Comments on "Estimating Causal Effects" by Maldonado and Greenland

Glenn Shafer
Graduate School of Management, Rutgers University
gshafer@andromeda.rutgers.edu
April 27, 2001

This article explains the counterfactual theory of causation, avoiding details and technicalities but providing a clear explanation of most of the terminology that is used when the theory is applied to epidemiology. At the end of the article, the authors mention that some people "reject counterfactuals as a foundation for casual inference." The editor has asked me, as one of those people, to explain the difficulties I see with the counterfactual theory. I will try to do so at the same non-technical level at which the article is written.

Although the authors begin their history of the counterfactual approach with a quotation from David Hume, they would probably agree that speculation about "what might have been" is as old as the human ideas of blame and regret. No doubt the objections to such speculation are equally as old. When your mother tells you that you would have avoided your cold by wearing a jacket, you may object that the result of wearing or not wearing a jacket was not predictable and perhaps not in any sense determined. If you could have acted differently in the matter of the jacket, you and others could have acted differently in other respects, many of which might also have impinged on your health. Who is to say who would have done what had you worn a jacket?

Epidemiologists are usually concerned with the effects of public-health risks on whole populations, and we might hope that the average effect of an exposure on a population might be well defined even when the effect on individuals is not, because of the averaging-out of other unpredictable factors. However, as the authors make clear, the counterfactual approach, as it has been developed in the statistical and epidemiological literature in recent decades, insists on the assumption that the effects on individuals are well defined. In this article, for example, they assume that it is determined whether a given individual will fall ill regardless of exposure. So the argument between the advocates of counterfactuals (such as the authors) and the dissenters (such as myself) really does boil down to the ancient argument between those who insist on always giving meaning to a might-have-been and those who demur.

What is the alternative to the counterfactual approach? The obvious alternative is a predictive approach. Using this approach, we say that A causes B in a strong sense if we can predict, using a method of prediction that proves consistently correct, that B will happen if we do A and will not happen if we do not do A. Weaker senses of causation can be expressed using probabilities; we say that the action A is a probabilistic cause of B if it raises the probability of B. This requires an objective concept of probability; it must be verified that B consistently happens more often when A is performed than when it is not, regardless of other factors.

As I explain in my 1996 book, *The Art of Causal Conjecture*, the practical aspects of causal inference (different ways of defining causal effects, ideas of confounding, etc.) can be handled by the predictive approach just as well as by the counterfactual approach. And the predictive approach has a decisive philosophical advantage: it makes clear that the concept of

causality has an empirical basis, independent of arbitrarily imagined might-have-beens. I say more about this in my article "Causality and responsibility," and my recent book with Vladimir Vovk elaborates a foundation for probability theory that can be used to support the predictive approach.

The reader might suspect that the predictive approach and the counterfactual approach say the same thing in different ways. The advocates of the counterfactual approach insist, however, on points that cannot be reconciled with the predictive approach. They begin by insisting on the word *counterfactual*. The very word places us in the situation where A has already been performed and so not(A) is counter to the facts. The counterfactual theory insists that there should be a well-defined answer, in this situation, to the question of what would have happened if A had not been performed. The predictive theory, on the other hand, considers only what can be predicted before the choice between A and not(A) is made. Later, this situation will be in the past, but it will never be in the subjunctive. If no definite prediction is possible about whether B will happen if A is not performed—if only probabilities can be given or not even that—and then A is performed, then there will be no answer as to whether B would have happened had A not been performed.

At the end of their introduction, the authors indicate that they are willing to consider probabilities: "under a stochastic model, the quantities we discuss are probabilities or expected values". They then cite two articles by one of the authors, Sander Greenland. They go out of their way, however, to deny causal meaning to the consistency across populations that would be needed to make probabilistic predictions meaningful. I have not been able to understand how the articles by Greenland resolve this contradiction.

Here are some comments that may broaden the picture painted by the authors' citations of literature on counterfactuals outside statistics and epidemiology. David Hume's counterfactual definition was only one of several definitions of cause that he formulated in *An Enquiry Concerning Human Understanding*. David Lewis, a philosopher at Princeton University, is cited as developing the counterfactual definition of causality currently used in the statistics literature, but in conversations with myself and other statisticians Professor Lewis has repeatedly disavowed this interpretation of his work, and during decades following the 1973 book cited here, he and his students have published numerous articles devoted to developing an empirical understanding of causality that would be consistent with the predictive approach I have sketched. The authors quite appropriately cite two physicists who favor the counterfactual approach, but their confident assertion that counterfactual analysis cuts through the fog in physics, juxtaposed with the name of Richard Feynman, should not be allowed to obscure the fact that Feynman never advocated the counterfactual approach and that many physicists explicitly oppose it; see for example Layzer (1991).

David Layzer (1991) *Cosmogenesis : The Growth of Order in the Universe*. Oxford University Press.
Glenn Shafer (1996). *The Art of Causal Conjecture*. MIT Press.
Glenn Shafer (2001). Causality and responsibility. Cardozo Law Review.
Glenn Shafer and Vladimir Vovk (2001). *Probability and Finance: It's Only a Game.* Wiley.