


## Cause and Association: Missing the Forest for the Trees

 See also Galea and Vaughan, p. 602; Hernán, p. 616; Ahern, p. 621; Chiolero, p. 622; Glymour and Hamad, p. 623; Jones and Schooling, p. 624; and Hernán, p. 625.

In his commentary, Hernán (p. 616) delineates the dangers of avoiding the term “causal”—the “C-word”—in describing the findings from observational research studies. His message seems to run counter to a central tenet of graduate studies in public health: “association is not causation.” This statement is certainly true and undisputed by Hernán. However, it has become a mantra, at potentially considerable cost, which Hernán addresses in a thoughtful case for resurrecting the C-word in discussing observational study results.

We have considered Hernán’s commentary from the perspective of instructors of an integrated introductory course in epidemiology and biostatistics for master of public health (MPH) students, which we have taught for the past five years. Clearly, such a course must aim to have its students distinguish between association and causation. At first blush, we might reject Hernán’s recommendation. However, it is critical to ensure that students understand the goals of public health research; we agree wholeheartedly that as public health professionals, our goal is to identify not just correlates but actual causes of disease, and to take action. This begs the question: can we accurately convey

that although our analysis may be limited to identifying associations, the paramount objective in public health and biomedical research studies is to assess causation? We think the answer is yes.

Our own class centers on assessing potential causes through a hierarchy of study designs that balance rigor and aptness<sup>1</sup> interwoven with an array of analytic techniques. We point our students to the excellent text by Hulley et al.,<sup>2</sup> with chapter 9 focusing on ways to enhance causal inference under observational study designs. Table 9.1 poses five possible explanations for observing an association between an exposure and an outcome in an observational study: chance, bias, effect-cause, confounding, and cause-effect; it thus highlights four explanations in addition to the last explanation, cause-effect, the causal hypothesis being evaluated. Chance refers to random error and the possibility of a spurious association between exposure and outcome. Bias refers to systematic error, also leading to a spurious association. Effect-cause, or reverse causation, underscores for students that an association between exposure and outcome may be real (versus spurious), but opposite to the

anticipated direction. Confounding, to which Hernán rightly devotes considerable attention, is also a real association, but not causal with respect to the primary exposure of interest. In our experience, students find this exposition of alternate, noncausal explanations both logical and accessible.

This structure, introduced at the beginning of the course and referenced repeatedly throughout, provides a straightforward way for students to engage with the notions of causation and association, without censorship of the C-word. Following the Hulley model, we describe how statistical methods can be used to reveal an association between two variables in a given data set. We then reinforce the notion that such an association may result from five different scenarios, only one of which is the hypothesis that the selected exposure causes the outcome in the context of a well-operationalized causal question. With four possibilities in addition to cause-and-effect, the course marches through a variety of design and analytic methods that allow us to

winnnow through the potential noncausal explanations for the observed association. In addition to its intuitive appeal, this approach also provides motivation for the various study designs and analytic methods we wish students to learn.

Although the message of “association is not causation” must remain, we agree that we in academia may have overstated the case, thereby doing a disservice to our students and the field. It is certainly possible, and desirable, to bring discussion of cause back into the literature on observational studies, and it may just lead to better science. **AJPH**

Melissa D. Begg, ScD  
Dana March, PhD, MPH

### CONTRIBUTORS

Both authors contributed equally to this editorial.

### REFERENCES

1. Susser M. Some principles in study design for preventing HIV transmission: rigor or reality. *Am J Public Health*. 1996; 86(12):1713–1716.
2. Hulley SB, Cummings SR, Browner WS, Grady DG, Newman TB. *Designing Clinical Research*. 4th ed. Philadelphia, PA: Lippincott Williams & Wilkins; 2013.

### ABOUT THE AUTHORS

Melissa D. Begg is with the Department of Biostatistics, Mailman School of Public Health, and Office of the Provost, Columbia University, New York, NY. Dana March is with the Department of Epidemiology, Mailman School of Public Health.

Correspondence should be sent to Melissa D. Begg, Professor & Vice Provost, Columbia University, 535 West 116th St, Low Memorial Library MC4341, New York, NY 10027 (e-mail: mdb3@columbia.edu). Reprints can be ordered at <http://www.ajph.org> by clicking the “Reprints” link.

This editorial was accepted February 1, 2018.  
doi: 10.2105/AJPH.2018.304366