

The C-Word: Scientific Euphemisms Do Not Improve Causal Inference From Observational Data

Causal inference is a core task of science. However, authors and editors often refrain from explicitly acknowledging the causal goal of research projects; they refer to causal effect estimates as associational estimates.

This commentary argues that using the term “causal” is necessary to improve the quality of observational research.

Specifically, being explicit about the causal objective of a study reduces ambiguity in the scientific question, errors in the data analysis, and excesses in the interpretation of the results. (*Am J Public Health*. 2018;108:616–619. doi:10.2105/AJPH.2018.304337)

Miguel A. Hernán, MD, DrPH



See also Galea and Vaughan, p. 602; Begg and March, p. 620; Ahern, p. 621; Chiolerio, p. 622; Glymour and Hamad, p. 623; Jones and Schooling, p. 624; and Hernán, p. 625.

You know the story:

Dear author: Your observational study cannot prove causation. Please replace all references to causal effects by references to associations.

Many journal editors request authors to avoid causal language,¹ and many observational researchers, trained in a scientific environment that frowns upon causality claims, spontaneously refrain from mentioning the C-word (“causal”) in their work. As a result, “causal effect” and terms with similar meaning (“impact,” “benefit,” etc.) are routinely avoided in scientific publications that describe nonrandomized studies. Instead, we see terms like “association” and others that convey a similar meaning (“correlation,” “pattern,” etc.), or the calculatedly ambiguous “link.”

The proscription against the C-word is harmful to science because causal inference is a core task of science, regardless of whether the study is randomized or nonrandomized. Without being able to make explicit references to causal effects, the goals of many observational studies can only be expressed in a roundabout way. The resulting ambiguity impedes a frank discussion about methodology because the methods used to estimate causal effects are not the same as those used to estimate associations.

Confusion then ensues at the most basic levels of the scientific process and, inevitably, errors are made.

We need to stop treating “causal” as a dirty word that respectable investigators do not say in public or put in print. It is true that observational studies cannot definitely prove causation, but this statement misses the point, as discussed in this commentary.

OF COURSE “ASSOCIATION IS NOT CAUSATION”

Suppose we want to know whether daily drinking of a glass of wine affects the 10-year risk of coronary heart disease. Because there are no randomized trials of long-term alcohol drinking, we analyze observational data by comparing the risk of heart disease across people with different levels of alcohol drinking over 10 years. Say that this analysis yields a risk ratio of heart disease of 0.8 for one

glass of red wine per day versus no alcohol drinking. For simplicity, disregard measurement error and random variability—that is, suppose the 0.8 comes from a very large population so that the 95% confidence interval around it is tiny.

The risk ratio of 0.8 is a measure of the association between wine intake and heart disease. Strictly speaking, it means that drinkers of one glass of wine have, on average, a 20% lower risk of heart disease than individuals who do not drink. The risk ratio of 0.8 does not imply that drinking a glass of wine every day lowers the risk of heart disease by 20%. It is possible that the kind of people who drink a glass of wine per day would have a lower risk of heart disease even if they didn’t drink wine because, for example, they have high enough incomes to buy, besides wine, nutritious food and to take time off to exercise, or have better access to preventive health care.

In other words, the risk ratio of 0.8 may be an unbiased

ABOUT THE AUTHOR

Miguel A. Hernán is with the Departments of Epidemiology and Biostatistics, Harvard T. H. Chan School of Public Health, and the Harvard-MIT Division of Health Sciences and Technology, Boston, MA.

Correspondence should be sent to Miguel Hernán, Department of Epidemiology, Harvard T. H. Chan School of Public Health, 677 Huntington Ave, Boston, MA 02115 (e-mail: miguel_hernan@post.harvard.edu). Reprints can be ordered at <http://www.ajph.org> by clicking the “Reprints” link.

This article was accepted January 14, 2018.

doi: 10.2105/AJPH.2018.304337

measure of the association between wine and heart disease, but a biased (confounded) measure of the causal effect of wine on heart disease. Importantly, we knew this before conducting the study. That observational effect estimates may be confounded is not a scientific statement, it is a logical one. The sentence “Your observational effect estimate may be seriously confounded” can never be proven wrong, regardless of how much data are available for confounding adjustment. In fact, the sentence is in the same logical category as “You can die in the next five years.” Sadly, both quoted statements are always logically possible, no matter what data you have at this moment.

We all agree: confounding is always a possibility and therefore association is not necessarily causation. One possible reaction is to completely ditch causal language in observational studies. This reaction, however, does not solve the tension between causation and association; it just sweeps it under the rug.

CONFLATING THE MEANS AND THE ENDS

In our example, the driving question for the research is whether modifying wine intake can reduce the incidence of heart disease. That is, the primary scientific aim of this observational study is to quantify “the causal effect of wine on heart disease,” not “the association between wine and heart disease” or (gasp) “the link between wine and heart disease.”

In a parallel universe, we might be able to estimate this causal effect by conducting a randomized trial in which large numbers of people are randomly assigned to different levels of

wine intake and forced to comply over 10 years. Fortunately, in our world, such a trial is considered unethical and hence forbidden. Unfortunately, that means that observational analyses become our best chance to quantify the long-term causal effect of wine on chronic diseases. And yet all we can estimate from the data are associations that may not reflect causation. The analysis of the observational study is necessarily associational, even though the goal of the observational study is causal.

Interestingly, the same is true of randomized trials. All we can estimate from randomized trials data are associations; we just feel more confident giving a causal interpretation to the association between treatment assignment and outcome because of the expected lack of confounding that physical randomization entails. However, the association measures from randomized trials cannot be given a free pass. Although randomization eliminates systematic confounding, even a perfect randomized trial only provides probabilistic bounds on “random confounding”—as reflected in the confidence interval of the association measure—and many randomized trials are far from perfect.

Hence we need to use causal language to accurately describe the aims of our research, be it based on randomized trial or observational data. Avoiding the word “causal” in a scientific paper or grant application makes it impossible to express the research aims unambiguously. In the wine example, our goal is to estimate the causal effect of wine on heart disease. Therefore, the term “causal effect” is appropriate in the title and Introduction section of our article when describing our aim, in the Methods section when describing which causal effect we are trying to estimate

through an association measure, and in the Discussion section when providing arguments for and against the causal interpretation of our association measure.² The only part of the article in which the term “causal effect” has no place is the Results section, which should present the findings without trying to interpret them.

Without causally explicit language, the means and ends of much of observational research get hopelessly conflated. As Rothman put it more than 30 years ago,

Some scientists are reluctant to speak so blatantly about cause and effect, but in statements of hypothesis and in describing study objectives such boldness serves to keep the real goal firmly in focus and is therefore highly preferable to insipid statements about ‘association’ instead of ‘causation.’^{3(p77)}

Carefully distinguishing between causal aims and associational methods is not just a matter of enhancing scientific communication and transparency. Eliminating the causal-associational ambiguity has practical implications for the quality of observational research too.

BETTER CAUSAL QUESTIONS

Associational questions are easy to formulate and straightforward to answer when data are available. Are you interested in the association between drinking one glass of wine daily and heart disease at a certain time in a certain population? Just compare the risk of heart disease between individuals who drink one glass of wine daily and those who did not drink wine. The statistical analysis

is trivial if the data are accurately measured in the population of interest.

Causal questions, on the other hand, are not always easy to formulate. If we want to estimate the causal effect of drinking one glass of wine daily on heart disease, we first need to explain what we mean by “the causal effect of drinking one glass daily on heart disease.” A helpful approach is to define the causal effect in our population as the causal effect that would have been observed in a hypothetical trial in which individuals in our population had been randomly assigned to either drinking one glass of wine or no wine drinking for some period (say, 10 years). Of course, such a trial is infeasible (and unethical), but that is beside the point.^{4–6} The point is that an observational analysis can be guided by defining the causal effect in a hypothetical trial as the inferential target.

In other words, a causal analysis in observational data can be viewed as an attempt to emulate a hypothetical trial—the target trial.⁷

Specifying the target trial is a useful device to sharply define a causal question in an observational analysis, and to better understand the data that are necessary for emulating the target trial.⁷ A key advantage of specifying the target trial is that it forces investigators to consider the intervention of interest and the time period during which it takes place. For example, a target trial could assign people to 10 years of daily wine drinking from age 55 years, or to 20 years of wine drinking from age 30 years. Each of these target trials answers a different causal question and therefore, if the trials were actually conducted, they would result in different causal effects. For the same reason,

emulating each of these trials using observational data will require a different set of data on treatments and outcomes. For example, to emulate the second target trial, we will need longitudinal data on wine intake and coronary heart disease since age 30 years. The explicit consideration of the causal goal of the research, via specification of the target trial, facilitates the scientific discussion about data requirements for causal inference.

BETTER CONFOUNDING ADJUSTMENT

If the goal of the observational analysis is causal, adjustment for confounding is generally necessary. In our wine example, the risk ratio of 0.8 may be partly or fully explained by access to preventive care and socioeconomic status, which are correlates of both moderate wine drinking and heart disease. Therefore, if the data analysis does not incorporate adjustment for factors that predict both wine drinking and heart disease, we will suspect that the association measure is confounded and therefore we will be reluctant to interpret it as a causal effect measure.

On the other hand, if the goal of the observational analysis is purely associational, no adjustment for confounding is necessary. Remember, if we just want to quantify the association between wine and heart disease, we simply compute it from the data. If we want to develop a predictive model for heart disease, we include covariates (like wine drinking and number of doctor visits in the last year) that predict heart disease, not only

confounders. In associational or predictive models, we do not try to endow the parameter estimates with a causal interpretation because we are not trying to adjust for confounding of the effect of every variable in the model. Confounding is a causal concept that does not apply to associations. There is no such thing as a “spurious association” unless we use the term to mean an association that cannot be causally interpreted—but then the goal of the analysis would be causal, not associational.

By contrast, in a causal analysis, we need to think carefully about what variables can be confounders so that the parameter estimates for treatment or exposure can be causally interpreted. Automatic variable selection procedures may work for prediction, but not necessarily for causal inference.⁸ Selection algorithms that do not incorporate sufficient subject-matter knowledge may select variables that introduce bias in the effect estimate,^{9–14} and ignoring the causal structure of the problem may lead to apparent paradoxes.^{15–17} Also, note that the parameters for the confounders cannot be causally interpreted because we do not adjust for the confounders of the confounders, and the adjustment variables may include mediators of the confounder effects, including the treatment itself.¹⁸

Many readers will correctly point out that there is no guarantee that a causal model incorporates all the confounders and therefore there is no guarantee that the parameter estimate for treatment can be causally interpreted, even approximately. We have gone full circle. Surely there is no guarantee the parameter estimate for treatment can be causally

interpreted, but we can have an informed scientific discussion about it only if we have first acknowledged the causal goal of the analysis.

CONCLUSIONS

The lack of clarity regarding the goals of the research has often been justified by the questionable validity of causal inferences from observational data. However, this argument simply conflates the aims and the methods of scientific research. An association measure from an observational analysis may be a biased estimate of a causal effect, but being explicit about the goal of the analysis is a prerequisite for good science. Do we want to estimate the association measure or the causal effect measure? Do we want to determine whether “the sort of people who drink a glass of red wine daily have a lower risk of heart disease” or do we want to determine whether “drinking a glass of red wine daily lowers the risk of heart disease”? Associational inference (prediction) or causal inference (counterfactual prediction)?

The answer to this question has deep implications for (1) how we design the observational analysis to emulate a particular target trial and (2) how we choose confounding adjustment variables. Each causal question corresponds to a different target trial, may require adjustment for a different set of confounders, and is amenable to different types of sensitivity analyses. It then makes sense to publish separate articles for various causal questions based on the same data. By contrast, no target trial or confounder selection is necessary in associational analyses.

Arguably, the biggest disservice of traditional statistics to science was to make “causal” into a dirty word,¹⁹ the C-word that researchers have learned to avoid. Glossing over associational and causal goals in many statistics courses and textbooks has led to widespread confusion among users of statistics. In a perfect example of cognitive dissonance, scientific journals often publish articles that avoid ever mentioning their obviously causal goal. It is time to call things by their name. If your thinking clearly separates association from causation, make sure your writing does too. **AJPH**

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health grant A1102634.

Sander Greenland provided insightful criticisms to an earlier version of this commentary.

REFERENCES

1. Ruich P. The use of cause-and-effect language in the JAMA Network journals. *AMA Style Insider*. 2017. Available at: <http://amastyleinsider.com/2017/09/19/use-cause-effect-language-jama-network-journals>. Accessed January 14, 2018.
2. Savitz DA. Re: “Associations are not effects.” *Am J Epidemiol*. 1991;134(4):442–444.
3. Rothman KJ. *Modern Epidemiology*. Boston, MA: Little, Brown and Company; 1986.
4. Hernán MA. Does water kill? A call for less casual causal inferences. *Ann Epidemiol*. 2016;26(10):674–680.
5. Kaufman JS. There is no virtue in vagueness. *Ann Epidemiol*. 2016;26(10):683–684.
6. Robins JM, Weissman M. Counterfactual causation and streetlamps. What is to be done? *Int J Epidemiol*. 2016;45(6):1830–1835.
7. Hernán MA, Robins JM. Using big data to emulate a target trial when a randomized trial is not available. *Am J Epidemiol*. 2016;183(8):758–764.
8. Robins JM, Greenland S. The role of model selection in causal inference from nonexperimental data. *Am J Epidemiol*. 1986;123(3):392–402.
9. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15(5):615–625.

10. 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.
11. Greenland S, Pearl J. Adjustments and their consequences—collapsibility analysis using graphical models. *Int Stat Rev*. 2011; 79(3):401–426.
12. Greenland S, Neutra R. Control of confounding in the assessment of medical technology. *Int J Epidemiol*. 1980;9(4): 361–367.
13. Greenland S, Pearl J, Robins JM. Causal diagrams for epidemiologic research. *Epidemiology*. 1999;10(1):37–48.
14. Greenland S, Pearl J, Robins JM. Confounding and collapsibility in causal inference. *Stat Sci*. 1999;14(1):29–46.
15. Hernández-Díaz S, Schisterman EF, Hernán MA. The birth weight “paradox” uncovered? *Am J Epidemiol*. 2006;164(11): 1115–1120.
16. Hernán MA, Clayton D, Keiding N. The Simpson’s paradox unraveled. *Int J Epidemiol*. 2011;40(3):780–785.
17. Snoep JD, Morabia A, Hernández-Díaz S, Hernán MA, Vandenbroucke JP. Commentary: a structural approach to Berkson’s fallacy and a guide to a history of opinions about it. *Int J Epidemiol*. 2014; 43(2):515–521.
18. Westreich D, Greenland S. The table 2 fallacy: presenting and interpreting confounder and modifier coefficients. *Am J Epidemiol*. 2013;177(4):292–298.
19. Pearl J. *Causality: Models, Reasoning, and Inference*. 2nd ed. New York, NY: Cambridge University Press; 2009.