Commentary

ENVIRONMENTAL EPIDEMIOLOGY

OPEN

Multiple mediators approach to study environmental chemicals as determinants of health disparities

Andrea Bellavia[a,b], Ami R. Zota[c], Linda Valeri[d,e], Tamarra James-Todd[a,f,g*]

Abstract: A major goal of health disparities research is to identify and intervene upon modifiable risk factors that help explain the observed associations between social factors and adverse health outcomes. To this end, statistical methods incorporating mediation analysis have shown promise, as they quantify the contribution of an intermediate variable in an exposure–outcome association. A growing body of literature suggests that environmental chemicals can contribute to health disparities. However, evaluating environmental chemicals as an important component of health disparities introduces methodological complexities that may make standard mediation approaches inadequate. Specific to environmental health is the issue of evaluating both the source and biomarker of the environmental toxicant to calculate the proportion of the disparity that would remain had we intervened on the modifiable factors. Recent methodological developments on multiple mediators can improve efforts to integrate both source and biomarker of exposure into epidemiological studies of health disparities. We illustrate a conceptual framework and present how mediation techniques can be used to address environmental health disparities questions. With this, we provide a methodological tool that has the potential to advance this growing field, while simultaneously informing public health prevention and policy surrounding the impact of environmental factors on health disparities.

Keywords: Health disparities; Mediation; Social determinants of health; Behavior; Race; Environmental chemicals; Biomarkers

Introduction

Health disparities are defined as differences that systematically have an adverse effect on the health of less-advantaged populations.[1] Disadvantaged groups are generally defined in terms of sociological, economic, or anthropological constructs, such as race/ethnicity; religion; socioeconomic status; disability; sexual orientation; gender or gender identity.[1] In addition, constructs such as age, mental health status, culture, and geographic locations can define vulnerable populations. These factors are often immutable, which makes it possible to identify disadvantaged or vulnerable populations, but makes developing recommendations or interventions more challenging.[2] To develop effective public health recommendations and interventions, identification of proximal risk factors that may help to explain why these groups show enhanced vulnerability to a specific set of diseases or health outcomes is critical.[3]

A growing body of literature suggests that environmental factors may be important contributors to the biological pathways leading to disparities in health outcomes.[4–7] However, little has been done to quantify the proportion of disparate conditions/diseases that may be attributed to a given environmental toxicants, such as environmental chemicals, where biomarker data exist. Part of this gap in the literature is due to the fact that this quantification would require not only a marker of the environmental toxicant but also information on its modifiable sources and whether these differ between vulnerable and nonvulnerable populations. This information is crucial from a public health perspective as interventions or recommendations aimed at

What this study adds

A growing body of literature suggests that environmental factors may contribute to the biological pathways leading to disparities in health outcomes, but statistical frameworks for environmental health disparities are not well-established. In this commentary, we conceptualize a model of interest in environmental health disparities and present how mediation analysis techniques can be used to simultaneously evaluate biomarkers of environmental toxicants, as well as their modifiable sources. The potential impact of this information would inform public health prevention and policy. As such, this commentary provides a methodological tool for future research that could advance the growing field of environmental health disparities.

[a]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts; [b]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts; [c]Department of Environmental and Occupational Health, Milken Institute School of Public Health, George Washington University, Washington, DC; [d]Psychiatric Biostatistics Laboratory, McLean Hospital, Belmont, Massachusetts; [e]Department of Psychiatry, Harvard Medical School, Boston, Massachusetts; [f]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts; [g]Division of Women's Health, Department of Medicine, Connors Center for Women's Health and Gender Biology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts

Sponsorships or competing interests that may be relevant to content are disclosed at the end of the article.

This work was supported by National Institute of Environmental Health Sciences (grant number: R00ES019881 [A.Z.] and R01ES026166 [T.J.T.]).

Data code to simulate data and replicate analyses is included in the Supplementary Material; http://links.lww.com/EE/A9

SDC Supplemental digital content is available through direct URL citations in the HTML and PDF versions of this article (www.epidem.com).

*Corresponding author. Address: Department of Environmental Health, Harvard T.H. Chan School of Public Health, 665 Huntington Avenue, Building. 1, 14th Floor, Boston, MA 02120. Tel.: 617 432 6460; fax: 617 525 7746. E-mail address: tjtodd@hsph.harvard.edu (T. James-Todd).

Environmental Epidemiology (2018) 2:e015

Received: 9 February 2018; Accepted 17 April 2018

Published online 30 May 2018

DOI: 10.1097/EE9.0000000000000015

reducing the disparity would do better to target the modifiable factors (i.e., sources of environmental toxicants) rather than the nonmodifiable factors (e.g., an individual's immutable characteristics, such as race/ethnicity).

Mediation analysis is an increasingly popular statistical method to investigate the contribution of third variables in explaining an exposure–outcome association.[8] By providing insights into the pathway underlying statistical associations, this method has been useful for understanding health disparities[9–14]; yet, it has been underutilized in environmental health disparities. Recent developments have extended the mediation analysis framework to incorporate several methodological complexities that may be present when investigating environmental factors. For example, methods for multiple mediators have been developed and may be used to simultaneously incorporate in the same statistical framework environmental factors together with their modifiable sources.[15,16] The aim of this commentary is to illustrate how methods for mediation analysis, particularly the use of multiple mediation analysis, may aid in addressing environmental health disparities research questions. For this, we will conceptualize the model of interest in a nontechnical review of the relevant regression-based statistical approaches for mediation analysis. In addition, we will provide several relevant examples to conceptualize the use of this analytic technique to address current environmental health research questions.

Conceptual model

The Figure presents a basic conceptual model to describe environmental health disparities, while the Table presents several examples of environmental health disparities research questions. Let *X* be an immutable or difficult to modify social factor (e.g., race/ethnicity) and *Y* be the health outcome unequally distributed across subgroups of *X* (e.g., cardiovascular disease [CVD]).[17] The goal in environmental health disparity research is to evaluate the contribution of environmental factors to this *X*-*Y* association. This occurs if a potential environmental chemical *E* (e.g., a phthalate metabolite)[4,18] is both unequally distributed over levels of *X* and is a risk factor of *Y*. When planning interventions to reduce the disparity (e.g., CVD having a higher prevalence among racial minorities), an additional step is to identify modifiable sources of *E*, in this case *B* (e.g., diet, such as fast food consumption).[19] Such sources, *B*, are also unequally distributed by *X* and may be independent risk factors for *Y*. In the case of the present example, fast food consumption is likely unequally distributed by race/ethnicity and an independent risk factor of CVD.

In statistical terms, both *B* and *E* can be seen as potential mediators of the *X*-*Y* association. Specifically, there are four possible pathways through which the health disparity *X*-*Y* (in the present example race/ethnicity and CVD) is generated: (1) over-exposure to the modifiable factor *B* in the disadvantaged group leading to higher levels of the harmful environmental factor *E* ($X \rightarrow B \rightarrow E \rightarrow Y$ or, in this example, race/ethnicity → fast food consumption → phthalates → CVD); (2) over-exposure to *E* due to other nonidentified factors independent of *B* ($X \rightarrow E \rightarrow Y$ or race/ethnicity → phthalates → CVD); (3) over-exposure to *B* which affects the outcome through other pathways not including *E* ($X \rightarrow B \rightarrow Y$ or race/ethnicity → fast food consumption → CVD); (4) other pathways that are independent to both *B* and *E*. In the next sections, we present and illustrate how mediation techniques involving multiple mediators can be adopted to statistically evaluate environmental health disparities and to what extent these pathways can be identified. In the Supplemental Content; http://links.lww.com/EE/A9, we also provide an illustration of such methods in a simulated dataset, together with guided Stata code for their implementation.

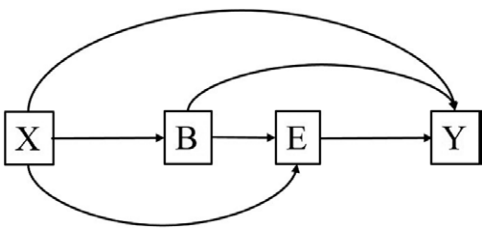

Figure. Conceptual model to describe environmental health disparities. The figure describes the contribution of an environmental factor (*E*) and its modifiable source (*B*) in the health disparity *X*-*Y*, where *X* is an immutable or difficult to modify social factor (e.g., race/ethnicity, gender, SES) and *Y* the health outcome known to be unequally distributed across subgroups of *X*.

Mediation analysis with a single mediator

Mediation models with a single mediator could be used to evaluate the contribution of the biomarker *E* alone, for example, the contribution of phthalates in explaining racial/ethnic disparities in CVD. In brief, mediation analysis allows decomposing the total effect of an *exposure X* on a given *outcome Y* into a *direct effect* of the exposure and an *indirect effect* that acts through a *mediator E*. [20] Results from mediation analysis are often presented in terms of *proportion mediated*, calculated as the ratio between the indirect and the total effect. In the context of health disparity research, the proportion mediated may be interpreted as the extent to which the disparity can be attributed to the specific mediator of interest.[14] Direct and indirect effect can also be defined in terms of potential outcomes (counterfactual approach).[21] Of particular interest are the effect of *X* on *Y* after fixing the mediator to a predefined value (controlled direct effect: CDE) and the effect of *X* on *Y* that only operates by changing *E* (natural indirect effect: NIE). Controlled effects are useful to retrieve information about the result of potential interventions, while natural effects provide information on the pathway through which the disparity is generated. In the context of health disparities, the CDE represents the proportion of disparity that would remain if we were to intervene on the mediator and can be referred to as counterfactual disparity measure[10] or residual disparity.[22] In the context of our example, the counterfactual disparity measure will provide a quantifiable estimate of the proportion of the disparity that would remain between race/ethnicity and CVD if we intervened on phthalates. Both parametric and nonparametric methods to estimate direct and indirect effects within the standard and counterfactual approaches have been presented and are available in all major statistical software.[21,23] Conditions and assumptions for the identification of such effects are reviewed in the Supplemental Content; http://links.lww.com/EE/A9.

Mediation analysis with multiple mediators

In practical situations (see the Table for examples), a model to describe environmental health disparities should also include upstream sources or factors to which interventions could be developed. For example, if we identified phthalates as a contributor to racial/ethnic disparities in CVD, we could target their modifiable sources such as fast-food consumption in potential interventions or recommendations. Ideally, we would like to quantify the effect of intervening on *B* in reducing the disparity of interest.

Parametric frameworks to estimate multiple mediation models can be used. These frameworks can also be extended to include exposure–mediator interactions, as well as to allow situations in which one of the two mediators is expected to be sequential to the other (e.g. *B* and *E* in Figure).[15,16,24] Introducing multiple and possibly sequential mediators, however, makes identification and interpretation of effects more challenging. The CDE (i.e., the direct arrow from *X* to *Y* in Figure, not

Table.

Examples of Health Disparities (*X*-*Y*) Where the Joint Contribution of Environmental Factors (*E*) and Their Modifiable Sources (*B*) Might be Hypothesized From the Literature

Disparity of Interest	Social/Anthropological Construct (*X*)	Modifiable Sources of Environmental Factors (*B*)	Biomarkers of Environmental Toxicants (*E*)	Health Outcomes (*Y*)
1. African American race is associated with higher prevalence of cardiovascular disease*	Race/ethnicity	Fast food consumption	Phthalates	Cardiovascular disease
2. Living longer in the United States is associated with higher risk of GDM in Asian immigrants	Acculturation	Occupational exposures (e.g., nail salons)	Toluene formaldehyde	Gestational diabetes
3. African American race is associated with higher risk of premature puberty	Race/ethnicity	Hair product use	Environmental phenols	Age at menarche
4. Poverty is associated with reduced IQ	Socioeconomic status	Housing and household factors (e.g., furniture)	PBDEs	Neurodevelopment
5. Women have higher prevalence of asthma	Gender	Cleaning products	Volatile organic compounds	Asthma

*Illustrative examples provided in the manuscript refer to the first line of this Table. These can be extended to any other environmental health disparity in which modifiable sources and biomarker of environmental exposures can be identified.

IQ indicates intelligent quotient; GDM, gestational diabetes mellitus; PBDE, polybrominated diphenyl ethers.

going through *B* or *E*) can be estimated and identified under the same assumptions presented in the context of a single mediator (Supplemental Content; http://links.lww.com/EE/A9) and retains its interpretation as the effect when both *B* and *E* are set to a referent value. In disparity terms, the CDE represents the proportion of disparity that would remain if we were to intervene on both mediators by fixing them to a specific arbitrary value. When mediators are sequential, the CDE represents the proportion that would remain after a hypothetical intervention on *B* and all other nonidentified sources of *E*. In our example, this would represent the CVD disparity that would remain after intervening on fast-food consumption and other unknown sources to reduce phthalate exposure.

The total NIE, representing the proportion of disparity that is jointly due to any pathway including *B*, *E*, or both can be identified and estimated under the classical assumptions (Supplemental Content; http://links.lww.com/EE/A9). This joint effect can be further decomposed into the sum of three path-specific NIE, capturing the effect of different mechanisms contributing to the disparity: (1) the effect going through *E*, but not *B* (NIE_E); (2) the effect going through *B*, but not *E* (NIE_B); (3) the effect going through both *B* and *E* (NIE_{EB}). In the model presented in Figure (sequential mediators), *B* also acts as a confounder of the mediator–outcome association, thus making the identification of path-specific effects more challenging. In such setting, it is only possible to identify the joint NIE (e.g., the effect due to fast-food consumption on phthalate exposure), the NIE_B (e.g., the proportion due to fast-food consumption alone), and the NIE_E (e.g., the proportion due to phthalate exposure alone). Identification and estimation of path-specific effects such as NIE_{EB} (e.g., the proportion of effect due to phthalate exposure that would be reduced by intervening on fast-food consumption), on the other hand, are not straightforward and may require defining effects in terms of randomized interventional analogues, involving advanced estimation techniques such as g-formulas and marginal structural models.[25–27] We refer to previous publications for these and other identification and estimation procedures of all possible direct and indirect effects in the context of multiple mediators.[15,25,28,29]

Final remarks

In this commentary, we presented and illustrated classical and novel methods for mediation analysis as a possible quantitative approach to evaluate environmental health disparities specific to situations where source information and biomarker data are available. We discussed how methods for mediation can help to quantify the proportion of disparity due to a specific environmental factor (NIE) and the proportion of disparity that would remain by interventions aimed at reducing the environmental toxicant (CDE). We have discussed how extensions to incorporate multiple sequential mediators can prove useful when interventions target modifiable sources of the environmental toxicant rather than the biomarkers alone. Further, we discussed the conditions under which the estimated effects maintain their policy-oriented relevance.

Environmental health disparities have been relatively understudied in the context of chemical exposures, and a methodological framework to address how environmental chemicals could impact health disparities has not been well established. One of the reasons for this gap is that a conceptual model for environmental health disparities requires the simultaneous inclusion of biomarkers of environmental toxicants and their modifiable sources, thus making standard mediation techniques, commonly used in health disparities research, inadequate. Here, we have discussed how recent developments of methods for multiple mediators can be used to incorporate some of these additional complexities and the extent to which measures of interest can be estimated. While we primarily focused this commentary on the simultaneous evaluation of environmental chemicals and their sources, there are several additional features that recent methodological developments allow us to take into account. For example, in the Supplemental Content; http://links.lww.com/EE/A9, we briefly review how nonlinearities, repeated measurements, and multiple independent mediators can also be included in the conceptual model presented. Future work should incorporate additional topics, including joint disparities.[30]

To evaluate all factors involved in a conceptual model for environmental health disparities, it is important that questions of mediation are addressed in diverse study populations, as either primary or secondary analyses. As such, these factors should be taken into account from the initial phases of study design. For instance, evaluating mediated effects requires estimating a larger number of parameters, and a power calculation based on the estimation of a total effect does not generally extend to the estimation of direct, indirect, and interactive effects.[31–33] Moreover, the sample size required to detect significant associations may rapidly grow as multiple interactions or mediators are simultaneously taken into account. Studies aimed at investigating mediation should be designed to assure a temporal sequence in exposure, mediators, and outcomes. While studies of diverse populations are not easily accessible, particularly in environmental health research, we encourage an increased focus on these topics.

The ultimate goal of health disparities research is to identify causal effects to develop interventions aimed at reducing the disparity. In recent years, approaches based on counterfactuals of potential outcomes have been the most widely used and developed methods for causal inference in observational studies, especially in the context of mediation analysis.[28,34] However,

applying the counterfactual approach in the context of health disparities research may not be straightforward as causal effects are defined in terms of hypothetical interventions. This has brought some discussion as to whether it is reasonable to speak of a causal effect of nonmodifiable factors such as race or gender.[22,35] In addition, it has been shown that when mediation analysis is used with nonmodifiable exposures such as the social/anthropological constructs, it is not straightforward to justify the assumptions required for the causal interpretation of effects.[36] Despite recent studies discussing the analytical and conceptual issues that are required to interpret the direct effect as the magnitude of disparity that would remain if a mediator was changed,[10,22] our general recommendation is that causal interpretation of statistical results—even in the most ideal situation—should always be taken with caution.

In conclusion, our work provides a conceptual framework and illustrates the methodological tools for future research in environmental health disparities. Using a multiple mediation approach to address these questions has the potential to advance this growing field, while simultaneously informing public health prevention and policy.

References

1. Dehlendorf C, Bryant AS, Huddleston HG, Jacoby VL, Fujimoto VY. Health disparities: definitions and measurements. *Am J Obstet Gynecol*. 2010;202(3):212.
2. Williams DR, Costa MV, Odunlami AO, Mohammed SA. Moving upstream: how interventions that address the social determinants of health can improve health and reduce disparities. *J Public Health Manag Pract JPHMP*. 2008;14(Suppl):S8.
3. Gehlert S, Sohmer D, Sacks T, Mininger C, McClintock M, Olopade O. Targeting health disparities: a model linking upstream determinants to downstream interventions. *Health Aff (Millwood)*. 2008;27(2):339–49.
4. James-Todd TM, Chiu Y-H, Zota AR. Racial/ethnic disparities in environmental endocrine disrupting chemicals and women's reproductive health outcomes: epidemiological examples across the life course. *Curr Epidemiol Rep*. 2016;3(2):161–80.
5. Gee GC, Payne-Sturges DC. Environmental health disparities: a framework integrating psychosocial and environmental concepts. *Environ Health Perspect*. 2004;112(17):1645–53.
6. Zota AR, Shamasunder B. The environmental injustice of beauty: framing chemical exposures from beauty products as a health disparities concern. *Am J Obstet Gynecol*. 2017;217(4):418–e1.
7. McDonald JA, Tehranifar P, Flom JD, Terry MB, James-Todd T. Hair product use, age at menarche and mammographic breast density in multiethnic urban women. *Environ Health*. 2018;17(1):1.
8. MacKinnon DP. *Introduction to Statistical Mediation Analysis*. Routledge; New York 2012.
9. Oakes JM, Naimi AI. Mediation, interaction, interference for social epidemiology. *Int J Epidemiol*. 2016;dyw279.
10. Naimi AI, Schnitzer ME, Moodie EEM, Bodnar LM. Mediation analysis for health disparities research. *Am J Epidemiol*. 2016;184(4):315–24.
11. Valeri L, Chen JT, Garcia-Albeniz X, Krieger N, VanderWeele TJ,Coull BA. 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):83–9.
12. Chatterjee R, Yeh H-C, Shafi T, Anderson C,Pankow JS, Miller ER, et al. Serum potassium and the racial disparity in diabetes risk: the Atherosclerosis Risk in Communities (ARIC) Study. *Am J Clin Nutr*. 2011;93(5):1087–1091.
13. Ibfelt E, Kjær SK, Johansen C, Høgdall C, Steding-Jessen M, Frederiksen K, Frederiksen BL, Osler M, Dalton SO. Socioeconomic position and stage of cervical cancer in Danish women diagnosed 2005–2009. *Cancer Epidemiol Biomarkers Prev*. 2012:cebp-1159.
14. Hystad P, Carpiano RM, Demers PA, Johnson KC, Brauer M. Neighbourhood socioeconomic status and individual lung cancer risk: Evaluating long-term exposure measures and mediating mechanisms. *Soc Sci Med*. 2013;97:95–103.
15. VanderWeele TJ, Vansteelandt S. Mediation analysis with multiple mediators. *Epidemiol Methods*. 2014;2(1):95–115.
16. Daniel RM, De Stavola BL, Cousens SN, Vansteelandt S. Causal mediation analysis with multiple mediators. *Biometrics*. 2015;71(1):1–14.
17. Pool LR, Ning H, Lloyd-Jones DM, Allen NB. Trends in racial/ethnic disparities in cardiovascular health among US adults from 1999–2012. *J Am Heart Assoc*. 2017;6(9):e006027.
18. James-Todd TM, Huang T, Seely EW, Saxena AR. The association between phthalates and metabolic syndrome: the National Health and Nutrition Examination Survey 2001–2010. *Environ Health*. 2016;15(1):52.
19. Zota AR, Phillips CA, Mitro SD. Recent fast food consumption and bisphenol A and phthalates exposures among the US population in NHANES, 2003–2010. *Environ Health Persp*. 2016;124(10):1521.
20. Baron RM, Kenny DA. The moderator–mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol*. 1986;51(6):1173.
21. Valeri L, VanderWeele TJ. Mediation analysis allowing for exposure–mediator interactions and causal interpretation: theoretical assumptions and implementation with SAS and SPSS macros. *Psychol Methods*. 2013;18(2):137–50.
22. VanderWeele TJ, Robinson WR. On the causal interpretation of race. *Epidemiology*. 2014;25(6):937–8.
23. Imai K, Keele L, Tingley D, Yamamoto T. Causal mediation analysis using R. In: Vinod HD, eds. *Advances in Social Science Research Using R*. New York: Springer; 2010:129–154.
24. Bellavia A, Valeri L. Decomposition of the total effect in the presence of multiple mediators and interactions. *Am J Epidemiol*. 2017(epub ahead of print).
25. Vansteelandt S, Daniel RM. Interventional effects for mediation analysis with multiple mediators. *Epidemiology*. 2017 Mar;28(2):258.
26. Naimi AI, Cole SR, Kennedy EH. An Introduction to G Methods. *Int J Epidemiol*. 2016:dyw323.
27. VanderWeele TJ, Vansteelandt S, Robins JM. Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. *Epidemiology*. 2014;25(2):300.
28. VanderWeele T. *Explanation in Causal Inference: Methods for Mediation and Interaction*. Oxford University Press; 2015.
29. Jackson JW, VanderWeele TJ. Decomposition analysis to identify intervention targets for reducing disparities. arXiv preprint arXiv:1703.05899. 20.17. Available at: https://arxiv.org/abs/1703.05899.
30. Jackson JW, Williams DR, VanderWeele TJ. Disparities at the intersection of marginalized groups. *Social Psychiatr Psychiatr Epidemiol*. 2016;51(10):1349–59.
31. Bind MA, Vanderweele TJ, Coull BA, Schwartz JD. Causal mediation analysis for longitudinal data with exogenous exposure. *Biostat Oxf Engl*. 2016;17(1):122–34.
32. Greenland S. Tests for interaction in epidemiologic studies: A review and a study of power. *Stat Med*. 1983;2(2):243–51.
33. Fritz MS, MacKinnon DP. Required sample size to detect the mediated effect. *Psychol Sci*. 2007;18(3):233–9.
34. VanderWeele TJ. The role of potential outcomes thinking in assessing mediation and interaction. *Int J Epidemiol*. 2016;dyw280.
35. Glymour C, Glymour MR. Commentary: race and sex are causes. *Epidemiology*. 2014;25(4):488–90.
36. Naimi AI, Kaufman JS. Counterfactual theory in social epidemiology: reconciling analysis and action for the social determinants of health. *Curr Epidemiol Rep*. 2015;2(1):52–60.