Moving Beyond Speculation

Quantifying Biases in Neighborhood Health Effects Research

Ana V. Diez-Roux

In this issue of EPIDEMIOLOGY, Chaix et al[1] use a novel approach to investigate the possible impact of neighborhood differences in study participation on estimates of the association between neighborhood characteristics and prevalence of disease. Although the possibility of differential participation by neighborhood characteristics is often noted in the discussion of study results, its role in biasing measures of the associations of interest is rarely investigated empirically.

Building on previous work showing how the selection bias generated by differential participation can be viewed as a form of collider bias,[2] the authors use a clever approach to illustrate how unmeasured, geographically varying factors related to participation and diabetes could affect estimates of the association of neighborhood education with diabetes. If both neighborhood education and these unmeasured factors are related to participation, they will become associated in analyses that condition on participation, even if they are marginally unassociated in the full population. If the unmeasured factors are also associated with diabetes, estimates of the association of neighborhood education with diabetes that do not account for these unmeasured factors will be biased.

As we often teach our students, very few, if any, epidemiologic studies can be completely free of bias. The issues are how large the bias is and to what extent the bias changes the conclusions that can be drawn from the study. In the study reported by Chaix et al, the bias is small and does not qualitatively change the conclusions: the point estimate for the prevalence odds ratio for the lowest versus the highest neighborhood education category changes from 1.56 to 1.50 or 1.44 after bias correction. Given all the other limitations of these types of analyses, it would be a mistake to draw much more than qualitative conclusions regarding the direction and very approximate size of the association. Of course, the size and direction of the bias depend on the directionality and strength of the associations shown in Figure 1 of the paper by Chaix and colleagues. This could be very different in different contexts. Although in their case the bias was small, Chaix et al provide a useful analytic framework for investigating this type of bias in other studies when appropriate data are available.

Aside from quantifying the magnitude and direction of the bias, it is interesting to consider whether anything can be learned substantively from the investigation of this type of bias. The unspecified neighborhood-level factors that contribute to the bias have several unique features: they are unassociated with neighborhood education in the full population and are positively associated with both participation and diabetes. Based on existing knowledge on the associations between neighborhood SES and other neighborhood built- and social-environment features and on the usual predictors of participation and diabetes, it is hard to imagine what these omitted neighborhood variables might be. The authors decline to even speculate on this, and yet hypothesizing on what could be driving these

From the Department of Epidemiology, School of Public Health and Center for Social Epidemiology and Population Health, University of Michigan, Ann Arbor, MI.

Correspondence: Ana V. Diez-Roux, Department of Epidemiology, School of Public Health and Center for Social Epidemiology and Population Health, University 2nd Floor, University of Michigan, Ann Arbor, MI 4810. E-mail: adiezrou@umich.edu.

ISSN: 1044-3983/11/2201-0040

DOI: 10.1097/EDE.0b013e318200d122

patterns would make both the search for and the need to account for these factors more compelling, especially in light of the weak bias actually observed. In addition to omitted neighborhood-level factors, omitted or mismeasured individual-level variables that vary over space (and are related to participation and diabetes) could also contribute to this bias. Neighborhood differences in participation could also result in biased estimates of the association of individual-level variables with diabetes (not just neighborhood-level variables), even if neighborhood factors are marginally unassociated with the individual-level characteristics. This could be especially important in the many individual-level studies that ignore geographic context.

As noted by Chaix et al, another important source of collider bias is that which would arise when both the neighborhood characteristic and the outcome (diabetes) are related to participation.[3] Under these circumstances, neighborhood characteristics and diabetes could become associated among participants even if there is no association in the full population. Moreover, this type of bias could result in substantial underestimates of the true association between neighborhood characteristics and health, making it appear as though there is no association, when in fact there is one. Given that health may substantially affect participation in health studies, this type of bias may be much more pervasive and larger than that investigated by the authors.

An interesting conclusion of the analyses of Chaix et al is that the bias could be avoided if specific measures of the neighborhood constructs (or of omitted individual-level variables that covary with space) related to participation were included in the analyses. This further emphasizes the importance of improved measurement of neighborhood constructs and adjustment for individual-level variables. The relative impact of this type of bias is not clear, relative to other methodological problems faced by neighborhood-health-effects researchers, including major misspecification of the potentially causal neighborhood-level variables. The relative impact of this type of bias compared with other large methodological problems faced by neighborhood health effects researchers (such as major misspecification of the casual neighborhood-level variables and residual confounding by individual-level factors) remains an unanswered question. Further investigation of participation-related biases as suggested by Chaix et al may help us better understand how large or small this problem may be.

More generally, the analyses of Chaix et al highlight the utility of drawing causal diagrams to help identify the most appropriate analytic approach and also to better understand the possible consequences of the approach taken under different scenarios.[4] There is no substitute for thoughtful formulation of specific research questions and careful consideration of the relevant neighborhood and individual-level constructs, and their hypothesized relations. Like other areas of epidemiology, the study of neighborhood health effects faces numerous methodological challenges. Our job as epidemiologists is to do the best we can to draw reasonable conclusions from messy data. Moving beyond speculation on the many possible methodological problems to empirical examinations of the role of bias (as illustrated by Chaix et al) is a welcome and much-needed addition to the field.

ABOUT THE AUTHOR

ANA V. DIEZ-ROUX is Professor of Epidemiology and Director of the Center for Social Epidemiology and Population Health at the University of Michigan. She has published extensively on the study of neighborhood health effects. Her current work focuses on longitudinal analyses of neighborhoods and health and on the application of systems approaches to better understand how neighborhoods affect health.

REFERENCES

1. Chaix B, Billaudeau N, Thomas F, et al. Neighborhood effects on type II diabetes: correcting bias from neighborhood effects on participation. *Epidemiology*. 2011;22:18–26.
2. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. *Epidemiology*. 2004;15:615–625.
3. Cole SR, Platt RW, Schisterman EF, et al. Illustrating bias due to conditioning on a collider. *Int J Epidemiol*. 2010;39:417–420.
4. Fleischer NL, Diez Roux AV. Using directed acyclic graphs to guide analyses of neighbourhood health effects: an introduction. *J Epidemiol Community Health.* 2008;62:842–846.