



Published in final edited form as:

Epidemiology. 2010 May ; 21(3): 383–388. doi:10.1097/EDE.0b013e3181d61eeb.

Negative Controls: A Tool for Detecting Confounding and Bias in Observational Studies

Marc Lipsitch^{1,2,3}, Eric Tchetgen Tchetgen^{1,3,4}, and Ted Cohen^{5,1,3}

¹ Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115

² Department of Immunology and Infectious Diseases, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115

³ Center for Communicable Disease Dynamics, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115

⁴ Department of Biostatistics, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115

⁵ Division of Global Health Equity, Brigham and Women's Hospital, Boston MA 02115

Abstract

Non-causal associations between exposures and outcomes are a threat to validity of causal inference in observational studies. Many techniques have been developed for study design and analysis to identify and eliminate such errors. Such problems are not expected to compromise experimental studies, where careful standardization of conditions (for laboratory work) and randomization (for population studies) should, if applied properly, eliminate most such non-causal associations. We argue, however, that a routine precaution taken in the design of biological laboratory experiments—the use of “negative controls”—is designed to detect both suspected and unsuspected sources of spurious causal inference. In epidemiology, analogous negative controls help to identify and resolve confounding as well as other sources of error, including recall bias or analytic flaws. We distinguish two types of negative controls (exposure controls and outcome controls), describe examples of each type from the epidemiologic literature, and identify the conditions for the use of such negative controls to detect confounding. We conclude that negative controls should be more commonly employed in observational studies, and that additional work is needed to specify the conditions under which negative controls will be sensitive detectors of other sources of error in observational studies.

Epidemiologists seek to distinguish the causal effect of exposure A on outcome Y from associations due to other mechanisms (Figure 1). Non-causal associations may be classified into three categories (in addition to chance)¹: mismeasurement (eg, recall bias), confounding, and biased selection of individuals into the analysis.

In experimental biology, the manipulation of experimental conditions prevents many of the noncausal associations that arise in observational studies. Nonetheless, experimental biologists routinely question whether they have correctly inferred causal relationships from

Address for correspondence: Marc Lipsitch, Department of Epidemiology, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115 USA, mlipsitc@hsph.harvard.edu, (617) 432-4559, (617) 566-7805 fax.

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

the results of their experiments. Biologists employ “negative controls” as a means of ruling out possible noncausal interpretations of their results. We describe the use of negative controls in experiments, highlight some examples of their use in epidemiologic studies, and define the conditions under which negative controls can detect confounding in epidemiologic studies. Although the particular threats to causal inference are different in experimental and observational sciences, the use of negative controls is a valuable means of identifying noncausal associations and can complement other epidemiologic methods for improving causal inference.

Experimental biology: threats to causal inference and the use of negative controls

One might imagine that the experimental method would circumvent most threats to the validity of causal inference that occur in observational studies. For example, consider the hypothesis that a particular cytokine—a chemical involved in signaling in the immune system—enhances the killing of a species of bacteria by neutrophils, a class of white blood cells.² An experiment is devised in which neutrophils, bacteria, and growth medium are mixed together. In condition 1, the cytokine is added, and in condition 2, some inert substance such as saline solution is added. After incubation, the bacteria are enumerated and the number of live bacteria compared between conditions 1 and 2.

If the investigator finds fewer live bacteria in condition 1 than in condition 2, the finding is consistent with the hypothesis that the cytokine enhanced neutrophil-mediated killing. Nonetheless, concern remains that something other than cytokine-aided, neutrophil-mediated killing may be responsible. For example, perhaps there is a contaminant in the cytokine preparation that directly kills bacteria, or perhaps the cytokine itself kills bacteria, or perhaps some other unintended difference between the treated and untreated conditions (e.g., temperature or pH) caused the differential survival of the bacteria.

Each of these unintended differences is broadly similar to a confounder – a characteristic associated with the exposure (presence or absence of the cytokine) and causes the outcome (differences in bacterial counts), thereby causing a spurious association between the presence of the cytokine and differences in bacterial counts.

Experimental biologists address such concerns in two ways. The first is to attempt to eliminate unwanted differences between the compared groups (in the design) and to measure and account for any unavoidable differences (in the analysis). For example, a researcher would make all conditions (dilution protocols, incubators, etc.) identical between the two conditions except for the variable of interest (i.e. the presence/absence of the cytokine). Replication of the experiment reduces the likelihood that some chance factor was systematically different between the two experimental arms. Sometimes experimental variation nonetheless remains. When experimental variation cannot be eliminated by these approaches, experimentalists may control for this variation by matching or statistical adjustment for the day on which an assay was performed. In experimental studies of population health outcomes (clinical trials), analogous precautions include randomization (to assure an expectation of baseline exchangeability between groups)³, use of multiple individuals in each treatment group (replication), and analytic adjustment for measured confounders.

The second general approach is to perform negative controls: to repeat the experiment under conditions in which it is expected to produce a null result and verify that it does indeed produce a null result. Several strategies are employed to design negative controls,² such as:

- Leave out an essential ingredient. In the absence of neutrophils, there should be the same number of bacteria with or without the cytokine; if a contaminant (or the cytokine itself) is killing bacteria without involving neutrophils, this negative control should produce fewer bacteria with the cytokine than without.
- Inactivate the (hypothesized) active ingredient. Specific antibodies that neutralize the cytokine (but would have no effect on a contaminant) can be added to the preparation; killing should not occur if the cytokine is responsible for the effect.
- Check for an effect that would be impossible by the hypothesized mechanism. Suppose there were a species of bacteria that was completely impervious to the actions of neutrophils. The experiment could be repeated with this species, rather than the species of interest, to confirm there is no difference between condition 1 and condition 2. This would help to rule out the possibility of some non-neutrophil-mediated effect of the cytokine preparation on bacteria.

As with the list of non-causal explanations for an experimental result, the list of possible negative controls is almost endless, and judgment is required to assess how many such non-causal explanations are plausible and which negative controls are of greatest value in ruling out key threats to valid inference. Peer reviewers of biological experiments usually require some negative controls to validate experimental results.

Examples of the use of negative controls in epidemiology

In an epidemiologic study to assess whether an association between a risk factor A and an outcome Y is likely to be causal, it is common to address the possibility of confounding by measured variables L by adjusting for them, using such techniques as restriction, stratification, multivariate modeling, matching, inverse-probability weighting, or g-estimation.⁴

Epidemiologists also sometimes use negative controls to detect confounding and other sources of incorrect causal inference. This approach has been elegantly applied to the debate over vaccination of the elderly and effects on “pneumonia or influenza hospitalization/pneumonia/influenza on all-cause mortality. Observational studies in elderly persons have shown that vaccination against influenza is associated with a remarkably large reduction in one’s risk of pneumonia/influenza hospitalization and also in one’s risk of all-cause mortality in the following season, after adjustment for measured covariates that indicate health status.⁵ However, older age is associated with a less robust immune response to influenza vaccination, and ecological data suggest that the benefits measured in observational studies far exceed the corresponding benefits expected at the population level when influenza vaccination rates have increased among the elderly.⁶ Importantly, both outcomes are nonspecific, in the sense that they have unknown and time-varying contributions from influenza. This is obviously true for all-cause mortality, and it is also true for pneumonia/influenza hospitalization, since the cause of respiratory infection is often not ascertained, and many pneumonia cases are caused by agents other than influenza. The large degree of protection against these outcomes observed in individual level studies, combined with the lack of measurable vaccine effect in ecological studies, have led to a suspicion that uncontrolled confounding has exaggerated the impact of influenza vaccination on mortality and on pneumonia/influenza hospitalization in the elderly.^{6–8}

To test this hypothesis, Jackson et al⁷ reproduced earlier estimates of the protective effect of influenza vaccination, but then repeated the analysis for two sets of negative control outcomes, and showed that the protective effect was observed even in circumstances where the vaccine could not have caused the protection. For the first negative control outcome, the authors⁷ used the fact that vaccination often begins in autumn, while influenza transmission

is often minimal until winter. Thus they could assess the risk of pneumonia/influenza hospitalization and all-cause mortality among vaccinated vs. unvaccinated persons both before, during and after influenza season. The only biologically plausible mechanism by which influenza vaccine could protect against mortality or pneumonia/influenza hospitalization is by preventing influenza or its consequences; therefore, Jackson and colleagues⁷ reasoned that if the effect measured in prior studies were causal, it should be most prominent during influenza season. If instead it were due to confounding, then the protective effect should be observable immediately after vaccination but before influenza season. In a cohort study analyzed with a Cox proportional hazards model, despite efforts to control for confounding, they observed that the protective effect was actually greatest before, intermediate during, and least after influenza season. They concluded that this is evidence that confounding, rather than protection against influenza, accounts for a substantial part of the observed “protection.” The use of this negative-control outcome approach is formally similar to the “leave-out-an-essential-ingredient” control described above, as influenza is essential in the proposed causal pathway.

Second, Jackson et al.⁷ postulated that the protective effects of influenza vaccination, if real, should be limited to outcomes plausibly linked to influenza. In contrast, if the relationship were due to an uncontrolled confounder, then the same “protection” might be observed for irrelevant outcomes. They repeated their analysis, but substituted hospitalization for injury or trauma as the endpoint. They found that influenza vaccination was also “protective” against injury or trauma hospitalization. This, too, was interpreted as evidence that some of the protection observed for pneumonia/influenza hospitalization or mortality was due to inadequately controlled confounding. This second negative control outcome is formally similar to the “check-for-an-effect-impossible-by-the-hypothesized-mechanism” approach described above.

Epidemiologists also sometimes use negative control exposures to examine whether observed associations are causal. An example is the inclusion in questionnaires of irrelevant variables, sometimes called “probe variables,” to assess if recall bias may be responsible for an observed association between a self-reported exposure and an outcome. A recent study⁹ tested the association between multiple sclerosis (MS) and a variety of common childhood infections assessed by self-report. The investigators found statistically significant positive associations of MS with a recalled history of five different viral infections. Suspecting that cases may recall prior medical events more often or with more certainty than controls, the investigators’ questionnaire also included several childhood medical events not plausibly associated with MS, such as broken limbs, tonsillectomy, and concussions. In the absence of a causal association, any measured association with these probe variables would suggest recall bias for the variables of interest. The authors found that the magnitude of association with these irrelevant exposures was comparable to the magnitude observed for each of the self-reported infections except one (infectious mononucleosis) that had a much stronger association. They concluded that, after accounting for recall bias, only infectious mononucleosis showed a specific association with MS.

Another application of negative controls has been to expose “immortal time bias,” a form of selection bias that produces spurious associations between observed variables. Suissa and Ernst¹⁰ suspected that the reported benefits of nasal corticosteroids in preventing asthma resulted from this form of bias, in which exposed persons are credited with time at risk during which the event cannot occur, and thus exposed persons have an artificially low event rate. Inclusion of the “immortal” time is dependent on both being exposed during that time and on not having the outcome during that time¹⁰; hence a (negative) association is induced between exposure and outcome. To demonstrate such bias, the authors repeated prior analyses but restricted the exposed class to persons with a single annual dose of

corticosteroids— a dose far too low to have plausible biological effect (i.e. a negative control exposure). They found that even this very modest exposure was associated with substantial protection against asthma, suggesting that the previous analytic approach was inappropriate. In this case, the investigators already suspected what form of bias was operating, and used the analysis to prove their point. In principle, the original investigators could have done such an analysis to test for bias.

Choice of negative controls to detect confounding in epidemiology

Negative controls have been used to detect confounding (the influenza vaccine example⁷), recall bias, (the MS example⁹), and selection bias (the nasal corticosteroid example¹⁰). Furthermore, it may be possible to specify how negative controls should be designed to aid in detecting biased causal inferences resulting from each of these mechanisms, and also perhaps to detect other forms of analytical errors. In this section, we focus on the conditions under which negative controls in epidemiology can detect confounding.¹

The essential purpose of a negative control is to reproduce a condition that cannot involve the hypothesized causal mechanism, but is very likely to involve the same sources of bias that may have been present in the original association. If a contaminant (source of bias) were responsible for the effect of the cytokine on bacteria, it should have its effect even when the hypothesized mechanism of the effect (through neutrophils) is prevented through neutralization of the cytokine or through omission of neutrophils from the experiment. If an uncontrolled confounder (general good health or healthful practices) is responsible for the protection observed from influenza vaccine against mortality or pneumonia/influenza hospitalization, the same confounder might be associated with other outcomes that are not plausibly prevented by influenza vaccination.

This description suggests a general principle for the selection of negative controls to detect residual confounding. Ideally, a negative control outcome (N) should be an outcome such that the set of common causes of exposure A and outcome Y should be as identical as possible to the set of common causes of A and N (Figure 2). To the extent that the set of unobserved common causes of A and Y overlaps with the set of unobserved common causes (U) of A and N, we call the negative control outcome N “U-comparable” to Y. If N and Y are U-comparable outcomes (i.e. with an identical set of common causes that are associated with A), and assuming that N is not caused by A, an association A-N when analyzed according to the same procedure used to analyze A-Y would indicate bias in the association A-Y. If N and Y are perfectly U-comparable and N is not caused by A, then a null finding of A-N implies that the A-Y association is not likely biased by the pathways examined through this negative control.

Negative control outcomes in practice will be only approximately U-comparable, at best. Thus it is possible that the observed association between A and N is caused by some uncontrolled confounder U₂, which is not a confounder of the AY association; hence, finding an unexpected association between A and N does not prove unequivocally that the A-Y association is biased. In the example of using death or hospitalization from injury as a negative control outcome for death or pneumonia/influenza hospitalization, one could argue that there may be some common causes of vaccination and injury that are not causes of all-cause death or pneumonia/influenza hospitalization. Such common causes (we cannot think of a plausible one) would create an association in the negative control analysis of vaccination and injury, even if the primary analyses of vaccination and death or pneumonia/influenza hospitalization were unconfounded— thus making the negative control detect bias even where none exists. On the other hand, if N is associated only with some, but not all, of the uncontrolled confounders of the association between A and Y, it is possible that A and N

will appear unassociated despite the presence of uncontrolled confounding between A and Y. In the influenza vaccine example, one could argue that there are common causes of vaccination and death or pneumonia/influenza hospitalization = that are not causes of injury-related outcomes. Such a common cause (say, an aversion to vaccination that makes an individual less likely to get the pneumococcal vaccine) would be undetectable by this particular negative control. Despite these limitations, negative controls have value in alerting the analyst to possible residual confounding.

In principle, the measured confounders L of the A-Y relationship need not be causes of N as well, since a properly specified model that accounted for the confounding by L of A-Y would not be misled if such confounding were absent for A-N. In practice, the ideal negative control outcome should nonetheless be one with incoming arrows as similar as possible to those of Y, including the incoming arrows from L. This is true, first, because it is difficult in practice to imagine an outcome N that lacks association with known confounders L, but has an association with uncontrolled (or even unknown) confounders similar to that of U-Y. In addition, because negative controls may be useful in detecting residual confounding by measured confounders L or analytic errors, it would be beneficial to have the L-N relationship be as similar as possible, quantitatively, to the L-Y relationship. In eAppendix 1 (<http://links.lww.com>), we describe the analytic basis for use of a U-comparable negative control outcome.

A negative control exposure B should be an exposure such that the common causes of A and Y are as nearly identical as possible to the common causes of B and Y (Fig. 3). To the extent that the set of unobserved common causes U of A and Y overlaps with the set of unobserved common causes of B and Y, we call the negative control exposure B “U-comparable” to A. If A and B are perfectly U-comparable and B does not cause Y, then an association B-Y when analyzed according to the same model used to analyze A-Y would indicate bias in the association A-Y. If A and B are perfectly U-comparable and B does not cause Y, then a null finding of A-N means that the A-Y association is unbiased. We are not aware of an example of the use of a negative control exposure to detect confounding in this sense. In the influenza vaccination example, one might hypothesize that whatever residual confounders U (e.g., poor health status) made one less likely to get influenza vaccine (A) and more likely to die of influenza or pneumonia (Y), might also make one less likely to get other vaccines, such as booster tetanus vaccine (B). Because tetanus does not cause pneumonia, tetanus vaccine receipt might be an appropriate negative control exposure for such a study. In the previous section, we mentioned the use of “probe variables” as negative controls to detect recall bias that might lead MS patients to over-report a history of childhood infections. Recall bias, a form of reverse causation, has a different causal structure from confounding,¹ and we do not outline here the causal requirements for negative controls to detect reverse causation.

In observational settings, the comparability between exposure A and negative control exposure B will be only approximate. As in the case of negative control outcomes, this approximate comparability means that B and Y may be associated even when A-Y is unbiased; this would occur if there is some other confounder U₂ linking B and Y that does not confound A-Y. Similarly, if A and B are only approximately comparable, it is possible for B and Y to show no association yet for A-Y to be biased, if the confounder biasing A-Y does not connect B to Y. An analytic basis for the use of negative control exposures is given in eAppendix 2 (<http://links.lww.com>).

In a cohort study, in which multiple exposures and outcomes are measured on each person, it is relatively straightforward to analyze negative control exposures and outcomes, assuming that suitable variables have been measured. In a case-control study, the use of negative control exposures is similarly straightforward because negative control exposures can be

added to the set of exposure variables collected for each subject. If a case-control study is nested within a cohort, irrelevant outcomes can be selected and analyzed. A stand-alone case-control study presents some logistical problems for implementing negative-control outcomes. This might require a second case-control study in which “cases” include some irrelevant but comparable outcome to the cases in the main study. This difficulty is reduced if multiple control groups are used, as is occasionally done for other reasons.^{11,12}

A useful contrast can be drawn between variables that can serve as negative controls and those that can be used as instruments.^{13–15} An instrumental variable is any variable that is connected causally to A but free of any of the confounding connections to Y from which A suffers. In contrast, a negative control outcome is connected to A through all possible confounding routes but not causally. Similarly, a negative control exposure is connected to Y through all possible confounding routes but not causally. Figure 3 depicts an instrumental variable Z that satisfies the necessary conditions of an instrument^{16,17} while the variable B is an ideal negative exposure candidate.

Discussion

We propose that negative controls should be applied more commonly in epidemiologic studies, as in laboratory experiments, and with the same goals: to detect uncontrolled confounding or other sources of bias that create a spurious causal inference.¹ The routine use of negative controls in experimental biology allows the detection of both suspected and unsuspected sources of bias. The challenge of deriving valid causal inference is at least as great in observational studies as in experiments. In other social sciences, negative control outcomes are sometimes recommended for use with observational as well as experimental studies,¹⁸ to compensate for limited sample size and possible imbalance between treatment arms.

A.B. Hill proposed specificity of association as one guideline for assessing causal inferences.¹⁹ Hill argued that causal inferences were more credible if the exposure (in his example, nickel mining) was associated with only certain types of outcomes (death from lung and nose cancer but not death from other cancers), and if the outcome was associated with one kind of exposure (nickel mining) but not many others. Hill himself, as well as more recent authors,^{16·20·21} have been ambivalent about this particular guideline. Weiss²² has argued that specificity of outcome and exposure may, in certain cases, lend credibility to causal inference, especially if there is a strong hypothesis of why the outcome (or exposure) should be specific to the cause. Both Hill’s and Weiss’s arguments are related to the ideas of negative controls; we suggest that informative tests of specificity of association are those that meet the criteria we have outlined for negative control exposures or outcomes. Their value will vary depending on the plausibility of the claim that the control considered is U-comparable to the exposure or outcome of interest.

Subject matter knowledge is required for the choice of negative controls, just as it is for the design of appropriate strategies to adjust for confounders. If an investigator identifies negative controls based on incorrect causal assumptions, the analysis involving negative controls may be misleading. If a causal association between two variables A-N is thought to be implausible and is used as a negative control for a study of some other association A-Y, then finding an association between A and N will erroneously suggest bias in the association AY.

A properly selected negative control is a sensitive, but blunt, tool to probe the credibility of a study. The “failure” of a negative control – the finding of an association that is judged not to be plausibly causal -- does not identify what form of bias is operating. In particular, as we

demonstrate in eAppendix 3 (<http://links.lww.com>), the magnitude of bias due to uncontrolled confounding cannot generally be inferred from the magnitude of a detected A-N (or B-Y) non-null association, without extra assumptions based on firm scientific understanding. Furthermore, such additional subject matter knowledge (or suspicion about the source of analytic errors) is necessary to determine where bias is likely to have arisen.

We have defined precisely the conditions under which negative controls are capable of detecting the existence and direction of bias due to uncontrolled confounders. We have argued by example that negative controls can also aid in detecting recall bias (reverse causation) or selection bias. Epidemiologists must weigh these potential benefits of employing negative controls against the increased cost associated with the measurement of additional variables, and the possibility that the assumptions under which the negative control variables were selected are faulty.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Murray Mittleman, Molly Franke, Justin O'Hagan, and Hsien-Ho Lin for helpful discussion.

Funding: Supported by NIH 5U01GM076497 and 1U54GM088558 (Models of Infectious Disease Agent Study) to ML.

References

- Hernan MA, Hernandez-Diaz S, Robins JM. A structural approach to selection bias. *Epidemiology* 2004;15(5):615–25. [PubMed: 15308962]
- Lu YJ, Gross J, Bogaert D, Finn A, Bagrade L, Zhang Q, Kolls JK, Srivastava A, Lundgren A, Forte S, Thompson CM, Harney KF, Anderson PW, Lipsitch M, Malley R. Interleukin-17A mediates acquired immunity to pneumococcal colonization. *PLoS Pathog* 2008;4(9):e1000159. [PubMed: 18802458]
- Fisher, RA. *The Design of Experiments*. Edinburgh: Oliver and Boyd; 1935.
- Greenland, S. Introduction to Regression Modeling. In: Rothman, KJ.; Greenland, S.; Lash, TL., editors. *Modern Epidemiology*. 3. New York: Lippincott, Williams and Wilkins; 2008. p. 418–455.
- Rivetti D, Jefferson T, Thomas R, Rudin M, Rivetti A, Di Pietrantonj C, Demicheli V. Vaccines for preventing influenza in the elderly. *Cochrane Database Syst Rev* 2006;3:CD004876. [PubMed: 16856068]
- Simonsen L, Taylor RJ, Viboud C, Miller MA, Jackson LA. Mortality benefits of influenza vaccination in elderly people: an ongoing controversy. *Lancet Infect Dis* 2007;7(10):658–66. [PubMed: 17897608]
- Jackson LA, Jackson ML, Nelson JC, Neuzil KM, Weiss NS. Evidence of bias in estimates of influenza vaccine effectiveness in seniors. *Int J Epidemiol* 2006;35(2):337–44. [PubMed: 16368725]
- Simonsen L, Reichert TA, Viboud C, Blackwelder WC, Taylor RJ, Miller MA. Impact of Influenza Vaccination on Seasonal Mortality in the US Elderly Population. *Arch Intern Med* 2005;165(3):265–72. [PubMed: 15710788]
- Zaadstra BM, Chorus AM, van Buuren S, Kalsbeek H, van Noort JM. Selective association of multiple sclerosis with infectious mononucleosis. *Mult Scler* 2008;14(3):307–13. [PubMed: 18208871]
- Suissa S, Ernst P. Bias in observational study of the effectiveness of nasal corticosteroids in asthma. *J Allergy Clin Immunol* 2005;115(4):714–9. [PubMed: 15805989]

11. Bell CM, Bajcar J, Bierman AS, Li P, Mamdani MM, Urbach DR. Potentially unintended discontinuation of long-term medication use after elective surgical procedures. *Arch Intern Med* 2006;166(22):2525–31. [PubMed: 17159020]
12. Harris AD, Samore MH, Lipsitch M, Kaye KS, Perencevich E, Carmeli Y. Control-Group Selection Importance in Studies of Antimicrobial Resistance: Examples Applied to *Pseudomonas aeruginosa*, Enterococci, and *Escherichia coli*. *Clin Infect Dis* 2002;34(12):1558–63. [PubMed: 12032889]
13. Brookhart MA, Wang PS, Solomon DH, Schneeweiss S. Evaluating short-term drug effects using a physician-specific prescribing preference as an instrumental variable. *Epidemiology* 2006;17(3):268–75. [PubMed: 16617275]
14. Hernan MA, Robins JM. Instruments for causal inference: an epidemiologist’s dream? *Epidemiology* 2006;17(4):360–72. [PubMed: 16755261]
15. Martens EP, Pestman WR, de Boer A, Belitser SV, Klungel OH. Instrumental variables: application and limitations. *Epidemiology* 2006;17(3):260–7. [PubMed: 16617274]
16. Glymour, MM.; Greenland, S. Causal diagrams. In: Rothman, KJ.; Greenland, S.; Lash, TL., editors. *Modern Epidemiology*. 3. Philadelphia: Lippincott, Williams & Wilkins; 2008.
17. Greenland S. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol* 2000;29(4):722–9. [PubMed: 10922351]
18. Cook, TD.; Campbell, DT. *Quasi-Experimentation: Design & Analysis Issues for Field Settings*. Boston: Houghton-Mifflin; 1979.
19. Hill AB. The Environment and Disease: Association or Causation? *Proc R Soc Med* 1965;58:295–300. [PubMed: 14283879]
20. Aschengrau, A.; Seage, GR, III. *Essentials of Epidemiology in Public Health*. Boston: Jones and Bartlett; 2008.
21. Rothman, KJ.; Greenland, S.; Poole, C.; Lash, TL. Causation and Causal Inference. In: Rothman, KJ.; Greenland, S.; Lash, TL., editors. *Modern Epidemiology*. 3. Philadelphia: Lippincott, Williams and Wilkins; 2008. p. 5-31.
22. Weiss NS. Can the “specificity” of an association be rehabilitated as a basis for supporting a causal hypothesis? *Epidemiology* 2002;13(1):6–8. [PubMed: 11805580]

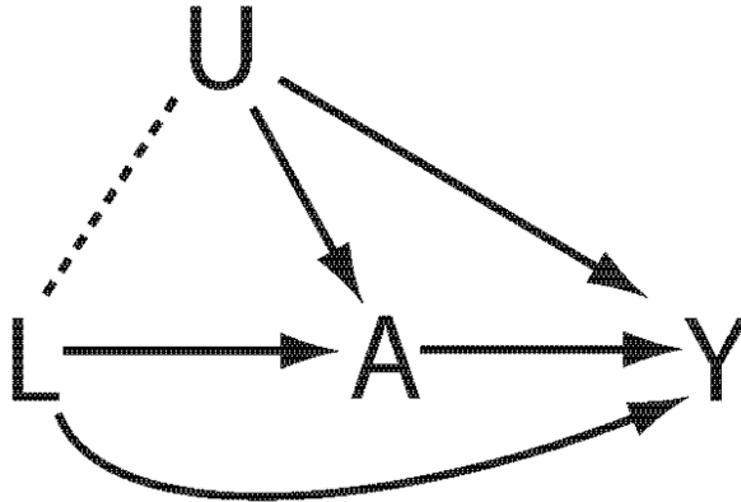


FIG 1. Causal diagram for the effect of an exposure of interest (A) on an outcome of interest (Y), with confounders L (assumed measured) and U (assumed uncontrolled) that cause both A and Y. The dashed line between L and U indicates that either may cause the other, and they may share common causes.

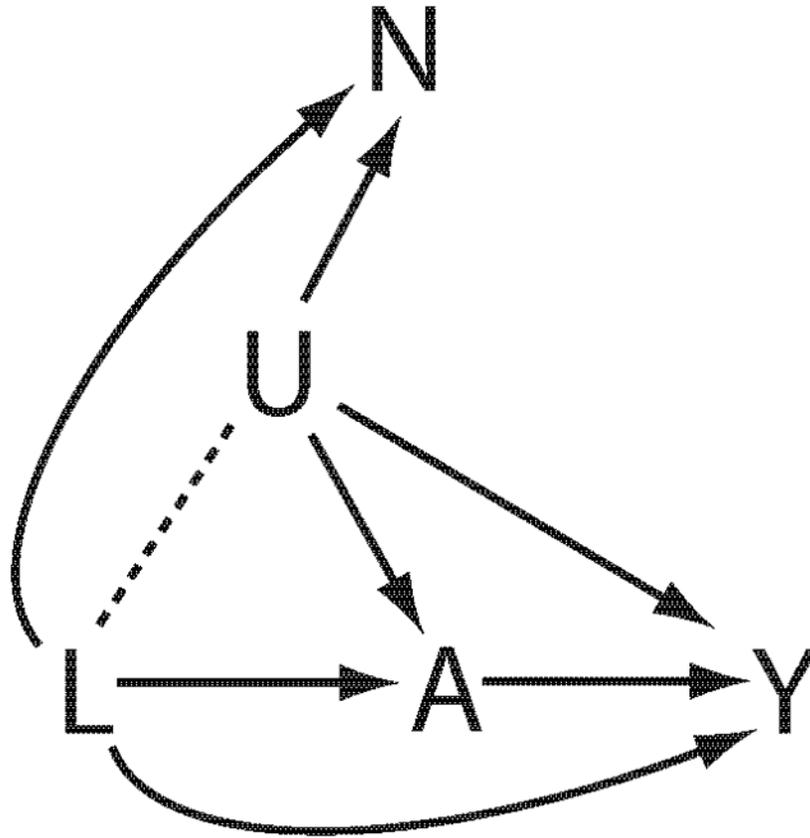


FIG 2. Causal diagram showing an ideal negative control outcome N for use in evaluating studies of the causal relationship between exposure A and outcome Y. N should ideally have the same incoming arrows as Y, except that A does not cause N; to the extent this criterion is met, N is called U-comparable to Y.

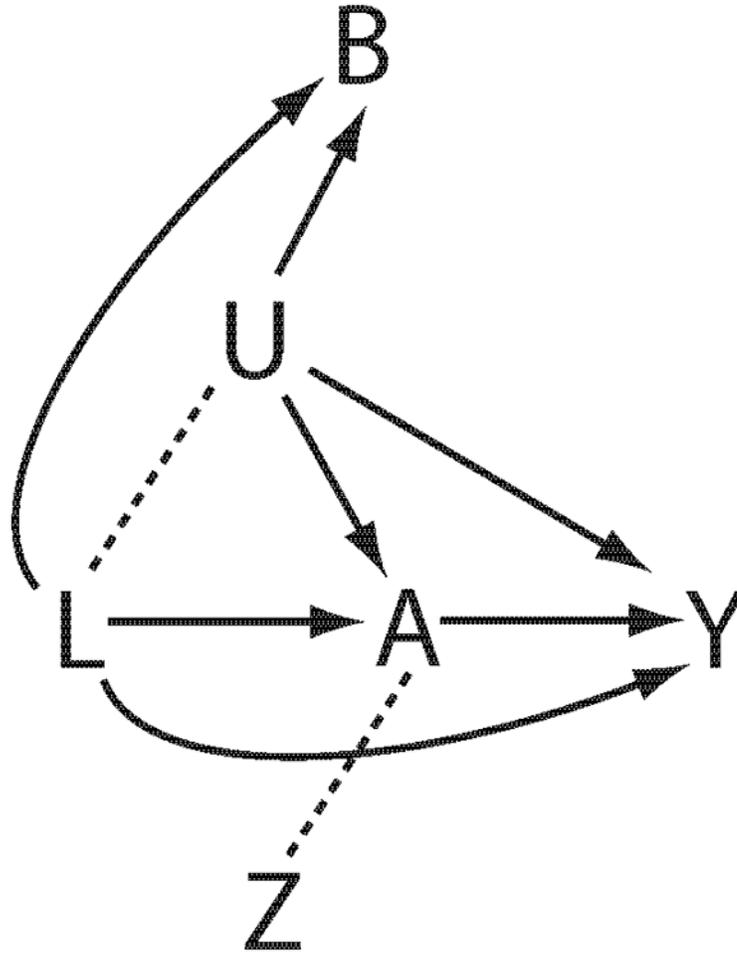


FIG 3. Causal diagram showing an ideal negative control exposure B for use in evaluating studies of the causal relationship between exposure A and outcome Y. B should ideally have the same incoming arrows as A; to the extent this criterion is met, B is called U-comparable to A. Z is an instrumental variable of the A-Y relationship and is depicted to illustrate the difference between an instrumental variable and a negative control variable.