

Invited Commentary: Causal Inference Across Space and Time—Quixotic Quest, Worthy Goal, or Both?

Jessie K. Edwards*, Catherine R. Lesko, and Alexander P. Keil

* Correspondence to Dr. Jessie K. Edwards, University of North Carolina, Department of Epidemiology, 2101 McGavran-Greenberg Hall, CB 7435, Chapel Hill, NC 27599 (e-mail: jessedwards@unc.edu).

Initially submitted February 27, 2017; accepted for publication March 17, 2017.

The g-formula and agent-based models (ABMs) are 2 approaches used to estimate causal effects. In the current issue of the *Journal*, Murray et al. (*Am J Epidemiol.* 2017;186(2):131–142) compare the performance of the g-formula and ABMs to estimate causal effects in 3 target populations. In their thoughtful paper, the authors outline several reasons that a causal effect estimated using an ABM may be biased when parameterized from at least 1 source external to the target population. The authors have addressed an important issue in epidemiology: Often causal effect estimates are needed to inform public health decisions in settings without complete data. Because public health decisions are urgent, epidemiologists are frequently called upon to estimate a causal effect from existing data in a separate population rather than perform new data collection activities. The assumptions needed to transport causal effects to a specific target population must be carefully stated and assessed, just as one would explicitly state and analyze the assumptions required to draw internally valid causal inference in a specific study sample. Considering external validity in important target populations increases the impact of epidemiologic studies.

agent-based models; causal inference; decision analysis; individual-level models; mathematical models; medical decision making; Monte Carlo methods; parametric g-formula

Abbreviation: ABM, agent-based model.

Epidemiology guides decisions related to public health, which are necessarily made with respect to specific target populations. Such decisions may involve choosing between 2 (or more) alternative intervention strategies, one of which may be to continue under the status quo. Ideally, these decisions should optimize the expected intervention impact in the group of people who would be affected by the decision (the target population).

With information on the counterfactual outcome distributions under the candidate interventions in the target population, it is trivial for a decision maker to choose between intervention strategies by maximizing some utility function (1). Counterfactual outcome distributions are not directly observable; 2 approaches to estimate counterfactual outcome distributions from epidemiologic data include the parametric g-formula (2) and ABMs (3).

We commend Murray et al. (4) for their thoughtful article probing the differences in these approaches. In their notable work, the authors demonstrated that applying the g-formula to data from the target population of interest outperformed an ABM that was parameterized at least partially from other sources. Murray et al. show that both approaches work well when parameterized entirely using data from the target population (their “base scenario”). This is unsurprising because, as the authors point out, the same parametric models can be used for the parametric g-formula and an ABM, and the 2 approaches are mathematically equivalent when all parameters are estimated in a single data source. However, in external target populations defined as high or low risk, the ABM, which estimated parameters entirely or partially within the base scenario, suffered from a set of limitations outlined in their work. Crucially, the performance of the g-formula in

the external target populations was judged after refitting the g-formula to new data from each target population.

The analysis of Murray et al. was enlightening with respect to the performance of each approach according to how they are commonly used. However, comparing ABMs and the g-formula in this way may conflate the performance of the methods with the availability of data in settings in which they are typically applied. As the authors note, the g-formula is typically applied to estimate causal effects in a target population with complete data, while ABMs often use information from one or more study samples to estimate a causal effect in an external target population (for which complete data are not available). It is not surprising that, given a choice between data sets, more complete data will usually yield more accurate results.

With complete information on confounders, exposure, and outcome in the target population, users of both g-formula and ABM approaches would likely leverage this information to estimate causal effects. In the absence of complete information on the target population, Murray et al. aptly describe the limitations of the ABM when used to make inferences across populations. The authors posit that the g-formula is not as susceptible to concerns about transportability because inferences from the parametric g-formula are typically limited to populations and time periods similar to the population and time period for which complete data are available. Of course, this is to say that the g-formula is not susceptible to concerns about transportability because it is not often used for transportability. How, then, are we to learn about the effects of interventions in populations for which we have incomplete data?

For epidemiology to be useful to decision makers, we, as epidemiologists, cannot abdicate a responsibility to provide evidence to inform necessary decisions in target populations. Public health decisions are often somewhat urgent—for example, delaying a decision in order to collect complete information in the target population may leave an unsatisfactory status quo in place. The cost of delaying a decision can be measured by reductions in quality of life or life span. Thus, someone making a decision related to public health must weigh the costs of collecting additional data (human and monetary) against the costs and probability of making a suboptimal choice using incomplete evidence.

Transporting causal effects has long been the domain of agent-based modelers striving to inform decision making by estimating the effects of real or imagined interventions in target populations using inputs from multiple data sources. Although ABMs suffer from limitations that arise from obtaining estimates from a mix of data sources, they have provided answers to questions on which traditional epidemiology was often silent. In contrast, traditional epidemiologic studies have typically ended at estimation of an internally valid causal effect for a study sample with complete data, rather than estimation of an externally valid causal effect for a target population of interest.

As Murray et al. highlighted, an internally valid causal effect estimate from a study or trial in one population will not necessarily equal the expected causal effect in another population (5–7). Happily, recent work has focused on articulating the data and assumptions required to generalize or transport results from a single epidemiologic study to a separate target

population (5, 6, 8). This work has given rise to a suite of methods to generalize or transport results from trials or observational studies to external target populations under a set of conditions. Methods such as the “transport formula” (7) and a variety of reweighting approaches have been developed in recent years to transport or generalize results from study samples that are not representative of the target population (9–13).

Transportability requires data and assumptions beyond those needed for internal validity (5–8, 14, 15). As the authors note, obtaining information needed to transport a causal effect, including the additional covariates needed in the study sample and all necessary covariates in the target population, may sometimes be difficult or impossible. However, some options are available. For some scientific questions in some target populations, one can obtain information on the joint distribution of covariates in the target population from publicly available databases, such as census data, the US Agency for International Development (USAID)’s Demographic and Health Surveys, the Center for Disease Control and Prevention (CDC)’s Behavioral Risk Factor Surveillance System, or chronic disease registries. With the rise of “big data,” epidemiologists may be able to take advantage of emerging data streams as sources of information on variables important for transporting causal effects for some causal questions. In the absence of information on one or more covariates in the target population, sensitivity analyses could be conducted by simulating different plausible distributions of these variables to estimate best- and worst-case scenarios for how different the causal effects expected in the target population may be from those observed in a separate study sample.

The assumptions required for transportability are strong (6, 8, 14). Indeed, one may question the value of any outputs from such an exercise. But what are the alternatives?

Without full data from the target population, the epidemiologist could restrict inference to the population for which she has data. As Murray et al. demonstrate, limiting inference to the population and time period represented by a specific study obviates the need for additional assumptions required to transport causal effects (and protects the epidemiologist from bias that can arise if they are not met). However, restricting inference to populations and time periods for which complete study data already exist limits the utility and relevance of epidemiology. Epidemiology should be more than a historical exercise describing causal effects that happened in a specific study population in the past. In much of epidemiology, the purpose of understanding causal effects is to inform future decisions, which requires transporting our results, at the very least, to a future version of our study population (16).

If we, as epidemiologists, absolve ourselves of responsibility for inference in the absence of complete data in the target population, then the duty falls to decision makers who may be left with no alternative but to implicitly transport published causal effects to target populations by assuming that the causal effect is constant across populations. Or the published effect estimates may be informally discounted or inflated for the target population to vaguely reflect some prior knowledge. In contrast to informal approaches, quantitative methods to transport causal effects lend rigor and transparency to the process of making inferences outside a given study population.

Epidemiologic data are (almost) always imperfect and incomplete (17). Just as we explicitly state the assumptions we make about unmeasured confounding to draw inference about causal effects in a specific study sample, we should also be explicit about the assumptions we make to infer causal effects in specific populations external to our studies. We can only fully evaluate the utility of epidemiologic results for improving health by gauging how well our work meets all of these assumptions. Without such intellectual rigor, we are left with either limiting inferences to settings with complete data or informal—and possibly misguided—attempts to broadly apply results that are highly dependent on the context in which they were derived. A valuable lesson of the results reported by Murray et al. (4) is that users of ABMs should consider the assumptions necessary to transport and synthesize findings across populations to improve the validity of their results. Conversely, those who are already steeped in the assumptions of causal inference can increase the impact of their results by generalizing or transporting causal effects to settings where they may be used as inputs into decisions rather than simply estimating internally valid causal effects where complete data exist (18).

Decision makers seek “evidence-based interventions” (19). The most straightforward way to provide evidence about a decision in a target population is to conduct a trial in that population or use statistical approaches that emulate a trial using observational data from that population, such as the g-formula (20). But the resources required to conduct such studies are finite and not optimally distributed. Thus, decisions about public health interventions often require transporting causal effects from one population to another as a way to inform public health actions in the face of incomplete information. Epidemiologists can improve public health by engaging in this process.

ACKNOWLEDGMENTS

Author affiliations: Department of Epidemiology, Gillings School of Global Public Health, University of North Carolina, Chapel Hill, North Carolina (Jessie K. Edwards, Alexander P. Keil); and Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland (Catherine R. Lesko).

This work was supported in part by the National Institutes of Health (grants K01AI125087, DP2-HD-08-4070, U01 HL121812, and U24 OD023382).

Conflict of interest: none declared.

REFERENCES

1. Aristotle. *Topica*. Forster ES, trans. Cambridge, MA: Harvard University Press; 1989.

2. Robins J. A new approach to causal inference in mortality studies with a sustained exposure period—application to control of the healthy worker survivor effect. *Math Model*. 1986;7(9–12):1393–1512.
3. Beck JR, Pauker SG. The Markov process in medical prognosis. *Med Decis Making*. 1983;3(4):419–458.
4. Murray EJ, Robins JM, Seage GR 3rd, et al. A comparison of agent-based models and the parametric g-formula for causal inference. *Am J Epidemiol*. 2017;186(2):131–142.
5. Hernán MA, VanderWeele TJ. Compound treatments and transportability of causal inference. *Epidemiology*. 2011;22(3):368–377.
6. Lesko C, Buchanan AL, Westreich D, et al. Generalizing study results: a potential outcomes perspective [published online ahead of print March 24, 2017]. *Epidemiology*. (doi:10.1097/EDE.0000000000000664).
7. Bareinboim E, Pearl J. A general algorithm for deciding transportability of experimental results. *J Causal Inference*. 2013;1(1):107–134.
8. Bareinboim E, Pearl J. Causal inference and the data-fusion problem. *Proc Natl Acad Sci USA*. 2016;113(27):7345–7352.
9. Cole SR, Stuart EA. Generalizing evidence from randomized clinical trials to target populations: the ACTG 320 trial. *Am J Epidemiol*. 2010;172(1):107–115.
10. Buchanan A, Hudgens M, Cole S, et al. Generalizing evidence from randomized trials using inverse probability of sampling weights (Paper 77). *Pharmacy Practice Faculty Publications*. 2016.
11. Westreich D, Edwards JK, Stuart EA, et al. Transportability of trial results using inverse odds of sampling weights. *Am J Epidemiol*. In press.
12. Bengtson AM, Pence BW, Gaynes BN, et al. Improving depression among HIV-infected adults: transporting the effect of a depression treatment intervention to routine care. *J Acquir Immune Defic Syndr*. 2016;73(4):482–488.
13. Lesko CR, Cole SR, Hall HI, et al. The effect of antiretroviral therapy on all-cause mortality, generalized to persons diagnosed with HIV in the USA, 2009–11. *Int J Epidemiol*. 2016;45(1):140–150.
14. Pearl J, Bareinboim E. External validity: from do-calculus to transportability across populations. *Stat Sci*. 2014;29(4):579–595.
15. Petersen ML. Compound treatments, transportability, and the structural causal model: the power and simplicity of causal graphs. *Epidemiology*. 2011;22(3):378–381.
16. Hoggatt KJ, Greenland S. Commentary: extending organizational schema for causal effects. *Epidemiology*. 2014;25(1):98–102.
17. Edwards JK, Cole SR, Westreich D. All your data are always missing: incorporating bias due to measurement error into the potential outcomes framework. *Int J Epidemiol*. 2015;44(4):1452–1459.
18. Westreich D, Edwards JK, Rogawski ET, et al. Causal impact: epidemiological approaches for a public health of consequence. *Am J Public Health*. 2016;106(6):1011–1012.
19. Victora CG, Habicht JP, Bryce J. Evidence-based public health: moving beyond randomized trials. *Am J Public Health*. 2004;94(3):400–405.
20. 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.