



PERGAMON

Social Science & Medicine 58 (2004) 1953–1960

SOCIAL  
SCIENCE  
&  
MEDICINE

www.elsevier.com/locate/socscimed

Commentary

# Estimating neighborhood health effects: the challenges of causal inference in a complex world

Ana V. Diez Roux\*

*Department of Epidemiology, Center for Social Epidemiology and Population Health, University of Michigan, 1214 South University, 2nd Floor, Ann Arbor, MI 48104, USA*

## Introduction

Oakes raises a series of important questions on the validity of past work on neighborhood health effects and suggests directions the field should take (Oakes, 2003). Indeed the limited nature of the evidence linking neighborhoods to individual-level outcomes, and its many methodological problems, have been noted in health and other fields (Diez Roux, 2001; Duncan & Raudenbush, 1999; Furstenberg & Hughes, 1997; Macintyre, Ellaway, & Cummins, 2002; Sampson et al., 2002; Tienda, 1991). The complex methodological issues inherent in the estimation of causal neighborhood effects are nevertheless worth reiterating and elaborating on as Oakes does. I would also posit that many of the issues Oakes raises are common to epidemiology generally (and to observational studies in other fields) and are not necessarily specific to research on neighborhood effects (or to so-called “social epidemiology” as he sometimes implies). Perhaps one of the problems is that in research on neighborhoods and health (as in epidemiology generally) the word “effects” is often used loosely, leading to the impression that the associations reported are always valid and precise estimates of generalizable causal parameters, when of course they are not. I will begin with some general comments on points raised by Oakes when he describes “a causal model for neighborhood effects”. I will then summarize agreements and disagreements with the methodological obstacles Oakes notes, and with the approaches he proposes.

## The selection or confounding issue

Undoubtedly, the selection issue (the fact that persons may be selected into neighborhoods based on individual attributes which are themselves related to health) is the

key problem in observational studies of neighborhood effects. Epidemiologists have attempted to account for this by controlling for often numerous individual-level variables. One of the problems, as Oakes notes, is that when numerous covariates are included it is likely that sparse data will be found in many cross-tabulated cells, resulting in a greater reliance on assumptions which often cannot be directly tested. Mismeasured and omitted individual-level variables may result in non-exchangeability of individuals across neighborhoods even when measured individual-level covariates are included (the omnipresent “residual confounding” problem in epidemiologic studies). These are familiar problems frequently alluded to in neighborhood effects research (Duncan, Connell, & Klebanov, 1997; Duncan & Raudenbush, 1999; Macintyre, Ellaway, & Cummins, 2002; Sampson et al., 2002; Tienda, 1991). Although the propensity score approach (Joffe & Rosenbaum, 1999; Rubin, 1997) has certain advantages when attempting to control for multiple factors related to selection into treatment groups (in this case into types of neighborhoods) it does not resolve the problems of mismeasured or omitted variables. The combination of propensity score matching with sensitivity analyses may allow better assessment of the potential impact of omitted variable confounding on the associations estimated (Harding, 2002).

A nagging, and perhaps even more complex problem is that many of these individual-level variables (including personal or family socioeconomic indicators as well as disease risk factors) may be mediators rather than confounders (or simultaneously mediators and confounders) of neighborhood effects. In addition, persons of low individual-level socioeconomic position may be at increased risk of disease in part because of the types of neighborhoods they live in. These issues raise questions regarding what factors should and should not be controlled for in estimating neighborhood effects. The limitations of using standard multivariate adjustment factors to estimate direct and indirect effects (Robins &

\*Tel.: +1-734-615-9204; fax: +1-734-998-0006.

E-mail address: adierou@umich.edu (A.V. Diez Roux).

Greenland, 1992) or to control for factors which are simultaneously mediators and confounders (Robins, 1989) have been noted and methods have been proposed to better account for these complex, and more realistic, causal webs (Robins, Blevins, Ritter, & Wulfsohn, 1992; Robins, Hernan, & Brumback, 2000). Nevertheless, the application of these methods is still limited. The extent to which they will be useful in the analysis of the purely observational data available in many studies of neighborhoods and health remains to be determined, but they are certainly worth pursuing in neighborhood effects research. Longitudinal studies which examine the relationship between neighborhood attributes (or changes in neighborhood attributes over time) and the development of individual-level characteristics over the lifecycle may also shed light on the confounder vs. mediator question (Duncan & Raudenbush, 1999).

### Building the “causal multilevel model for neighborhood effects”

Oakes proposes a series of steps to construct what he refers to as “a causal multilevel model for neighborhood effects”. The steps he proposes are logical, based on existing multilevel analysis literature, and have often been followed (either explicitly or implicitly) by neighborhood effects researchers. As in model building generally, there is no single way to build a “causal multilevel model”, and using a given approach does not guarantee that the estimates obtained will be valid estimates of causal parameters. Although using the most appropriate model for the research question at hand is of course important, ultimately models are simply tools that help us describe the data. Inferring causality is a much more complicated process and requires more than statistical models.

Many approaches to model building in multilevel analysis (including the one described by Oakes) begin by examining estimates of between neighborhood variability (the variance of the random effect  $U_g$ , in Oakes’s notation) in order to determine whether “neighborhood effects” (i.e. the effects of specific neighborhood attributes on individual-level outcomes) are worth examining. Although this approach appears logical, several caveats are in order. First, the power to detect the variance components (for example, the variance of random effects or random coefficients) is affected by the number of groups and the number of persons per group in a different manner than the power to detect fixed effects. For example, Snijders and Bosker (1999) show that for a fixed total sample size, the standard error of the association between a group-level variable and an individual-level outcome (e.g. the association between neighborhood availability of recreational spaces and the physical activity of individuals) may be minimized by

sampling many groups with relatively few observations per group. On the other hand, for relatively low intraclass correlations (common in social science research), small group sizes may result in large standard errors for the intraclass correlation coefficient estimated (see Snijders & Bosker, 1999) Chapter 10 for details). Thus, a given study may have insufficient power to detect between neighborhood variance and yet have sufficient power to detect the fixed effect of a specific neighborhood attribute. In fact, the test of the association of a specific neighborhood-level variable with the individual-level outcome will often have more power than the more diffuse test that the between neighborhood variance is zero. Thus, one should be wary of concluding that associations between neighborhood characteristics and individual-level variables are not worth examining because in a given study the variance of the random neighborhood effect is not statistically significant.<sup>1</sup>

Another important point is that the estimate of the variance of the random effects is conditional on level 1 variables in the model. The absence of between neighborhood variability in the null model does not necessarily imply that significant between neighborhood variability will not emerge once relevant individual-level variables are accounted for. But what if no significant between neighborhood variation remains after all relevant individual-level variables are included? The problem here is that, if the estimates of the level 1 coefficients are confounded by an omitted neighborhood-level variable, the estimate of the neighborhood level variance may be biased downward. Therefore, the only way to test a specific hypothesis about a neighborhood effect is to explicitly test that hypothesis rather than rely on estimates of neighborhood-level variance to determine whether or not the hypothesis is worth testing. Thus, in the presence prior theory and a specific hypothesis regarding the effects of a given neighborhood attribute, it may make sense to test that hypothesis, regardless of the between neighborhood variance estimate obtained. Of course, in the absence of prior

<sup>1</sup> Most research to date is likely to have grossly underspecified neighborhood contexts due to the use of very crude census proxies for the relevant areas. This will obviously affect estimates of between neighborhood variance in outcomes. It is also possible that the effects of individual-level variables vary significantly across neighborhoods (i.e. the variance of the random coefficients is large) suggesting the presence of cross-level interactions even if the variance of random neighborhood effects or intercepts is estimated as 0 (Raudenbush & Wilms, 1995). Raudenbush and Wilms (1995) discuss how in the presence of cross-level interactions, group-level (e.g. school-level or neighborhood) effects can influence within group variability, further highlighting the limitations of basing inferences regarding group effects exclusively on the presence of between group variability.

theory, the failure to reject the null hypothesis regarding between neighborhood variance should discourage testing for the effects of multiple neighborhood level variables.<sup>2</sup>

In addition, as noted by Duncan and Raudenbush (1999) “effect sizes that program evaluators commonly viewed as medium or even large translate into small proportions of variance in individual outcomes “explained” by neighborhood membership and into small intraneighborhood correlations.” (Duncan & Raudenbush, 1999, p. 33). Therefore small within neighborhood correlations (even correlations close to 0) do not necessarily imply that the effects of neighborhood variables on individual-level outcomes are not worth investigating. A final caveat is that estimation procedures for multilevel logistic, Poisson, and survival models are an area of ongoing research (Brown & Prescott, 1999; Goldstein & Rasbash, 1996; Raudenbush & Bryk, 2002; Rodriguez & Goldman, 1995). Epidemiologists should be aware of the difficulties inherent in estimating the variance components in many of these models. For all the reasons outlined above, we should be wary of drawing inferences regarding the presence of neighborhood effects (or the need to investigate a specific hypothesis regarding the causal effect of a specific neighborhood attribute on a specific individual-level outcome) based on the presence or absence of significance between neighborhood variability in multilevel models.

Oakes also discusses the interpretation of coefficients associated with the level 1 variables and argues that “interest in, or interpretation of, level 1 effects is inconsistent with an etiological neighborhood effects methodology”. If the implication is that the inclusion of multiple, intercorrelated individual-level variables (included only for the purpose of controlling for individual-level confounding) limits the usefulness of interpreting their “independent effects” I agree. However, one of the strengths of multilevel models is that one can simultaneously examine associations of individual-level and group-level variables with the outcome. In the full multilevel model, the coefficients for individual-level variables quantify how each variable is associated with the outcome after adjusting for the group characteristics included and conditional on group random effects. Interpreting the individual-level associations as “causal” is of course subject to the same caveats as interpreting any adjusted association from an observational study as causal, but there is nothing inherent about the multilevel approach which precludes studying potentially causal associations of individual level variables with outcomes.

<sup>2</sup>I am greatly indebted to Steven Raudenbush and Harvey Goldstein for valuable insights on many of the issues discussed in this section. Any errors are my own.

A side issue worthy of comment is the problem of centering the level one variables in multilevel models. In the multilevel context, level 1 variables may be centered around the grand mean or around the group mean. Of particular relevance to the estimation of level 2 effects, centering around the group mean eliminates the possibility that the individual level variable (which was group-mean centered) will contribute to between group differences. Hence, contextual effects may be overestimated because they are in fact unadjusted for the compositional differences in the individual-level variable across groups. Centering may also affect estimation of the variance components (for details see Kreft, de Leeuw, & Aiken, 1995; Raudenbush & Bryk, 2002, Chapter 5).

Although it is often assumed that cross-level interactions are specific to the multilevel model, cross-level interactions may be included in standard models with individuals as the units of analysis. For example, a study of individuals could include in a standard regression equation both the individual-level variable, a neighborhood characteristic, and the interaction between both. If necessary, residual within neighborhood correlations can be accounted for using marginal models (or covariance pattern models) without resorting to multilevel models. A disadvantage of this approach is that one would not obtain estimates of between and within neighborhood variability, and the sources of the within neighborhood correlation would not be explicitly modeled as they are in the multilevel approach. As described elsewhere (Zeger, Liang, & Albert, 1988), there are also subtle differences between multilevel models and marginal models in the interpretation of regression coefficients. In multilevel models, the identification and interpretation of interactions raises issues similar to those described for interactions at the individual-level (including, for example, the fact that the presence of statistical interactions depends on whether effects are assumed to be additive or multiplicative).

Another important issue raised by Oakes pertains to the interpretation of the coefficient associated with the neighborhood (aggregate) socioeconomic status (SES) measure. It is possible as Oakes suggests that this coefficient is essentially picking up unmeasured differences in the individual-level SES of the residents of different neighborhoods. This interpretation assumes that these unmeasured individual-level SES characteristics are shared by persons within a given neighborhood.

The utility of using the neighborhood SES measure to control for unspecified interindividual SES differences is however questionable, since the aggregate SES measure of the neighborhood or area in which a person resides is often a poor proxy for the analogous individual-level construct (Diez Roux et al., 2001; Geronimus & Bound, 1998). An alternative interpretation of the neighborhood

SES variable (which underlies its use in neighborhood effects research), is that this variable is serving as a proxy for a variety of specific features of neighborhoods potentially relevant to health. Unpacking these specific features of neighborhoods and testing hypotheses about their relationship to specific health outcomes is what is needed to draw inferences regarding causal neighborhood health effects.

### Methodological obstacles

#### *Social stratification confounds comparisons*

Oakes correctly notes that if “there is (approximately) complete confounding between the background attributes of persons in a given neighborhood and (approximately) complete separation between background attributes of people in other neighborhoods” it will be impossible to identify neighborhood effects. But the extent to which individual and neighborhood characteristics are so strongly associated is an empirical question that can be examined in the data. Although neighborhood characteristics and individual-level SES are obviously associated, it is also possible that persons of similar family income live in very different types of neighborhoods. If sufficient variability in individual-level variables across categories of neighborhood attributes is documented then the estimation of adjusted associations is possible and meaningful. Rigorous studies of neighborhood effects will begin by carefully determining whether there is enough overlap between the neighborhood characteristic of interest and the relevant individual-level confounders to allow “separation” of their effects.

Oakes cites work on the estimation of teacher effects on children’s achievement after accounting for the social background of the students as an example of the appropriate use of the multilevel approach (Raudenbush & Willms, 1995). The example cited by Oakes differs from existing neighborhood effects research in that the specific research question pertains to estimation of the effects of school characteristics *after controlling for school socioeconomic context* (which may affect school outcomes through norms and peer influence) as well as student individual-level characteristics. Here there is a clear justification for including both types of variables because norms and peer influence (proxied by school aggregate SES measures) are hypothesized to be important to school outcomes, independently of the characteristics of schools (and vice versa). A similar approach could be used in neighborhood effects research, if one hypothesizes that neighborhood socioeconomic context is relevant to the outcome *independently* of other specific features of neighborhoods. Just like schools have teachers (with varying characteristics),

neighborhoods have specific features which may make them more or less healthy. If the specific question warrants it, it may be appropriate to include both aggregate SES measures and specific features of neighborhoods in the same model. A limitation of this approach however, is that correlations between these variables may be high, making their independent effects unidentifiable. It is worth emphasizing that Raudenbush and Willms (1995) include the school socioeconomic context variable in their model not as a way to control for selection (or unmeasured individual-level variables) but because it is hypothesized that socioeconomic context has an independent effect on school achievement (through its relationship with norms and peer influence). The main (and important) advantage of research on school effects is that the relevant context (school) and the potential attributes of interest are much better defined (and specified in the research) than they are in neighborhood effects research, where both the relevant “context” and its attributes remain poorly defined and consequently underspecified. However, both neighborhood effects and school effects research (including the example cited by Oakes) face similar problems regarding the need to control for residual confounding or selection effects related to individual (or family) socioeconomic position (since schools are as segregated by SES as neighborhoods are). The key issues in both types of research are essentially the same.

#### *Emergent contexts are endogenous*

As posited by Oakes, a crucial problem in research on neighborhood characteristics and health is that neighborhood effects are by definition endogenous to the compositional characteristics of neighborhoods. Another way to state this is to say that neighborhood characteristics are determined by the individual characteristics of the residents. As Oakes says “there is no exogenous intervention causing them”. Hence, any attempt to separate out the “independent” effects of neighborhood characteristics and individual-level socioeconomic indicators is futile because the former is endogenous to the latter. But is it true that all features of neighborhoods potentially related to health are indeed determined by the characteristics of individuals who reside in them?

Oakes’s specific focus appears to be on situations where the average dependent variable for a group (or neighborhood) is included as a predictor of individual level outcomes in order to study how the mean group outcome affects the individual-level outcome (Manski’s endogenous effects, i.e. when “the propensity of an individual to behave in some way varies with the prevalence of the behavior in the group” (Manski, 1995, p. 127), also referred to as contagion in

epidemiology).<sup>3</sup> Manski's "reflection problem" (cited by Oakes) arises out of the presence of the mean outcome of the group as a regressor in an equation that also includes the effects of context (e.g. the mean SES of individuals in the group) as well as individual characteristics or other group-level variables (for example, institutional environments). Manski shows that the endogenous effects (in Manski's terminology) are unidentifiable (in particular it is not possible to distinguish endogenous and contextual effects, although in some situations it may be possible to determine whether an overall social effect (defined as endogenous plus contextual effects), is present, for more details see Manski, 1995, Chapter 7). But this situation, in which a researcher attempts to separate out endogenous, contextual, and individual-level (or other group-level effects) is not typical of neighborhood effects research to date.

Once again, perhaps an important source of the confusion regarding the endogeneity of neighborhood effects is that health researchers have to date mostly relied on measures of neighborhood socioeconomic characteristics as proxies for neighborhood attributes, raising the specter of endogeneity because by definition aggregate measures are constructed by summarizing the characteristics of individuals residing in the neighborhood.<sup>4</sup> But few would argue that it is say, median area income *itself*, which is related to physical activity, but rather *something for which median income is serving as a proxy* (for example, features of the built environment like walkability). The use of these proxies (rather than direct measures of the neighborhood attributes of interest) has been a major limitation. Recent interest in conceptualizing and directly measuring neighborhood attributes is a promising trend in the field. And although even specific neighborhood attributes may result in part from the individual attributes of residents (for example, dietary habits of residents may influence neighborhood availability of healthy foods), it would be hard to argue that these attributes are fully endogenous and that exogenous factors related to policies and macro-economic forces do not play a significant role. Of course, some would argue that all "contexts" are endogenous to individuals (or are reducible to the

attributes of individuals). This would invalidate the notion of "neighborhood effects" (and in the extreme even social effects) altogether. But this is another and much broader debate, which, ultimately, is not about methodology. Nevertheless, Oakes is correct in pointing out the complexities of the problem. If neighborhood contexts are indeed causally related to health, the pathways involved are likely to be complex and involve reciprocal causation (or jointly dependent or endogenous variables) and feedback loops. Observational studies (and the analytical methods commonly used in epidemiology) are not always adequate and will not allow full elucidation of these pathways; other complementary approaches (both quantitative and qualitative) will be necessary.

#### *Extrapolation*

Oakes correctly questions whether people from two different neighborhoods are "exchangeable", even after controlling for the individual-level covariates usually included in contextual or multilevel models. In the absence of the exchangeability assumption, one cannot conclude that the associations observed between neighborhood characteristics and health reflect a causal effect of neighborhoods on health. This is essentially a reformulation of the problem of uncontrolled confounding due to omitted or mismeasured individual-level variables. A related (although I would argue distinct issue) is that, if there is little overlap between individual-level SES characteristics across neighborhoods, adjustment for these variables may not lead to a valid estimate of the "neighborhood" effect because it will be based purely on extrapolation (often with little or no support in the data, as illustrated in Fig. 2 in Oakes). This problem, however, is not unique to multilevel neighborhood effects models, and may occur in any adjusted comparison of two or more categories. The limitations of adjustment strategies in the presence of distributions with little overlap are well-established. A careful researcher will examine whether there is sufficient overlap in the distributions of a covariate across groups before estimating any covariate-adjusted associations. Of course, limitations of control strategies (including, for example, situations in which adjustment for a variable may actually increase confounding) remain (Greenland & Robins, 1986; Lieberman, 1985) as is true in epidemiology generally.

#### *Disequilibria*

Is it true that "SUTVA [the stable unit treatment value assumption] is violated by most notions of neighborhood effects" (Oakes, 2003, p. XXX)? SUTVA implies that that the response for a given unit depends on the treatment assigned to that unit but not on the

<sup>3</sup> It is important to note that Manski's concept of endogenous effects refers to one of the hypotheses proposed to explain why individuals belonging to the same group behave similarly. It is related but is not synonymous with the more general concept of endogenous variables (or jointly dependent variables) used in economics.

<sup>4</sup> Similar arguments which allude to the interrelatedness of individual and aggregate measures are raised to critique the interpretation of coefficients from multilevel models including both individual-level and aggregate SES measures. See Greenland (2002). A review of multilevel theory of ecologic analyses *Stat Med*, 21, 389–395.

treatment assigned to other units (Rubin, 1986). Moving many poor persons to a rich neighborhood will change the “rich” neighborhood, and therefore the treatment (if treatment is defined as moving to the high mean income neighborhood) given to one person will affect the treatment received by others (because the treatment, i.e. exposure to a high mean neighborhood income of the rich neighborhood, will itself change as more and more poor people are moved into the rich neighborhood). Under this scenario SUTVA is violated because the treatment received by others may affect the response for a given individual. As Oakes notes, the “rich” neighborhood itself will ultimately change becoming a poor neighborhood. But determining what happens when groups of poor people move *en masse* to rich neighborhoods is not what most studies of neighborhood effects are trying to get at. Rather they are trying to determine whether attributes of neighborhoods (such as availability of recreational resources, violence levels, or social cohesion) are related to health. Again, the confusion regarding whether SUTVA is violated may stem in part from misunderstanding of (and problems inherent in) the use of aggregate neighborhood socio-economic characteristics as surrogate measures of the real neighborhood attributes of interest.

Some of the questions that studies of neighborhood effects are trying to answer may violate SUTVA but others may not. Why would one person’s exposure to a residential environment with many opportunities for recreation affect another person’s exposure? Why would a given individual’s response to this neighborhood exposure depend on the exposure of others? It is possible, of course that the neighborhood prevalence of physical activity modifies the relationship between neighborhood physical environment and physical activity. In this case SUTVA is violated unless neighborhood prevalence of physical activity (and its interaction with the exposure variable) is accounted for. Violation of SUTVA is more likely when contagion processes are present, and these are not properly accounted for in the analysis. But not all hypotheses regarding neighborhood health effects involve contagion processes. Even if potential violation of SUTVA exists for a specific research problem, it may be more productive to evaluate the sensitivity of results to this violation and determine if anything can be done to minimize it, rather than discard observational studies of neighborhood effects altogether.

#### *Community trials as an alternative*

No one would disagree that randomized community trials have much to contribute to the investigation of neighborhood health effects. But, as noted by Oakes, they have their own set of limitations, not least of which is that for many neighborhood attributes of interest it

may not be logistically feasible or ethical to design a randomized community trial. They also have their own set of methodological problems as illustrated by the literature on community intervention trials to prevent cardiovascular disease (Fortmann et al., 1995; Koepsell, Diehr, Cheadle, & Kristal, 1995; Murray, 1995). Oakes suggests that community trials encourage epidemiologists to consider “practicable” interventions, as opposed to indulging themselves by considering what things might be like, *if the world were just a little different*. Feasibility in a community trial is not analogous to feasibility in the real world of course. A more fundamental problem is that what is “practicable” is often a matter of perspective. For example, in today’s world it often appears more “practicable” to change a person’s genes than to modify the way society is organized. But perhaps it may be good for the public’s health to consider “what *could* be done *if the world were just a little different*”, as Oakes says. Isn’t this what the application of science to human problems is all about? Of course, elucidating *what could be done if the world were just a little different* may require many approaches, including (but not limited to) community trials.

#### **Where does this leave observational studies of neighborhood effects?**

Have contextual and multilevel studies of neighborhood health effects concluded that the observed associations are causal? No. Should the conditional measures of associations they have generated be taken literally as valid and definite measures of causal effects parameters? Of course not, as most rigorous studies of neighborhood effects note. But although still extremely crude, taken together with other data (plus many of our daily living experiences which are after all, an important source of hypotheses for research), the associations report suggest that there is something there worth looking into. So where do we go from here? Oakes recommendations are consistent with those of others in the field: the need to combine focused qualitative together with quantitative approaches, the need to take advantage of interventions, and the need for methods that deal better with complex causality and reciprocal relationships between factors. Oakes does not elaborate on what he means specifically by “dependency-oriented theory and methodology”. If what is implied is that we need to develop theories and methods for the study of complex causal paths with multiple reciprocal and non-linear relationships that take into account dynamic relationships between individuals and between neighborhoods, I would agree wholeheartedly. Of course this is easier said than done. Just as over simplification limits our ability to understand, the inability to simplify can be paralyzing. As Levins has so aptly stated “The art of

research is the sensitivity to decide when a useful and necessary simplification has become and obfuscating simplification.” (Levins, 1996, p. 105).

I would disagree with Oakes that given the problems he outlines, issues of model specification and measurement are moot. In fact, they are an integral part of the causal inference problem in neighborhood effects research. The testing of theory-based, specific hypotheses regarding the processes through which specific area attributes may affect specific health outcomes (which necessarily includes better definition of relevant areas and the relevant variables, and improved measurement) are fundamental to the broader process of drawing causal inference. Ultimately this implies better specification of the counterfactual contrast of interest. Issues related to residual confounding, the role of individual level variables, the need to account for reciprocal effects, and better attention to study design have been discussed in this context (Diez Roux, 2001; Macintyre, Ellaway & Cummins, 2002; Sampson et al., 2002). In the absence of theory and at least some evidence (even if observational and limited) regarding the specific processes involved, intervention studies will be impracticable or uninformative.

The associations reported in the epidemiologic literature on neighborhood health effects are what they are: measures of conditional associations under certain assumptions. People of course can argue about what they mean. I doubt however that better inference on neighborhood health effects will flow exclusively from refining quantitative approaches (although this would of course be desirable). It will most likely come from combinations of evidence drawn from interventions, natural experiments, more focused observational work, and qualitative studies. An unfortunate consequence of the popularity of multilevel modeling approaches is that neighborhood effects research has come to be perceived by many as synonymous with multilevel analysis which of course (and fortunately!) it is not. A more fundamental question is how much evidence we need in order to conclude that health policy should include interventions on residential environments together with current individual-based strategies. But this would be the subject of another debate.

## References

- Brown, H., & Prescott, R. (1999). *Applied mixed models in medicine*. New York: Wiley.
- Diez Roux, A. V. (2001). Investigating neighborhood and area effects on health. *American Journal of Public Health, 91*, 1783–1789.
- Diez Roux, A. V., Kiefe, C. I., Jacobs Jr., D. R., Haan, M., Jackson, S. A., Nieto, F. J., Paton, C. C., & Schulz, R. (2001). Area characteristics and individual-level socioeconomic position indicators in three population-based epidemiologic studies. *Annals of Epidemiology, 11*, 395–405.
- Duncan, G., Connell, J., & Klebanov, P. (1997). Conceptual and methodological issues in estimating causal effects of neighborhoods and family conditions on individual development. In J. Brooks-Gunn, G. Duncan, & J. Aber (Eds.), *Neighborhood poverty: Vol I. Context and consequences for children* (pp. 219–250). New York: Russell Sage.
- Duncan, G., & Raudenbush, S. W. (1999). Assessing the effects of context in studies of child and youth development. *Educational Psychologist, 34*, 29–41.
- Fortmann, S. P., Flora, J. A., Winkleby, M. A., Schooler, C., Taylor, C. B., & Farquhar, J. W. (1995). Community intervention trials: Reflections on the Stanford five-city project experience. *American Journal of Epidemiology, 142*, 576–585.
- Furstenberg, F., & Hughes, M. (1997). The influence of neighborhoods on children’s development. In J. Brooks-Gunn, G. Duncan, & J. Aber (Eds.), *Neighborhood poverty: Vol II. Policy implications for studying neighborhoods* (pp. 23–47). New York: Russell Sage.
- Geronimus, A. T., & Bound, J. (1998). Use of census-based aggregate variables to proxy for socioeconomic group: Evidence from national samples. *American Journal of Epidemiology, 148*, 475–486.
- Goldstein, H., & Rasbash, J. (1996). Improved approximations for multilevel models with binary responses. *Journal of the Royal Statistical Society A, 159*, 505–513.
- Greenland, S. (2002). A review of multilevel theory for ecologic analyses. *Statistic in Medicine, 21*, 389–395.
- Greenland, S., & Robins, J. M. (1986). Identifiability, exchangeability, and epidemiological confounding. *International Journal of Epidemiology, 15*, 413–419.
- Harding, D. (2002). Counterfactual models of neighborhood effects: The effect of neighborhood poverty on high school dropout and teenage pregnancy.
- Joffe, M. M., & Rosenbaum, P. R. (1999). Invited commentary: Propensity scores. *American Journal of Epidemiology, 150*, 327–333.
- Koepsell, T. D., Diehr, P. H., Cheadle, A., & Kristal, A. (1995). Invited Commentary: Symposium on community intervention trials. *American Journal of Epidemiology, 142*, 594–599.
- Kreft, I., de Leeuw, J., & Aiken, L. (1995). The effect of different forms of centering in hierarchical linear models. *Multivariate Behavioral Research, 30*, 1–21.
- Levins, R. (1996). Ten propositions on science and antisience. *Social Text, 14*, 101–111.
- Lieberman, S. (1985). *Making it count: The improvement of social research and theory*. Berkeley, CA: University of California Press.
- Macintyre, S., Ellaway, A., & Cummins, S. (2002). Place effects on health: How can we conceptualise, operationalise and measure them? *Social Science & Medicine, 55*, 125–139.
- Manski, C. F. (1995). *Identification problems in the social sciences*. Cambridge, MA: Harvard University Press.
- Murray, D. M. (1995). Design and analysis of community trials: Lessons from the minnesota heart health program. *American Journal of Epidemiology, 142*, 569–575.
- Oakes, J. M. (2003). The (mis)estimation of neighborhood effects: Causal inference for a practicable social

- epidemiology. *Social Science & Medicine* doi:10.1016/j.socscimed.2003.08.004.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods*. London: Sage.
- Raudenbush, S. W., & Willms, J. D. (1995). The estimation of school effects. *Journal of Educational and Behavioral Statistics*, 20, 307–335.
- Robins, J. (1989). The control of confounding by intermediate variables. *Statistics in Medicine*, 8, 679–701.
- Robins, J. M., Blevins, D., Ritter, G., & Wulfsohn, M. (1992). G-estimation of the effect of prophylaxis therapy for *Pneumocystis carinii* pneumonia on the survival of AIDS patients. *Epidemiology*, 3, 319–336.
- Robins, J. M., & Greenland, S. (1992). Identifiability and exchangeability for direct and indirect effects. *Epidemiology*, 3, 143–155.
- Robins, J. M., Hernan, M. A., & Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11, 550–560.
- Rodriguez, G., & Goldman, N. (1995). An assessment of estimation procedures for multilevel models with binary responses. *Journal of Royal Statistical Society A*, 158, 73–89.
- Rubin, D. B. (1986). Which ifs have causal answers? *JASA*, 81, 961–962.
- Rubin, D. B. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine*, 127, 757–763.
- Sampson, R. J., Morenoff, J. D., & Gannon-Rowley, T. (2002). Assessing “neighborhood effects”: Social processes and new directions in research. *Annual Review of Sociology*, 28, 443–478.
- Snijders, T. A. B., & Bosker, R. (1999). *Multilevel analysis: An introduction to basic and advanced multilevel modeling*. London: Sage.
- Tienda, M. (1991). Poor people and poor places: Deciphering neighborhood effects on poverty outcomes. In J. Huber (Ed.), *Macro-micro linkages in sociology* (pp. 244–262). Newbury Park: Sage.
- Zeger, S. L., Liang, K. Y., & Albert, P. S. (1988). Models for longitudinal data: A generalized estimating equation approach. *Biometrics*, 44, 1049–1060.