I. A new criterion for confounder selection. *Biometrics*. 2011; 67(4):1406–1413. [PubMed: 21627630]
24. Rubin DB. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. *Stat Med*. 2007; 26(1):20–36. [PubMed: 17072897]
25. Hill J. Discussion of research using propensity-score matching: comments on 'A critical appraisal of propensity-score matching in the medical literature between 1996 and 2003' by Peter Austin, *Statistics in Medicine*. *Stat Med*. 2008; 27(12):2055–2061. discussion 2066–9. [PubMed: 18446836]
26. Rubin DB. Estimating causal effects from large data sets using propensity scores. *Ann Int Med*. 1997; 127(8 Pt 2):757–763. [PubMed: 9382394]

27. Cepeda MS, Boston R, Farrar JT, Strom BL. Comparison of logistic regression versus propensity score when the number of events is low and there are multiple confounders. *Am J Epidemiol*. 2003; 158(3):280–287. [PubMed: 12882951]
28. Pearl J. Remarks on the method of propensity score. *Stat Med*. 2009; 28:1415–1424. [PubMed: 19340847]
29. Myers JA, Rassen JA, Gagne JJ, et al. Myers et al. respond to "understanding bias amplification". *Am J Epidemiol*. 2011; 174(11):1228–1229.
30. D'Avolio LW, Farwell WR, Fiore LD. Comparative effectiveness research and medical informatics. *Am J Med*. 2010; 123(12 Suppl 1):e32–e37. [PubMed: 21184865]
31. Rassen JA, Glynn RJ, Brookhart MA, Schneeweiss S. Covariate selection in high-dimensional propensity score analyses of treatment effects in small samples. *Am J Epidemiol*. 2011; 173(12):1404–1413. [PubMed: 21602301]
32. Vanderweele TJ, Arah OA. Bias formulas for sensitivity analysis of unmeasured confounding for general outcomes, treatments, and confounders. *Epidemiol*. 2011; 22(1):42–52.
33. Vanderweele TJ. Sensitivity analysis: distributional assumptions and confounding assumptions. *Biometrics*. 2008; 64(2):645–649. [PubMed: 18482060]
34. Schneeweiss S. Sensitivity analysis and external adjustment for unmeasured confounders in epidemiologic database studies of therapeutics. *Pharmacoepidemiol Drug Saf*. 2006; 15(5):291–303. [PubMed: 16447304]
35. Rosenbaum PR. Sensitivity analysis for m-estimates, tests, and confidence intervals in matched observational studies. *Biometrics*. 2007; 63(2):456–464. [PubMed: 17688498]
36. Rosenbaum PR. Sensitivity analysis for matched case-control studies. *Biometrics*. 1991; 47(1):87–100. [PubMed: 2049516]
37. Greenland S. Useful methods for sensitivity analysis of observational studies. *Biometrics*. 1999; 55(3):990–991. [PubMed: 11315040]
38. Greenland S. Basic methods for sensitivity analysis of biases. *Int J Epidemiol*. 1996; 25(6):1107–1116. [PubMed: 9027513]
39. Brumback BA, Hernán MA, Haneuse SJ, Robins JM. Sensitivity analyses for unmeasured confounding assuming a marginal structural model for repeated measures. *Stat Med*. 2004; 23(5):749–767. [PubMed: 14981673]
40. Arah OA, Chiba Y, Greenland S. Bias formulas for external adjustment and sensitivity analysis of unmeasured confounders. *Ann Epidemiol*. 2008; 18(8):637–646. [PubMed: 18652982]

Figure 1.

A causal graph illustrating confounding from the unmeasured variable $U2$. Conditioning on the measured variable (C_0), as indicated by the box around the variable, removes confounding from $U1$. Measured confounders are often proxies for unmeasurable constructs. For example, family history of heart disease is a measured variable indicating someone's risk for cardiovascular disease ($U1$).

Figure 2.

Hypothetical causal diagram illustrating Mtype collider stratification bias. Formulary policy ($U1$) influences treatment with CCB (A_0) and treatment for erectile dysfunction (F_0).

Unmeasured alcohol use ($U2$) influences impotence and erectile dysfunction treatment (F_0) and acute liver disease (Y_1). In this example there is no effect of antihypertensive treatment on liver disease, but antihypertensive treatment and liver disease would be associated when adjusting for medical treatment of erectile dysfunction. The box around F_0 , represents adjustment and the conditional relationship is represented by the dotted arrow connecting $U1$ and $U2$.

Figure 3. Bias is amplified (Z-bias) when an instrumental variable (Z_0) is added to a model with unmeasured confounders (U_1).