

Biometrics. Author manuscript; available in PMC 2012 December 1.

Published in final edited form as:

Biometrics. 2011 December; 67(4): 1406–1413. doi:10.1111/j.1541-0420.2011.01619.x.

A new criterion for confounder selection

Tyler J. VanderWeele and

Departments of Epidemiology and Biostatistics, Harvard School of Public Health 677 Huntington Avenue, Boston, MA 02115, Phone: 617-432-7855; Fax: 617-4321884

Ilva Shpitser

Department of Epidemiology, Harvard School of Public Health 677 Huntington Avenue, Boston, MA 02115

Tyler J. VanderWeele: tvanderw@hsph.harvard.edu

Abstract

We propose a new criterion for confounder selection when the underlying causal structure is unknown and only limited knowledge is available. We assume all covariates being considered are pretreatment variables and that for each covariate it is known (i) whether the covariate is a cause of treatment, and (ii) whether the covariate is a cause of the outcome. The causal relationships the covariates have with one another is assumed unknown. We propose that control be made for any covariate that is either a cause of treatment or of the outcome or both. We show that irrespective of the actual underlying causal structure, if any subset of the observed covariates suffices to control for confounding then the set of covariates chosen by our criterion will also suffice. We show that other, commonly used, criteria for confounding control do not have this property. We use formal theory concerning causal diagrams to prove our result but the application of the result does not rely on familiarity with causal diagrams. An investigator simply need ask, "Is the covariate a cause of the treatment?" and "Is the covariate a cause of the outcome?" If the answer to either question is "yes" then the covariate is included for confounder control. We discuss some additional covariate selection results that preserve unconfoundedness and that may be of interest when used with our criterion.

Keywords

Causal inference; confounding; covariate selection; directed acyclic graphs

1. Introduction

Control for confounding is one of the central challenges in the analysis of treatment effects from observational data. Effort is often made to collect data on a number of potentially confounding variables. A recent series of letters and exchanges between Shrier (2008), Rubin (2008, 2009), Pearl (2009) and Sjölander (2009) raised the question as to whether one should always adjust, either by regression or propensitiy score analysis (Rosenbaum and Rubin, 1983) for all pretreatment covariates. Pearl, Shrier and Sjölander described an example in which one obtained consistent estimates of a causal effect without adjusting for a pretreatment variable but inconsistent estimates of a causal effect when adjusting. Rubin (2009) argued that in practice one would essentially never want to refrain from adjusting for

Supplementary Materials

The supplementary materials containing the proof of the theorem and propositions may be accessed at the Biometrics website (http://www.biometrics.tibs.org).

a pretreatment covariate. In this paper we propose a new criterion for confounder control that we believe will aid researchers in determining which covariates to control for as confounders and will help mediate between the positions of Pearl, Shrier, Sjölander and Rubin.

One existing proposal for the problem of confounder selection involves the use of causal diagrams (Pearl, 1995, 2009). Pearl's "backdoor path criterion" (Pearl, 1995) provided a simple graphical criterion to assess the adequacy of controlling for a particular covariate set. However, the use of this result in practice presupposes that the structure of a causal diagram is known. Often this will be implausible. In particular, for a specific treatment variable of interest and a specific outcome, to apply Pearl's results one needs to have knowledge not simply on whether each covariate affects the treatment and the outcome but also on how each of the covariates are causally related to one another. In a number of analyses in the biomedical and social sciences, this knowledge is not available.

In this paper we give a result which will aid researchers in determining which covariates to control for as confounders when the underlying causal structure is not completely known. In particular we will consider a general setting, common in epidemiologic and biomedical research, in which for each covariate it is known whether the covariate is a cause of treatment and whether the covariate is a cause of the outcome, but it is not known how the covariates are all causally related to one another. We assume all the covariates are pretreatment covariates and we propose a new criterion for confounding control which consists of controlling for any covariate that is either a cause of treatment or a cause of the outcome or both. We show that irrespective of the true causal structure is and irrespective of whether there are important unobserved variables, if there exists some subset of the observed covariates that suffice to control for confounding, then the set obtained by applying our criterion will also constitute a set that suffices. Other frequently employed criteria do not have this property; among the standard criteria that do not have this property are: "control for all variables that are causes both of the treatment and of the outcome other than through treatment i.e. all common causes" (cf. Glymour et al., 2008) and "control for all pretreatment covariates" (Rubin, 2009).

Most analyses in epidemiology and the social science addressing questions of causal inference rely on the assumption of conditionally ignorable treatment assignment (Rosenbaum and Rubin, 1983). One way to construe our result would be that it provides a criterion for confounder selection to help ensure the plausibility of the conditionally ignorable treatment assignment assumption.

The proof we give of the result employs causal diagrams but familiarity with causal diagrams are not needed to *apply* the result. To apply the result an investigator simply needs to ask the question, for each covariate, "Is the covariate a cause of the treatment?" and "Is the covariate a cause of the outcome?" If the answer to either question is "yes" then the covariate is included for confounder control. The paper has thus been structured to the extent possible so that a reader unfamiliar with causal diagrams will still find the material accessible. Thus, technical material concerning causal diagrams and all proofs are relegated to appendices.

The remainder of the article is organized as follows. Section 2 reviews relevant notation and concepts from causal inference; section 3 presents our new confounding criterion and main result and discusses its relevance to the assumption of conditional ignorability. Section 4 gives some additional covariate selection results about the preservation of unconfoundedness (i.e. conditional ignorability) that may be of interest when used in conjunction with our new confounding criterion. Section 5 illustrates our confounder

selection criterion and our main result with a simulation. Section 6 offers some concluding remarks.

2. Causal Inference and Conditional Ignorability

We will use A to denote treatment, Y to denote the outcome, C to denote the set of measured pretreatment covariates and X to denote some subset of C. Within the potential outcomes framework (Rubin 1974, 1978, 1990), we let Y_a denote the potential outcome for Y if treatment A were set, possibly contrary to fact, to the value a. If treatment is binary then $E(Y_1) - E(Y_0)$ is generally referred to as the average causal effect of treatment. Note that the potential outcomes notation Y_a presupposes that an individual's potential outcome does not depend on the treatments received by other individuals. This assumption is sometimes referred to as SUTVA, the stable unit treatment value assumption (Rubin, 1990) or as a no-interference assumption (Cox, 1958).

We use the notation $E \perp F | G$ to denote that E is independent of F conditional on G. For a treatment variable A and outcome Y, we say that the treatment assignment is conditionally ignorable given X (or unconfounded given X) if $Y_a \perp A | X$. Conditional ignorability given covariates X means that within strata of X, treatment gives no information on the distribution of the potential outcomes. If the effect of A on Y is conditionally ignorable given X then the causal effect can be consistently estimated by:

$$E(Y_1) - E(Y_0) = \sum_{x} \{ E(Y|A=1, X=x) - E(Y|A=0, X=x) \} P(X=x).$$

Generally X will contain many variables some of which may be continuous and it is thus impractical to stratify on X. As a result, investigators often rely either on a regression model for E(Y|A=a, X=x) or use propensity score techniques (Rosenbaum and Rubin, 1983, 1985; Hernán and Robins, 2006) to estimate causal effects.

Pearl, in the context of causal diagrams (Pearl, 1995) gave a graphical criterion to assess the conditional ignorability assumption. Causal diagrams represent structural relationships amongst variables. Pearl's criterion is referred to as the "backdoor path criterion." On a causal diagram, a back-door path from some variable A to another variable Y is a path to Y which begins with an edge into A. For a treatment variable A and outcome Y, a set of variables X is said to satisfy the backdoor path criterion with respect to A, Y if no variable in X is a descendant of A and if X blocks all back-door paths from A to Y. See the Appendix for a review of causal diagrams. Pearl (1995) showed that if X satisfies the backdoor path criterion with respect to A, Y then the treatment assignment is conditionally ignorable given X (i.e. unconfounded given X) i.e. $Y_a \perp A|X$. Pearl's backdoor path criterion is a sufficient condition for $Y_a \perp A|X$; it is not necessary. We have given a complete graphical criterion characterizing when conditional ignorability holds on causal diagrams elsewhere (Shpitser, VanderWeele and Robins, 2010) but the result will not be needed here.

One way to think about a causal diagram is as a very general data generating process. A causal diagram corresponds to a set of non-parametric structural equations. If *X* satisfies the backdoor path criterion with respect to (*A*, *Y*) then treatment assignment will be conditionally ignorable given *X* irrespective of what the specific functional form or distribution of the variables actually are. See Appendix 1 or Pearl (1995) for further details on causal diagrams. The back door criterion provides a simple graphical criterion for assessing confounding when the structure of the causal diagram is known. However, in many cases the details of such causal structures will not be known and one must rely on criteria that can be assessed without such knowledge.

3. A New Criterion for Confounder Selection

We assume that the set C consists only of pretreatment covariates. We assume that for treatment A and outcome Y and observed covariate set C there is some unknown causal diagram G governing the causal relationships amongst these variables; the graph G may involve unobserved variables not in the set of measured covariates C. We assume that the researcher does not know the underlying causal structure but does know for each covariate $C_i \in C$ whether (i) C_i is a cause of A and whether (ii) C_i is a cause Y. Such knowledge is frequently presupposed in epidemiologic studies. Note, however, we do not assume that knowledge is available on how each covariate in C may be related to other covariates in C.

A criterion which is sometimes employed in epidemiologic studies consists of controlling for covariates that are causes both of A and of Y other than through A. For example, Subramanian et al. (2007) informally describe the task of confounding control as one of "adjusting for all common causes of the exposure and the outcome." Likewise, Glymour et al. (2008) note that, "The conventional approach to control for confounding is to identify the possible common causes of the exposure of interest and the outcome and then condition on these common causes, for example, through stratification, restriction, or statistical adjustment in a regression model," and then go on to describe more formal ideas of confounding control using causal diagrams. The criterion "control for all covariates that are common causes of the treatment and the outcome" is generally not articulated as a formal principle but is sometimes used in practice. We will refer to this criterion for confounder selection as the "common cause criterion." Another criterion which is sometimes used is to simply control for all pretreatment covariates for which data is available (Rubin, 2009). We will refer to this as the "pretreatment criterion." We instead propose an alternative criterion which consists of controlling for all covariates that satisfy (i) the covariate is a cause of treatment, or (ii) the covariate is a cause of the outcome, or both. Because of the disjunction, this is also equivalent to (i) the covariate is a cause of treatment, or (ii*) the covariate is a cause of the outcome not through treatment, or both. We will refer to this criterion as the "disjunctive cause criterion."

Example 1 below will demonstrate that there are cases in which the assumption of conditional ignorability holds for some subset of the measured covariates but it does not hold for the set of covariates selected by the common cause criterion or by the pretreatment criterion. The central result in this paper, stated in the following theorem, demonstrates that this situation effectively does not arise with the disjunctive cause criterion. In particular we show that if any subset of the observed covariates suffices to block all backdoor paths from treatment to outcome then the set selected by the disjunctive cause criterion will also suffice. The proof of the result, along with a slightly more general version of the result, is given in the Online Appendix.

Theorem 1

Let C be a set of measured pretreatment covariates. Let $S \subseteq C$ be the subset of C whose elements are either causes of A or of Y or of both. If there is any set $W \subseteq C$ such that W blocks all backdoor paths from A to Y then S does also and thus $Y_a \coprod A|S$.

Theorem 1 states that if there is any subset W of the observed covariates C that suffices to block all backdoor paths from treatment to outcome then the subset of C which consists of those observed covariates that are causes of A or Y or both will also and thus will suffice to control for confounding (i.e. for conditional ignorability). There may be multiple subsets of the measured covariates that suffice to control for confounding; in that case, the set S selected by the disjunctive cause criterion will then constitute one such subset (the subset it selects may but need not be a subset or superset of a particular set W). An interesting and

intuitive implication of Theorem 1 is that if there exists some (possibly unknown) subset of the measured covariates that blocks all backdoor paths then no bias is introduced by discarding variables that are neither causes of treatment or outcome.

Note no knowledge of the theory concerning causal diagrams is required for an investigator to apply our criterion. An investigator simply need ask, for each covariate, the question "Is the covariate a cause of the treatment?" and "Is the covariate a cause of the outcome?" If the answer to either question is "yes" then the covariate is included for confounder control. Theorem 1 does require that there exist some subset of the observed covariates that suffices to control for confounding; of course if no such subset exists then no criterion for confounder selection will be able to identify such a set.

The following example illustrates the common cause criterion, the pretreatment criterion and the disjunctive cause criterion and shows that the first two of these can fail in the sense of there being a subset of the observed covariates that suffice to control for confounding but neither the common cause criterion nor the pretreatment criterion in fact select such a set.

Example 1

Suppose that the true underlying causal diagram is that given in Figure 1 and that data is available on (A, Y, C_1, C_2, C_3) but not on (U_1, U_2, U_3) . Suppose further that the structure of the diagram is unknown and that it is only known that (C_1, C_2, C_3) are pretreatment covariates and that for each C_i , i = 1, 2, 3 it is known whether C_i is a cause of A and whether C_i is a cause of A other than through A. Thus, specifically, it is known that (i) C_1 is a cause of A and a cause of A other than through A, (ii) C_2 is neither a cause of A nor of A (though it is associated with both A and A0 and (iii) A1 is a cause of A2 but it is not a cause of A3 other than through A3.

The pretreatment criterion would suggest that control be made for (C_1, C_2, C_3) . The common cause criteria would suggest that control need be made only for C_1 since neither C_2 nor C_3 are causes of Y other than through A. The disjunctive cause criterion would suggest that control be made for (C_1, C_3) ; note C_2 would be discarded by the disjunctive cause criterion as it is not a cause of A or Y. In this example, the pretreatment criterion fails because the backdoor path $A \leftarrow U_1 \rightarrow C_2 \leftarrow U_2 \rightarrow Y$ from A to Y is unblocked conditional on (C_1, C_2, C_3) since C_2 is a "collider" on this path (see Appendix). This phenomenon is similar to the examples given by Shrier (2008), Pearl (2009), and Sjölander (2009). It is illustrated by a simulation in the next section. The common cause criterion also fails because the backdoor path $A \leftarrow C_3 \leftarrow U_3 \rightarrow Y$ from A to Y is unblocked conditional on C_1 alone. The disjunctive cause criterion does however succeed in identifying a set of covariates, namely (C_1, C_3) , that blocks all backdoor paths from A to Y. For this set of covariate, (C_1, C_3) , the conditional ignorability assumption will hold. Theorem 1 shows that this is the case more generally. If any subset of the measured pretreatment covariates can block all backdoor paths then the set selected by the disjunctive cause criterion will also.

Result 1 effectively mediates between the positions of Shrier (2008), Pearl (2009), and Sjölander (2009) on the one hand and Rubin (2009) on the other. Following Rubin (2009), we believe it is generally implausible that an investigator will know the complete underlying causal structure for the data generating process. Causal inference instead proceeds by relying on the assumption of conditional ignorability given some set of covariates *X*. However, by using our disjunctive cause criterion one avoids, if at all possible with the measured covariates, the types of structures described by Shrier (2008), Pearl (2009), and Sjölander (2009) which can occasionally lead to bias when conditioning on a pretreatment covariate.

4. Additional Covariate Selection Results

The disjunctive cause criterion selects all measured covariates that are causes of the treatment or the outcome. As seen in Example 1, controlling for covariates such as C_3 that are causes of the treatment but are not causes of the outcome except through treatment can sometimes block backdoor paths from the treatment to the outcome. However, one disadvantage of the disjunctive cause criterion is that it will also include in the set of covariates any variable that is a cause of A but completely unrelated to the outcome Y (e.g. an instrument for A). It is well known that control for such variables tends to increase the standard error of the estimates of treatment effects (e.g. Robinson and Jewell, 1991; Schisterman et al., 2009). To help overcome this potential disadvantage we state two propositions related to forward and backward selection. These two propositions are in principle independent of the main approach of the paper, but can be utilized to complement the application of the proposed criterion. The results essentially state that if some specified set of covariates (e.g. the set S in Theorem 1) suffice to control for confounding and if either a forward or backward selection procedure correctly discards variables independent of the outcome conditional on other variables then the final set selected by the procedure will also suffice to control for confounding. The first proposition follows immediately by iterated application of a result of Robins (1997), stated as a lemma in the Online Appendix, and relates to backward selection. It does not rely on theory of causal diagrams at all.

Proposition 1

Suppose that for some set S, $Y_a \perp A \mid S$ and that under some ordering of the elements of S, $(S_1, ..., S_n)$, and for some k, $Y \perp S_{i+1} \mid (A, S_1, ..., S_i)$ for i = k, ..., n-1, then $Y_a \perp A \mid (S_1, ..., S_k)$.

Proposition 1 suggests iteratively discarding variables unassociated with the outcome. One could likewise iteratively discard variables unassociated with treatment (Robins, 1997) and only associated with the outcome. However, for purposes of efficiency controlling for variables that are associated only with the outcome tends to increase the power of statistical tests for treatment effects at least when used in conjunction with regression (e.g. Robinson and Jewell, 1991). In practice one may thus not want to discard such variables. Backward selection under Proposition 1 discards only variables associated with treatment.

The second result relates to forward selection and requires an assumption of faithfulness for causal diagrams (see Appendix).

Proposition 2

Suppose that for some set S, $Y_a \perp A|S$, that the distribution of (Y, A, S) is faithful to the underlying causal diagram G, and that under some ordering of the elements of S, $(S_1, ..., S_n)$, and for some k, $Y \perp S_i|(A, S_1, ..., S_k)$ for i = k + 1, ..., n then $Y_a \perp A|(S_1, ..., S_k)$.

If the disjunctive cause criterion is used to select some initial set S of covariates to control for as confounders and if the sample size is sufficiently large so that independence tests can correctly identify which covariates are independent of Y, then Proposition 1 implies that if a backward selection procedure correctly iteratively discards covariates from S that are independent of Y conditional on treatment and the other covariates then the resulting covariate set will still suffice to control for confounding (i.e. will ensure conditional ignorability). Proposition 2, moreover, implies that if the underlying data generating mechanism is a causal diagram such that faithfulness holds then if a forward selection procedure iteratively identifies covariates (beginning with just A as an initial set) that are not independent of Y and then, after some number k covariates have been selected, correctly concludes that each of the remaining covariates are independent of Y, conditional on A and the covariates already selected, then the set of covariates resulting from this forward

selection procedure will also suffice to control for confounding (i.e. will ensure conditional ignorability). These two results, used along with the disjunctive confounding criterion, allow for the discarding of covariates that may be a cause of treatment but unrelated to the outcome. This could potentially circumvent the issue of increased standard errors due to the control of such variables. Because Proposition 2 requires that the distribution of (Y, A, S) is faithful to graph and because this cannot be verified if the underlying causal diagram is unknown, we would in practice recommend backward selection procedures over forward selection procedures.

A final note of warning is important here. Theorem 1 guarantees that if any subset of the measured covariates suffice to control for confounding then the set selected by the disjunctive cause criterion will suffice also. When no subset of measured covariates suffice to control for confounding, the set selected by the disjunctive cause criterion will of course not suffice (since no set will). In this setting, however, the disjunctive cause criterion has the potential to lead to greater bias than the application of other criteria. This is because the disjunctive cause criterion may result in the inclusion covariates that are causes of treatment and completely unrelated to the outcome i.e. instruments for treatment. Wooldridge (2009) and Pearl (2010) have shown that when bias due to unmeasured confounding is present, control for an instrument can amplify the existing confounding bias. We return to this point in the simulations below and in the concluding remarks.

5. Simulations

We illustrate our main result with a simulation based on the diagram in Figure 1. We generated 500 simulated datasets and a sample size of 2,000 was used for each of these datasets. Each simulated data was generated using the following structural equations corresponding to the structural relationships indicated in Figure 1.

$$\begin{split} U_i \sim N(0,1) & \text{ for } i{=}1,2,3 \\ C_1 \sim N(0,1) \\ C_2{=}1{+}2U_1{+}3U_2{+}\varepsilon_{c_2} & \text{ where } \varepsilon_{c_2} \sim N(0,0.5) \\ C_3{=}\frac{1}{5}{+}\frac{4}{5}U_3{+}\varepsilon_{c_3} & \text{ where } \varepsilon_{c_3} \sim N(0,0.5) \\ A{=}1[\log it^{-1}\{\frac{1}{4}(3U_1{+}C_1{+}2C_3)\}{>}\varepsilon_A] & \text{ where } \varepsilon_A \sim U(0,1) \\ Y{=}3{+}3A{+}2C_1{+}4U_2{+}4U_3{+}\varepsilon_y & \text{ where } \varepsilon_y \sim N(0,1) \end{split}$$

We consider estimation based both on linear regression models and propensity score matching (Rosenbaum and Rubin, 1985). Propensity score matching is implemented with 1:1 matching using an R package (Sekhon, 2010) that implements the estimators of Abadie and Imbens (2006) to calculate standard errors. The propensity score is constructed using a logistic regression model with all quadratic and product terms included in the model.

Under the pretreatment criterion, control would be made for (C_1, C_2, C_3) and a linear regression model could be fit:

$$E(Y|a, c_1, c_2, c_3) = \beta_0 + \beta_1 a + \beta_2 c_1 + \beta_3 c_2 + \beta_4 c_3$$
.

The estimate $\hat{\beta}_1$ for the effect of A from this regression, averaged over the 500 datasets, is 1.75 and the 95 percent confidence limits, averaged over the 500 datasets, are (1.46, 2.04). None of the individual 95% confidence intervals contained the true value of the causal effect of A on Y of 3. Similar results are obtained with propensity score matching on (C_1, C_2, C_3) : The estimate of the causal effect from propensity score matching, averaged over the 500

datasets, is 1.76 and the 95 percent confidence limits, averaged over the 500 datasets, are (1.35, 2.16). None of the individual 95% confidence intervals contained the true value of the causal effect of *A* on *Y* of 3.

Under the common cause criterion, control would be made only for C_1 and a linear regression model could be fit:

$$E(Y|a,c_1)=\beta_0+\beta_1a+\beta_2c_1$$

The estimate $\hat{\beta}_1$ for the effect of A from this regression, averaged over the 500 datasets, is 4.37 and the 95 percent confidence limits, averaged over the 500 datasets, are (3.87, 4.88). None of the individual 95% confidence intervals contained the true value of the causal effect of A on Y of 3. Similar results are obtained with propensity score matching on C_1 : The estimate of the causal effect from propensity score matching, averaged over the 500 datasets, is 4.37 and the 95 percent confidence limits, averaged over the 500 datasets, are (3.75, 4.99). The individual 95% confidence intervals in 4 of the 500 simulated data sets contained the true value of the causal effect of A on Y of 3.

Under the disjunctive cause criterion control would be made for (C_1, C_3) and a linear regression model could be fit:

$$E(Y|a, c_1, c_3) = \beta_0 + \beta_1 a + \beta_2 c_1 + \beta_3 c_3$$

The estimate $\hat{\beta}_1$ for the effect of A from this regression, averaged over the 500 datasets, is 3.00 and the 95 percent confidence limits, averaged over the 500 datasets, are (2.58, 3.42). The 95% confidence intervals in 472 of the 500 simulated datasets (94.4%) contained the true value of the causal effect of A on Y of 3. Similar results are obtained with propensity score matching on (C_1, C_3) : The estimate of the causal effect from propensity score matching, averaged over the 500 datasets, is 2.99 and the 95 percent confidence limits, averaged over the 500 datasets are (2.46, 3.53). The 95% confidence intervals in 481 of the 500 simulated datasets (96.2%) contained the true value of the causal effect of A on Y of 3.

We now present a second example to illustrate that (i) if there does not exist any subset of measured covariates that suffice to control for confounding then application of the disjunctive cause criterion can lead to more severe bias than other confounder selection criterion (though all will be biased in this setting since no sufficient set exists) and (ii) that in settings in which some sufficient set exists, the use of backwards selection of Proposition 1 after the application of the disjunctive cause criterion can improve precision of the estimate of the causal effect. Suppose that the true unknown underlying causal structure was that given in Figure 2 with structural equations given by:

$$\begin{split} U_i \sim N(0,1) & \text{ for } i{=}1,2,3 \\ C_i \sim N(0,1) & \text{ for } i{=}1,3 \\ C_2{=}1{+}2U_1{+}3U_2{+}\varepsilon_{c_2} & \text{ where } \varepsilon_{c_2} \sim N(0,0.5) \\ A{=}1[\log it^{-1}\{\frac{1}{4}(3U_1{+}2U_3{+}C_1{+}5C_3)\}{>}\varepsilon_{_A}] & \text{ where } \varepsilon_{_A} \sim U(0,1) \\ Y{=}3{+}3A{+}2C_1 - 2U_2{+}4U_3{+}\varepsilon_{_Y} & \text{ where } \varepsilon_{_Y} \sim N(0,1). \end{split}$$

Suppose that all that was known was that (i) C_1 was a cause of A and of Y, (ii) C_2 was neither a cause of A nor Y and (iii) C_3 was a cause of A only. The set selected by the common cause criterion would be C_1 ; the set selected by the disjunctive cause criterion would be (C_1, C_3) and the set selected by the pretreatment criterion would be (C_1, C_2, C_3) .

Note that here, no subset of the measured covariates will suffice to control for confounding since there will always be confounding due to U_3 . All estimates will be biased; however the magnitude of these biases may vary. Both the pretreament criterion and the disjunctive cause criterion select a variable that is in fact an instrument for A, namely C_3 , and as shown by Wooldridge (2009) and Pearl (2010), control for such variable can amplify bias due to unmeasured confounding already present.

We once again generate 500 simulated datasets each with a sample size of 2,000. The true causal effect of A on Y is 3. We report estimates from regression using the covariates sets selected by the various criteria; estimates from propensity score matching were similar and are not reported. Under the common cause criterion, the estimate and 95% confidence interval for the effect of A from a regression controlling for C_1 , averaged over the 500 datasets, is 4.39, CI = (3.99, 4.79). Under the disjunctive cause criterion, the estimate and 95% confidence interval for the effect of A from a regression controlling for (C_1, C_3) , averaged over the 500 datasets, is 4.73, CI = (4.28, 5.17). Under the pretreatment criterion, the estimate and 95% confidence interval for the effect of A from a regression controlling for (C_1, C_2, C_3) , averaged over the 500 datasets, is 5.37, CI = (4.96, 5.78). In each case, none of the individual 95% confidence intervals for any of the 500 datasets contained the true value of the causal effect of A on Y of 3. However, the bias under the common cause criterion, 4.39 - 3.00 = 1.39 was amplified under the disjunctive cause criterion (4.73 - 3 = 1.73) and under the pretreatment criterion (5.37 - 3 = 2.37) due to controlling for the instrument C_3 .

As a final illustrative simulation, suppose that for the causal diagram in Figure 2, that the coefficient for U_3 was in fact 0 (with all other relationships the same) so that U_3 was no longer a confounder for the effect of A on Y. Suppose also that the sample size of each dataset is now only 200 rather than 2,000. The common cause, disjunctive cause and pretreatment criteria would once again select C_1 , (C_1, C_3) and (C_1, C_2, C_3) , respectively. Since U_3 would no longer be a confounder both C_1 and (C_1, C_3) would constitute sets that sufficed to control for confounding; however, the disjunctive cause would once again select C₃, an instrument. Under the common cause criterion, the estimate and 95% confidence interval for the effect of A from a regression controlling for C_1 , averaged over the 500 datasets, is 3.00, CI = (2.38, 3.63). Under the disjunctive cause criterion, the estimate and 95% confidence interval for the effect of A from a regression controlling for (C_1, C_3) , averaged over the 500 datasets, is 3.01, CI = (2.31, 3.70). Both the common cause criterion and the disjunctive cause criterion give unbiased estimates but the confidence intervals under the set selected by the disjunctive cause criterion are in general now somewhat wider than those under the set selected by the common cause criterion because the disjunctive cause criterion selects an instrument. Applying backwards selection under Proposition 1, however, to the set selected by the disjunctive cause criterion would discard C_3 from the set of confounders as it is unassociated with Y conditional on A and C_1 , thus resulting in the same set selected by the common cause criterion.

6. Concluding Remarks

We have shown that the disjunctive confounding criterion we introduced in this paper has the attractive property that if any subset of the observed covariates suffice to block all backdoor paths from the treatment to the outcome then the subset selected by the disjunctive cause criterion will also suffice and thus will control for confounding for the effect of the treatment on the outcome (i.e. will ensure conditional ignorability). The criterion can be applied when the structure of the underlying causal structure is unknown provided that knowledge is available on whether each covariate is a cause of the treatment or the outcome. Knowledge is not needed on which covariates are related to other covariates. In many

epidemiologic and biomedical applications, subject matter experts have intuitive knowledge of whether each covariate is a cause of the treatment or the outcome.

In current practice, many applications assume that treament assignment is conditionally ignorable (i.e. unconfounded) given the set of measured pre-treatment covariates C. It is, however, possible that treatment assignment is not conditionally ignorable given C but is conditionally ignorable given a subset of C. The practical relevance of the result we have presented here is that if a researcher has information on whether each covariate is a cause of treatment or the outcome then this information can be used to construct a subset of C such that if any subset suffices for conditional ignorability by blocking all backdoor paths then the one constructed by our criterion will as well. The criterion thus helps ensure the plausability of the conditional ignorability assumption.

Future research could further examine the consequences of using various confounder selection criteria in contexts in which no set of observed covariates suffice for conditional ignorability. This may be especially relevant when considering whether to control for proxies of confounders; it is known that in some cases conditioning on a misclassified confounder can increase bias even if misclassification is non-differential (Brenner, 1993). Future work could characterize settings in which conditioning on a proxy of a confounder will reduce bias and how such characterizations may be combined with confounder selection criteria. Future work could perhaps also consider settings in which partial information is available on which covariates affect which others and whether refinements to the disjunctive confounding criterion might address biases that may arise in the application of the criterion in settings in which there is in fact no subset that suffices to control for confounding.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the editor, associate editor and two referees for helpful comments. The research was funded by NIH grants ES017876 and HD060696.

References

Abadie A, Imbens GW. Large sample properties of matching estimators for average treatment effects. Econometrica. 2006; 74:235–267.

Brenner H. Bias due to non-differential misclassification of polytomous confounders. Journal of Clinical Epidemiology. 1993; 46:57–63. [PubMed: 8433115]

Cox, DR. Planning of Experiments. New York: John Wiley & Sons; 1958.

Dawid AP. Influence diagrams for causal modelling and inference. Int Statist Rev. 2002; 70:161-189.

Geiger D, Verma TS, Pearl J. Identifying independence in Bayesian networks. Networks. 1990; 20:507–534.

Glymour MM, Weuve J, Chen JT. Methodological challenges in causal research on racial and ethnic patterns of cognitive trajectories: Measurement, selection, and bias. Neuropsychology Review. 2008; 18:194–213. [PubMed: 18819008]

Hernán MA, Robins JM. Estimating causal effects from epidemiological data. Journal of Epidemiology and Community Health. 2006; 60:578–586. [PubMed: 16790829]

Lauritzen, S. Graphical Models. Oxford University Press; Oxford: 1996.

Lauritzen SL, Dawid AP, Larsen BN, Leimer HG. Independence properties of directed Markov fields. Networks. 1990; 20:491–505.

Pearl J. Causal diagrams for empirical research. Biometrika. 1995; 82:669-688.

Pearl, J. Causality: Models, Reasoning, and Inference. Cambridge: Cambridge University Press; 2009.

- Pearl, J. On a class of bias-amplifying covariates that endanger effect estimates. In: Grunwald, P.; Spirtes, P., editors. Proceedings of the Twenty-Sixth Conference on Uncertainty in Artificial Intelligence. AUAI; Corvallis, OR: 2010. p. 417-424.
- Richardson TS, Spirtes P. Ancestral graph Markov models. Annals of Statistics. 2002; 30:962–1030.
- Robins, JM. Causal inference from complex longitudinal data. In: Berkane, M., editor. Latent Variable Modeling and Applications to Causality. Lecture Notes in Statistics. NY: Springer Verlag; 1997. p. 69-117.
- Robins, JM. Semantics of causal DAG models and the identification of direct and indirect effects. In: Green, P.; Hjort, NL.; Richardson, S., editors. Highly Structured Stochastic Systems. New York: Oxford University Press; 2003. p. 70-81.
- Robinson L, Jewell NP. Some surprising results about covariate adjustment in logistic regression models. Int Stat Rev. 1991; 59:227–240.
- Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. Biometrika. 1983; 70:41–55.
- Rosenbaum PR, Rubin DB. Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. Am Statist. 1985; 39:33–38.
- Rubin DB. Estimating causal effects of treatments in randomized and nonrandom-ized studies. Journal of Educational Psychology. 1974; 66:688–701.
- Rubin DB. Bayesian inference for causal effects: The role of randomization. Annals of Statistics. 1978; 6:34–58.
- Rubin DB. Formal modes of statistical inference for causal effects. Journal of Statistical Planning and Inference. 1990; 25:279–292.
- Rubin DB. The design versus the analysis of observational studies for causal effects: parallels with the design of randomized trials. Statistics in Medicine. 2007; 26:20–36. [PubMed: 17072897]
- Rubin DB. Author's reply (to Ian Shrier's Letter to the Editor). Statistics in Medicine. 2008; 27:2741–2742.
- Rubin DB. Author's reply (to Judea Pearl's and Arvid Sjölander's Letters to the Editor). Statistics in Medicine. 2009; 28:1420–1423.
- Schisterman EF, Cole SR, Platt RW. Overadjustment bias and unnecessary adjustment in epidemiologic studies. Epidemiology. 2009; 20:488–495. [PubMed: 19525685]
- Sekhon JS. Matching: Multivariate and propensity score matching with balance optimization. Journal of Statistical Software. 2010 in press.
- Shrier I. Letter to the editor. Statistics in Medicine. 2008; 27:2740–2741. [PubMed: 18069729]
- Shpitser I, Pearl J. Dormant independence. Proceedings of the Twenty-Third Conference on Artificial Intelligence. 2008:1081–1087.
- Shpitser, I.; VanderWeele, TJ.; Robins, JM. Proceedings of the 26th Conference on Uncertainty and Artificial Intelligence. AUAI Press; Corvallis, WA: On the validity of covariate adjustment for estimating causal effects; p. 527-536.
- Sjölander A. Letter to the editor. Statistics in Medicine. 2009; 28:1416–1420. [PubMed: 19340845]
- Spirtes, P.; Glymour, C.; Scheines, R. Causation, Prediction and Search. New York: Springer-Verlag; 1993.
- Subramanian, SV.; Glymour, MM.; Kawachi, I. Identifying causal ecologic effects on health: a methodological assessment. In: Galea, S., editor. Macrosocial Determinants of Health. Vol. Chapter 15. Springer Media; 2007. p. 301-331.
- Verma, T.; Pearl, J. In: Shachter, R.; Levitt, TS.; Kanal, LN., editors. Causal networks: Semantics and expressiveness; Proceedings of the Fourth Workshop on Uncertainty in Artificial Intelligence; Amsterdam: Elesevier; 1988. p. 352-359.p. 69-76.Reprinted in Uncertainty in Artificial Intelligence
- Verma, T.; Pearl, J. Technical Report R-150. Department of Computer Science, University of California; Los Angeles: 1990. Equivalence and synthesis of causal models.

Wooldridge, J. Should instrumental variables be used as matching variables?. Michigan State University Tech Rep. 2009.

https://www.msu.edu/~ec/faculty/wooldridge/current%20research/treat1r6

Appendix

We review definitions and results concerning causal directed acyclic graphs. A directed graph consists of a set of nodes and directed edges amongst nodes. A path is a sequence of distinct nodes connected by edges regardless of arrowhead direction; a directed path is a path which follows the edges in the direction indicated by the graph's arrows. A directed acyclic graph is a directed graph such that there is no node with a sequence of directed edges back to itself. The nodes with directed edges into a node A are said to be the parents of A; the nodes into which there are directed edges from A are said to be the children of A. We say that node A is ancestor of node B if there is a directed path from A to B; if A is an ancestor of B then B is said to be a descendant of A. If X denotes a set of nodes then An(X) will denote the ancestors of X. A node is said to be a collider for a particular path if it is such that both the preceding and subsequent nodes on the path have directed edges going into that node. A path between two nodes, A and B, is said to be blocked given some set of nodes C if either there is a variable in C on the path that is not a collider for the path or if there is a collider on the path such that neither the collider itself nor any of its descendants are in C. For disjoint sets of nodes A, B and C, we say that A and B are d-separated given C if every path from any node in A to any node in B is blocked given C. Directed acyclic graphs are sometimes used as statistical models to encode independence relationships amongst variables represented by the nodes on the graph (Lauritzen, 1996). We will use the notation $A \perp \!\!\! \perp B|C$ to denote that Ais conditionally independent of B given C. The variables corresponding to the nodes on a graph are said to satisfy the global Markov property for the directed acyclic graph if for any disjoint sets of nodes A, B, C we have that $A \perp \!\!\!\perp B \mid C$ whenever A and B are d-separated given C. The distribution of some set of variables V on the graph are said to be faithful to the graph if for all disjoint sets A, B, C of V we have that $A \perp \!\!\! \perp B \mid C$ only when A and B are dseparated given C.

Directed acyclic graphs can be interpreted as representing causal relationships. Pearl (1995) defined a causal directed acyclic graph as a directed acyclic graph with nodes $(X_1, ..., X_n)$ corresponding to variables such that each variable X_i is given by its non-parametric structural equation $X_i = f_i(pa_i, \varepsilon_i)$ where pa_i are the parents of X_i on the graph and the ε_i are mutually independent. The non-parametric structural equations encode counterfactual relationships amongst the variables represented on the graph. The equations themselves represent one-step ahead counterfactuals with other counterfactuals given by recursive substitution (see Pearl, 2009, for further discussion). For example, if A is a parent of Y and if there are no intermediates between A and Y included on the graph then if $f_Y(pa_Y, \varepsilon_Y)$ is the non-parametric structural equation for Y then for an individual ω , the counterfactual $Y_a(\omega)$ is given by $Y_a(\omega) = f_y(a, \widetilde{PA}(\omega), \varepsilon_y(\omega))$ where \widetilde{PA} denotes the parents of Y other than A on the graph i.e. A is fixed to level a and the other parents of Y are left at their actual level $\widetilde{PA}(\omega)$. The requirement that the ε_i be mutually independent is essentially a requirement that there is no variable absent from the graph which, if included on the graph, would be a parent of two or more variables (Pearl, 1995, 2009). A causal directed acyclic graph defined by nonparametric structural equations satisfies the global Markov property as stated above (cf. Verma and Pearl, 1988; Geiger et al., 1990; Lauritzen et al., 1990; Pearl, 2009). Further discussion of the causal interpretation of directed acyclic graphs can be found elsewhere (Spirtes et al., 1993; Pearl, 1995, 2009; Dawid, 2002; Robins, 2003).

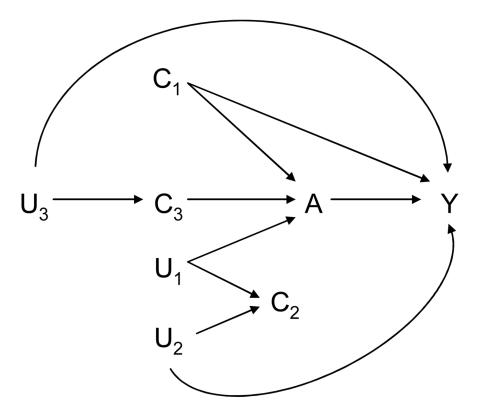


Figure 1. Unknown causal diagram illustrating confounding criteria.

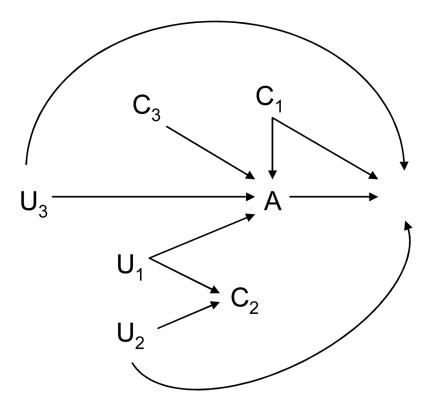


Figure 2. Unknown causal diagram illustrating conditioning on an instrument, C_3 .