Title

Invited Commentary: Making Causal Inference More Social and (Social) Epidemiology More Causal.

Journal

American Journal of Epidemiology, 189(3)

ISSN

0002-9262

Authors

Jackson, John W
Arah, Onyebuchi A

Publication Date

2020-03-02

DOI

10.1093/aje/kwz199

Peer reviewed

American Journal of Epidemiology

Vol. 189, No. 3
DOI: 10.1093/aje/kwz199
Advance Access publication:
October 1, 2019

Invited Commentary

Invited Commentary: Making Causal Inference More Social and (Social) Epidemiology More Causal

John W. Jackson* **and Onyebuchi A. Arah**

* Correspondence to John W. Jackson, Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, 615 North Wolfe Street, Room E6543, Baltimore, MD 21205 (e-mail: john.jackson@jhu.edu).

Initially submitted July 29, 2019; accepted for publication August 1, 2019.

A society's social structure and the interactions of its members determine when key drivers of health occur, for how long they last, and how they operate. Yet, it has been unclear whether causal inference methods can help us find meaningful interventions on these fundamental social drivers of health. Galea and Hernán propose we place hypothetical interventions on a spectrum and estimate their effects by emulating trials, either through individual-level data analysis or systems science modeling (*Am J Epidemiol*. 2020;189(3):167–170). In this commentary, by way of example in health disparities research, we probe this "closer engagement of social epidemiology with formal causal inference approaches." The formidable, but not insurmountable, tensions call for causal reasoning and effect estimation in social epidemiology that should always be enveloped by a thorough understanding of how systems and the social exposome shape risk factor and health distributions. We argue that one way toward progress is a true partnership of social epidemiology and causal inference with bilateral feedback aimed at integrating social epidemiologic theory, causal identification and modeling methods, systems thinking, and improved study design and data. To produce consequential work, we must make social epidemiology more causal and causal inference more social.

agent-based models; causal inference; decomposition; disparities; equity; microsimulation; policy analysis; population health; social epidemiology; social exposome; systems science

A society's structure and the interactions of its members shape the context in which we live and thereby determine who is exposed to, among others, persistent stress, environmental toxins, effective medical care, and preventive efforts. They also determine when these drivers of health occur, how long they last, and how they operate. As such, the social features of society are among the most fundamental drivers of population health. Epidemiologists, having long recognized this (1, 2), are rightly concerned with asking how we might advance population health by addressing these features of society (3, 4).

CAUSAL INFERENCE IN SOCIAL EPIDEMIOLOGY

As Galea and Hernán argue in their timely proposal (5), this is an endeavor that the development and application of methods from the field of causal inference can, and should, continue to help with. In some cases, it is straightforward to use causal inference methods to evaluate the effects of complex interventions such as a social policy or health care delivery model. Even when interventions are not randomized, rolled out in a staggered schedule, or lack a control group, the field of causal inference has provided a way forward (6). In each case, the intervention is not only well defined but exists and has been observed in the real world. As epidemiologists, however, our primary charge is not only to evaluate well-defined interventions once conceived and enacted but also to inform new ones by elucidating the causal structures and factors that shape population health.

It is here—this fundamental decision of what new interventions should address and how—that the fields of social epidemiology and causal inference have been at an impasse. Galea and Hernán (5) argue that this deadlock is an illusion, primarily due to misconceptions about what the standards for causal inference are, what answers it can provide, and whether it only provides answers for exposures that can be (experimentally) manipulated. They clarify that causal inference with social exposures is on the same ground as any other, that it answers questions not about causal ontology but about effects of specific actions, and that those actions need

not be feasible by today's technology or political landscape but do need to be sufficiently well-defined to predict their consequences in a hypothetical experiment. They posit that emulating an experiment through the sculpting of individual-level data or the engineering of complex systems models ultimately will be helpful because doing so will evaluate actions that could improve population health.

In this commentary, we aim to probe this "closer engagement of social epidemiology with formal causal inference approaches" (5) by considering the application of causal inference methods in health disparities research, an area where many epidemiologists and statisticians have already broken ground (7–17). In doing so, we point out remaining tensions under Galea and Hernán's framework (5) and discuss what is needed for progress. Before proceeding, we would like to clarify that our parenthetical treatment of the word "social" in the title conveys our opinion that much of our discussion applies to epidemiology generally.

FACING THE LIMITS OF CAUSAL INFERENCE

Evidence-based interventions that address health and health care disparities are often complex and address several barriers that impede marginalized groups' ability to prevent poor health or to access, navigate, and effectively benefit from medical care (18). These interventions are often guided by a conceptual model for how social forces such as structural racism (19) and discrimination (20, 21) erect these barriers and lead to worse profiles of risk factors and subsequent negative health outcomes. Although the architecture and theories behind these models vary, they tend to focus on risk factors known to be associated with the outcome of interest and are overrepresented among socially marginalized groups.

Causal decomposition methods (17) aim to inform the evidence base for such interventions. They are used to ask a simple question: If we were able to change how one or more risk factors were distributed in the population, what patterns of health inequity would we observe? If outcomes moved toward being more equitable, an intervention designed to remove inequity in the targeted risk factor may have promise for improving equity in outcomes. The methods attempt to build into the causal estimand a substantively meaningful measure of disparity to be affected by the intervention on the risk factor, rather than an artefactual disparity measure defined by a statistical modeling procedure (17). Moreover, the methods base causal inference squarely around the risk factor of interest rather than the social status. The assumptions of no unmeasured confounding, positivity, and consistency all revolve around the targeted risk factor.

Arguably, because these decomposition methods are used to study a modifiable risk factor rather than the social status as a causal target for intervention, they are used to pursue questions toward the right of Galea and Hernán's (5) spectrum of experimental specification. Like all modern causal methods, they rely on observed outcomes to make statements about counterfactuals under a hypothetical intervention. The bridge from observations to counterfactuals is supported by the consistency assumption (22) under the potential-outcomes framework and the causal Markov assumption (23) under the structural causal model framework. Unfortunately, the types of interventions that have been evaluated tend to be system preserving. That is, they tend to assume that the way in which outcomes are distributed, given the risk factor and covariates do not change as a result of being intervened upon rather than being observed. In the case of equity, history is replete with examples of explicit interventions to address inequities being met with highly adaptive barriers that manifest in unforeseen ways.

This is a sobering limitation. To be fair, system-preserving assumptions lurk behind applications of causal methods to evaluate even seemingly less vague interventions such as maintaining adherence to treatment guidelines during medical care. Achieving such goals would generally require supports from multiple institutions, stakeholders, and, perhaps, shifts in societal values. The successful implementation and sustainability of any hypothetical intervention will depend on cofactors that may not exist at the time of the study and, if they do, are seldom measured or incorporated into our modeling and inference.

All is not lost here. In our view, what we obtain from causal methods is a step away from association toward causation through ruling out alternate explanations. If we see that the differential distribution of a risk factor is associated with disparities, causal methods help us understand whether this association is due to the differential distribution of a third unmeasured factor (confounding). Likewise, they help us rule out alternate explanations involving differential participation or follow-up (selection bias) or mismeasurement of outcomes (information bias). If we can specify a sufficiently well-defined intervention, we may also learn that the association we see is not clouded by differential versions of the hypothetical intervention. When we cannot rule out these explanations, we can use bias analysis to bound or quantify our uncertainty (24, 25).

Nonetheless, causal methods from the potential-outcomes and structural causal-model frameworks are often intimately tied up with the present or the past and thus provide effects under "ideal" conditions. For effective intervention development, their answers must be enveloped with a clear and historically informed understanding of how the systems that drive risk-factor distributions operate (26). Our results cannot be effective without social, economic, and political theory and input from stakeholders who are actively involved in the development, assessment, and implementation of interventions, including the targets of such interventions: the patients and community members themselves.

EXTENDING CAUSAL INFERENCE TO MAKE IT MORE SOCIALLY ENGAGED

When observational data are insufficient to predict the effects of hypothetical actions, Galea and Hernán (5) suggest we combine data and expert knowledge and gain traction through simulations. An emerging strand of literature outlines a path forward by connecting results that underlie causal inference methods (e.g., the g-formula, front-door formula, and others) with those from systems science (e.g., agent-based modeling, microsimulation, and so forth) (27,

28). This mapping essentially uses models to inform the likelihood of a simulated person or population transitioning from one state to another (e.g., from unexposed to exposed). A contribution of causal inference is to require that these densities condition not just exposure but also confounders and any necessary causal partners.

Although these important steps have been taken, systems science remains a relatively uncharted territory for social epidemiology (29). How are we to map causal identification criteria to agent decision-making and agent-agent interaction to inform agent-based simulation models? How do we map such criteria to the impact agents exert on their own and others' environments, and how those environments, in turn, constrain and expand opportunities for health? Moreover, how do we design systems models in ways that align with theoretical frameworks such as ecosocial theory (30), a praxis such as critical race theory (31), or an intersectional understanding of marginalization? (32, 33).

Relatedly, this may be a welcome opportunity to re-examine whether our data definition and collection paradigms are up to the task of informing such models and estimating ambitious counterfactuals (34). For example, much of health-equity research relies on cross-sectional panel studies that are useful for assessing disease burden, health care use, and carrying out policy evaluations but lack the linkages needed to model trajectories of marginalized groups. Cohort studies overcome this but often have limited survey value and do not always contain detailed assessments of relevant psychosocial stressors and life experiences.

We also need to consider expanding the type of information we capture directly or through linkages. Disparities in education, employment, housing, and the legal system underlie disparities in risk factors. Data from these sectors may provide relevant important intervention targets and confounders. From a life-course perspective, early life experiences have lasting effects but are not prospectively captured by cohorts that recruit in adulthood. From a contextual perspective, measures with greater spatial and temporal resolution may help us detect environmental features that have high leverage. From a health care perspective, administrative data that measure social status variables well, along with characteristics of providers, the quality of their communication with patients, and the nonmedical interventions they prescribe may be critical for informing equity in prognosis. Finally, we need to build trust, engage communities, and sample populations in ways that ensure the most vulnerable are adequately represented. If we realize our methodological goals without improving the data they rely on, our inferences may have limited impact on informing effective strategies to achieve health equity.

To move forward, we challenge social epidemiology and causal inference researchers to consider the following: How can we work together to define, specify, and evaluate hypothetical interventions of consequence and ultimately translate them into actual interventions, even in the face of uncertainty? We suspect this will require a truly symbiotic partnership where causal inference becomes more engaged with social theory and social epidemiology leverages the best of causal rigor—a marriage that may push both beyond their current frontiers. As a byproduct, we would have clearer guidelines for reasoning with and applying causal principles in our studies along with better study design, execution, interpretation and, most importantly, strong evidence for intervention development.

CONCLUSION

A social epidemiology of consequence calls for robust causal reasoning, modeling, and inference. Although causal inference has made incredible theoretical strides, it has yet to engage social epidemiology as practically and fully as it could. We applaud Galea and Hernán for finding an inclusive common ground to encourage greater dialogue. Imagining better counterfactuals that are critically relevant to social epidemiology must go beyond the simple application of causal inference methods and systems science models. Social epidemiology and causal inference must partner every step of the way to yield impact and scale. We believe they need each other more than either may realize.

ACKNOWLEDGMENTS

Author affiliations: : Department of Epidemiology, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland (John W. Jackson); Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland (John W. Jackson); Center for Health Equity, Johns Hopkins University (John W. Jackson); Center for Health Disparities Solutions, Johns Hopkins Bloomberg School of Public Health (John W. Jackson); Department of Epidemiology, Fielding School of Public Health, University of California, Los Angeles, Los Angeles, California (Onyebuchi A. Arah); Department of Statistics, UCLA College of Letters and Science, Los Angeles, California (Onyebuchi A. Arah); and Department of Public Health, Aarhus University, Aarhus, Denmark (Onyebuchi A. Arah).

J.W.J. was supported by the National Heart, Lung, and Blood Institute (grant K01HL145320). J.W.J. also benefited from facilities and resources provided by the Center for Health Disparities Solutions at the Johns Hopkins Bloomberg School of Public Health, which receives core support (grant U54MD000214) from the National Institute on Minority Health and Health Disparities. O.A.A. was supported by the National Center for Advancing Translational Science (grant UL1TR001881), the Eunice Kennedy Shriver National Institute of Child Health and Human Development (grant R01HD072296), and the Norwegian Research Council's BEDREHELSE program (grant 273823). O.A.A. also benefited from facilities and resources provided by the California Center for Population Research at University of California, Los Angeles, which receives core support (grant R24HD041022) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

Conflict of interest: none declared.

REFERENCES

1. Cassel J. The contribution of the social environment to host resistance: the fourth Wade Hampton frost lecture. *Am J Epidemiol*. 1976;104(2):107–122.
2. Jones-Eversley SD, Dean LT. After 121 years, it's time to recognize W.E.B. Du Bois as a founding father of social epidemiology. *J Negro Educ*. 2018;87(3):230–245.
3. Arah OA. On the relationship between individual and population health. *Med Health Care Philos*. 2009;12(3): 235–244.
4. Arah OA. On the evaluative space for measuring public healht performance. In: Dawson A, ed. *The Philosophy of Public Health*. Farnham, Surrey, UK: Ashgate Publishing Limited; 2009:49–62.
5. Galea S, Hernán M. Win-win: reconciling social epidemiology and causal inference. *Am J Epidemiol*. 2020; 189(3):167–170.
6. Spiegelman D. Evaluating public health interventions: 2. Stepping up to routine public health evaluation with the stepped wedge design. *Am J Public Health*. 2016;106(3): 453–457.
7. Greiner DJ, Rubin DB. Causal effects of perceived immutable characteristics. *Rev Econ Stat*. 2011;93(3):775–785.
8. VanderWeele TJ, Robinson WR. On the causal interpretation of race in regressions adjusting for confounding and mediating variables. *Epidemiology*. 2014;25(4):473–484.
9. Jackson JW, VanderWeele TJ. Intersectional decomposition analysis with differential exposure, effects, and construct. *Soc Sci Med*. 2019;226:254–259.
10. Valeri L, Chen JT, Garcia-Albeniz X, et al. The role of stage at diagnosis in colorectal cancer black-white survival disparities: a counterfactual causal inference approach. *Cancer Epidemiol Biomarkers Prev*. 2016;25(1):83–89.
11. Naimi AI, Schnitzer ME, Moodie EEM, et al. Mediation analysis for health disparities research. *Am J Epidemiol*. 2016;184(4):315–324.
12. Howe CJ, Dulin-Keita A, Cole SR, et al. Evaluating the population impact on racial/ethnic disparities in HIV in adulthood of intervening on specific targets: a conceptual and methodological framework. *Am J Epidemiol*. 2018;187(2): 316–325.
13. Howe CJ, Robinson WR. Survival-related selection bias in studies of racial health disparities: the importance of the target population and study design. *Epidemiology*. 2018; 29(4):521–524.
14. Mayeda ER, Banack HR, Bibbins-Domingo K, et al. Can survival bias explain the age attenuation of racial inequalities in stroke incidence?: A simulation study. *Epidemiology*. 2018;29(4):525–532.
15. Jackson JW. Explaining intersectionality through description, counterfactual thinking, and mediation analysis. *Soc Psychiatry Psychiatr Epidemiol*. 2017;52(7):785–793.
16. Jackson JW. On the interpretation of path-specific effects in health disparities research. *Epidemiology*. 2018;29(4): 517–520.
17. Jackson JW, VanderWeele TJ. Decomposition analysis to identify intervention targets for reducing disparities. *Epidemiology*. 2018;29(6):825–835.
18. Purnell TS, Calhoun EA, Golden SH, et al. Achieving health equity: closing the gaps in health care disparities, interventions, and research. *Health Aff. (Millwood)*. 2016; 35(8):1410–1415.
19. Bailey ZD, Krieger N, Agénor M, et al. Structural racism and health inequities in the USA: evidence and interventions. *Lancet*. 2017;389(10077):1453–1463.
20. Williams DR, Mohammed SA. Racism and health I: pathways and scientific evidence. *Am Behav Sci*. 2013;57(8): 1152–1173.
21. Williams DR, Mohammed SA. Racism and health II: a needed research agenda for effective interventions. *Am Behav Sci*. 2013;57(8):1200–1226.
22. VanderWeele TJ. Concerning the consistency assumption in causal inference. *Epidemiology*. 2009;20(6):880–883.
23. Pearl J. *Causality: Models, Reasoning and Inference*. 2nd ed. New York, NY: Cambridge University Press; 2009:1–484.
24. Arah OA. Bias analysis for uncontrolled confounding in the health sciences. *Annu Rev Public Health*. 2017;38(1):23–38.
25. Arah OA. Analyzing selection bias for credible causal inference: when in doubt, DAG it out. *Epidemiology*. 2019; 30(4):517–520.
26. Hicken MT. Invited commentary: fundamental causes, social context, and modifiable risk factors in the racial/ethnic inequalities in blood pressure and hypertension. *Am J Epidemiol*. 2015;182(4):354–357.
27. Murray EJ, Robins JM, Seage GR, et al. A comparison of agent-based models and the parametric g-formula for causal inference. *Am J Epidemiol*. 2017;186(2):131–142.
28. Arnold KF, Harrison WJ, Heppenstall AJ, et al. DAG-informed regression modelling, agent-based modelling and microsimulation modelling: a critical comparison of methods for causal inference. *Int J Epidemiol*. 2018;48(1): 243–253.
29. Nianogo RA, Arah OA. Agent-based modeling of noncommunicable diseases: a systematic review. *Am J Public Health*. 2015;105(3):e20–e31.
30. Krieger N. Theories for social epidemiology in the 21st century: an ecosocial perspective. *Int J Epidemiol*. 2001; 30(4):668–677.
31. Ford CL, Airhihenbuwa CO. The public health critical race methodology: praxis for antiracism research. *Soc Sci Med*. 2010;71(8):1390–1398.
32. Bauer GR. Incorporating intersectionality theory into population health research methodology: challenges and the potential to advance health equity. *Soc Sci Med*. 2014; 110:10–17.
33. Jackson JW, Williams DR, VanderWeele TJ. Disparities at the intersection of marginalized groups. *Soc Psychiatry Psychiatr Epidemiol*. 2016;51(10):1349–1359.
34. Howe CJ. Reducing HIV racial/ethnic disparities: what's good data got to do with it? *Epidemiology*. 2017;28(2): 221–223.