Multilevel Modeling

1. Learning Objectives

After reviewing this chapter, readers should better be able to:

- Recognize a research problem requiring a multilevel modeling approach;
- Describe the technical and substantive advantages of multilevel models;
- Explain the basic principles of multilevel modeling using graphical, verbal, and statistical language for a range of multilevel models;
- Develop a variety of models that enable quantitative assessment of contextual effects; and
- Apply multilevel models to a research problem according to a well-articulated research strategy.

2. Introduction

Describing area-based differences in health outcomes has a long history (Macintyre and Ellaway, 2003). We know that places differ in terms of their average health achievements, but do places make a difference?

This question has received systematic attention in the last decade or so (Duncan, Jones et al., 1993; Macintyre, Maciver et al., 1993; Diez Roux, 2001; Pickett and Pearl, 2001; Kawachi and Berkman, 2003; Lynch, Smith et al., 2004; Subramanian and Kawachi, 2004). Besides its resonance with the move to look at "upstream" determinants of health and recognizing that health behaviors and outcomes need to be understood within their context (Susser and Susser, 1996a; Susser and Susser, 1996b; Berkman and Kawachi, 2000; Kawachi and Berkman, 2003; Wilkinson and Marmot (Eds), 2003; Link and Phelan, 1995; Jones and Moon, 1987; Moon, Subramanian et al., 2005), a major impetus for examining the role of contexts in explaining health variations comes from the advances in quantitative methods, in particular those related to multilevel statistical methods (Bryk and Radenbush, 1992; Goldstein, 1995). In this chapter, we review and provide an overview of the basic principles of multilevel modeling as applied to public health research (Subramanian, Jones et al., 2003; Moon, Subramanian et al., 2005; Blakely and Subramanian, 2006).

3. Multilevel Framework

A contextual perspective to raising and answering research questions is intrinsically *multilevel*, i.e., factors that affect health are viewed as simultaneously operating at the *level of individuals* and at the *level of contexts*.

> **The term multilevel relates to the levels of analysis in public health research, which usually, but not always, consists of individuals (at lower level) who are nested within spatial units (at higher levels).**

The term 'multilevel' has also been used to advocate a multidisciplinary perspective of public health (Anderson, 1999). In this chapter, however, 'multilevel' refers to an analytical perspective, that is, in relation to the levels of analysis in research, which involves taking a multidisciplinary perspective on the questions of epidemiologic interest.

Multilevel methods, meanwhile, consist of statistical procedures that are pertinent when:

1. The observations that are being analyzed are correlated or clustered along spatial, non-spatial, or/and temporal dimensions; or
2. The causal processes are thought to operate simultaneously at more than one level; and/or
3. There is an intrinsic interest in describing the variability and heterogeneity in the population, over and above the focus on average relationships (Diez Roux, 2002; Subramanian, Jones et al., 2003; Subramanian, 2004; Subramanian, 2004).

It is clear that individuals are organized within a nearly infinite number of levels of organization, from the individual up (e.g., families, neighborhoods, counties, states, regions), from the individual down (e.g., body organs, cellular matrices, DNA), and for overlapping units (e.g., area of residence and work environment).
Therefore it is necessary that links should be made between these possible levels of analysis (Susser, 1998; McKinlay and Marceau, 2000).

3. Multilevel Framework

Figure 1 identifies a typology of designs for data collection and analyses (Blakely and Woodward, 2000; Kawachi and Subramanian, 2006) where the rows indicate the level or unit at which the outcome variable is being measured, i.e., at the individual level (*y*) or the aggregate, or ecological, level (*Y*), and the columns indicate whether the exposure is being measured at the individual level (*t*) or the ecological level (*T*).

Figure 1: Typology of Studies

		Exposure	
		t	T
Outcome	y	$\{y, t\}$ Traditional risk factor study	$\{y, T\}$ Contextual study
	Y	$\{Y, t\}^{(A)}$	$\{Y, T\}$ Ecological study

Note: $^{(A)}$ This type of study is impossible to specify as it stands. Practically speaking, it will either take the form of $\{Y, T\}$, i.e., ecological study, where T will now simply be central tendency of t. Or, if dis-aggregation of Y is possible, so that we can observe y, then it will be equivalent to $\{y, t\}$.

- **Study-type** $\{y, t\}$ is most commonly encountered when the researcher aims to link exposure measured at the individual level (e.g., diet) to individual health outcomes (e.g., BMI). Study-type $\{y, t\}$ not only ignores ecological effects (either implicitly or explicitly), but with its individualistic focus resonates with the notion of health as solely a matter of individual responsibility (Moon, Subramanian et al., 2005).

- Conversely, **study-type** $\{Y, T\}$ - referred to as an "ecological study" – may seem intuitively appropriate for research on population health and ecological exposures.
- However, **study-type** $\{Y, T\}$ conflates the genuinely ecological and the 'aggregate' or compositional (Moon, Subramanian et al., 2005), and precludes the possibility of testing heterogeneous contextual effects on different types of individuals.

Ecological effects reflect predictors and associated mechanisms operating solely at the contextual level. The search for such measures and their scientific validation and assessment is an area of active research (Raudenbush, 2003).

Aggregate effects, in contrast, equate the effect of a neighborhood with the sum of the individual effects associated with the people living within the neighborhood. In this situation the interpretative question becomes particularly relevant. If common membership in a neighborhood by a set of individuals brings about an effect that is over and above those resulting from individual characteristics, then there may indeed be an ecological effect (i.e., the whole may be more than the sum of its parts). If this is not the case, then it is individual factors that matter, not ecological effects.

3. Multilevel Framework

Exercise 1: BMI Neighborhood Differences

We wish to know whether the neighborhood level differences in body mass index (BMI) is due to clustering of individual characteristics or is due to characteristics associated with neighborhoods.

Click on one of the four table cells that constitute the appropriate typology for this study

		Exposure	
		t	T
Outcome	y	$\{y, t\}$ Traditional risk factor study	$\{y, T\}$ Contextual study
	Y	$\{Y, t\}^{(A)}$	$\{Y, T\}$ Ecological study

Note: $^{(A)}$ This type of study is impossible to specify as it stands. Practically speaking, it will either take the form of $\{Y, T\}$, i.e., ecological study, where T will now simply be central tendency of t. Or, if dis-aggregation of Y is possible, so that we can observe y, then it will be equivalent to $\{y, t\}$.

Answering the question in Exercise 1 requires a study-type $\{y, T\}$, i.e., in which an ecological exposure (e.g., proportion in poverty) is linked to an individual outcome (BMI). A more complete representation would be type $\{y, x, T\}$, whereby we have an individual outcome, individual confounders (x), and ecologic exposure reflecting a multilevel structure of individuals nested within ecologies. When the ecological exposure is an aggregate measure of individual characteristics, such as percent poverty, it is obvious that information on both individual poverty and neighborhood percent poverty is required to test for an ecological effect. However, multilevel data are essential even if the ecological variable is a structural feature such as neighborhood presence of fast food outlets, because people with individual level disadvantage are likely to be overrepresented in places with structural risk factors.

3. Multilevel Framework

A fundamental motivation for study- type $\{y, x, T\}$, is to distinguish "neighborhood differences in health" from "the difference a neighborhood makes to individual health outcomes" (Moon, Subramanian et al., 2005). Stated differently, ecological effects on the individual outcome can only be ascertained after individual factors that reflect the composition of the neighborhood have been controlled. Indeed, compositional explanations for ecological variations in health are common, to paraphrase the methodologist Gary King,

> ***...if we really understood [health variations], we would not need to know much of contextual effects*** **(King, 1997).**

This is an important challenge for researchers interested in understanding ecologic effects. It nonetheless makes intuitive sense to test for the possibility of ecological effects, besides anticipating that the impact of individual level, compositional factors may vary by context. Thus, unless contextual variables are considered, their direct effects and any indirect mediation through compositional variables remain unidentified. Moreover, composition itself has an intrinsic ecologic dimension; the very fact that individual (compositional) factors may 'explain' ecologic variations serves as a reminder that the real understanding of ecologic effects is complex. The multilevel framework with its simultaneous examination of the characteristics of the individuals at one level and the context or ecologies in which they are located at another level offers a comprehensive framework for understanding the ways in which places can affect people (contextual) or, alternatively, people can affect places (composition).

4. Multilevel Methods and Analyses

The existence of multilevel data structures is neither random nor ignorable; for instance, individuals differ but so do the neighborhoods.

Differences among neighborhoods could either be directly due to the differences among individuals who live in them; or groupings based on neighborhoods may arise for reasons less strongly associated with the characteristics of the individuals who live in them.

Importantly, once such groupings are established, even if their establishment is random, they will tend to become differentiated. This would imply that the group (e.g., neighborhoods) and its members (e.g., individual residents) can exert influence on each other, suggesting different sources of variation (e.g., individual-induced and neighborhood-induced) in the outcome of interest and thus compelling analysts to consider independent variables at the individual and at the neighborhood level.

Ignoring this multilevel structure of variations does not simply risk overlooking the importance of neighborhood effects; it has implications for statistical validity.

Example 1: Importance of Multilevel Analysis

In an influential study of progress among primary school children, Bennett (1976), using single-level multiple regression analysis, claimed that children exposed to a 'formal' style of teaching exhibited more progress than those who were not. The analysis while recognizing individual children as units of analysis ignored their grouping into teachers/classes. In what was the first important example of multilevel analysis using social science data, Aitkin, Anderson et al., (1981) reanalyzed the data and demonstrated that when the analysis accounted properly for the grouping of children (at lower level) into teachers/classes (at higher levels), the progress of formally taught children could not be shown to significantly differ from the others.

What was occurring in this example was that children within any one class/teacher, because they were taught together, tended to be similar in their performance thereby providing much less information than would have been the case if the same number of children had been taught separately. More formally, the individual samples (e.g., children) were correlated or clustered. Such clustered samples do not contain as much information as simple random samples of similar size. As was shown by Aitkin (Aitkin, Anderson et al., 1981), ignoring this autocorrelation and clustering results in increased risk of finding differences and relationships where none exist.

4. Multilevel Methods and Analyses

Clustered data also arise as a result of sampling strategies. For instance, while planning large-scale survey data collection, for reasons of cost and efficiency, it is usual to adopt a multistage sampling design. A national population survey, for example, might involve a three-stage design, with regions sampled first, then neighborhoods, and then individuals. A design of this kind generates a three-level hierarchically clustered structure of individuals at level-1, which are nested within neighborhoods at level-2, which in turn are nested in regions at level-3. Individuals living in the same neighborhood can be expected to be more alike than they would be if the sample were truly random. Similar correlation can be expected for neighborhoods within a region.

Much documentation exists on measuring this "design effect" and correcting for it. Indeed, clustered designs (e.g., individuals at level-1, nested in neighborhoods at level-2, nested in regions at level-3) are often a nuisance in traditional analysis. However, individuals, neighborhoods, and regions can be seen as distinct structures that exist in the population that should be measured and modeled.

While the conventional approach to such correlated data structures is to treat the clustering as a nuisance, multilevel models view such hierarchical structures as a feature of the population and one that is of substantive interest. Indeed, "once you know that hierarchies exist, you see them everywhere" (Kreft and de Leeuw, 1998).

Exercise 2: Data Structure

Below are a series of statements on the possible data structures, some are hierarchical and some are not – match the most appropriate category to each statement.

- Single Level Data Structure
- Two Level Data Structure
- Three Level Data Structure
- Non Hierarchical, multiple membership
- Cross Classified Model

Statement	Category
Cross-sectional data on a random sample of individuals.	
Cross-sectional data on a random sample of individuals from a random sample of households from a random sample of neighborhoods.	
Cross-sectional data on a random sample of individuals from a random sample of neighborhoods.	
Cross-sectional data on a multiple responses for a random sample of children in a single classroom.	
Cross-sectional data on a random sample of children belonging to a random sample of teachers, with the possibility of children being taught by multiple teachers.	
Cross-sectional data a random sample of individuals drawn from a random sample of neighborhoods and worksites.	
Data on repeated measures for an individual.	

5. Desiderata for Multilevel Research

Some core concepts are intrinsic to adopting a multilevel perspective and are discussed in this section, including:

- Contextual and Compositional Sources of Variation;
- Contextual Heterogeneity;
- Individual Heterogeneity;
- Individual/Contextual Interaction;
- Multiple Hierarchical Contexts;
- Changing People/ Changing Places;
- Interrelated Outcomes; and
- Overlapping Contexts.

Contextual and compositional sources of variation Evidence for variations in poor health between different settings or contexts can be due to factors that are intrinsic to, and are measured at, the contextual level. In other words, the variation can be due to what can be described as contextual, area, or ecological effects. Alternatively, variations between places may be compositional, i.e., certain types of people who are more likely to be in poor health due to their individual characteristics happen to live in the same places.

The Research Question here is not whether variations between different settings exist (they usually do), but what is their source, i.e., are the variations across places compositional or contextual?

The notions of contextual and compositional sources of variation have general relevance and they are applicable whether the context is administrative (e.g., political boundaries), temporal (e.g., different time periods), or institutional (e.g., schools or hospitals). The research question focused on this core concept would be: are there significant contextual differences in health between settings (such as neighborhoods), after taking into account the individual compositional characteristic of the neighborhood?

Contextual heterogeneity

Beyond disentangling the contextual and compositional sources of variation, contextual differences may be complex such that it may not be the same for all types of people. For example, while neighborhood contexts may matter for the health outcomes of one population group (e.g., low social class), it may not have any influence upon the health status of other groups (e.g., high social class).

The Research Question in this case is: are the contextual neighborhood differences in poor health different for different types of population groups?

Individual heterogeneity

Within particular contexts, one group's health experience may be more or less variable than the other, over and above the average differences. For example, people of low social class, in addition to being contextually heterogeneous, may experience more variability compared to other groups.

The Research Question is: are individual differences in poor health different for different types of population groups?

Individual/contextual interaction

Contextual differences, in addition to people's characteristics, may also be influenced by the different characteristics of neighborhoods. Stated differently, individual differences may interact with context. For example, poor people (individual characteristic) may experience different levels of health depending upon the poverty levels (place characteristic) of the area in which they live.

The Research Question of interest is: what is the average relationship between individual poor health and neighborhood-level socioeconomic characteristics, and does the effect of neighborhood-level socioeconomic characteristics on individual health differ for different types of individuals based on their demographic and socioeconomic characteristics?

5. Desiderata for Multilevel Research

Multiple hierarchical contexts

Contextual settings themselves can be conceptualized and measured at multiple levels such that individual health experiences are not simply influenced by people's proximate environment (e.g., neighborhoods) but also their macro ecologic settings (e.g., states). Moreover, neighborhoods rarely exist in a vacuum, and considering their broader contextual settings can be vital given the functional interconnectedness between geographic levels. An analysis of health should consider both the immediate contextual setting of people (e.g., neighborhoods) but also the macro contextual settings to which both people and neighborhoods belong (e.g., states).

The Research Question is: what additional contextual levels are relevant for the health outcome in question?

Interrelated outcomes

Health outcomes themselves are often interrelated. For instance, people often engage simultaneously in high-risk behaviors, such as smoking and excess drinking. In addition, each of these behaviors may have a qualitative (yes or no) and a quantitative (how much) aspect. For instance, whether a person smokes or not may reveal nothing about the number of cigarettes smoked. There may be neighborhoods where few people smoke, but those who do smoke heavily; an average figure would be very misleading.

An ideal modeling approach should allow consideration of multiple responses, and allow us to ask: are neighborhoods with high proportion of smokers also high in the intensity of cigarettes smoked

Overlapping contexts

Not only are contexts multiple, they may also overlap. For instance, health behaviors, such as smoking, may be influenced not only by the neighborhoods in which people live, but also by their work environment. Clearly, workplaces and residential neighborhoods need not be nested neatly within one another.

Here, the relevant question is: what is the relative contribution to health outcomes of different contextual settings that may not be nested within each other, but overlap (e.g., neighborhoods and workplaces)? A related situation is where individual health behaviors are influenced not only by the characteristics of the neighborhood in which they live but also by characteristics of adjoining areas.

5. Desiderata for Multilevel Research

Exercise 3: Multilevel Perspective

Below are several hypotheses and core concepts intrinsic to adopting a multilevel perspective. Select which core concept addresses each of the hypotheses below.

- Disentangling contextual and compositional sources of variation
- Examining individual/contextual interaction
- Examining contextual heterogeneity
- Examining individual heterogeneity

Concept	Multilevel Hypothesis
	There is variation in health by geographic units, and it is independent of the variation in age, gender, race and SES by geographic units.
	The effect of gender on BMI varies by neighborhood.
	While women have higher BMI than men, on average, they are also more variable.
	The effect of neighborhood safety (a contextual attribute) on BMI varies by gender (an individual attribute) such that it is stronger for women than men.

6. Multilevel Data Structures

Hierarchies

It is well known that once groupings are created (consisting of individuals), even if their origins are essentially 'random,' individuals end up being influenced by their group membership. Such groupings can be spatial (e.g., areas) or non-spatial (e.g., communities). Hierarchies are one way of representing the dependent or correlated nature of the relationship between individuals and their groups. Thus, for instance, we can conceptualize a two-level structure of many level-1 units (e.g., individuals) nested within fewer level-2 groups (e.g., neighborhoods/places) as illustrated in Figure 2. Since individual outcomes are anticipated as being dependent upon the neighborhoods in which they live, responses within a neighborhood are more alike than different. When dependency is anticipated in the 'population' or 'universe,' they represent population-based or naturally occurring hierarchies.

Figure 2: Two-level Structure

The importance of identifying and specifying the 'higher' levels is critical for multilevel research. Researchers must a priori specify why they think that there will be variation in the outcome at these levels over and above variation at the individual-level. Such thinking naturally leads to considerations of which levels to include in the model. For example, do we expect variation at the level of small neighborhoods (e.g., census blocks) or larger neighborhoods (e.g., census tracts)? The most common multilevel model is a two-level hierarchic nested modeling with many level-1 units within a smaller number of level-2 units, as exemplified in Figure 2. A multilevel structure can be cast, with great advantage, to incorporate a range of circumstances where one may anticipate clustering (Subramanian, Jones et al., 2003).

6. Multilevel Data Structures

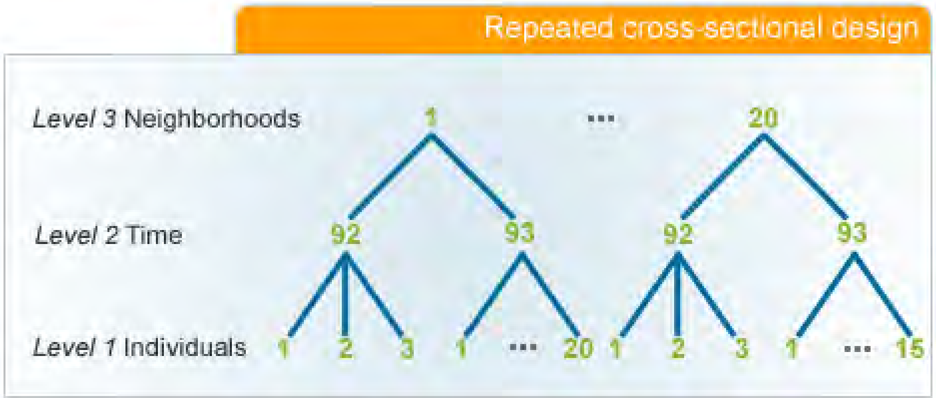
Besides extending the two-level structure to a three-level structure of, for example, individuals (level-1) within neighborhoods (level-2) within counties (level-3), a number of other data structures can be thought to be special cases of multilevel models. For instance, health outcomes and behaviors as well as their causal mechanisms are rarely stable and invariant over time, producing data structures that involve repeated measures.

Two possibilities arise depending on the unit that is repeatedly measured. When individuals are repeatedly measured within a panel design, the outcomes taken at different times form level-1. The same outcomes measured over different times are nested within individuals at level-2, which in turn nest within higher-level units (e.g., neighborhoods). This structure is shown in Figure 3(a) and allows the assessment of individual change within a contextual setting. The other possibility is a repeated cross-sectional survey, where places are monitored at regular time intervals (repeatedly measuring places over time).

Figure 3a: Panel Design

The structure would then be: individuals at level-1, time/years within places at level-2, and places at level-3, as shown in Figure 3(b). Such a structure permits an investigation of trends within geographic settings controlling for their compositional make-up. Multilevel models could be used to explore what sorts of individuals and what sorts of places have changed with respect to health outcomes.

Figure 3b: Cross-Sectional Design



When different responses/outcomes are correlated this lends itself to a multivariate multilevel data structure in which level-1 are sets of response variables measured on individuals at level-2, nested in neighborhoods at level-3. The key feature is that the set of responses (outcomes) is nested within individuals. The response could be a set of outcomes that relate to, for instance, different aspects of health behavior (e.g., smoking and drinking). Crucially, such responses could be a mixture of 'quality' (do you smoke/do you drink) and 'quantity' (how many/how much). A multilevel structure on different aspects of health behavior could include measurements (e.g., smoking and drinking, both at level-1), nested within individuals (at level-2), within neighborhoods (at level-3).

The substantive benefit of this approach is that it is possible to assess whether different types of behavior are related to individual characteristics in the same or different ways. Moreover, the residual co-variances at level-2 and level-3 measure the 'correlation' of behaviors between individuals and between places. Additionally, we can ascertain whether neighborhoods that are high for one behavior are also high for another; and whether neighborhoods with high prevalence of smoking, for instance, are also high in terms of the number of cigarettes smoked. Technical benefits flow in terms of efficiency if the response is correlated and if there are many missing responses, as in matrix sample designs. Figure 3(c) presents a structure where the responses at level-1 capture different aspects of health behaviors and Figure 3(d) portrays the idea of 'mixed' (quality and quantity) responses on a particular aspect of health behavior.

Figure 3c: Multivariate Responses

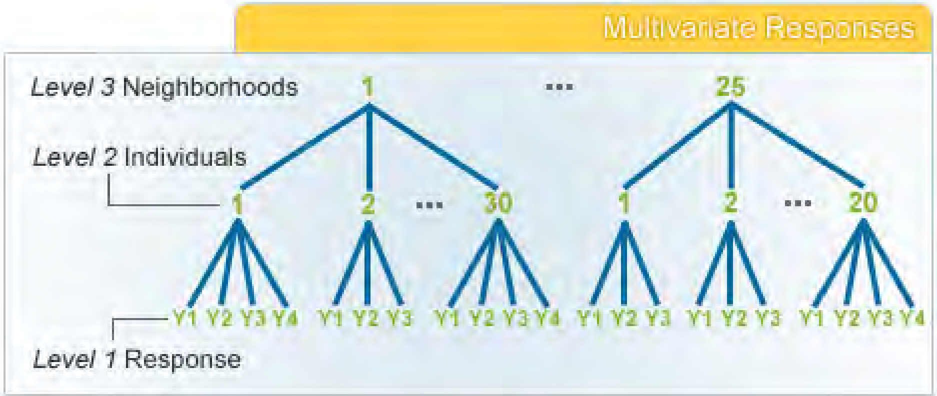
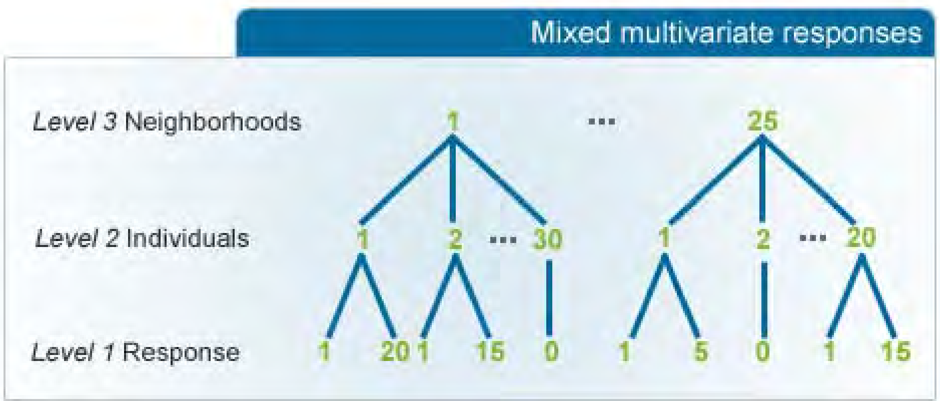


Figure 3d: Mixed Multivariate Responses

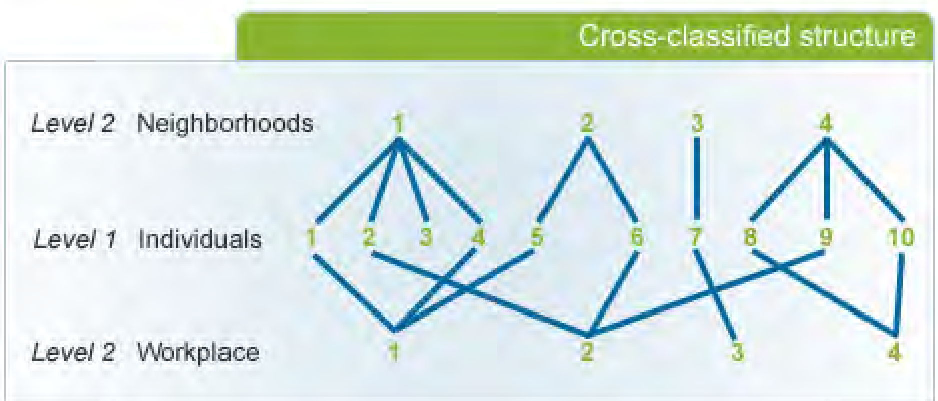


6. Multilevel Data Structures

Crossed Structures

All the previous examples are strictly hierarchical in that all level-1 units can belong to one and only one level-2 unit. However, data structures can be 'non-hierarchical.' Individuals live their lives in a number of overlapping settings such as neighborhoods, work place, home, etc. Such contexts do not always lend themselves to a neat hierarchical structure. Instead, the different settings may overlap at the same level, thus producing a crossed structure. While the importance of such structures has long been recognized, it is only recently that it has become technically and computationally tractable (Goldstein, 1994; Jones, Gould et al., 1998). The 'quasi-hierarchical' format employed within cross-classified multilevel models enables an assessment of the relative importance of a number of different, overlapping contexts after allowing for the differential composition of each. As such models identify contexts that have a confounding influence, they also ascertain which contexts have the greatest significance.

Figure 3e: Cross-Classified Structure



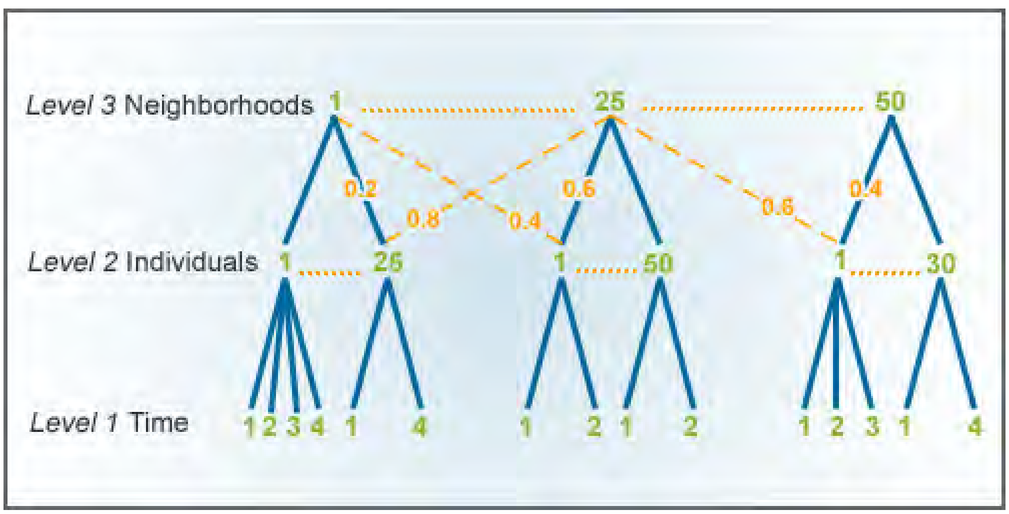
For example, a cross-classified model of health behavior (e.g., smoking) could be formulated with individuals at level-1 and both residential neighborhoods and workplaces at level-2, as shown in Figure 3(e). If account is not taken of this cross-classified structure, what may appear to be between-work place variation could actually be between-neighborhood variation and vice versa.

A related structure occurs where for a single level-2 classification (e.g., neighborhoods), level-1 units (e.g., individuals) may belong to more than one level-2 unit and these are also referred to as multiple membership designs. The individual can be considered to belong simultaneously to several neighborhoods with the contributions of each neighborhood being weighted in relation to its distance (if the interest is spatial) from the individual.

While for the purpose of clarity and ease of understanding we have discussed each of the multilevel structures separately, readers are urged to think about these structures in an integrated manner. For instance, a structure can be a combination of more than one of the designs discussed above, as shown in Figure 4.

Figure 4: Repeated Measurements

Multilevel Structure of Repeated Measurements of Individuals over Time Across Neighborhoods with Individuals Having Multiple Memberships to Different Neighborhoods Across the Time Span



Source: Subramanian 2004 Subramanian SV. The relevance of multilevel statistical models for identifying causal neighborhood effects. Social Science and Medicine. 2004;58:1961-1967.

Time measurements (level-1) are nested within individuals (level-2) who are in turn nested in neighborhoods (level-3). Importantly, individuals are assigned different weights for the time spent in each neighborhood.

Example 2: Multiple Membership

Individual 25 moved from neighborhood 1 to neighborhood 25 during the study time-period t1-t2, spending 20% of her time in neighborhood 1 and 80% in her new neighborhood. This multiple-membership panel design could allow control of changing context as well as changing composition, besides enabling a consideration of weighted effects of proximate contexts (Langford, Bentham et al., 1998). So, for example, the geographical distribution of disease can be seen not only as a matter of composition and the immediate context in which an outcome occurs, but also as a consequence of the impact of nearby contexts, with nearer areas being more influential than more distant ones. Goldstein, (2003) provides an elegant and comprehensive classification schema.

7. Levels and Variables

The higher levels that were discussed in the previous section (e.g., neighborhoods) can be considered as variables in a regression equation with an indicator variable specified for each neighborhood.

> **Put differently, why are variables such as gender, ethnicity/race, or social class not a level?**

Neighborhoods are treated as a level because they are a population of units from which we have observed a random sample. This enables us to draw generalizations for a particular level (e.g., neighborhoods) based on an observed sample of neighborhoods. On the other hand, gender, for instance, is not a level because it is not a sample out of all possible gender categories. Rather, it is an attribute of individuals. Thus, male or female in our gender example are 'fixed' discrete categories of a variable with the specific categories only contributing to their respective means. They are not a random sample of gender categories from a population of possible gender groupings.

The situation becomes less clear when the study includes all individuals in the population, and hence also includes all neighborhoods, ethnic/race, gender, and social class groups. Such a study design arises when census data is linked to mortality data, e.g. (Blakely, Salmond et al., 2000). Why might we still consider neighborhoods here as levels, but not ethnicity/race? First, it is more efficient to model neighborhoods as a random variable given the (likely) large number of neighborhoods. Second, we would usually wish to ascribe a fixed effect to each ethnic group, but not each neighborhood. Rather, we wish to model an ecologic attribute such as social capital at the neighborhood-level.

It is possible to consider 'levels' as 'variables.' Thus, when neighborhoods are considered as a variable, they are typically reflective of a fixed classification. While this may be useful in certain circumstances, doing so robs the researcher of the ability to generalize to all neighborhoods (or 'population' of schools) and inferences are only possible for the specific neighborhoods observed in the sample.

8. A Graphical Introduction

One of the main attractions of multilevel models for public health research is their ability to allow relationships to vary across different contextual settings.

We illustrate this with a two-level structure consisting of individuals at level-1 nested within neighborhoods at level-2 with a single continuous outcome (e.g., poor health score) and a single continuous individual (compositional) predictor (e.g., age) centered about its mean.
Figure 5 illustrates a range of hypothetical graphical models for representing this data structure. In Figure 5(a), the poor health/age relationship is shown as a straight line with a positive slope: older people generally have poorer health. This model conceptualizes health status only in terms of an individual's age, and the neighborhood context is ignored.

Figure 5a: Fixed Intercept, Fixed Slope

Fixed intercept, Fixed slope

This is remedied in Figure 5(b), in which the relationships in each of the neighborhoods (six here, but typically more) is represented by a separate line at a varying distance from the general underlying relationship, as shown by the thicker line. The parallel lines imply that while the poor health/age relationship in each neighborhood is the same, some neighborhoods have uniformly higher levels of poor health than others.

Figure 5b: Fixed Intercept, Fixed Slope

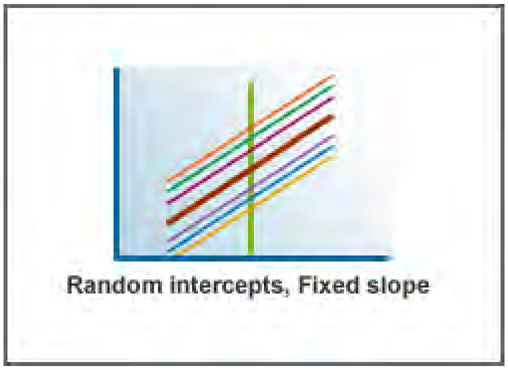
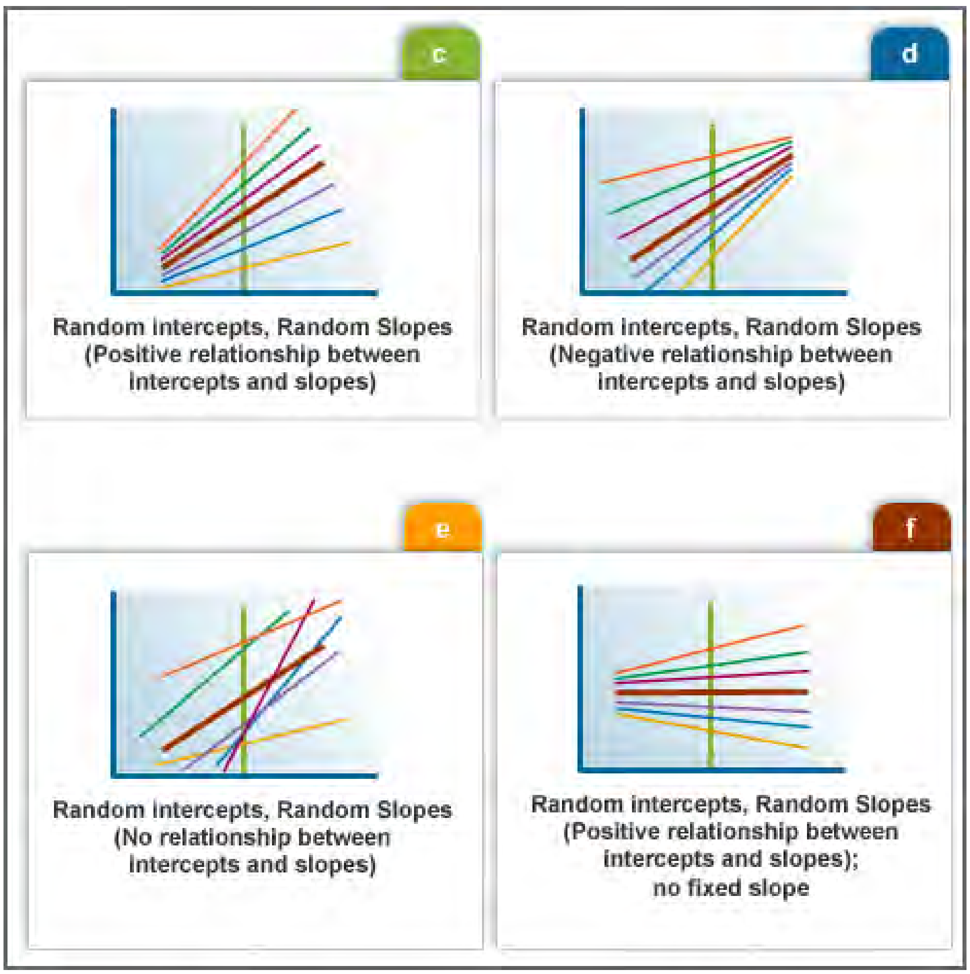
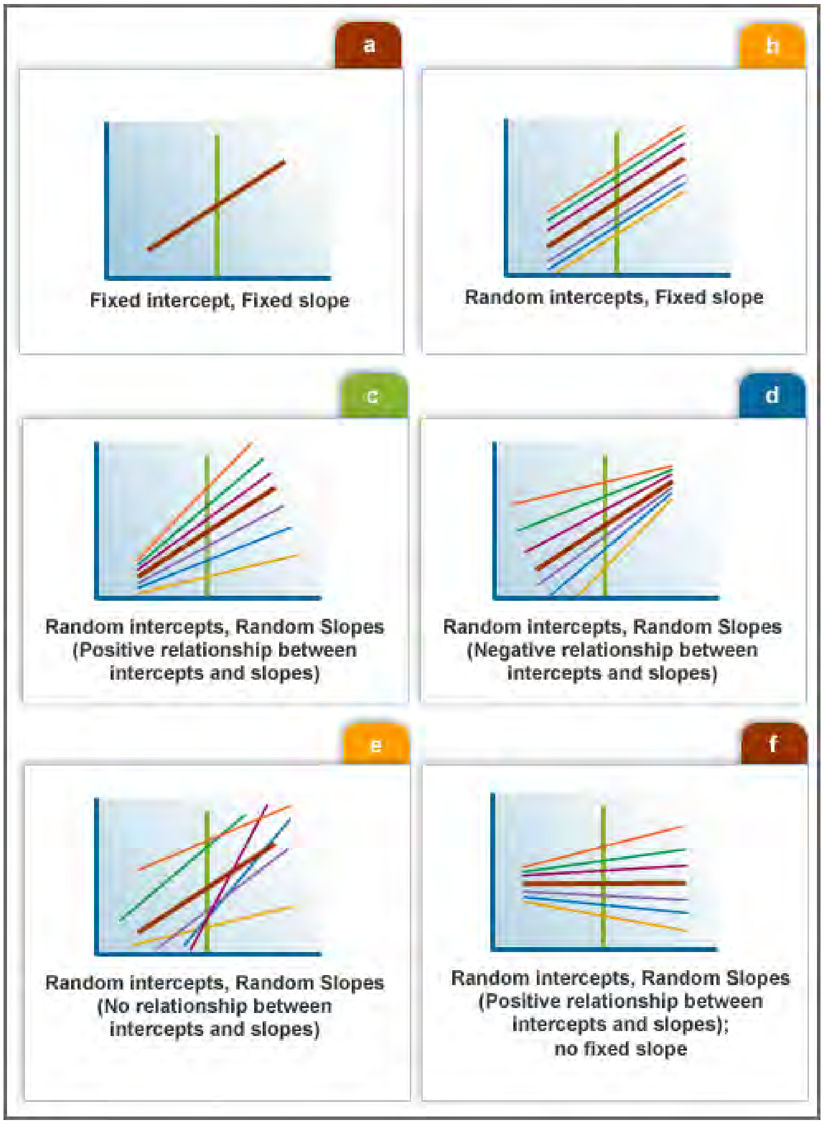


Figure 5c-f: Intercepts, Slopes and Relationships



8. A Graphical Introduction

Figure 5: Different Patterns



The different patterns in Figure 5 are achieved by allowing the average (fixed) 'intercept' and the average (fixed) 'slope' to vary (be random) across neighborhoods. Multilevel models specify the different intercepts and slopes for each context as coming from a distribution at a higher level.

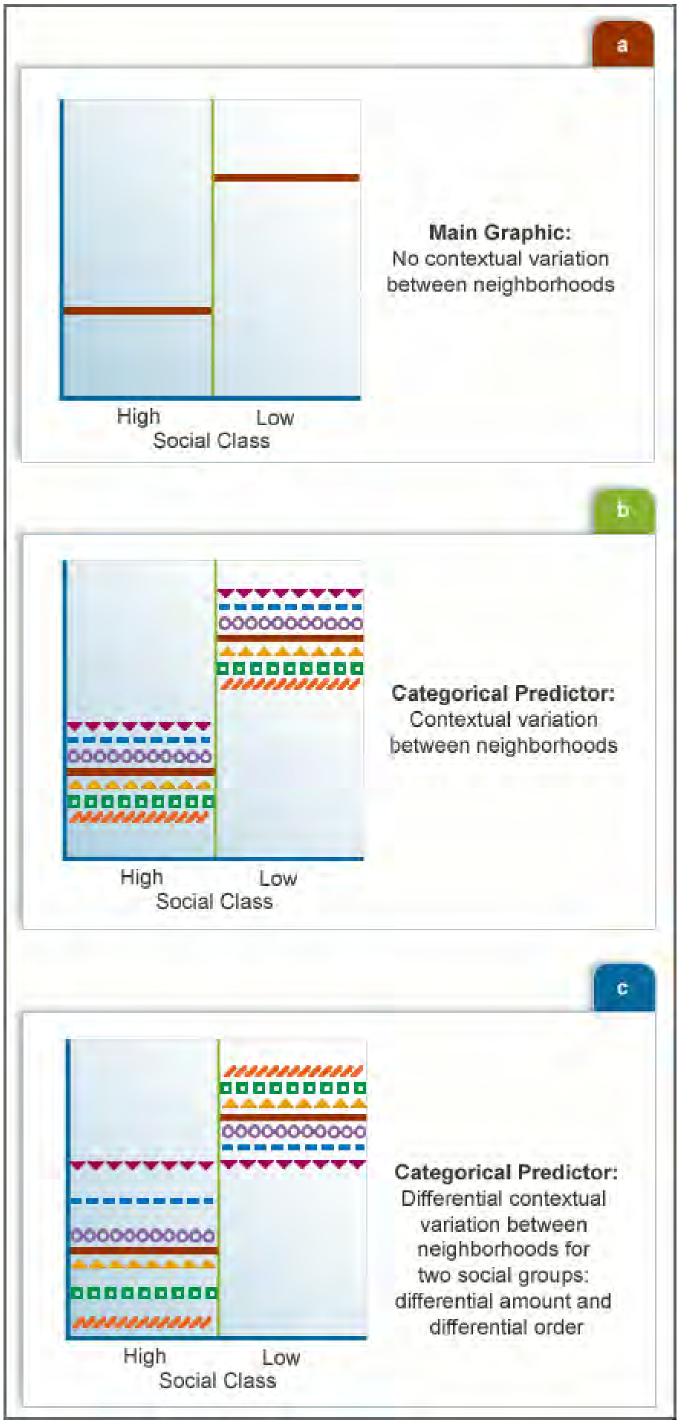
The different forms of relationships represented in Figures 5(c)-(f) are a result of how the intercepts and slopes are associated. Graphical models represented in Figures 5(c)-(f) are also called random-slopes models, since the patterns are achieved by allowing the fixed slope to vary across neighborhoods. Figure 5(b), meanwhile, is the simplest form of multilevel modeling and is referred to as a random-intercepts model, as only intercepts are allowed to vary across neighborhoods.

For instance in Figure 5(c), the relationship between poor health and age is strongest in neighborhoods (a steeper slope) where poor health rates are quite high for average age groups (a high intercept). Stated differently, there is a positive association between the intercepts and the slopes. In Figure 5(d) high intercepts are shown to be associated with shallower slopes, that is, a negative association between the slopes and the intercepts. The complex criss-crossing in Figure 5(e) results from a lack of pattern between the intercepts and the slopes such that the health achievement rates of a neighborhood at average age tells us nothing about the direction and magnitude of the poor health/age relationship. The distinctive feature of Figure 5(f) results from the slopes varying around zero. In other words, while typically there is no poor health/age relationship, in some neighborhoods the slope is positive; in others it is negative. In this case, a single-level model would reveal no relationship whatsoever between poor health and age, and as such the ‘average’ relationship would not occur anywhere.

8. A Graphical Introduction

While the illustrative example in Figure 5 was based on a continuous predictor (e.g., age), one can conceptualize the notion of varying relationships in the presence of categorical predictors as well. Figure 6 illustrates the interpretation of neighborhood heterogeneity with categorical predictors. We consider social class as two category individual variables, high social class and low social class, and these are shown on the horizontal axis (x) with the response being a continuous score of poor health (y-axis) in Figure 6.

Figure 6: Neighborhood Heterogeneity



8. A Graphical Introduction

Figure 6(a) presents the simplest outcome: differences between social groups but no variation between neighborhoods. With only one fixed average for each group, it shows an individual-level model in which the same relationship is fitted to all neighborhoods.

Figure 6a: No Contextual Variation

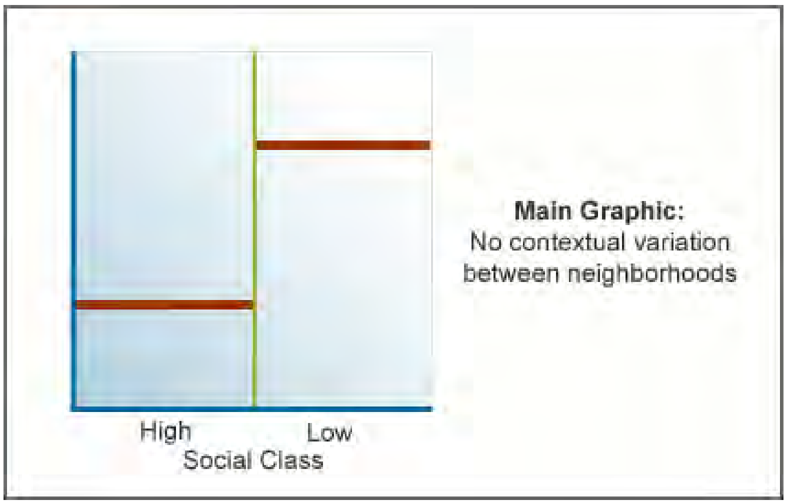
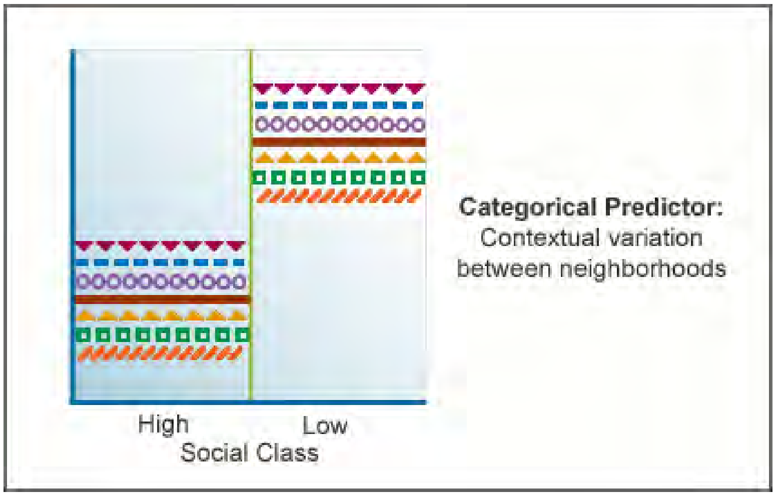


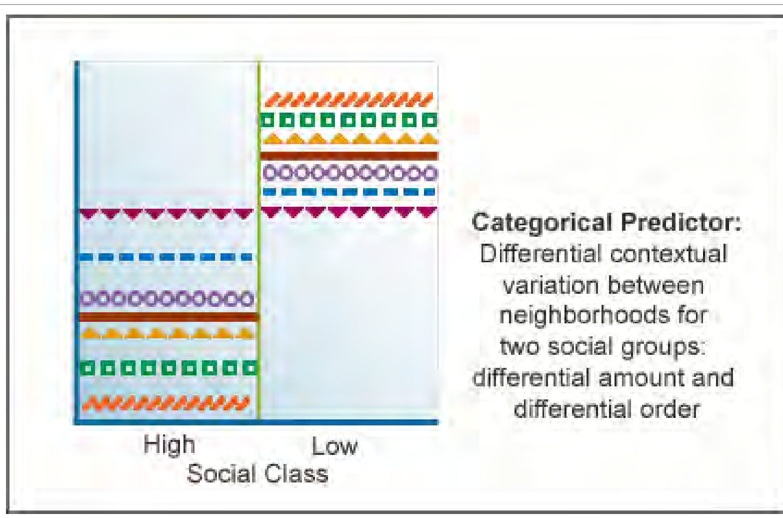
Figure 6(b) represents a two-level model with each of six neighborhoods having its own poor health/social class relationship. The thick solid lines represent the average poor health rates for the two groups, while the symbol-lines (one for each neighborhood) represent the variation between neighborhoods around the average line. Since the individual relationship between social class and poor health is also shown in the model, the graph implies that the variation between neighborhoods is not due solely to the varying social composition of neighborhoods and is, therefore, contextual. The neighborhood differences, however, are assumed to be simple such that neighborhoods that are high for one group are also high for the other and vice versa (similar to the 'random-intercepts' model). Thus, while there is a (contextual) geography of poor health, it can be summarized in one map.

Figure 6b: Contextual Variation



We can, however, anticipate the neighborhood variation to be significantly different for the two groups. This difference consists of two dimensions. First, the amount (range) of neighborhood variation can be different for the two groups. In Figure 6(c), those in the high social class category tend on average to have lower chances of being in poor health, but the neighborhood variation is relatively large as compared to the low social class. For the low social class, it is the reverse: a higher probability of being in poor health, on average, but smaller variation between neighborhoods.

Figure 6c: Contextual Variation



The second aspect of the neighborhood difference relates to the ordering. Thus, neighborhoods that are high for one group may be low for the other and vice versa, as shown in Figure 6(c). An attractive feature of multilevel models – one that is commonly used in health research – is their ability to model contextual differences as a function of characteristics that relate to neighborhoods, in addition to individual characteristics. At the same time, the nature and type of interactions between individual characteristics and neighborhood characteristics can also be assessed.

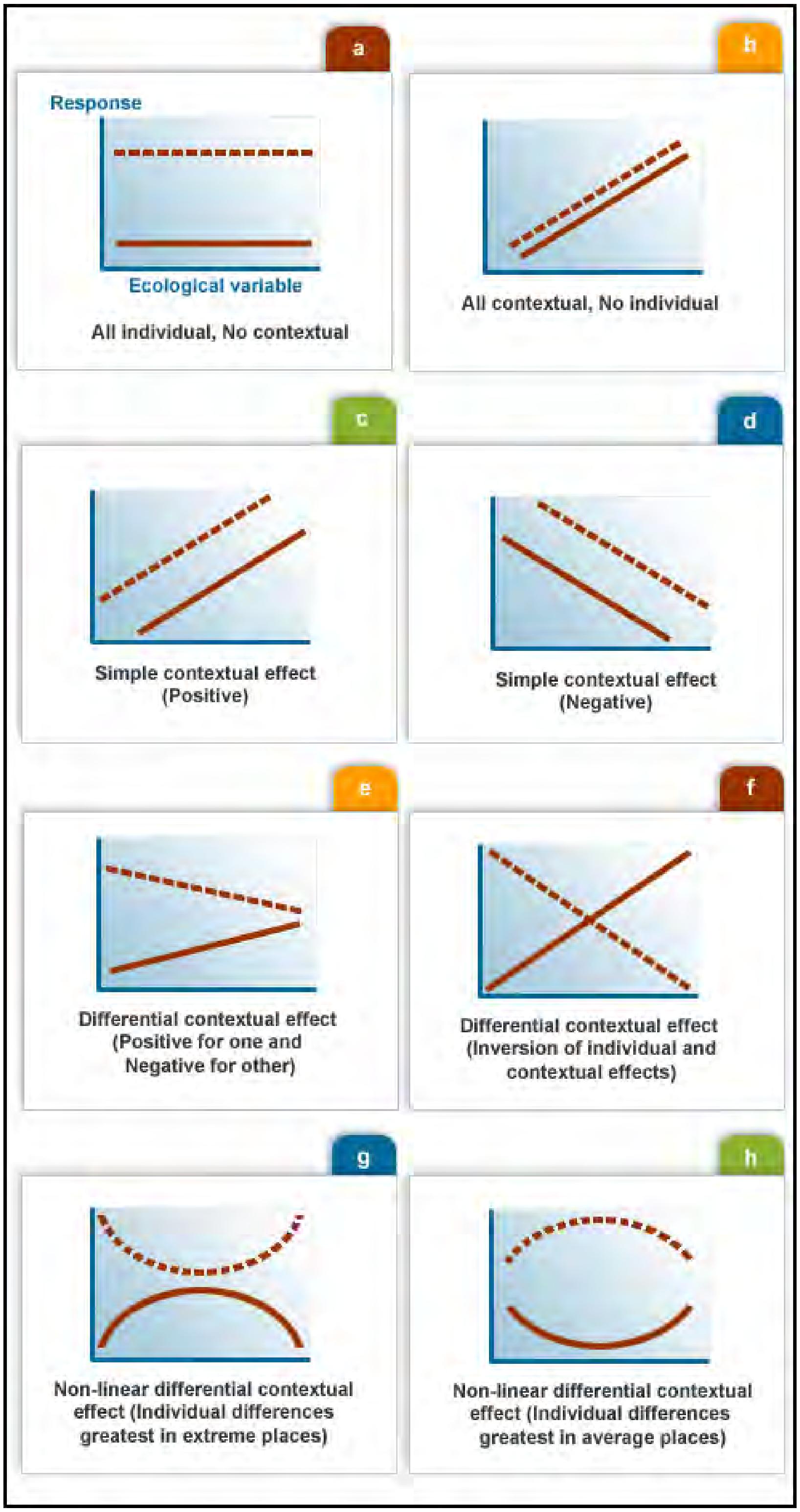
8. A Graphical Introduction

We illustrate the idea of such cross-level interactions by building on our running example of a two-level model (individuals at level-1 within neighborhoods at level-2) with the response being a score for poor health for each individual. We consider the categorical individual predictor, social class (with high social class as a reference and low social class specified as a contrast indicator variable), and a continuous neighborhood-level contextual predictor (e.g., socioeconomic deprivation index). Figure 7 portrays a range of hypothetical graphical models. In Figures 7(a)-(h), y-axis represents the poor health score and the x-axis shows the neighborhood socioeconomic deprivation index. The dashed-line represents low social class, and the solid-line represents high social class.

Figure 7(a) shows marked differences between high social class and low social class but no contextual effect for neighborhood socioeconomic deprivation (all individual, no contextual). Figure 7(b) represents the converse: small difference between the two social groups but a large contextual effect of socioeconomic deprivation (all contextual, no individual). The parallel lines in Figure 7(c) and 7(d) show both individual and contextual effects. In Figure 7(c), the neighborhood socioeconomic deprivation is shown to have a detrimental effect on the health of the individuals and the reverse is shown in Figure 7(d).
The key point is that the contextual effect of socioeconomic deprivation is seen to be the same for both high social class and low social class. Put differently, while neighborhood socioeconomic deprivation explains the prevalence of poor health, it does not account for the inequalities in health between the social class groups. In Figure 7(e) contextual effects are different for different groups. They are shown as positive for high social class and negative for low social class, such that in neighborhoods with highest level of socioeconomic deprivation health inequalities are minimum.

Thus, neighborhood-level socioeconomic deprivation is not only related to average health achievements but also shapes social inequalities in health. Figure 7(f) represents the case where contextual effects are strong enough to invert the individual effects. Figures 7(g) and 7(h) show models in which non-linear terms are of importance, such that the smallest or largest group inequalities in health are found respectively at ‘average’ levels of socioeconomic deprivation and not at the extreme levels of neighborhood socioeconomic deprivation.

Figure 7: Hypothetical Graphical Models



9. Specifying and Interpreting Models

Suppose we are interested in studying the variation in a poor health score, as a function of certain individual and neighborhood predictors. Let us assume that the researcher collected data on a sample of 50 neighborhoods and, for each of these neighborhoods, a random sample of individuals. We then have a two-level structure where the outcome is a poor health score, *y*, for individual *i* in neighborhood *j*. We will restrict this example to one individual-level categorical predictor, poverty, x_{1ij}, coded as 0 if not poor and 1 poor, for every individual *i* in neighborhood *j*; and one neighborhood predictor, w_{1j}, a socioeconomic deprivation index in neighborhood *j*.

Multilevel models operate by developing regression equations at each level of analysis. In the illustration considered here, models would have to be specified at two levels, level-1 and level-2. The model at level-1 can be formally expressed as:

$$1)\ y_{ij}=\beta_{0j}+\beta_1\chi_{ij}+e_{0ij}$$

In this level-1 model, β_{0j} (associated with a constant, X_{0ij}, which is a set of 1s, and therefore, not written) is the mean poor health score for the j^{th} neighborhood for the non-poor group; β_1 is the average differential in health score associated with individual poverty status (X_{1ij}) across all neighborhoods. Meanwhile, e_{oij} is the individual or the level-1 residual term. To make this a genuine two-level model we let β_{0j} become a random variable, with an assumption that:

$$2)\ B_{0j} = B_0+u_{0j}$$

where u_{oj} is the random neighborhood-specific displacement associated with the overall mean poor health score (B_0) for the non-poor group. Since we do not allow, at this stage, the average differential for the poor and non-poor group (B_1) to vary across neighborhoods, u_{oj} is assumed to be same for both groups. The equation in (2) is then the level-2 between-neighborhood model.

9. Specifying and Interpreting Models

It is worth emphasizing that the 'neighborhood effect', u_{0j}, can be treated in one of two ways. One can estimate each one separately as a fixed effect (i.e., treat them as a variable; with 50 neighborhoods there will be 49 additional fixed parameters to be estimated). Such a strategy may be appropriate if the interest is in making inferences about just those neighborhoods. Or, on the other hand, neighborhoods could be treated as a (random) sample from a population of neighborhoods (which might include neighborhoods in future studies if one has complete population data), if the interest is in making inferences about the variation between neighborhoods in general. Adopting this multilevel statistical approach makes u_{0j} a random variable at level-2 in a two-level statistical model.

Substituting the level-2 model (equation 2) into level-1 model (equation 1) and grouping them into fixed and random-part components (the latter shown in brackets) yields the following combined, also referred to as *random-intercepts* or *variance components*, model:

$$3)\ y_{ij}=\beta_0+\ \beta_1 x_{ij}+(u_{0j}+e_{0ij})$$

We have now expressed the response *yij* as the sum of a fixed part and a random part. Assuming a normal distribution with a 0 mean, we can estimate a variance at level-1 (σ^2_{e0}: the between-individual within-neighborhood variation) and level-2 (σ^2_{u0}: the between-neighborhood variation), both conditional on fixed individual poverty differences in poor health score.

It is the presence of more than one residual term (or the structure of the random part more generally) that distinguishes the multilevel model from the standard linear regression models or analysis of variance type analysis. The underlying random structure (variance-covariance) of the model specified in equation (3) is:

$$Var[u_{0j}] \sim N(0,\ \sigma^2_{u0}\ )\ ;\ Var[e_{0ij}] \sim N(0,\ \sigma^2_{e0}\ );\ and\ Cov[u_{0j},\ e_{0ij}\ ]=\ 0.$$

It is this aspect of the regression model that requires special estimation procedures in order to obtain satisfactory parameter estimates (Goldstein, 2003).

9. Specifying and Interpreting Models

The model specified in equation (3) with the above random structure is typically used to partition variation according to the different levels, with the variance in y_{ij} being the sum of σ^2_{u0} and σ^2_{e0}. This leads to a statistic known as intra-class correlation, or intra-unit correlation, or more generally variance partitioning coefficient (Goldstein, Browne et al., 2002), representing the degree of similarity between two randomly chosen individuals within a neighborhood. This can be expressed as:

$$p = \frac{\sigma^2_{u0}}{\sigma^2_{u0} + \sigma^2_{e0}}$$

We can expand the random structure in equation (3) by allowing the fixed effect of individual poverty (β_1) to randomly vary across neighborhoods in the following manner:

$$4)\ y_{ij} = \beta_{0j} + \beta_{1j}x_{1ij} + e_{0ij}$$

At level-2, there will now be two models:

$$5)\ \beta_{0j} = \beta_{0j} + u_{oj}$$

$$6)\ \beta_{ij} = \beta_1 + u_{1j}$$

Substituting the level-2 models in equations (5) and (6) into the level-1 model in equation (4) gives:

$$7)\ y_{ij} = \beta_0 + \beta_1 x_{1ij} + (u_{0j} + u_{1j}x_{1ij} + e_{0ij})$$

Across neighborhoods, the mean poor health score for non-poor is β_0, the mean poor health score for the poor is $\beta_0 + \beta_1$, and the mean 'poverty-differential' is β_1. The poverty differential is no longer constant across neighborhoods, but varies by the amount u_{1j} around the mean, β_1. Such models are also referred to as random-slopes or random coefficient models. These models have a much more complex variance-covariance structure than before:

$Var\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N(0, \begin{bmatrix} \sigma^2_{u0} & \\ \sigma_{u01} & \sigma^2_{u1} \end{bmatrix})$; and $Var[e_{0ij}] \sim N(0, \sigma^2_{e0})$.

9. Specifying and Interpreting Models

With this formulation, it is no longer straightforward to think in terms of a summary intraclass correlation statistic ρ, as the level-2 variation is now a function of an individual predictor variable, x_{1ij} (Goldstein, Browne et al., 2002). In our example when x_{1ij} is a dummy variable, we will have two variances estimated at level-2: one for non-poor, which is *σ2u0*; and one for poor, which is $\sigma^2_{u0} + 2\sigma_{uou1}x_{1ij} + \sigma^2_{u1}x^2_{1ij}$.

That is, level-2 variation will be a quadratic function of the individual predictor variable when *xij* is a continuous predictor. Thus the notion of 'random intercepts and slopes,' while intuitive, is not entirely appropriate. Rather, what these models are really doing is modeling variance as some function (constant, quadratic, or linear) of a predictor variable (Subramanian, Jones et al., 2003).

Building on the above perspective of modeling the variance-covariance function (as opposed to 'random intercepts and slopes'), we can extend the concept to modeling variance function at level-1. It is extremely common to assume that the variance is 'homoskedastic' in the random part at level-1 (σ^2_{e0};equation (7)), and, indeed, researchers seldom report whether this assumption was tested or not. One strategy would be to model the different variances for poor and non-poor of the following form:

$$8)\ y_{ij} = \beta_0 + \beta_1 x_{1ij} + (u_{0j} + u_{1j}\ x_{1ij} + e_{1ij}x_{1ij} + e_{2ij}x_{2ij})$$

where x_{1ij}= 0 for non-poor, 1 for poor, and the new variable x_{2ij} = 1 for non-poor, 0 for poor, with $Var[e_{1ij}] = \sigma^2_{e1}$ giving the variance for poor, and $Var[e_{2ij}] = \sigma^2_{e2}$ giving the variance for non-poor, and $Cov[e_{1ij},\ e_{2ij}] = 0$. There are other parsimonious ways to model level-1 variation in the presence of a number of predictor variables (Goldstein, 2003; Subramanian, Jones et al., 2003). With this specification, we do not have an interpretation of the random level-1 coefficients as 'random slopes' as we did at level-2. The level-1 parameters, σ^2_{e1} and σ^2_{e2}, describe the complexity of level-1 variation, which is no longer homoskedastic (Goldstein, 2003).

Anticipating and modeling heteroskedasticity or heterogeneity at the individual level may be important in multilevel analysis as there may be cross-level confounding –- what may appear to be neighborhood heterogeneity (level-2), to be explained by some ecological variable, could in fact be due to a failure to take account of the between-individual (within-neighborhood) heterogeneity (level-1).

9. Specifying and Interpreting Models

An attractive feature of multilevel models – one that is perhaps most commonly used in research – is their utility in simultaneously modeling neighborhood and individual characteristics, and any interaction between them. We will consider the underlying level-2 model related to equation (8), which is exactly the same as specified in equations (5) and (6), but now including a level-2 predictor: *w1j*, the deprivation index for neighborhood *j*:

$$9)\ \beta_{0j} = \beta_0 + a_1 w_1 + u_{0j}$$

$$10)\ \beta_{1j} = \beta_1 + a_2 w_{1j} + u_{1j}$$

Note that the separate specification of micro and macro models correctly recognizes that the contextual variables (w_{1j}) are predictors of between-neighborhood differences.

The extension of micro model (8) will now be:

$$11)\ y_{ij} = \beta_0 + \beta_1 x_{1ij} + a_1 w_{1j} + a_2 w_{1j} + x_{1ij} + (u_{0j} + u_{1j} x_{1ij} + e_{1ij} x_{1ij} + e_{2ij} x_{2ij})$$

The combined formulation in equation (11) highlights an important feature, the presence of an interaction between a level-2 and level-1 predictor $(w_{1j} \cdot x_{1ij})$, represented by the fixed parameter a_2. Now, a_1 estimates the marginal change in health score for a unit change in the neighborhood deprivation index for the non-poor, and a_2 estimates the extent to which the marginal change in health score for unit change in the neighborhood deprivation index is *different* for the poor.

More generally, this formulation has a direct translation to the assessment of social inequalities in health (Subramanian, Jones et al., 2003). For instance, evidence for an interaction between an ecologic predictor and an individual predictor suggests that the effect of the ecologic predictor on the individual outcome is different at different levels of the individual predictor variable. Vice versa, it would also mean that the individual-based inequalities in health would be different at different levels of ecologic disparities. This multilevel statistical formulation allows *cross-level effect modification*, or interaction between individual and neighborhood characteristics, to be robustly specified and estimated.

To summarize, multilevel models are concerned with modeling, at different levels, both the average and the variation around the average. To accomplish this they consist of two sets of parameters: those summarizing the average relationship(s), and those summarizing the variation around the average at both the level of individuals and neighborhoods. A fundamental point that needs to be emphasized from the discussion above is that it is *not* the neighborhood-specific values that are estimated by multilevel models. Rather they estimate the variance and the covariance.

10. Modeling Contextual Effects

It is worth drawing parallels between a simple multilevel or a random-intercepts model (3) and the conventional OLS or fixed-effects regression model. Consider the fixed-effects model, whereby the neighborhood effect is estimated by including a dummy for each neighborhood, as shown below:

$$12)\ y_{ij} = \beta_0 + \beta x_{ij} + \beta N_j + (e_{0ij})$$

where N_j is a vector of dummy variables for $N - 1$. neighborhoods. The key conceptual difference between the fixed and the random-effects approach to modeling contexts is that while the fixed part coefficients are estimated separately, the random part differentials (u_{0j}) are conceptualized as coming from a distribution (Goldstein, 2003).

This conceptualization results in three practical benefits (Jones and Bullen, 1994):

1. *Pooling information* between neighborhoods, with all the information in the data being used in the combined estimation of the fixed and random part; in particular, the overall regression terms are based on the information for all neighborhoods;
2. *Borrowing strength*, whereby neighborhood-specific relations that are imprecisely estimated benefit from the information for other neighborhoods; and
3. *Precision-weighted estimation*, whereby unreliable neighborhood-specific fixed estimates are differentially down-weighted or shrunk toward the overall city-wide estimate. A reliably estimated within-neighborhood relation will be largely immune to this shrinkage.

10. Modeling Contextual Effects

The random-effects and the fixed-effects estimates for each neighborhood, meanwhile, are related (Jones and Bullen, 1994). The neighborhood-specific random intercept (β_{0j}) in a multilevel model is a weighted combination of the specific neighborhood coefficient in a fixed effects model (β^{*}_{0j}) and the overall multilevel intercept (β_0), in the following way: $\beta_{0j} = w_j\beta^{*}_{0j} + (1 - w_j)\, \beta_0$, with the overall multilevel intercept being a weighted average of all the fixed intercepts: $\beta_0 = (\Sigma\, w_j\beta^{*}_{0j})/\Sigma\, w_j$.

Each neighborhood weight is the ratio of the true between-neighborhood parameter variance to the total variance, which additionally includes sampling variance resulting from observing a sample from the neighborhood. Consequently, the weights represent the reliability or precision of the fixed terms:

$$w_j = \frac{\sigma^2_{uo}}{\upsilon^2_j + \sigma^2_{uo}}$$

, where the random sampling variance of the fixed parameter is:

$$\upsilon^2_j = \frac{\sigma^2_e}{n_j},$$

with n_j being the number of observations within each neighborhood.

When there are genuine differences between the neighborhoods and the sample sizes within a neighborhood are large, the sampling variance will be small in comparison to the total variance. As a result, the associated weight will be close to 1, with the fixed neighborhood effect being reliably estimated, and the random-effect neighborhood estimate close to the fixed-neighborhood effect. As the sampling variance increases, however, the weight will be less than 1 and the multilevel estimate will increasingly be influenced by the overall intercept based on pooling across neighborhoods. Shrinkage estimates allow the data to determine an appropriate compromise between specific estimates for different neighborhoods and the overall fixed estimate that pools information across places over the entire sample (Jones and Bullen, 1994).

Importantly, the fixed effects approach to modeling neighborhood differences using cross-sectional data is not a choice for a typical multilevel research question. This is especially true where there is an intrinsic interest in an exposure measured at the level of neighborhood such as the one specified in model (2); in such instances, a multilevel modeling approach is a necessity. This is because the dummy variables associated with the neighborhoods (measuring the fixed effects of each neighborhood) and the neighborhood exposure is perfectly confounded, and, as such, the latter is not identifiable (Fielding, 2004). Thus, the fixed effects specification to understand neighborhood differences is unsuitable for the sort of complex questions which multilevel modeling can address.

Exercise 4: Fixed or Random Strategies

Below are two statements. For each statement, select whether the appropriate strategy is fixed or random.

Statement A:
The researcher wishes to quantify the association between density of fast food restaurants in a neighborhood and individual BMI.

Fixed Random

Statement B:
The researcher wishes to ascertain the effect of individual SES on individual BMI.

Fixed Random

11. Multiple Spatial Contexts

Much of the existing accounts of multilevel methods have been largely restricted to two-level structures, which typically put individuals at level-1 and places at level-2. In this section, we extend the model to consider the multiplicity of spatial levels in public health. For instance, in the US, geographical units such as block groups (BGs), census tracts (CTs), counties, or states may each exert a differential influence on health in the population.

Despite this, most research examining the effects of context on health has conceptualized contextual effects at only one level of geography.

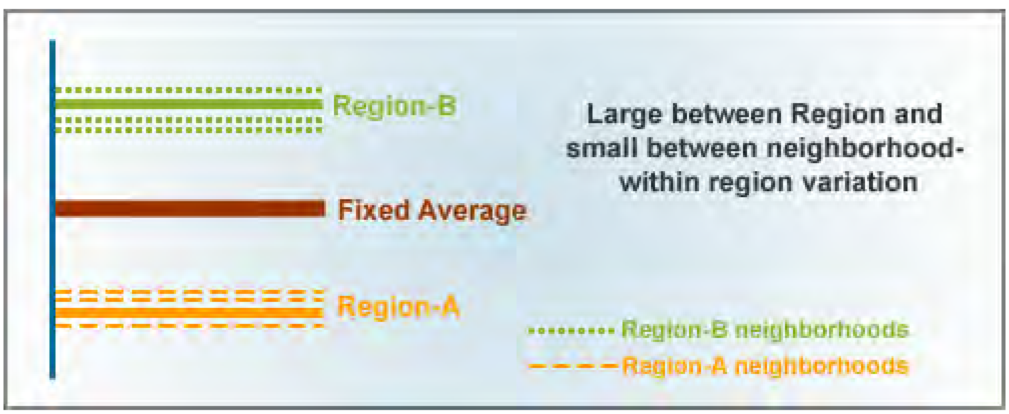
Multiple hierarchical geographic levels may be needed to explain the mechanisms by which context at different levels affects health. The multiplicity of geographic levels raises a fundamental issue in design: determining the number of levels necessary to analyze a particular health outcome and the relative importance of different levels. Consider, for example, a hierarchy of different geographic levels, in which BGs are nested within CTs that in turn are nested within counties within states. A fallacy would occur if poor health has a strong dependence at the BG level, but the analysis only considers the CT level, thereby resulting in incorrect inference at the individual level and the CT level.

11. Multiple Spatial Contexts

To appreciate the importance and implications of including this additional spatial level (at level-3), a series of graphical typologies is developed (Subramanian, Duncan et al., 2001). For the purpose of clarity and ease of understanding, we start with the simple case, shown in Figure 8, in which we assume that the differences between places (at both the spatial levels) are the same for both the social class groups. We continue with the use of the term 'neighborhoods' to represent level-2 spatial units, and introduce the term 'regions' to represent level-3 spatial units.

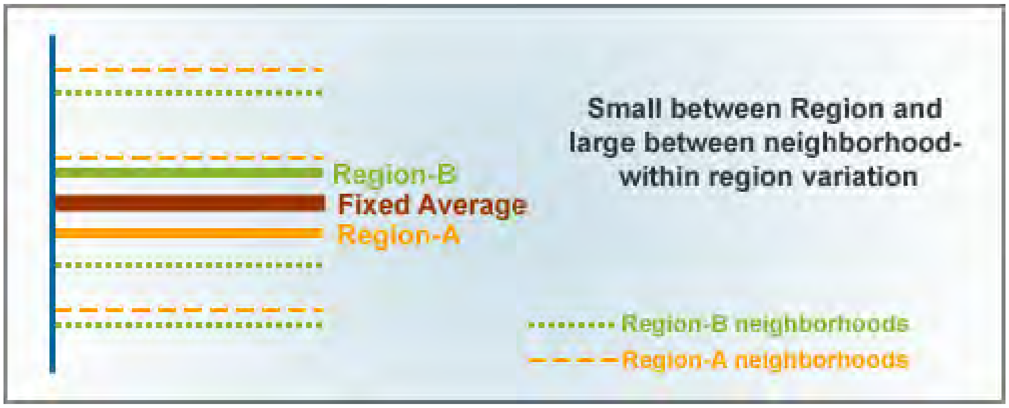
In Figure 8, the y-axis represents the individual score for poor health, the solid thick line representing the fixed average, the thinner solid lines representing the regions, while the dashed and the dotted lines represent neighborhoods within those regions.
In Figure 8(a) it can be seen that while regions vary significantly around the average line, such that one is high (Region-B) and one is low (Region-A), the neighborhoods within each lie close to their respective region lines. This suggests that there is no need to include neighborhoods as a level, and that a structure of individuals nested within regions is sufficient to capture the main source of geographic variation.

Figure 8a: Large and Small Variation



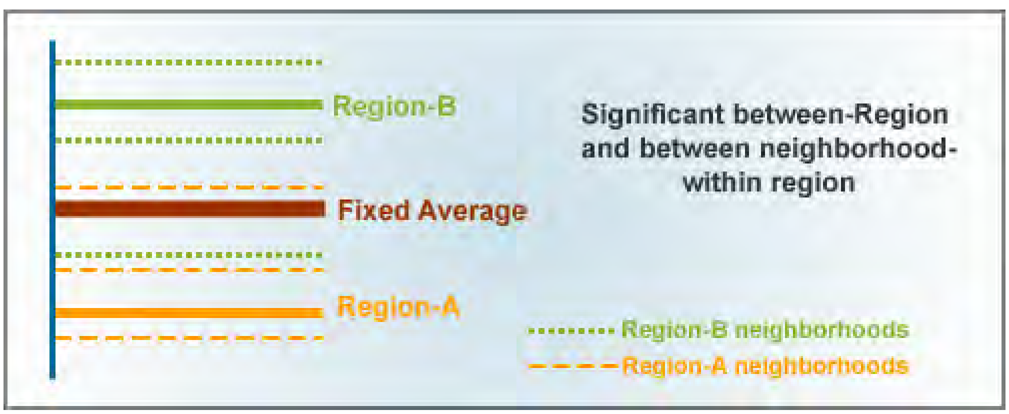
In Figure 8(b), the converse is portrayed; while the differences between regions are insignificant (i.e., they are grouped close to the overall average line), those between neighborhoods are substantial. This would suggest the greater importance of neighborhood level compared to region level.

Figure 8b: Small and Large Variation



Finally, Figure 8(c) anticipates a situation with significant variation at both region and neighborhood levels. While the relative importance of each might vary, both levels need to be included in an empirical model.

Figure 8c: Region and Neighborhood



11. Multiple Spatial Contexts

Ascertaining the relative importance of different spatial scales can, after taking into account (individual) compositional effects, provide important clues to the level 'at which the action lies.' A multilevel framework is ideally and readily suited to this task. Thus, underlying Figure 8 is a multilevel model based on a three-level structure of individuals (level-1) nested within neighborhoods (level-2) nested within regions (level-3). The micro model can be written as:

$$13)\ y_{ijk} = \beta_{0jk} + \beta_1 x_{1ijk} + e_{0ijk}$$

with an additional subscript to represent the regions. In addition, there would be a macro model at the neighborhood level (level-2):

$$14)\ \beta_{0jk} = \beta_{0k} + u_{0jk}$$

where, β_{0k}, is the poor health proportion for region k; and u_{0jk} is the differential for the j^{th} neighborhood in the region. There would also be a macro model at the region level (level-3):

$$15)\ \beta_{0k} = \beta_0 + \upsilon_{0k}$$

where, β_0 is the average poor health score; and υ_{0k} is the differential for the k^{th} region, to form an overall three-level 'random-intercepts' model:

$$16)\ y_{ijk} = \beta_0 + \beta_1 x_{1ijk} + (\upsilon_{0k} + u_{0jk} + e_{0ijk})$$

Depending on the relative size of the neighborhood and region level variance terms (σ^2_{uo}) and ($\sigma^2_{\upsilon o}$), respectively, that summarizes the place-specific differentials at each level, this model would produce one of the patterns shown in Figure 8.

Exercise 5: Creating Models for Three Hypotheses

You will be presented with three hypotheses. Using the characters provided, review each hypothesis and create the correct model to test it.

Consider the following, where:
neighborhood is *j* and individual is *i*
BMI is *y*
regression parameters is β

Characters:

$\boxed{y_{ij}}$ $\boxed{\beta}$ $\boxed{\beta_0}$ $\boxed{(FF_j)}$ $\boxed{(Age_{ij})}$ $\boxed{(u_{0j} + e_{0ij})}$

Hypothesis 1 of 3:
Neighborhood variation in BMI is conditional on age.

$\boxed{} = \boxed{} + \boxed{}\boxed{} + \boxed{}$

Hypothesis 2 of 3:
Neighborhood density of fast food restaurants is associated with individual BMI, conditional on individual age.

$\boxed{} = \boxed{} + \boxed{}\boxed{} + \boxed{}\boxed{} + \boxed{}$

Hypothesis 3 of 3:
Neighborhood density of fast food restaurants is associated with individual BMI, but its effect varies by age.

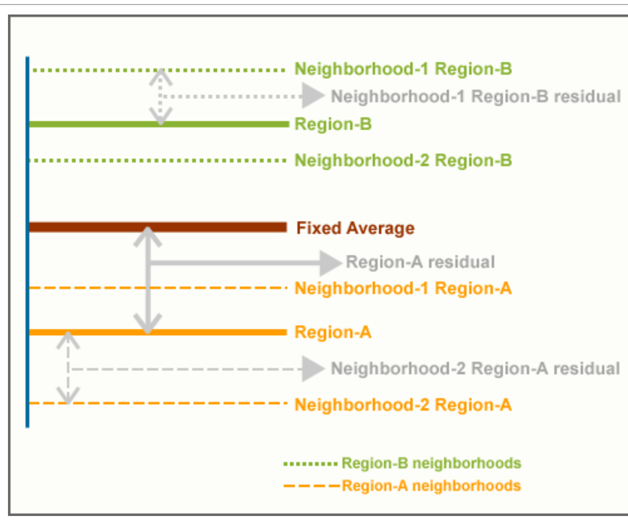
$\boxed{} = \boxed{} + \boxed{}\boxed{} + \boxed{}\boxed{} + \boxed{}\boxed{}\boxed{} + \boxed{}$

12. Multilevel Residual Mapping

While it is the variances that are estimated in a multilevel model at each of the specified levels, it is possible to estimate place-specific (posterior) residuals at each of the contextual levels. Residual mapping is an extremely useful application of multilevel models, especially when the interest lies in simultaneous multiple geographies, and when all the units at each of the geographic level can be observed in the analysis (e.g., the census) (Subramanian, Duncan et al., 2001). In order to appreciate this, Figure 9 unpacks the way in which residuals are constructed when there are two spatial levels.

The region-specific residuals ($u0k$) at level-3 represent the difference from the fixed average line, $\beta 0$. For example, REGION-A will have a negative residual given its lower rate of poor health compared to the overall average. REGION-B, in contrast, will have a positive residual, given its high rate compared to the average. Neighborhood-specific residuals ($u0jk$) at level-2, meanwhile, are measured as the difference from their respective regions to which they belong (hence, the subscript jk) and not as a difference from the fixed average.
Consider NEIGHBORHOOD-1 in REGION-A in Figure 9. From a conventional perspective, this neighborhood would be considered a 'healthy place' (given that it is below the average). From a multilevel perspective, however, this neighborhood would be considered an 'unhealthy place' given the healthy context of the region (low rate) to which it belongs, and as such, it would have a positive neighborhood residual. Such ideas are extremely useful in social policy (Goldstein and Spiegelhalter, 1996).

Figure 9: Multilevel Residual Mapping



12. Multilevel Residual Mapping

Example 3

Consider NEIGHBORHOOD-2 in REGION-A, and NEIGHBORHOOD-2 in REGION-B in Figure 9. Both neighborhoods are seen to be performing well with low rates of poor health (negative neighborhood residuals). The similarity in neighborhood effects is, however, occurring in entirely different contexts and as such may be telling quite a different story. While low rates in NEIGHBORHOOD-2 in REGION-A are being achieved within a favorable context (a low-rate region), in NEIGHBORHOOD-2 in REGION-B they are occurring within an unfavorable context (a high-rate region).

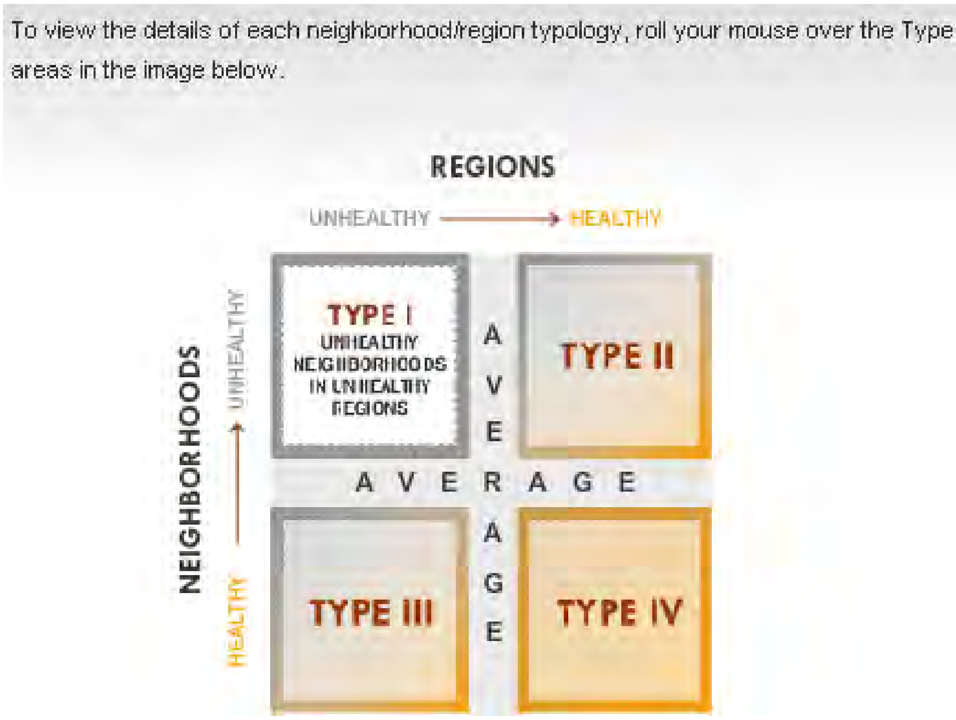
As this example illustrates, we have a nuanced way of evaluating and monitoring the performance of particular places.

One possibility, as shown in Figure 10, is to have a simple four-fold typology of neighborhood health performance:

- TYPE-I: Unhealthy neighborhoods in unhealthy regions;
- TYPE-II: Unhealthy neighborhoods in healthy regions;
- TYPE-III: Healthy neighborhoods in unhealthy regions; and
- TYPE-IV: Healthy neighborhoods in healthy regions.

The purpose of such typologies is not simply methodological, but substantive and practical. For instance, TYPE-I neighborhoods are doubly disadvantaged as ‘unhealthy’ neighborhoods in ‘unhealthy’ regions, while TYPE-IV neighborhoods suggest a ‘virtuous’ reinforcement of contextual advantage (‘healthy’ neighborhoods in ‘healthy’ regions). For TYPE-II and TYPE-III neighborhoods, meanwhile, contextual advantage at one level offsets disadvantage at the other. Determining the ‘cut-off’ points’ for what can be considered ‘healthy’ and ‘unhealthy’ is critical and care must be taken while identifying specific places, an issue to which we shall return later in this chapter. Nonetheless, our aim here is to illustrate the potential of a multilevel approach for evaluative and monitoring exercises that are usually of interest for public health departments.

Figure 10: Neighborhood and Region Typology



13. Parameter Estimation

Maximum Likelihood Estimators *(MLE)* that provide population values that maximize the so-called *Likelihood Function (LF)* that gives the probability of observing the sample data, given the parameter estimates. *MLE*, therefore, are parameter estimates that maximize the probability of finding the sample data that we have actually found. The *MLE* are available using the Newton-Raphson Fisher Scoring, Iterative Generalized Least Squares, or the Expectation Maximization algorithms (Longford, 1993).

Computing the *MLE* requires an iterative procedure.

1. **At the beginning**, starting values for the various parameter estimates (usually based on the *Ordinary Least Squares* regression estimates) are generated.
2. **In the next step**, the computation procedure improves upon the starting values to produce better estimates producing *Generalized Least Squares (GLS)*.
3. **This step is repeated** (iterated) until the changes in the estimates between two successive iterations become very small indicating *convergence*, with the parameter estimates now being *MLE*.

Lack of convergence could suggest: a) model mis-specification in the fixed part; b) mis-specification of the variance-covariance structure (either too simple or too complex); and c) small sample sizes at different levels.

13. Parameter Estimation

Different varieties of *MLE* are used in the available software for multilevel modeling. One is the *Full Information Maximum Likelihood (FIML)* where both the regression coefficients and the variance components are included in the *LF*. The other is the *Restricted Maximum Likelihood (REML)* and here only the variance components are included in the *LF*. The difference is that *FIML* treats the estimates for the regression coefficients as known quantities when the variance components are estimated, while *REML* treats them as estimates that carry some amount of uncertainty (Bryk and Raudenbush, 1992; Goldstein, 1995). While *REML* is more realistic and is recommended, especially when the number of grouping is small (Bryk and Raudenbush, 1992), *FIML* is computationally less demanding and allows for greater comparison across different model specifications.

The *MLE* theory is based on several assumptions. Three that are critical from an applied perspective are:

- Random parameters at all levels are normally distributed;
- Level-2 random parameters are independent of the level-1 random parameters; and
- Sample size is large and tends to infinity.

In practice, these assumptions will at best only be met approximately. Violations of these assumptions could lead to bias of the estimators and incorrect standard errors. In recent years, however, Bayesian estimation using Gibbs sampling (Gilks, Richardson et al., 1996); quasi-likelihood estimation; together with bias correction procedure (Goldstein and Rasbash, 1996) have been developed as alternatives. For inference, interval estimates are obtained directly from Gibbs sampling and via large sample deviance statistics or bootstrapping for *LF* estimation.

14. Non-linear Multilevel Models

So far we have illustrated the methodological concepts by considering a continuous response variable that has a Normal distribution.

However, a large number of outcomes of interest in public health research are not continuous and do not have Gaussian (Normal) distributive properties.

While not discussed in detail here, multilevel models are capable of handling a wide range of responses and in this sense there exist 'generalized multilevel models' to deal with:

- *Binary outcomes*;
- *Proportions* (as logit, log-log, and probit models);
- *Multiple categories* (as multinomial and ordered multinomial models); and
- *Counts* (as poisson and negative binomial distribution models) (Leyland and Goldstein, 2001).

Indeed, all these outcomes can be modeled using any of the hierarchical and non-hierarchical structures previously discussed (Goldstein, 2003).

These models work, in effect, by assuming a specific, non-Normal distribution for the random part at level-1, while maintaining the Normality assumptions for random parts at higher levels. Consequently, much of the discussion presented in this chapter focusing at the neighborhood and region level (higher contextual levels) would continue to hold regardless of the nature of the response variable. It may, however, be noted that the computation of VPC, discussed earlier, is not as straightforward in complex non-linear models as it is in Normal models, and is an issue of applied methodological research (Goldstein, Browne et al., 2002; Subramanian, Jones et al., 2003).

We should also mention that there are research developments whereby multilevel a perspective has been extended to survival and event history models, meta-analysis, structural equation modeling, and factor analysis (Goldstein, 2003).

15. Spatially Aggregated Data

While we have so far discussed the multilevel structure in terms of individuals at level-1 and places at level-2, we argue that a similar framework of people within places can be established using routinely available aggregate data (e.g., census and mortality data). As is well-known, analyses of aggregated data confounds the micro scale of people and the macro scale of places. Although regrettable, this situation is usually tolerated owing to the other obvious attractions of these data sets (e.g., large, extensive coverage of places at multiple levels). A multilevel approach offers a solution to this problem (Subramanian, Duncan et al., 2001).

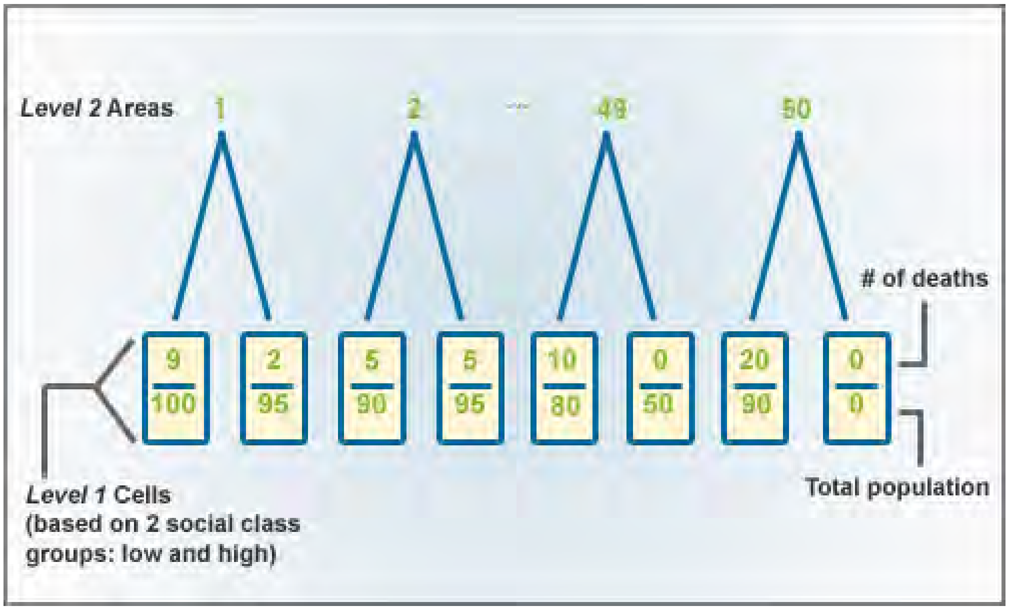
Table 1: Death Counts

Areas	Counts of Death out of total population	
	Low Social Class	High Social Class
1	9 out of 50	2 out of 50
2	5 out of 90	5 out of 95
-	-	-
49	10 out of 80	0 out of 50
50	20 out of 90	0 out of 0

Table 1 provides hypothetical data of deaths for two social groups in a format that is typical for spatially aggregated data.

Thus, in Area 1, 9 out of 50 in the low social class category died in a particular year; in Area 2, 5 out of 95 in the high social class category died, and so on. In this table, individuals are grouped as 'types' (low and high social class) and are represented as 'cells' of a table that contain counts of death for each social group in every area. Importantly, by using the compact, aggregated form of Table 1, data agencies can preserve individual confidentiality.

Figure 11: Social Class Data



Five points needs to be made about this table.

1. It is vital to note that underlying Table 1 is simply a set of individual records that happens to be presented in a tabular format, but can easily be changed into an individual record format.
2. Just as individuals nest within areas producing a two-level hierarchical data structure, so do the cells presented in Table 1, as shown in Figure 11.
3. Although the data here is cross-tabulated by only one individual characteristic, exactly the same principles apply when there is a greater degree of cross-tabulation.
4. If in an area there are no people of a particular type (e.g., missing high social class in Area 50 in Table 1) this poses no special problems as multilevel data structures can be unbalanced.
5. There are good reasons for invoking the notion of cells even when data is available in an individual record format since the amount of information, and therefore the associated computing time, can be reduced substantially.

Consequently, routinely available aggregated data can readily be adapted to a multilevel data structure with table cells at level-1 (representing the population groups) nested within places at level-2. The counts within each cell give the number of people with the outcome of interest (e.g., number of deaths) together with the ‘denominator’ (the total population). The proportion so formed becomes the response variable and the cell characteristics, meanwhile, are the individual predictor variables. Such a structure now lends itself to all the analytical capabilities that were discussed earlier (Subramanian, Duncan et al., 2001).

16. Interpreting a Multilevel Model

Multilevel Modeling Checklist

Were the neighborhoods specified correctly?

Researchers must make sure that the level-2 units are clearly defined and motivated in addition to constituting a random sample (with exchangeable properties) of all level-2 units (to avoid selection bias). In educational research, with schools or classes serving as a level-2 unit, the definition of level-2 units is usually straightforward. Similarly, institutional settings such as hospitals or clinics are more clearly defined than neighborhoods. It is also clear that all observed variation (net of individual characteristics) need not be systematically related to the (unobserved) neighborhood predictor. Rather, a part of it could be simply due to sampling variation due to the sample of neighborhoods. Epidemiologically, incorrect specification of neighborhood boundaries would probably cause an underestimation of neighborhood effects due to the introduction of non-differential misclassification bias. So, while a potential (and inevitable) source of error, it is unlikely to give rise to a spurious positive finding.

Was personal socio-economic position adequately controlled for?

Any substantive interpretation of σ^2_{u0} or a_1 is entirely dependent on the 'appropriate' specification of the other parts of the model: specifically, the fixed part of the model (i.e., individual-level and ecological-level confounders) as well as the random part specification of level-1.

Returning to our hypothetical example above and a substantive interpretation of the fixed-effect of neighborhood deprivation, the model that controlled just for personal income was probably inadequate. Personal socio-economic position is a complex and multidimensional construct that is often viewed as including income, educational attainment, social class, and (more recently) some measure of personal deprivation of hardship. Therefore, controlling for just one socio-economic factor is unlikely to fully adjust for personal socio-economic position. That is, residual confounding would remain. Nevertheless, the practice of adjusting for just one socio-economic factor before declaring neighborhood effects is not uncommon.

Even the model that fully adjusts for all the parallel individual-level variables that went into the construction of the neighborhood deprivation index may still be prone to residual confounding. First, for a fluctuating variable such as income, the average neighborhood income may be a better measure of your likely average income than personally declared (including reporting errors) income for the last year. Second, a measure or a personal socio-economic factor at one point in time does not capture dynamics over the lifecourse. While just one study, it is interesting to note that there was no association of neighborhood deprivation with mortality among a cohort of Scottish men after adjustment for social class at multiple points in the lifecourse (Davey Smith, Hart et al., 1997). Third, even if a satisfactorily full set of socio-economic factors across the life-course can be included in the analysis, the issue of (inevitable) measurement error of these covariates and resulting 'resonant confounding' of the neighborhood deprivation-mortality association remains (Marshall and Hastrup, 1996).

What about the interpretation of between-neighborhood variation, $\sigma^2 u0$?

It is a common finding that such variation is small in social epidemiology studies, and often not statistically significant – be it before or after the adjustment for individual-level covariates. However, direct interpretation of these variations for neighborhood-health research must be done cautiously, especially when the outcome is not linear (e.g., binary). Rather, the major utility of $\sigma^2 u0$ lies statistically in allowing robust and reliable fixed effect estimates of level-2 exposures on health (i.e., estimating the size and precision of direct ecological effects and cross-level effect modification). Within the framework of the 'context versus composition' question, undue focus on the random variation may be misleading, especially when we have only observational data. Importantly, moderately strong and statistically significant ecological effects are often found in the absence of statistically significant between-neighborhood variation (Merlo, 2003).

Were other individual-level confounders adequately controlled for?

This is a difficult, if not unanswerable, question for any observational study. By including further covariates, we gain from adjusting for further potential confounding. But, on the other hand, we risk including variables that are also on the causal pathway from neighborhood deprivation to health. For example, smoking may be patterned by one's context, meaning that adjusting for smoking is actually an attempt at quantifying the indirect ecological effect of neighborhood deprivation mediated by smoking.

This problem of variables that are both likely to be confounders and mediating variables in the association of an exposure with outcome is a perplexing problem in all observational studies. Short of conducting intervention studies on neighborhood deprivation, longitudinal studies with repeated measures of individual histories of changing neighborhood deprivation is one study design that may assist. That said, data-sets with both repeated measures and ecological-level variables are uncommon.

Do I need to use multilevel statistical methods?

This chapter has attempted to illustrate that multilevel statistical modeling meshes well with multilevel thinking. But there are other approaches to modeling clustered data (e.g., Generalized Estimating Equation, or GEE). However, it is beyond the scope of this chapter to canvass these options in detail. Briefly though, the GEE approach will often deliver the same result. The key difference between the GEE and the multilevel approach is that the latter models the random variation as being of intrinsic interest, rather than a nuisance to overcome. As such, the choice of strategy is really dependent upon the conceptual motivation of the researcher (Heagerty and Zeger, 2000).

What about endogeneity?

Since individuals do, to some extent, choose where to live, 'unobserved' individual or family factors can be mistaken for neighborhood effects. Similar 'unobserved' factors may also characterize neighborhoods and other spatial levels of analysis. This problem of endogeneity ---whereby an unobserved variable is related to a set of predictors and the response--- is only beginning to be addressed within the context of multilevel methods. The issue of endogeneity is even more complex in multilevel analysis because the unmeasured influences of omitted variables in the fixed part gets incorporated in the random part of the model, thereby violating the assumption of the independence of regressors and model disturbances (Rice, Jones et al., 1998). Three ways of dealing with this issue have been suggested in the multilevel literature. The first is to include data that actually "measure the crucial omitted variable" (Duncan and Raudenbush, 1999). The second is to apply specially developed multilevel instrumental variable estimation techniques (Spencer, 1998; Spencer and Fielding, 2000), i.e., the standard solution to endogeneity problems in single-level regression now extended to multilevel regression models. The third is to use a repeated measures, cross-classified structure, longitudinal fixed-effects model based on the nesting of panel observations for those who change neighborhoods within a cross-classified structure, with time varying covariates at each level of the analysis (Rasbash and Goldstein, 1994). This strategy, of course, is extremely data intensive and involves intensive computational demands. It is recommended that future applications be sensitive to the critical implication that the issue of endogeneity poses for multilevel research.

What about the sample of neighborhoods?

Even when one is not making uninformed predictions for specific contextual units, there are issues that need to be considered while making inferences for the population of contextual units. Simply because multilevel models treat the higher-level units as a sample drawn from a common population does not automatically mean that unconditional inferences are possible. There are crucial exchangeability judgments that are often neglected (Morris, 1995). Specifically, researchers need to ensure that the sample of neighborhoods comes from/can be exchanged with/is similar to the population that they wish to make inference about, with this being true for each specific neighborhood for which they have data. If there are reasons to believe that certain neighborhoods are truly independent or that they come from different populations they should not be regarded as exchangeable with the remaining random sample of neighborhoods and as such should be treated as fixed effects. While one option is to perform diagnostics after the model is fitted and/or conducting multilevel analysis with and without those neighborhoods that are believed to share exchangeable properties, a conceptually sound approach is to carefully plan the selection of neighborhoods at the design stage, analogous to the sampling of individuals in a survey (Draper, 1995).

17. Power and Sample Size

As we have emphasized, multilevel models are not about modeling each neighborhood separately; rather, the sample of neighborhoods is seen as one realization from a population of neighborhoods.

When designing a powerful multilevel study it is vital, therefore, to consider the importance of two things:

1. Determining sample sizes at the various levels of analysis; and
2. Ensuring the property of exchangeability.

We first discuss the issue of **sampling in multilevel analysis**.

It is vital that the study design has 'adequate' number of units at all the levels of analysis. Specifically, by increasing sample sizes at all levels, estimates and their standard errors become more accurate. The analysis of binomial data in particular requires larger samples than the analysis of normally distributed data (Hox, 2002). Determination of sample sizes at level-1 and level-2 units for efficiency, unbiasedness and consistency of parameter estimates is not entirely straight-forward and this is especially the case if we are interested in the random slopes component.

In a two-level random intercepts model, the sample design question is analogous to computing the effective sample size in two-stage cluster sampling, as given by (Kish, 1965). Effective sample size of a two-stage cluster sampling design, n_{eff}, is computed by: $n^{eff}=n/[1+(n_{clus}-1)\rho]$, where n is the total number of individuals in the study, that is, the actual sample size; n_{clus} is the number of individuals per neighborhood; and ρ is the intra-class correlation. However, the analogy is not straightforward for random slopes models, because the ICC for these models is a function of the independent variable.

Consensus has yet to be developed on the precise power calculations within multilevel models. Some argue for a sample of at least 30 groups with at least 30 individuals in each group (Kreft, 1996). This advice is considered sound provided the interest is largely in the fixed parameters.

Modification to this ‘rule’ is advised if interest is in estimating cross-level interactions and/or variance and covariance components (Hox, 2002). For the former, a 50/20 ‘rule’ is recommended (about 50 neighborhoods with at least 20 individuals per neighborhood) and for a variance-covariance components model about 100 neighborhoods with about 10 individuals per neighborhood is suggested.

Indeed, if this is the case then one has to be cautious about making neighborhood-specific predictions. These ‘rules’ take into account that there are costs attached to data collection, such that if the number of neighborhoods is increased, the number of individuals per neighborhood decreases (Snijders and Bosker, 1993; Snijders, 2001).

18. Summary

Current implementations of multilevel models have generally failed to exploit the full capabilities of the analytical framework (Subramanian, 2004; Leyland, 2005; Moon, Subramanian et al., 2005).

Much, if not all, of the current research linking neighborhoods and health is cross-sectional, and assumes a hierarchical structure of individuals nested within neighborhoods. This simplistic scenario ignores possibilities such as, for instance, the fact that an individual might move several times and as such reflect neighborhood effects drawn from several contexts, or that other competing contexts (e.g., schools, workplaces, hospital settings) may simultaneously contribute to contextual effects.

While a multilevel analytic approach provides a sound rationale for modeling ecologic effects, it obviously does not overcome the limitations intrinsic to any observational study design, single-level or multilevel. Recent discussions on identifying causal ecologic effects (Oakes, 2004), inappropriately conflate issues of study design with the relevance of multilevel models. The critical challenge for identifying neighborhood effects arises from the use of observational (and often cross-sectional) study design, and not from the use of any specific analytic technique. Arguably, multilevel models are appropriate analytical techniques for understanding ecologic effects, regardless of whether the data were generated through observation or were a result of an experiment (Subramanian, 2004).

Identifying true casual associations of ecologic exposures with health:

Careful study design, thorough analysis, and careful interpretation, are essential components of this work. However, there are some other pointe to consider as well.

1. Rely less on interpreting residual associations, and **model directly the ecological exposure**. The above example of a composite index of socio-economic deprivation is a classic example. It is difficult to interpret a residual association of such an index with individual health for the reasons listed above, and because it is not actually clear what properties of neighborhoods the index is actually capturing. Following the longstanding

exhortations of Macintyre and others (Macintyre, Maciver et al., 1993; Macintyre, Ellaway et al., 2002; Macintyre and Ellaway, 2003; Cummins, Macintyre et al., 2005), conceptualizing and directly measuring those characteristics of neighborhoods that are hypothesized to effect health is likely to be more rewarding in the long-run, albeit more difficult.

2. While often impossible to conduct, **intervention studies that actually change ecological or neighborhood characteristics should be seized upon** by researchers whenever possible.
3. Third, **longitudinal studies with repeated measurements of neighborhood characteristics** over peoples' lifecourses should also be sought out.
4. We need to **be cognizant of the limits of quantitative multilevel analysis** and empiricism more generally.

There is a deep, complex, and dynamic inter-relationship between people and context. Where you live influences who you are (e.g., employment opportunities), and who you are influences your neighborhood. It will not always be possible, nor correct, to decompose health variations to personal and contextual characteristics. Rather, we will also need qualitative and other social science approaches.

18. Summary

Multilevel models have several features that make them attractive for public health research. In this chapter, we have sought to explain and emphasize how these methods offer an extremely flexible yet unified framework for conceptualizing and investigating substantive ideas related to contextuality and heterogeneity. Both heterogeneity and the correlated nature of data structures are seen as the norm, not an aberration, and consequently multilevel methods do not ignore it or adjust for it, but rather anticipate and model it. In doing so, we show that these methods encourage and foster refinement in our thinking about ideas related to different levels of causation.

> **Multilevel methods constitute a conceptual and technical approach that can help re-conceptualize much of public health research.**

Specifically, they compel the researcher to reflect on the multilevel nature of causal processes; and raise questions that are not simply about fixed averages but rather about the variability and heterogeneity of populations. Indeed, multilevel methods are changing the way we think about 'individual effects' and 'contextual effects.' From an initial view of interpreting these effects in terms of 'who you are in relation to where you are,' multilevel methods encourage us to think along the lines of 'who you are depends upon where you are.' While the methodological capability of multilevel models to disentangle compositional (individual) and contextual effects related to spatial variations has been well demonstrated, the multilevel framework has also re-defined the construct of 'individual compositional explanations.' As Macintyre and Ellaway point out,

> *"your SES and income are partly a product of your place of upbringing, rather than being intrinsically personal attributes"* (Macintyre, 2000; Macintyre and Ellaway, 2000).

While being attractive, conceptually and technically, multilevel methods are also undoubtedly complex and should not be approached in simplistic terms. Simplistic use of complex methodological tools can lead to interpretive confusion and a potential overstatement of what may be validly concluded from a given piece of research (Draper, 1995). These last comments are not meant to discourage the use of multilevel methods. Rather, multilevel models can raise new research agendas and provide important insights into existing knowledge. At the same time, as one of the pioneers in this field reminds us, multilevel methods, like all statistical methods, need to be used “with care and understanding” (Goldstein, 1995).

19. References

Aitkin, M., Anderson, D., et al. (1981) Statistical modelling of data on teaching styles (with discussion). Journal of the Royal Statistical Society A 144: 148-161.

Anderson, N. B. (1999) Solving the puzzle of socioeconomic status and health: The need for integrated, multilevel, interdisciplinary research. Annals of the New York Academy of Sciences. 896: 302-312.

Bennett, N. (1976) Teaching styles and pupil progress. London: Open Books.

Berkman L.F., Kawachi I. (2000) Social epidemiology. New York: Oxford University Press. xxii, 391.

Blakely, T., Salmond, C., et al. (2000) Anonymous linkage of New Zealand mortality and Census data. Australian and New Zealand Journal of Public Health. 24: 92-95.

Blakely T., Subramanian, S.V. (2006) Multilevel studies. In: J.M. Oakes, J. S. Kaufman, (Eds). Methods in social epidemiology. San Francisco: Jossey-Bass. 316-340.

Blakely, T. A., Woodward, A. J. (2000) Ecological effects in multilevel studies. Journal of Epidemiology and Community Health. 54: 367-374.

Bryk, A. S., Raudenbush, S. W. (1992) Hierarchical linear models: Applications and data analysis methods. Newbury Park, UK: Sage Publications.

Cummins, S., Macintyre, S., et al. (2005) Measuring neighbourhood social and material context: Generation and interpretation of ecological data from routine and non-routine sources. Health & Place. 11(3): 249-260.

Davey Smith, G., Hart, C. et al. (1997) Lifetime socioeconomic position and mortality: prospective observational epidemiology. British Medical Journal. 314: 547-552.

Diez Roux, A. V. (2001) Investigating neighborhood and area effects on health. American Journal of Public Health. 91(11): 1783-9.

Diez Roux, A. V. (2002) A glossary for multilevel analysis. Journal of Epidemiology and Community Health. 56(8): 588-594.

Draper, D. (1995) Inference and hierarchical modeling in the social sciences. Journal of Educational and Behavioral Statistics. 20: 115-147.

Duncan G., Raudenbush, S. (1999) Assessing the effects of context in studies of child and youth development. Educational Psychologist. 34: 29-41.

Duncan, C., Jones, K., et al. (1993) Do places matter? A multilevel analysis of regional variations in health-related behaviour in Britain. Social Science and Medicine. 37(6): 725-733.

Fielding, A. (2004) The role of the Hausman test and whether higher level effects should be treated as random or fixed. Multilevel Modeling Newsletter. 16(2): 3-9.

Gilks, W., Richardson, S., et al. (1996) Markov Chain Monte Carlo in practice. London: Chapman and Hill.

Goldstein, H. (1994) Multilevel cross-classified models. Sociological methods and research. 22: 364-375.

Goldstein, H. (1995) Multilevel statistical models. London: Edward Arnold.

Goldstein, H. (2003) Multilevel statistical models. London: Edward Arnold.

Goldstein, H., Browne, W. J., et al. (2002) Partitioning variation in multilevel models. Understanding Statistics. 1: 223-232.

Goldstein, H., Rasbash, J. (1996) Improved approximations for multilevel models with binary responses. Journal of the Royal Statistical Society A 159: 505-513.

Goldstein, H., Spiegelhalter, D. (1996) League tables and their limitations: Statistical issues in comparisons of institutional performance (with discussion). Journal of the Royal Statistical Society Series A 159: 385-443.

Heagerty, P., Zeger, S. (2000) Marginalized multilevel models and likelihood inference (with discussion). Statistical Science. 15: 1-26.

Hox, J. (2002) Multilevel analysis: Techniques and applications. Mahwah, NJ: Lawrence Erlbaum Associates.

Jones, K., Bullen, N. (1994) Contextual models of urban house prices: A comparison of fixed- and random-coefficient models developed by expansion. Economic Geography. 70: 252-272.

Jones, K., Gould, M. I., et al. (1998) Multiple contexts as cross-classified models: The labor vote in the British general elections of 1992. Geographical Analysis. 30: 65-93.

Jones, K., Moon, G. (1987) Health, disease and society. London: Routledge.

Kawachi, I., Berkman, L. (Eds.). (2003) Neighborhoods and health. New York: Oxford Press.

Kawachi, I., Subramanian, S. V. (2006) Measuring and modeling the social and geographic context of trauma: A multilevel modeling approach. Journal of Traumatic Stress. 19(2): 195-203.

King, G. (1997) A solution to the ecological inference problem: Reconstructing individual behavior from aggregate data. Princeton: Princeton University Press.

Kish, L. (1965) Survey sampling. New York: John Wiley.

Kreft, I., de Leeuw, J. (1998) Introducing multilevel models. London: Sage Publications.

Kreft, I. G. G. (1996) Are multilevel techniques necessary? An overview including simulation studies. Los Angeles: Calfornia State University.

Langford, I. H., Bentham, G., et al. (1998) Multilevel modelling of geographically aggregated health data: A case study on malignant melanoma mortality and UV exposure in the European Community. Statistics in Medicine. (17): 41-57.

Leyland, A.H. (2005) Socioeconomic gradients in the prevalence of cardiovascular disease in Scotland: the roles of composition and context. Journal of Epidemiology and Community Health. 59 (9):799-803.

Leyland, A. H., Goldstein, H. (Eds.). (2001) Multilevel modelling of health statistics. Wiley Series in Probability and Statistics. Chichester: John Wiley & Sons Ltd.

Link B. G., Phelan, J. (1995) Social conditions as fundamental causes of disease. J Health Soc Behav. Spec No: 80-94.

Longford, N. (1993) Random coefficient models. Oxford, UK: Clarendon Press.

Lynch, J., Davey Smith, G., et al. (2004) Is income inequality a determinant of population health? Part 1. A systematic review. Milbank Q 82(1): 5-99.

Macintyre, S. (2000) The social patterning of health: Bringing the social context back in. Medical Sociology Newsletter. 26: 14-19.

Macintyre, S., Ellaway, A. (2000) Ecological approaches: Rediscovering the role of physical and social environment. In: L. F. Berkman and I. Kawachi (Eds). Social Epidemiology. New York: Oxford Press. 332-348.

Macintyre, S., Ellaway, A. (2003) Neighbourhoods and health: An overview. In: I. Kawachi and L. Berkman (Eds.). Neighborhoods and Health. New York: Oxford Press. 20-42.

Macintyre, S., Ellaway, A., et al. (2002) Place effects on health: How can we conceptualise, operationalise and meaure them? Social Science and Medicine. 55: 125-139.

Macintyre, S., Maciver, S. et al. (1993) Area, class and health: Should we be focusing on places or people? J Social Policy. 22(2): 213-234.

Marshall, J., Hastrup, J. (1996) Mismeasurement and the resonance of strong confounders: Uncorrelated errors. American Journal of Epidemiology. 143(10): 1069-1078.

McKinlay, J. B., Marceau, L. D. (2000) To boldly go. American Journal of Public Health. 90(1): 25-33.

Merlo, J. (2003) Multilevel analytical approaches in social epidemiology: Measures of health variation compared with traditional measures of association. Journal of Epidemiology and Community Health. 57: 550-552.

Moon, G., Subramanian, S. V., et al. (2005) Area-based studies and the evaluation of multilevel influences on health outcomes. In: A. Bowling and S. Ebrahim (Eds.). Handbook of Health Research Methods: Investigation, Measurement and Analysis. Berkshire, UK: Open University Press. 266-292.

Morris, C. (1995) Hierarchical models for educational data: An overview. Journal of Educational and Behavioral Statistics. 20: 190-199.

Oakes, M. (2004) The (mis)estimation of neighborhood effects: Causal inference for a practicable social epidemiology. Social Science and Medicine. 58: 1929-1952.

Pickett, K., Pearl, M. (2001) Multilevel analyses of neighbourhood socioeconomic context and health outcomes: A critical review. Journal of Epidemiology and Community Health. 55: 11-122.

Rasbash, J., Goldstein, H. (1994) Efficient analysis of mixed hierarchical and cross-classified random structures using a multilevel model. Journal of Educational and Behavioural Statistics. 19(4): 337-350.

Raudenbush, S. W. (2003) The quantitative assessment of neighborhood social environment. In: I. Kawachi and L. F. Berkman (Eds.). Neighborhoods and health. New York: Oxford University Press.

Rice, N., Jones, A., Goldstein, H. (1998) Multilevel models where the random effects are correlated with the fixed predictors: A conditioned iterative generalised least squares estimator (CIGLS). Multilevel Modelling Newsletter. 10: 10-14.

Snijders, T. A. B. (2001) Sampling. In: A. H. Leyland and H. Goldstein (Eds.). Multilevel modelling of health statistics. Chichester: John Wiley. 159-174.

Snijders, T. A. B., Bosker, R. J. (1993) Standard errors and sample sizes for two-level research. Journal of Educational Statistics. 18: 237-259.

Spencer, N. (1998) Consistent parameter estimation for lagged multilevel models. Hertfordshire: University of Hertfordshire Business School. UHBS 19 UHBS 19. 4-7 p.

Spencer, N., Fielding, A. (2000) An instrumental variable consistent estimation procedure to overcome the problem of endogenous variables in multilevel models. Multilevel Modelling Newsletter. 12: 4-7.

Subramanian, S. (2004) Multilevel methods, theory and analysis. In: N. Anderson (Ed.). Encyclopedia on Health and Behavior. Thousand Oaks, CA: Sage Publications. 602-608.

Subramanian, S., Jones, K., et al. (2003) Multilevel methods for public health research. In: I. Kawachi and L. Berkman (Eds.). Neighborhoods and Health. New York: Oxford Press. 65-111.

Subramanian, S. V. (2004) The relevance of multilevel statistical methods for identifying causal neighborhood effects. Social Science & Medicine. 58(10): 1961-1967.

Subramanian, S. V., Duncan, C., et al. (2001) Multilevel perspectives on modeling census data. Environment and Planning A. 33(3): 399-417.

Subramanian, S. V., Kawachi, I. (2004) Income inequality and health: What have we learned so far? Epidemiol Reviews. 26: 78-91.

Susser, M. (1998) Does risk-factor epidemiology put epidemiology at risk? Peering into the future. Journal of Epidemiology and Community Health. 52: 608-611.

Susser, M., Susser, E. (1996a) Choosing a future of epidemiology: I. eras and paradigms. American Journal of Public Health. 86: 668-673.

Susser, M., Susser, E. (1996b) Choosing a future of epidemiology: II. from black box to Chinese boxes and eco-epidemiology. American Journal of Public Health. 86: 674-677.

Wilkinson, R., Marmot, M., (Eds.). (2003) Social determinants of health: The solid facts, second edition. Copenhagen: World Health Organization Regional Office for Europe.

20. Author Biography

S V Subramanian, PhD, ('Subu' or 'Subra') has a PhD in geography with specialization in multilevel statistical methods. He also has a Masters in the field of development studies from the University of Delhi and was the recipient of the 1999-2000 MacArthur Leadership Program in Population and Development Studies based at the Harvard Center for Population and Development Studies.

The main focus of his research is on understanding how different contextual settings influence individual health outcomes and the population disparities in health achievements. He has specifically investigated the impact of income inequality and social capital on individual health outcomes. His work has demonstrated the need to explicitly consider a multilevel methodological framework while conceptualizing and estimating contextual effects on public health issues.

Currently, through a National Institutes of Health/National Heart, Lung, Blood Institute Career Development Award, he is investigating the role of neighborhood-level factors (neighborhood structural disadvantage, collective psychosocial characteristics, and physical environmental conditions) in explaining the occurrence and distribution of asthma. He is also conducting original collaborative research on the different ways in which health gets spatially and socially stratified in India, with particular focus on the socioeconomic distribution and determinants of nutritional status.

Besides his substantive interest in understanding place-effects on health, Subu is additionally conducting independent research on the methodological challenges to estimating causal effects of neighborhood factors on health.

Subu has published over 100 journal articles, book chapters, books, and working papers. He is an Assistant Editor for Social Science and Medicine; Editorial Consultant to The Lancet; Member of the Editorial Board for Health and Place, and BMC Public Health; and Book Review Editor for Economics and Human Biology.

Subu has extensive experience in teaching and training graduate students and post-doctoral fellows in multilevel statistical methods. He has lectured and conducted workshops on the concept and practical applications of multilevel models in Argentina, Australia, Austria, Belgium, Chile, India, Japan, Switzerland, the U.K., and the U.S. He and Professor Kelvyn Jones

(University of Bristol) have co-developed a training manual to assist researchers in the concept and application of multilevel models using the MLwiN program.

Acknowledgments

The work presented here draws upon earlier as well as on-going works. I am indebted to my generous and wonderful collaborators, Tony Blakely, William Browne, Nina Bullen, Craig Duncan, Maria Glymour, Harvey Goldstein, Myles Gould, Kelvyn Jones, Ichiro Kawachi, Graham Moon, Jon Rasbash, and Liz Twigg, from whom I have learned and continue to learn much about multilevel models. My work is supported by the National Institutes of Health Career Development Award (NHLBI 1 K25 HL081275-01), and the Robert Wood Johnson Foundation through their RWJ Health and Society Scholars.