



Next steps in understanding the multilevel determinants of health

A V Diez Roux

J Epidemiol Community Health 2008;62:957-959
doi:10.1136/jech.2007.064311

Updated information and services can be found at:
<http://jech.bmj.com/cgi/content/full/62/11/957>

These include:

References

This article cites 33 articles, 12 of which can be accessed free at:
<http://jech.bmj.com/cgi/content/full/62/11/957#BIBL>

1 online articles that cite this article can be accessed at:
<http://jech.bmj.com/cgi/content/full/62/11/957#otherarticles>

Rapid responses

You can respond to this article at:
<http://jech.bmj.com/cgi/eletter-submit/62/11/957>

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article

Notes

To order reprints of this article go to:
<http://journals.bmj.com/cgi/reprintform>

To subscribe to *Journal of Epidemiology and Community Health* go to:
<http://journals.bmj.com/subscriptions/>

Next steps in understanding the multilevel determinants of health

A V Diez Roux

Correspondence to:
Dr A V Diez Roux, Department of
Epidemiology, University of
Michigan, 1214 South University
2nd floor, Ann Arbor MI 48103,
USA; adiezrou@umich.edu

Accepted 31 January 2008

ABSTRACT

This commentary briefly summarises past work that has used multilevel analysis to investigate the multilevel determinants of health and outlines possible new directions in this area. Topics discussed include the need to (1) examine contexts other than neighbourhoods; (2) improve measurement of group-level constructs; (3) apply techniques more appropriate for causal inference from observational data; (4) analyse data from "natural experiments" involving exogenous variations in contextual characteristics; (5) examine dependencies between groups (such as spatial dependencies) more broadly and allow for reciprocal relations between individuals and contexts; and (6) contrast multilevel statistical models (or regression models generally) and complex systems models in the study of multilevel effects.

It has been approximately 10 years since the irruption of multilevel analysis in epidemiology and public health.¹⁻³ As is the case with many new techniques, multilevel analysis (or more specifically multilevel statistical models) was initially received with great enthusiasm and expectation, fundamentally because it allowed the empirical and quantitative investigation of group-level effects on health, avoiding both the methodologic limitations of ecologic studies and the often naïve and sometimes scientifically misleading methodologic individualism of traditional individual-based epidemiologic approaches.⁴ Although the presence of group-level effects has been posited in epidemiology for decades (and to many the explicit recognition of group-level or population-level phenomena is in fact one of the distinguishing features of epidemiology), the empirical investigation of these effects was often limited to ecologic studies and subject to critiques primarily related to limitations of these studies in separating true group-level from purely compositional effects. In contrast, multilevel analysis allowed a relatively simple operationalisation of the contributions of individual and group-level factors to both between-individual and between-group variability, showing how individual-level and group-level factors can contribute to variability at both levels, and transcending the artificial dichotomy of individuals and groups. Building on the concepts discussed by Geoffrey Rose over 20 years ago,⁵ multilevel analysis allowed for the possibility that different factors contribute to within-group and between-group variability, and permitted estimation of group-level effects after accounting for compositional differences across groups.

A comprehensive review of the results of studies applying multilevel analysis is beyond the scope of

this brief commentary, so only a very general assessment (with some illustrative but not exhaustive examples) is provided here. Although multilevel analysis is applicable to the study of a broad range of "groups" or contexts, the vast majority of applications in the health field have focused on geographically defined contexts, such as countries,⁶ states,⁷ counties,⁸ and most commonly "neighbourhoods" defined in various ways, most commonly by smaller administrative areas.⁹⁻¹¹ The types of group-level constructs investigated have included, for example, income inequality,¹² social capital,¹³ residential segregation,¹⁴ women's status,¹⁵ and neighbourhood characteristics such as neighbourhood disadvantage or other measures of neighbourhood social and physical environments.^{10 11 16 17} Most studies have used multilevel analysis to isolate associations of group-level factors with individual-level health outcomes after accounting for individual-level confounders (i.e. individual-level variables associated with the health outcomes and with group membership, and, therefore, with group characteristics). A smaller number have focused on the complementary objective of decomposing variance into between- and within-group components.¹⁸

Overall, the results of multilevel analyses published to date are consistent with main effects of a variety of group-level variables on individual-level outcomes that persist after controlling for individual-level variables. The strength of this main effect has varied substantially depending on the study and the research question investigated. The detection of these group effects is striking given their often very distal relationship to the health outcomes being studied, the misspecification of groups and group-level variables, and the often extensive adjustment for much better measured individual-level variables, many of which are mediators rather than true confounders of the group-level effects. Generally, the per cent of total variance in the individual-level health outcome that is between groups (as compared to within groups) has been small. However, this result must be viewed in light of the fact that the relevant "groups" are generally grossly misspecified, that partitioning variance is complex for health outcomes that are not continuous variables, and that even well established individual-level risk factors often explain only a very small amount of the observed inter-individual variability.

Perhaps even more important than the specific empirical results obtained to date, multilevel analysis has stimulated and promoted multilevel thinking generally within epidemiology, challenging researchers to begin to think very specifically

Research agenda

about the various levels of organisation, and the constructs at each level that are relevant to the health outcome they are studying. By allowing empirical analysis of multilevel data, multilevel analysis has also challenged epidemiologists to move beyond theory and speculation and operationalise and empirically test specific aspects of the theoretical models. However, exponential growth of use of multilevel analysis to investigate the multilevel determinants of health has also brought to the forefront limitations of work done to date. The remainder of this commentary discusses research needs in the investigation of multilevel determinants of health generally. It is important to note that investigation of the multilevel determinants involves much more than the fitting of multilevel statistical models. Thus, many of the issues noted apply to the general topic of investigating multilevel determinants regardless of whether multilevel statistical models are required.

MOVING BEYOND NEIGHBOURHOODS TO INVESTIGATE OTHER CONTEXTS

Due in part to practical reasons related to data availability, neighbourhoods have been one of the most common contexts investigated using multilevel analysis. Neighbourhoods are especially challenging to investigate due to (1) limited theory on the neighbourhood features that may be relevant to health; (2) difficulties in defining the spatial scale (and therefore neighbourhood definition) relevant to different health-related processes; (3) challenges in measuring neighbourhood level attributes and disentangling effects of correlated neighbourhood characteristics; and (4) complexities in accounting for confounding for individual-level variables related to the sorting of individuals into neighbourhoods in observational studies.¹⁹ Neighbourhoods may not be the most relevant contexts for many health outcomes. Nevertheless, application of multilevel analysis in epidemiology has become almost synonymous with investigation of neighbourhood health effects. An important need is, therefore, to expand research into health effects of other well-defined contexts (e.g. countries or other policy-relevant units, schools, workplaces) with modifiable features likely to be related to health. This will require thinking theoretically, but also specifically, about what contexts and features of contexts are likely to be most relevant to different health outcomes.

IMPROVING THE MEASUREMENT OF GROUP-LEVEL CONSTRUCTS

A second research need is the improved measurement of group-level constructs at different levels of organisation. In the absence of adequate measurement, not even the most sophisticated analytical techniques will allow convincing causal inference. The reliance on existing data has obligated researchers to use existing and often imperfect proxies for measures of the true group-level constructs they are interested in. This has limited the informativeness of the analyses and the inferences that can be drawn from them. Recent work in neighbourhood effects research has moved towards the development and validation of measures of neighbourhood-level constructs,^{20 21} but continued work on the measurement of a variety of contexts or ecologic settings is needed.

APPLYING NEW METHODS THAT MAY ENHANCE THE ABILITY TO DRAW CAUSAL INFERENCES REGARDING GROUP EFFECTS FROM OBSERVATIONAL DATA

The growing popularity of multilevel analysis and the publication of more and more results using this technique has also

highlighted the difficulties inherent in drawing causal inferences regarding the presence of group-level effects through analyses of observational data.²²⁻²⁴ Chief among these is the difficulty in fully accounting for unmeasured characteristics related to the sorting of individuals into different contexts. In parallel, there has been a growing discussion in epidemiology on the threats to causal inference in observational studies generally, and on the use of tools and new analytic techniques that may help researchers design their analyses and interpret their findings more appropriately. Consideration of how these approaches (including directed acyclic graphs, propensity score matching, instrumental variables, and marginal structural models)²⁵⁻²⁹ may be used to improve multilevel causal inference, the evaluation of their impact on results in real-life scenarios (through for example sensitivity analyses and simulations), and the contrast of results obtained using traditional approaches with those obtained using more sophisticated methods are important areas for new research.

TAKING ADVANTAGE OF NATURAL EXPERIMENTS

Another important area for new research is the study of the health impact of the changes in contextual characteristics, which are occurring in the real world all the time for reasons completely unrelated to health. For example, neighbourhoods are constantly changing as a result of policies or the organised efforts of residents. The study of the health consequences of these changes (which can be thought of as "natural experiments" often involving exogenous variations in contextual characteristics) may sometimes help address important limitations in our ability to draw causal conclusions from pure observational analyses. In addition, from the practical point of view, it could yield important insights into which interventions work and which do not work in the real world.

ACCOUNTING FOR DEPENDENCIES BETWEEN GROUPS

As noted above, many published multilevel analyses focus on geographic contexts. These applications have brought place and geography back into epidemiology, but they have also simplified space by fragmenting it into artificially independent units, and have for the most part ignored how areas interact and how features of adjacent areas affect residents of a given area. For example, the physical activity of a resident of a given area may be affected not only by local resources but also by resources in surrounding areas. In addition, the effect of the characteristics of a local area on residents may be modified by features of surrounding areas. For example, the effect of violence levels on physical activity may be greater when the area is also surrounded by high violence areas. In these cases, the outcome for a given individual may depend not only on the characteristics of their neighbourhood, but also on the characteristics of other neighbourhoods. This can be thought of as violation of the non-interference assumption at the neighbourhood-level. If this assumption is violated it will be necessary to use techniques which relax the assumption and allow for the modelling of spillover effects (which can be done in a multilevel context).³⁰ More generally, recent work has begun to explore alternative ways of incorporating space and interdependencies over space into the study of contextual effects.^{31 32}

USING METHODS APPROPRIATE FOR THE STUDY OF DYNAMIC SYSTEMS

Like other regression models used in epidemiology, multilevel models model health outcomes as a function of "independent"

New directions in the investigation of multilevel determinants

- ▶ Examine effects of contexts other than neighbourhoods.
- ▶ Improve measurement of relevant group-level variables.
- ▶ Apply techniques more appropriate for causal inference using observational data and evaluate impact on results in realistic scenarios.
- ▶ Seek out and analyse data from “natural experiments” involving exogenous variations in contextual characteristics over time, geographic regions, or other well-defined contexts such as workplaces or schools.
- ▶ Develop methodologic approaches that (1) examine dependencies between groups (such as spatial dependencies) and their sources more broadly; and (2) allow for and explore reciprocal relations and feedback loops between individuals and contexts.
- ▶ Contrast multilevel statistical models (or regression models generally) and complex systems models in the study of multilevel effects.

variables. They implicitly assume that these effects can be isolated from each other and do not allow for feedback loops or reciprocal interactions between groups and individuals, or between outcomes and predictors. In so doing, they simplify a process that occurs within a complex system. For example, they allow for the fact that neighbourhood features affect individuals but do not allow for the feedback mechanisms through which individuals shape neighbourhood features. Recent interest in applications of systems approaches to health^{35, 34} has highlighted the potential utility of methods such as agent-based models or dynamic systems models in understanding the determinants of health.^{35, 36} Applications of these approaches to multilevel questions, and research that contrasts the insights obtained from multilevel models and systems models would be important contributions to the field.

Ten years is not a long time. And yet in this relatively short period multilevel analysis has become a mainstream technique, to the point where it is used quite extensively, and even in situations where, strictly speaking, it is not necessary and simpler methods would suffice. The great contribution of multilevel analysis has been as a promoter of multilevel thinking and of the development and testing of multilevel questions in epidemiology. This is no small contribution. But it is important to return to the basic questions that motivated multilevel analyses and think of what modifications or complementary techniques are necessary to address these questions, so that the method will not become an end in itself but rather a tool, which together with other tools, will help us comprehend complex aspects of the world we want to understand, intervene on, and change in order to improve health.

Competing interests: None declared.

REFERENCES

1. **Duncan C**, Jones K, Moon G. Health-related behaviour in context: a multilevel modelling approach. *Soc Sci Med* 1996;**42**:817–30.
2. **Diez Roux AV**, Nieto FJ, Muntaner C, *et al*. Neighborhood environments and coronary heart disease: a multilevel analysis. *Am J Epidemiol* 1997;**146**:48–63.
3. **Diez Roux AV**. Multilevel analysis in public health research. *Annu Rev Public Health* 2000;**21**:171–92.
4. **Diez Roux AV**, Schwartz S, Susser E. Ecologic studies and ecologic variables in public health research. In: *The Oxford textbook of public health*. Detels R, McEwen J, Beaglehole R, Tanaka H, eds. London: Oxford University Press, 2002.
5. **Rose G**. Sick individuals and sick populations. *Int J Epidemiol* 1985;**14**:32–8.
6. **Chung H**, Muntaner C. Welfare state matters: a typological multilevel analysis of wealthy countries. *Health Policy* 2007;**80**:328–39.
7. **Kim D**, Kawachi I. U.S. state-level social capital and health-related quality of life: multilevel evidence of main, mediating, and modifying effects. *Ann Epidemiol* 2007;**17**:258–69.
8. **Muntaner C**, Li Y, Xue X, *et al*. County level socioeconomic position, work organization and depression disorder: a repeated measures cross-classified multilevel analysis of low-income nursing home workers. *Health Place* 2006;**12**:688–700.
9. **Chaix B**, Rosvall M, Merlo J. Neighborhood socioeconomic deprivation and residential instability: effects on incidence of ischemic heart disease and survival after myocardial infarction. *Epidemiology* 2007;**18**:104–11.
10. **Rundle A**, Diez Roux AV, Free LM, *et al*. The urban built environment and obesity in New York City: a multilevel analysis. *Am J Health Promot* 2007;**21**(4 Suppl):326–34.
11. **Winkleby M**, Sundquist K, Cubbin C. Inequities in CHD incidence and case fatality by neighborhood deprivation. *Am J Prev Med* 2007;**32**:97–106.
12. **Subramanian SV**, Kawachi I. Whose health is affected by income inequality? A multilevel interaction analysis of contemporaneous and lagged effects of state income inequality on individual self-rated health in the United States. *Health Place* 2006;**12**:141–56.
13. **Kim D**, Subramanian SV, Gortmaker SL, *et al*. US state- and county-level social capital in relation to obesity and physical inactivity: a multilevel, multivariable analysis. *Soc Sci Med* 2006;**63**:1045–59.
14. **Bell JF**, Zimmerman FJ, Almgren GR, *et al*. Birth outcomes among urban African-American women: a multilevel analysis of the role of racial residential segregation. *Soc Sci Med* 2006;**63**:3030–45.
15. **Chen YY**, Subramanian SV, Acevedo-Garcia D, *et al*. Women's status and depressive symptoms: a multilevel analysis. *Soc Sci Med* 2005;**60**:49–60.
16. **Sundquist K**, Theobald H, Yang M, *et al*. Neighborhood violent crime and unemployment increase the risk of coronary heart disease: a multilevel study in an urban setting. *Soc Sci Med* 2006;**62**:2061–71.
17. **Fone D**, Dunstan F, Lloyd K, *et al*. Does social cohesion modify the association between area income deprivation and mental health? A multilevel analysis. *Int J Epidemiol* 2007;**36**:338–45.
18. **Merlo J**, A splund K, Lynch J, *et al*. Population effects on individual systolic blood pressure: a multilevel analysis of the World Health Organization MONICA Project. *Am J Epidemiol* 2004;**159**:1168–79.
19. **Diez Roux AV**. Neighborhoods and health: where are we and where do we go from here? *Rev Epidemiol Sante Publique* 2007;**55**:13–21.
20. **Mujahid MS**, Diez Roux AV, Morenoff JD, *et al*. Assessing the Measurement Properties of Neighborhood Scales: From Psychometrics to Econometrics. *Am J Epidemiol* 2007;**165**:858–67.
21. **Moudon AV**, Lee C, Cheadle AD, *et al*. Attributes of environments supporting walking. *Am J Health Promot* 2007;**21**:448–59.
22. **Oakes JM**. The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Soc Sci Med* 2004;**10**:1929–52.
23. **Diez Roux AV**. Estimating neighborhood health effects: the challenges of causal inference in a complex world. *Soc Sci Med* 2004;**58**:1953–60.
24. **Subramanian SV**. The relevance of multilevel statistical methods for identifying causal neighborhood effects. *Soc Sci Med* 2004;**58**:1961–7.
25. **Cole SR**, Hernan MA. Fallibility in estimating direct effects. *Int J Epidemiol* 2002;**31**:163–5.
26. **Hernan MA**, Robins JM. Estimating causal effects from epidemiological data. *J Epidemiol Community Health* 2006;**60**:578–86.
27. **Hernan MA**, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 2000;**11**:561–70.
28. **Glymour M**. Using causal diagrams to understand common problems in social epidemiology. In: *Methods in social epidemiology*. Oakes MJ KJ, ed. San Francisco: Jossey Bass, 2006.
29. **Harding D**. Counterfactual models of neighborhood effects: the effect of neighborhood poverty on high school dropout and teenage pregnancy. *AJS* 2003;**109**:676–719.
30. **VanderWeele T**. Ignorability and stability assumptions in neighborhood effects research. *Stat Med* 2008;**27**:1934–43.
31. **Chaix B**, Leyland AH, Sabel CE, *et al*. Spatial clustering of mental disorders and associated characteristics of the neighbourhood context in Malmö, Sweden, in 2001. *J Epidemiol Community Health* 2006;**60**:427–35.
32. **Auchincloss AH**, Diez Roux AV, Brown DG, *et al*. Association of insulin resistance with distance to wealthy areas: the multi-ethnic study of atherosclerosis. *Am J Epidemiol* 2007;**165**:389–97.
33. **Pearce N**, Merletti F. Complexity, simplicity, and epidemiology. *Int J Epidemiol* 2006;**35**:515–9.
34. **Diez Roux AV**. Integrating social and biologic factors in health research: a systems view. *Ann Epidemiol* 2007;**17**:569–74.
35. **Gorman DM**, Mezic J, Mezic I, *et al*. Agent-based modeling of drinking behavior: a preliminary model and potential applications to theory and practice. *Am J Public Health* 2006;**96**:2055–60.
36. **Homer JB**, Hirsch GB. System dynamics modeling for public health: background and opportunities. *Am J Public Health* 2006;**96**:452–8.