

Is the Smog Lifting?

Causal Inference in Environmental Epidemiology

W. Dana Flanders^{a,b} and Michael D. Garber^a

We thank the Editor for this opportunity to comment on the accompanying article by Pearce et al.¹ In their commentary, Pearce et al. discuss methods that they argue have been and should be useful for making causal inferences in environmental epidemiology. They also cite part of an ongoing debate in the literature (their references 6–12) that criticizes some causal inference methods in epidemiology. In this commentary, we first point out many areas of agreement with Pearce et al. We then comment on several issues raised by their commentary. One issue concerns omission of important references, and four issues concern claims that were poorly supported or potentially misleading. We also noted how, with the consideration of a wider scope of methods and more nuanced interpretation of recommendations regarding emulation of randomized controlled trials (RCTs), many newer causal inference methods have been and should continue to be quite useful in environmental epidemiology, and, in fact, have been used to justify many of the methods Pearce et al. suggested.

Six main points of this article, at least at heart, are clear, documented, and convincing. In simple terms, these points are the following: (1) causal inference for certain questions in environmental epidemiology, such as studying the health effects of climate change, is difficult; (2) traditional methods have worked quite well for certain problems in environmental epidemiology; (3) additionally, at least five specific other methods (extensions) should be helpful in environmental epidemiology; (4) many methods can be useful in environmental epidemiology; (5) related to fourth point, the subgroup of methods as defined by Pearce et al. (RCT mimicking set of “causal inference” methods) does not include everything that might be useful; and (6) triangulation methods are one such additional method. The first and second points are supported by examples; the third point is also supported by examples for each method noted. Interestingly, all of the “extensions” have benefited by considerations from the causal inference tool box, as we elaborate below. The fourth and fifth points are perhaps self-evident, as we know of no subgroup of methods that covers everything or claims to. The last point is essentially an extension of the well-accepted Bradford Hill considerations, with additional considerations about the directions of potential bias. We agree with these six points.

The first issue we highlight concerns the omission of important references that are part of an ongoing debate about the merits of certain aspects of modern causal inference concepts, methods, and tools. Although we welcome the largely muted level of criticism in this commentary, Pearce et al. cite only one side (their references 6–12) of this debate, omitting many responses to and commentaries on these criticisms.^{2–9} To varying degrees,

Editor's Note: A related commentary appears on p. 311.

From the ^aDepartment of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA; and ^bDepartment of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA.

Correspondence: W. Dana Flanders, Department of Epidemiology, Rollins School of Public Health, Emory University, 1518 Clifton Road NE, Atlanta, GA 30322. E-mail: wflande@emory.edu. W.D.F. owns Epidemiologic Research & Methods, LLC, which does consulting work for pharmaceutical companies, environmental laboratories, and attorneys. The other author has no conflicts to report.

Copyright © 2019 Wolters Kluwer Health, Inc. All rights reserved.

ISSN: 1044-3983/19/3003-0317

DOI: 10.1097/EDE.0000000000000986

these responses and commentaries have clarified positions, provided strong rebuttals of certain criticisms, explained misunderstandings, pointed out a straw man fallacy, and offered conciliatory remarks.

Our second issue concerns Pearce et al.'s unqualified claim that "the term 'causal inference' is being used to denote a specific set of newly developed methods . . .," characterized as the "... RCT mimicking set of 'causal inference' methods, in contrast to the broader field of causal inference of which it is a part." This claim is unsupported and potentially misleading because, in the context of the ongoing debate, it could be taken to suggest that those who have contributed to modern methods of causal inference and who are also involved in the ongoing debate used the term "causal inference" in this way. On the contrary, most, likely all, who have contributed to modern methods of causal inference and who are also involved in the ongoing debate about the merits of certain causal inference methods, do not, in general, use the term "causal inference" to refer only to the restricted, narrow subgroup of methods described by Pearce et al. as "RCT mimicking" (Specifically, Greenland, Hernán, Pearl, Robins, and VanderWeele (alphabetically), all contributors to modern causal inference methods and involved in the ongoing debate, have used causal inference to refer to a much broader range of methods and thus do not generally use the term in the restricted way described by Pearce et al. Here are a few examples. In their book *Causal Inference*,² Hernán and Robins discuss the "context in which observational studies *cannot* often be conceptualized as conditionally randomized experiments . . ." Hernán also writes "Causal inference relies on transparency of assumptions and on *triangulation* of results from methods that depend on different sets of assumptions."¹⁰ Greenland coauthors the book *Modern Epidemiology*.¹¹ Chapter 2 on causation and causal inference includes an overview of the philosophy of scientific inference, with causal inference as a special case. Also included are the Bradford Hill "criteria." Bareinboim and Pearl¹² tackled the problems of combining different sources of information (data fusion) and addressing biases like confounding and selection bias to make causal inferences. Finally, VanderWeele has coauthored articles using methods, such as Mendelian randomization and meta-analysis for causal inference.¹³ He also states that "Inference to be best explanation is important in causal inference and diverse types of evidence can and should be used."⁷ He explicitly includes under the causal inference umbrella: the instrumental variable, regression discontinuity, and difference-in-difference methods.¹⁴) In other words, the definition put forth by Pearce et al. is consistent with neither the language used by these key contributors to modern methods, nor with the entirety of methods that those contributors have developed, used, and cited in their work.

Our third issue concerns Pearce et al.'s claim that: "This [modern causal inference movement, as defined by PVL] proposes that observational studies should mimic key aspects of randomized trials, since this allows them to be rooted in

counterfactual reasoning, which is said to formalize the natural way that humans think about causality." This claim is also potentially misleading because, in view of the ongoing debate²⁻⁹ and without qualification, the claim might be read as implying that all observational studies should mimic key aspects of randomized trials. Worse yet, it might be read as implying that it represented the position of those involved in the debate. These interpretations would be incorrect. We agree that Hernán strongly advocates attempting to emulate randomized experiments as a device to aid study design, a position justified in part because such emulation can help sharpen effect definitions as counterfactual contrasts of better-defined, possibly hypothetical interventions, thereby helping to reduce the vagueness in causal questions.^{4,15,16} But with a broader reading of Hernán's work and that of other causal inference contributors, three important additional observations emerge. First, these causal inference authors are making a conditional claim: if such emulation is possible, then many advantages will likely accrue. For example, as just noted the definition of the causal effect of interest may be clarified; additionally, an intervention of potential public health utility might be better identified and evaluated, and certain properties that can contribute to valid causal inference including exchangeability, consistency, and positivity might be better defined and more evaluable.^{2,15} See also Daniel et al.⁹ and Petersen and van der Laan.¹⁷ Second, many causal inference authors note that in some contexts (e.g., system-wide interventions), it may not be possible to conceptualize an RCT that can be usefully emulated by an observational study as part of the causal inference process.^{2,6,14,18} Third, these authors note that if one cannot emulate an RCT, then other methods are available and may apply, including those mentioned by Pearce et al. as useful in environmental epidemiology, such as instrumental variables, Mendelian randomization, and regression discontinuity and difference-in-differences designs.^{2,10,11,13,19}

As an alternative to the restrictive definition of causal inference methods used by Pearce et al., we suggest, in broad agreement with their overall message, that causal inference methods correspond to what the phrase suggests and should include any, useful, valid method, many of which are already being used in environmental epidemiology. The list would include the methods listed by Pearce et al. and others, some not mentioned or emphasized, such as G-computation, G-estimation, marginal structural models,²⁰⁻²² probabilistic causal models,²³ structural equation models, combining information from multiple sources in a structural equation framework,¹² negative control outcomes and exposures²⁴ with specific applications to air pollution and environmental epidemiology,²⁵⁻²⁷ directed acyclic graphs (DAGs), simulation modeling, and more. We specifically note the important contribution of causal graphs,²⁸ and in particular DAGs.²⁹ They were used in the development of methods to detect confounding with examples from environmental epidemiology in mind.²⁵⁻²⁷ More widely, they are commonly used to describe causal relationships and

might be viewed as providing a language for doing so efficiently. As with good notation and language, DAGs can aid in the thought process.³⁰ In addition to the noted observational approaches, simulations, projection studies, and agent-based modeling^{18,31,32} can be useful in environmental epidemiology, although more work remains in delineating and understanding conditions for valid effect estimation.^{33,34} In research on climate change and health, the projection of estimated effects under varying emissions and socioeconomic scenarios has become more common.^{35–37} These types of studies are useful for projecting the health burden attributable to climate change, but results are, of course, dependent on the initial estimate of effects.

Our fourth issue concerns Pearce et al.'s statement that: "We are not arguing that 'causal inference' methods that mimic randomized controlled trials are not useful; for example, they can improve individual studies with individual-level exposures that can be seen as interventions." This claim is also potentially misleading because it could be read as implying or suggesting that studies that emulate RCTs must involve randomization of individual-level exposures. In fact, the emulation applies more widely—one need to only consider group-randomized trials and appreciate that neither randomization nor the potential outcome framework precludes group-level exposure. This wider appreciation is particularly relevant for environmental epidemiology because, as Pearce et al. note, environmental exposures often "... affect individuals across entire communities."

Interestingly, many of the causal inference methods (e.g., "extensions of traditional approaches") specifically mentioned by Pearce et al. have benefited from and have been justified by some of the modern causal inference concepts, methods, and tools. For example, Balke and Pearl³⁸ used counterfactual models to derive identifiable limits for the magnitude of effect based on analyses of instrumental variables. Although the difference-in-differences design precedes the formulation of the more recent causal inference concepts, the major identifying assumption of difference in differences—parallel trends—appears now to be generally understood in econometrics as based on a counterfactual approach (see, e.g., Abadie³⁹ or Lechner⁴⁰). Moreover, the regression discontinuity design can be based on the assumption that potential outcomes have a continuous distribution at the threshold.⁴¹ Indeed, Imbens and Wooldridge, important contributors to the econometrics literature where instrumental variable analysis, difference in differences, and regression discontinuity have long been used, stated that: "... the Rubin potential outcomes framework is now the dominant framework."⁴²

In summary, we agree with much of Pearce et al.'s basic message as contained in at least six simplified main points. However, they omitted entirely the responses to and commentaries on the debate they cited. They also set up a restricted definition of "causal inference," claiming without qualification that it refers to a subset of RCT mimicking set of causal inference methods

and proceeded to make additional, unqualified characterizations about the subset they defined. Further, they omitted discussion of how RCT emulation can apply to group-level exposures although they note the importance of this in environmental epidemiology. These omissions, characterizations, and failure to qualify or note the conditional nature of potential advantages if a randomized trial can be emulated are, perhaps, unintentional. Regardless of intentionality, these distractions are unfortunate because the broad message—environmental epidemiology will benefit from use of a variety of causal inference methods—is one with which few would disagree including VanderWeele et al.,⁵ who pointed out these benefits more widely.

ACKNOWLEDGMENTS

We thank Tyler VanderWeele and Miguel Hernán for their helpful comments.

ABOUT THE AUTHORS

W. DANA FLANDERS is a professor of Epidemiology at Rollins School of Public Health. He teaches epidemiologic methodology and does research in several areas, including epidemiologic methodology, environmental epidemiology, and cancer epidemiology.

MICHAEL D. GARBER is a PhD student in Epidemiology at Rollins School of Public Health. He is interested in the effects of the environment on physical activity and injury and in epidemiologic methods.

REFERENCES

1. Pearce N, Vandenbroucke JP, Lawlor DA. Causal inference in environmental epidemiology: old and new. *Epidemiology*. 2019;30:311–316.
2. Hernán MA, Robins J. *Causal Inference*. Boca Raton, Fla: Chapman and Hall/CRC; 2018: In press.
3. Greenland S. For and against methodologies: some perspectives on recent causal and statistical inference debates. *Eur J Epidemiol*. 2017;32:3–20.
4. Hernán MA. Does water kill? A call for less casual causal inferences. *Ann Epidemiol*. 2016;26:674–680.
5. VanderWeele TJ, Hernán MA, Tchetgen Tchetgen EJ, Robins JM. Re: causality and causal inference in epidemiology: the need for a pluralistic approach. *Int J Epidemiol*. 2016;45:2199–2200.
6. Robins JM, Weissman MB. Commentary: counterfactual causation and streetlamps: what is to be done? *Int J Epidemiol*. 2016;45:1830–1835.
7. VanderWeele TJ. Commentary: on causes, causal inference, and potential outcomes. *Int J Epidemiol*. 2016;45:1809–1816.
8. VanderWeele TJ. On well-defined hypothetical interventions in the potential outcomes framework. *Epidemiology*. 2018;29:e24–e25.
9. Daniel RM, De Stavola BL, Vansteelandt S. Commentary: the formal approach to quantitative causal inference in epidemiology: misguided or misrepresented? *Int J Epidemiol*. 2016;45:1817–1829.
10. Swanson SA, Hernán MA. Commentary: how to report instrumental variable analyses (suggestions welcome). *Epidemiology*. 2013;24:370–374.
11. Rothman, K.J., Greenland, S. and Lash, T.L., *Modern epidemiology*. (Vol. 3). Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2008.
12. Bareinboim E, Pearl J. Causal inference and the data-fusion problem. *Proc Natl Acad Sci U S A*. 2016;113:7345–7352.
13. Song Y, Yeung E, Liu A, et al. Pancreatic beta-cell function and type 2 diabetes risk: quantify the causal effect using a Mendelian randomization approach based on meta-analyses. *Hum Mol Genet*. 2012;21:5010–5018.
14. VanderWeele TJ, Mathur MB, Chen Y. Outcome-wide longitudinal designs for causal inference: a new template for empirical studies. arXiv preprint arXiv:181010164 2018.

15. Hernán MA, Taubman SL. Does obesity shorten life? The importance of well-defined interventions to answer causal questions. *Int J Obes (Lond)*. 2008;32(suppl 3):S8–S14.
16. Hernán MA. Invited commentary: hypothetical interventions to define causal effects—afterthought or prerequisite? *Am J Epidemiol*. 2005;162:618–620.
17. Petersen ML, van der Laan MJ. Causal models and learning from data: integrating causal modeling and statistical estimation. *Epidemiology*. 2014;25:418–426.
18. Hernán MA. Invited commentary: agent-based models for causal inference—reweighting data and theory in epidemiology. *Am J Epidemiol*. 2014;181:103–105.
19. Pearl J. Does obesity shorten life? Or is it the soda? On non-manipulable causes. *J Causal Inference*. 2018;6.
20. Robins J. A new approach to causal inference in mortality studies with sustained exposure periods— application to control of the health worker survivor effect. *Math Model*. 1986;7:1393–1515.
21. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology*. 2000;11:550–560.
22. Robins JM, Blevins D, Ritter G, Wulfsohn M. G-estimation of the effect of prophylaxis therapy for *Pneumocystis carinii* pneumonia on the survival of AIDS patients. *Epidemiology*. 1992;3:319–336.
23. Pearl J. *Causality*. 2nd ed. Cambridge: Cambridge University Press; 2009.
24. Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology*. 2010;21:383–388.
25. Flanders WD, Klein M, Darrow LA, et al. A method to detect residual confounding in spatial and other observational studies. *Epidemiology*. 2011;22:823–826.
26. Flanders WD, Klein M, Strickland M, et al. A method of identifying residual confounding and other violations of model assumptions. *Epidemiology*. 2009;20:S44–S45.
27. Flanders WD, Klein M, Darrow LA, et al. A method for detection of residual confounding in time-series and other observational studies. *Epidemiology*. 2011;22:59–67.
28. Robins JM, Richardson T. Alternative graphical causal models and the identification of direct effects. In: Shrouf P, Keyes K, Ornstein K, eds. *Causality and Psychopathology: Finding the Determinants of Disorders and Their Cures*, Chapter 6. Oxford, UK: Oxford University Press; 2011:103–158.
29. Pearl J. Causal diagrams for empirical research (with discussion). *Biometrika*. 1995;82:669–710.
30. Iverson KE. Notation as a tool of thought. *ACM SIGAPL APL Quote Quad*. 2007;35:2–31.
31. Auchincloss AH, Diez Roux AV. A new tool for epidemiology: the usefulness of dynamic-agent models in understanding place effects on health. *Am J Epidemiol*. 2008;168:1–8.
32. Marshall BD, Galea S. Formalizing the role of agent-based modeling in causal inference and epidemiology. *Am J Epidemiol*. 2015;181:92–99.
33. Murray EJ, Robins JM, Seage GR, Freedberg KA, Hernán MA. A comparison of agent-based models and the parametric g-formula for causal inference. *Am J Epidemiol*. 2017;186:131–142.
34. Keyes KM, Tracy M, Mooney SJ, Shev A, Cerdá M. Invited commentary: agent-based models—bias in the face of discovery. *Am J Epidemiol*. 2017;186:146–148.
35. Ebi KL, Hallegatte S, Kram T, et al. A new scenario framework for climate change research: background, process, and future directions. *Clim Change*. 2014;122:363–372.
36. Sellers S, Ebi KL. Climate change and health under the shared socioeconomic pathway framework. *Int J Environ Res Public Health*. 2017;15:3.
37. Springmann M, Mason-D’Croz D, Robinson S, et al. Global and regional health effects of future food production under climate change: a modelling study. *Lancet*. 2016;387:1937–1946.
38. Balke A, Pearl J. *Counterfactual Probabilities: Computational Methods, Bounds and Applications*. Proceedings of the Tenth International Conference on Uncertainty in Artificial Intelligence, Seattle, Wash, 29–31 July 1994. San Francisco, Calif: Morgan Kaufmann Publishers Inc.; 1994:46–54.
39. Abadie A. Semiparametric difference-in-differences estimators. *Rev Econ Stud*. 2005;72:1–19.
40. Lechner M. The estimation of causal effects by difference-in-difference methods. *Found Trends Econometrics*. 2011;4:165–224.
41. Moscoe E, Bor J, Bärnighausen T. Regression discontinuity designs are underutilized in medicine, epidemiology, and public health: a review of current and best practice. *J Clin Epidemiol*. 2015;68:122–133.
42. Imbens GW, Wooldridge JM. Recent developments in the econometrics of program evaluation. *J Econ Lit*. 2009;47:5–86.