Health and neighbourhood

Neighbourhood influences on health

I Kawachi, S V Subramanian

Outstanding issues in the neighbourhood research agenda

Although multilevel studies help to tease apart contextual from compositional influences on health, they do not in themselves consider other threats to causal inference, particularly selection and endogeneity.[1] Endogeneity occurs when people choose to move to a particular neighbourhood—for example, one with cleaner air or medical amenities—because of an existing health problem (reverse causation). Endogeneity can also occur because of the presence of unobserved common prior causes of neighbourhood-level exposures and health outcomes (confounding)—for example, it is commonly supposed that the presence of fast-food outlets in a neighbourhood increases the risk of obesity for local residents. However, it is equally plausible that the decision of fast food franchises to open their businesses in particular locations occurs in response to the tastes of local residents. In this instance, taste for fatty food is an unobserved variable that is related to both the location of outlets as well as the risk of obesity. Generally speaking, epidemiological studies to date have seldom attempted to deal with these threats to causal inference.

Arguably, the problems we have described could be overcome by collecting data on a comprehensive range of unobserved variables and controlling for them. Alternatively, analysts could overcome some of the limitations of observational data by importing methods developed in other social sciences, such as instrumental variable estimation.[2] Instrumental variable estimation has long been used in economics. The goal is to manipulate the exposure of interest (eg, neighbourhood poverty) by identifying variables (instruments) that cause exogenous variation in that exposure—for example, the evacuation of the residents of New Orleans to hundreds of different communities across the United States in the wake of Hurricane Katrina could be considered a classic "instrument" (provided that the location of destination communities was a matter of a random lottery, which it seems to have been). Unfortunately, in this specific instance, any attempt to examine the health effects of different neighbourhood contexts appears to have been lost because government authorities did not keep records of where the residents were evacuated to.

Cutler and Glaeser[3] used instrumental variable estimation to examine the effects of residential segregation of racial and ethnic groups in the US on their schooling, employment and single parenthood rates. To control for the endogeneity of location choice, the researchers used instruments based on topography (number of rivers in a metropolitan area) and public finance characteristics of local governments that increase the benefits of segregation. Their results suggested that a one standard deviation decrease in segregation would eliminate one third of the black–white disparity in the outcomes examined.[3] More detailed descriptions of instrumental variable estimation procedures can be found elsewhere.[2 4 5]

A different approach that attempts to mimic the experimental study design is propensity score matching.[6] This approach attempts to estimate causal effects by contrasting outcomes between groups matched on the probability of being assigned to the exposure of interest (eg, the probability of living in a high-poverty neighbourhood), and is gaining recognition in social epidemiology.[7] The virtue of this approach lies in the way that it forces investigators to explicitly consider covariate imbalance across exposure groups and helps analysts to avoid "off-support inferences"—that is, attempts to estimate a causal effect based on non-existent data, for example, comparing outcomes between affluent people who live in poor neighbourhoods (no data) versus poor people who live in the same areas.[7] Put simply, matching strategies are a useful first step to even determine whether there exists power to answer the research question of interest. A limitation of the propensity score approach and other matching strategies is that they are entirely based on observed covariates. From a multi-level perspective, there also remains the issue of matching neighbourhood exposures on both individual and neighbourhood covariates. Researchers may not have an exhaustive list of covariates at individual and neighbourhood levels required to minimise confounding.

Finally, studies are needed in which specific aspects of neighbourhood contexts (eg, presence of sidewalks to promote physical activity) are directly manipulated through cluster community trials.[1] The Moving to Opportunity program in the US is an example of a large-scale randomised trial of neighbourhood contexts on health.[8] The trial, which randomised families moving from high-poverty to low-poverty neighbourhoods in five US cities, found some evidence for improved mental health and obesity prevalence among those who moved.

MEASUREMENT

A second major area of work that needs to be carried out in neighbourhood research is improving the conceptualisation and measurement of neighbourhood contexts. To date, most studies have relied on administrative data (such as the census) to define neighbourhood characteristics. Although multilevel studies suggest that living in a deprived neighbourhood increases the risks of adverse health outcomes, the mechanisms underlying this association remain largely unexplored. In other words, there is a need to "unpack" the specific exposures and pathways through which neighbourhood disadvantage leads to poor health outcomes.[9] Neighbourhood researchers must develop theoretically grounded approaches to measuring aspects of local physical, service and social environments that lie along the "chain of causation" from poverty to health outcomes.[10] Novel applications of Geographic Information System techniques as well as statistical methods to establish the validity and reliability of ecological exposures (ecometrics)[11] will undoubtedly improve the science. Indeed, identifying "true" neighbourhood differences also requires identifying true neighbourhoods, an aspect on which much of the applied work remains largely silent.[12]

MULTILEVEL ANALYSIS

Independent of issues related to design and measurement, the third topic that we highlight is exploiting the full capabilities of the multilevel analytical framework. Multilevel analysis is by now a well established method in social epidemiology.[13] However, existing studies have yet to fully exploit the capabilities of the approach. Much, if not all, of the current research linking neighbourhoods and health is cross-sectional, and assumes a hierarchical structure of individuals nested in neighbourhoods. This simplistic scenario ignores, for instance, the possibility that an individual might move

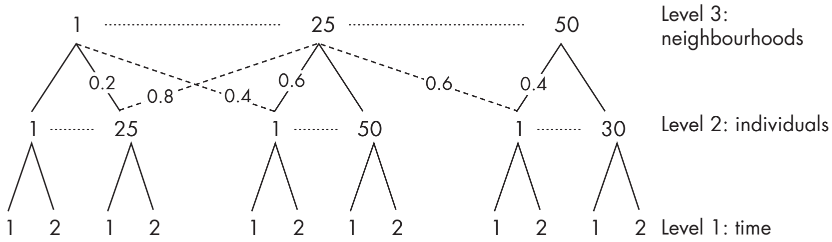

Figure 1 Multi-level structure of repeated measurements of individuals over time across neighbourhoods, with individuals having multiple membership randomised to different neighbourhoods across the time span.

several times, and as such reflects neighbourhood effects drawn from several contexts, or that other competing contexts (eg, schools, workplaces, hospital settings) may simultaneously contribute to contextual effects. Figure 1 provides a visual illustration of one complex, but realistic multi-level structure for neighbourhoods and health research, where time measurements (level 1) are nested within individuals (level 2) who are in turn nested within neighbourhoods (level 3). Importantly, individuals are assigned different weights for the time spent in each neighbourhood—for example, individual 25 was moved from neighbourhood 1 to neighbourhood 25 during the time period t1–t2, spending 20% of her time in neighbourhood 1 and 80% in her new neighbourhood. This multiple membership design would allow control of changing context as well as changing composition. Such designs could be extended to incorporate memberships to additional contexts, such as workplaces or schools. Although such analyses require high-quality longitudinal and context-referenced data, models that incorporate such realistic complexity[14] are likely to improve our understanding of true neighbourhood effects.

In summary, we have highlighted some outstanding issues in the neighbourhood research agenda. By incorporating the methods and approaches from economics, geography, sociology (among other disciplines) into public health, research on neighbourhood effects is poised to make a quantum leap in causal inference as well as usefulness for policy.

J Epidemiol Community Health 2007;**61**:3–4.
doi: 10.1136/jech.2005.045203

Authors' affiliations

I Kawachi, S V Subramanian, Department of Society, Human Development and Health, Harvard School of Public Health, Boston, Massachusetts, USA

Correspondence to: Professor I Kawachi, Harvard School of Public Health, 677 Huntington Avenue, Boston, MA 02115, USA

Funding: IK is supported in part by the MacArthur Network on SES and Health. SVS is supported by the National Institutes of Health Career Development Award (NHLBI 1 K25 HL081275).

Competing interests: None declared.

REFERENCES

1 **Oakes M**. The (mis)estimation of neighborhood effects: causal inference for a practicable social epidemiology. *Soc Sci Med* 2004;**58**:1929–52.
2 **Glymour M**. Natural experiments and instrumental variables analyses in social epidemiology. In: Oakes JM, Kaufman JS, eds. *Methods in social epidemiology*. New York: Jossey-Bass/Wiley, 2006.
3 **Cutler DM**, Glaeser EL. Are ghettos good or bad? *Q J Econ* 1997;**112**:827–72.
4 **Angrist JD**, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *J Am Stat Assoc* 1996;**91**:444–55.
5 **Greenland S**. An introduction to instrumental variables for epidemiologists. *Int J Epidemiol* 2000;**29**:722–9.
6 **Rosenbaum PR**, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;**70**:41–55.
7 **Oakes JM**, Johnson PJ. *Propensity score matching for social epidemiology*, In: Oakes JM, Kaufman JS, eds. *Methods in social epidemiology*. New York: Jossey-Bass/Wiley, 2006.
8 **Kling JR**, Liebman JB, Katz LF. Experimental analysis of neighborhood effects. *Econometrica* 2006 (in press).
9 **Kawachi I**, Berkman LF, eds. *Neighborhoods and health*. New York: Oxford University Press, 2003.
10 **Cummins S**, Macintyre S, Davidson S, *et al*. Measuring neighbourhood social and material context: generation and interpretation of ecological data from routine and non-routine sources. *Health Place* 2005;**11**:249–60.
11 **Raudenbush SW**. *The quantitative assessment of neighborhood social environment*, In: Kawachi I, Berkman LF, eds. *Neighborhoods and health*. New York: Oxfford University Press, 2003.
12 **Sampson RJ**, Morenoff JD. Assessing neighborhood effects: social processes and new directions in research. *Annu Rev Sociol* 2002;**28**:443–78.
13 **Subramanian SV**. The relevance of multilevel statistical models for identifying causal neighborhood effects. *Soc Sci Med* 2004;**58**:1961–7.
14 **Best NG**, Spiegelhalter DJ, Thomas A, *et al*. Bayesian analysis of realistically complex models. *J R Statist Soc A* 1996;**159**:232–342.
15 **O'Campo P**. Invited commentary: advancing theory and methods for multilevel models of residential neighborhoods and health. *Am J Epidemiol* 2003;**157**:9–13.
16 **Pickett KE**, Pearl M. Multilevel analyses of neighborhood socioeconomic context and health outcomes: a critical review. *J Epidemiol Community Health* 2001;**55**:111–22.