Commentary: Population versus individual level causal effects

Felix Elwert and Christopher Winship

We congratulate Maldonado and Greenland (MG henceforth) on an interesting and provocative paper. Aiming at epidemiological applications, MG identify the causal effect of changing a distribution of exposures to a target population on the population’s outcome distribution. Instead of applying a particular treatment to an individual, MG apply a distribution of treatments (the exposure distribution) to a population. By raising the unit of analysis from the individual to the population, MG depart in important respects from the standard model of counterfactual causal inference. Comparing MG’s model to the standard model we make two points: First, MG’s conceptualization of causal effects on the population level is valuable if the stable unit-treatment assumption (SUTVA) does not hold at lower levels, but the data requirements are steep. Second, as

In closing, we express our congratulations to Maldonado and Greenland for this contribution to the literature. Awareness of the foundations of causal inference in epidemiology has increased in recent years, and this is due in large part to the diligent efforts of Sander Greenland, James Robins, and their students. The present paper serves to provoke further discussion and insight, and to instruct a wider audience of epidemiologists. Through this ongoing process, we benefit our understanding thereby improving our science, and thus, our capacity to intervene upon and improve human health.

References

4 Copas JB. Randomization models for the matched and unmatched 2 x 2 tables. Biometrika 1973;60:467–76.
they mention, we emphasize that MG’s population level estimates generally cannot be interpreted as estimates of average causal effects (ACE) in the standard individual-level approach.

**Individual-level causal effects**

We remind the reader of the standard individual-level presentation of the counterfactual model of causal inference, also known as the Rubin Model. Here, a particular treatment, \( t \), is applied to a unit of analysis, \( i \), (e.g. a person). The causal effect of \( i \) on \( i \), \( \delta_i \), is defined as the difference between the outcome of the unit under treatment, \( Y(t)_i \), and the outcome of the same unit under control, \( Y(\bar{c})_i \).

\[
\delta_i = Y(t)_i - Y(\bar{c})_i.
\]

The ‘fundamental problem of causal inference’ is that \( Y(t)_i \) and \( Y(\bar{c})_i \) cannot be directly observed together, because every unit of analysis is placed either in treatment or in control condition, but not in both at the same time. Therefore direct estimation of causal effects is impossible. As in MG, the solution is to substitute for the counterfactual observation another unit of analysis, \( j \), which resembles \( i \) in all causally relevant respects other than treatment status.

Typically, we are not interested in the causal effect for a specific individual, but rather the average causal effect, ACE, in the study population:

\[
ACE = \frac{1}{n} \sum_{i=1}^{n} \delta_i = \bar{Y}(t) - \bar{Y}(\bar{c}) \text{ for } i = 1, ..., n.
\]

In completely randomized experiments, the standard estimator for this parameter subtracts the mean outcome of the units in the treatment group from the mean outcome of the units in the control group:

\[
AC\bar{E} = \bar{Y}(t) - \bar{Y}(\bar{c})
\]

This approach assumes that there is no interaction between units and that all treated units in the study receive identical treatments. Rubin terms this the ‘stable unit-treatment value assumption’ (SUTVA). The key virtue of randomization is to create balanced treatment and control groups that resemble each other across all causally relevant variables except treatment status. Techniques such as matching on propensity scores are available to achieve balance even in non-randomized observational studies.

**Population-level causal effects: utility and data requirements**

MG’s framework applies exposure distributions to target populations. Consequently, their unit of analysis is the population. This approach has merit, particularly when SUTVA does not hold within the population. Such situations occur frequently, e.g. in educational research where student test scores may be affected by tutoring their classmates received. Here one would want to use classes for units of analysis, rather than students.

Note, however, that the higher the unit of analysis, the more challenging the data requirements due to comparability of units of analysis, and identity of treatments.

The counterfactual model relies on the comparison of units of analysis that resemble each other in all causally relevant aspects except treatment status. To continue our educational example on the population (classroom) level, it would be necessary to find comparable classes, rather than comparable students. If SUTVA does not hold, this would not only involve comparable student populations, but also comparable dependencies between students within classes in order to ensure comparable peer effects.

The standard model further assumes that all units in the treatment group receive identical treatments. (Note that in a population level analogy to the standard individual-level model, a treatment group contains multiple target populations as units of analysis, each of which contains multiple individuals. Comparing a single target population to a single substitute would amount to working with a sample of \( N = 2 \).) If the treatment in question is an exposure distribution, as MG stipulate, identity of treatments across units (i.e. target populations) becomes much harder to assert. It depends on two aspects: (1) the exposure distribution’s marginal distribution, which records the relative frequency of exposure levels within a target population; and (2) the mapping of distinct exposures from the exposure distribution onto individuals within a target population. If the population is heterogeneous in its members, different mappings of the same exposure distribution will induce different outcomes. Thus, to assure identity of treatments, both the marginal exposure distribution and its mapping have to be held constant across target populations in the treatment group. Due to these challenges, it seems advisable to choose the smallest unit for which SUTVA still holds as unit of analysis.

**Dissimilarity of population-level causal contrasts and average causal effects**

MG remark that ‘not all population causal contrasts can be interpreted as averages of individual causal effects of exposure’ (p.1039 in their paper). We would like to go further and argue that MG’s population-level estimates will hardly ever represent average individual-level causal effects, because their approach generally does not sustain the conditions of a standard individual-level counterfactual analysis.

An example of the causal effect of smoking on lung cancer may convey the guiding intuition. Consider a population of 1000 men. Of these, 40% are highly susceptible to smoking-induced lung cancer and smoke, and 60% are minimally susceptible to cancer and do not smoke. The rate of lung cancer in this population is 40%. We want to estimate the effect of a change in the exposure distribution from 40% to 60% ever-smokers (similar to MG’s example on p.1039).

We identify a perfect substitute population of 1000 other men, 600 of whom smoke. However, all of these smokers are only minimally susceptible to lung cancer. In this population the cancer rate is 1%. MG’s measure of causal contrast would indicate that increasing the exposure to smoking has decreased the incidence of lung cancer, even though each individual member would suffer an increased risk of cancer by taking up smoking. The reason is that different individuals smoke in the two populations.

This result makes sense in MG’s approach, because it accurately identifies the population-level causal effect of having changed...
both the exposure distribution's marginal distribution and its mapping onto the target population. In the individual-level approach this result would be impossible, because the ACE cannot be negative if all $\delta_i$ are positive. MG's population level estimates and the standard individual-level ACE are not equivalent.

References


Commentary: Estimating causal effects

Glenn Shafer

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 causal 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, I, 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.