

MULTILEVEL ANALYSIS IN PUBLIC HEALTH RESEARCH

Ana V. Diez-Roux

Division of General Medicine, Columbia College of Physicians and Surgeons, and Division of Epidemiology, Joseph T. Mailman School of Public Health, Columbia University, New York, New York; e-mail: diezrou@medicine1.cpmc.columbia.edu

Key Words methods, contextual effects, random effects, social epidemiology, ecologic

■ **Abstract** Over the past few years there has been growing interest in considering factors defined at multiple levels in public health research. Multilevel analysis has emerged as one analytical strategy that may partly address this need, by allowing the simultaneous examination of group-level and individual-level factors. This paper reviews the rationale for using multilevel analysis in public health research, summarizes the statistical methodology, and highlights some of the research questions that have been addressed using these methods. The advantages and disadvantages of multilevel analysis compared with standard methods are reviewed. The use of multilevel analysis raises theoretical and methodological issues related to the theoretical model being tested, the conceptual distinction between group- and individual-level variables, the ability to differentiate “independent” effects, the reciprocal relationships between factors at different levels, and the increased complexity that these models imply. The potentialities and limitations of multilevel analysis, within the broader context of understanding the role of factors defined at multiple levels in shaping health outcomes, are discussed.

INTRODUCTION

The term multilevel analysis (or hierarchical modeling) has been used in the fields of education (4), demography (44, 71), and sociology (19) to describe an analytical approach that allows the simultaneous examination of the effects of group-level and individual-level variables on individual-level outcomes. Over the past few years, interest in the use of multilevel analysis to investigate public health problems (14, 23, 96) has grown. This growth has been stimulated in part by a resurgence of interest in the potential ecological-, macro-, or group-level determinants of health and the notion that variables referring to groups or to how individuals are related to each other within groups may be relevant to understanding the distribution of health outcomes (14, 22, 84, 93, 96). A second driving force in the use of multilevel

methods has been the accelerated development of the statistical methods themselves (as well as the accompanying software) and the recognition that they have applications in a broad range of circumstances involving nested data structures.

The availability of these complex statistical methods challenges public health researchers to articulate theories of the causes of disease that bring together factors defined at different levels. This will ensure that the method does not become an end in itself, but rather serves as a tool to investigate more sophisticated and hopefully more realistic models of disease causation. This paper (a) reviews the rationale for using multilevel analysis in public health research; (b) describes the fundamentals of the methods involved and how they compare with traditional methods; (c) highlights selected areas in which these methodologies have been applied in the literature; and (d) summarizes the potential and limitations of multilevel analysis in achieving a more comprehensive understanding of the determinants of health and disease. Although the focus of this review is on the use of multilevel analysis to investigate research questions involving groups and individuals nested within them, other applications are also briefly mentioned.

RATIONALE FOR THE USE OF MULTILEVEL ANALYSIS

The idea that individuals may be influenced by their social context is a key notion in the social sciences, and has led to much debate and empirical research on the interactions between attributes of groups and attributes of individuals (2, 19, 46, 47, 95). In contrast, despite the fact that health and disease occur in social contexts, research into the determinants of health has often been characterized by individualization, that is, explaining individual-level outcomes exclusively in terms of individual-level independent variables. The underlying assumption is that all disease determinants are best conceptualized (and consequently best measured) at the individual level. Group-level variables are used only as proxies for individual-level data when the latter are unavailable. Populations (or groups) are thought of as collections of independent individuals, rather than entities with properties that may affect individuals within them. Consequently, there is generally little interest in examining group-to-group variation per se. Although there has been abundant discussion in the epidemiologic literature of the fallacy inherent in using data at one level to draw inferences at another level (specifically of the ecological fallacy), until recently there has been relatively little discussion of the substantive problem of ignoring potentially important variables that are best conceptualized and measured at the group level. Just as studies examining differences between groups may need to take into account possible differences in group composition (i.e. characteristics of the individuals within them), studies of individuals may need to take into account differences in the properties of the groups to which individuals belong (14).

In explaining the occurrence of a given phenomenon, researchers can appeal to different types of theories, which may be more or less relevant depending on the

particular question being investigated (9, 19). In the simplest case, the outcome at one level is explained by independent variables that apply to the same level. This is the approach commonly taken in epidemiology when individual-level outcomes are explained in terms of individual-level variables (as in traditional cohort studies) or group-level outcomes are explained in terms of group-level variables (as in ecological studies aimed at drawing group-level inferences). In a second approach, the outcome at one level is explained in terms of variables defined at a lower level. This, for example, is the approach taken when differences in disease rates across groups are explained in terms of the characteristics of individuals composing the groups. A third approach is to explain the outcome at one level as a function of variables defined at a higher level, for example, when an individual-level outcome is explained exclusively as a function of the attributes of the group to which individuals belong. A fourth approach is to explain variation in the dependent variable at one level as a function of variables defined at various levels, plus interactions within and between levels. Multilevel analysis is one methodology that can be used to approximate the latter situation.

THE MULTILEVEL METHOD AND ITS DIFFERENCES WITH OTHER APPROACHES

In analyzing data corresponding to individuals nested within groups, researchers have several options. The first is to ignore group membership and focus exclusively on interindividual variation and on individual-level attributes. This approach has the drawback of ignoring the potential importance of group-level attributes in influencing individual-level outcomes. In addition, if outcomes for individuals within groups are correlated, the assumption of independence of observations is violated, resulting in incorrect standard errors and inefficient estimates (17). A second option is to focus exclusively on inter-group variation and on data aggregated to the group level. This approach eliminates the nonindependence problem mentioned above, but has the drawback of ignoring the role of individual-level variables in shaping the outcome. Both approaches essentially collapse all variables to the same level and ignore the multilevel structure. A third approach is to define separate regressions for each group. This approach allows regression coefficients to differ from group to group, but does not examine how specific group-level properties may affect individual-level outcomes or interact with individual-level variables. In addition, it is not practical when dealing with large numbers of groups or small numbers of observations per group. A fourth approach is to include group membership in individual-level equations in the form of dummy variables (as well as the interactions of group dummy variables with individual-level predictors). This approach is analogous to fitting separate regressions for each group and does not allow examination of exactly what group characteristics may be important in explaining the outcome. In addition, this approach treats the groups as unrelated

and ignores the fact that groups may be drawn from a larger population of groups with things in common.

Multilevel analysis differs from the approaches outlined above in that (a) it allows the simultaneous examination of the effects of group-level and individual-level predictors, (b) the nonindependence of observations within groups is accounted for, (c) groups or contexts are not treated as unrelated, but are seen as coming from a larger population of groups (23), and (d) both interindividual and intergroup variation can be examined (as well as the contributions of individual-level and group-level variables to these variations) (91). Thus, multilevel analysis allows researchers to deal with the micro-level of individuals and the macro-level of groups or contexts simultaneously (23).

THE STATISTICAL MODEL

The statistical models referred to here as multilevel models (36, 56, 71) have appeared in different literature under a variety of names including hierarchical linear models (4), random-effects or random-coefficient models (17, 58, 66), and covariance components models (13). Several publications on multilevel modeling have appeared in the educational, sociological, geographical, and health-related literature over the past few years (4, 5, 19, 23, 36, 56, 77–79, 104). A brief summary of the statistical method is presented below.

For multilevel analysis involving two levels (e.g. individuals nested within groups), the model can be conceptualized as a two-stage system of equations in which the individual variation within each group is explained by an individual-level equation, and the variation across groups in the group-specific regression coefficients is explained by a group-level equation. The case for a normally distributed dependent variable is illustrated below. The illustration focuses on the case of only one independent variable at the individual and one independent variable at the group level (although models can of course be extended to include as many independent variables as needed).

In the first stage, a separate individual-level regression is defined for each group. In this first stage, the units of analysis are individuals.

$$Y_{ij} = b_{0j} + b_{1j}I_{ij} + \varepsilon_{ij} \quad \varepsilon_{ij} \sim N(0, \sigma^2), \quad 1.$$

where Y_{ij} = outcome variable for i th individual in j th group (or context) and I_{ij} = individual-level variable for i th individual in j th group (or context).

Individual-level errors (ε_{ij}) within each group are assumed to be independent and normally distributed with a mean of 0 and a variance of σ^2 . The same regressors are generally used in all groups, but regression coefficients (b_{0j} and b_{1j}) are allowed to vary from one group to another (hence the subscript j for these coefficients).

In a second stage, each of the group- or context-specific regression coefficients defined in Equation 1 (b_{0j} and b_{1j} in this example) is modeled as a function of

group-level variables. In this second stage, the units of analysis are groups.

$$b_{0j} = \gamma_{00} + \gamma_{01} C_j + U_{0j} \quad U_{0j} \sim N(0, \tau_{00}); \quad 2.$$

$$b_{1j} = \gamma_{10} + \gamma_{11} C_j + U_{1j} \quad U_{1j} \sim N(0, \tau_{11}); \quad 3.$$

$$\text{cov}(U_{0j}, U_{1j}) = \tau_{01},$$

where C_j is a group-level or contextual variable.

The errors in the group-level equations (U_{0j} and U_{1j}), sometimes called macro errors, are assumed to be normally distributed with mean 0 and variances τ_{00} and τ_{11} respectively. The macro error term U_{0j} measures the unique deviation of the intercept of each group from the overall intercept, γ_{00} , after accounting for the effect of C_j . Analogously, the macro error term U_{1j} represents the deviation of the slope within each group from the overall slope, γ_{10} , also after accounting for the effect of C_j . τ_{00} , and τ_{11} are the variances of the group intercepts and group slopes, respectively (after accounting for the group-level variable C_j). τ_{01} represents the covariance between intercepts and slopes; for example, if τ_{01} is positive, as the intercept increases the slope increases. Thus, multilevel analysis summarizes the distribution of the group-specific coefficients in terms of two parts—a fixed part that is unchanging across groups (γ_{00} and γ_{01} for the intercept and γ_{10} and γ_{11} for the slope) and a random part (U_{0j} for the intercept and U_{1j} for the slope) that is allowed to vary from group to group. Macro errors are assumed to be independent across contexts and independent of the individual-level errors (ε_{ij}) (71, 103). As can be seen in Equations 2 and 3, in multilevel analysis the group-to-group variability in individual-level regression coefficients is itself summarized and modeled (23).

By including an error term in the group-level equations (Equations 2 and 3), these models allow for sampling variability in the group-specific coefficients (b_{0j} and b_{1j}) and also for the fact that the group-level equations are not deterministic (i.e. the possibility that not all relevant group-level variables have been included in the model) (69, 102). The underlying assumption, from a frequentist perspective, is that group-specific intercepts and slopes are random samples from a normally distributed population of group-specific intercepts and slopes (or that groups are a random sample from a population of groups). Equivalently, from a Bayesian perspective, the macro errors are assumed to be exchangeable. This means that based on one's prior knowledge, one would be indifferent to the permutation over the j contexts of the U_{0j} s and U_{1j} s within Equations 2 and 3, respectively (i.e. one would be indifferent to arbitrary mixing-up of the macro errors) (44, 71). The exchangeability assumption implies that the residual variation in group-specific coefficients across groups is unsystematic (19).

An alternative way to present the model fitted in multilevel analysis is to substitute Equations 2 and 3 in Equation 1 to obtain:

$$Y_{ij} = \gamma_{00} + \gamma_{01} C_j + \gamma_{10} I_{ij} + \gamma_{11} C_j I_{ij} + U_{0j} + U_{1j} I_{ij} + \varepsilon_{ij}. \quad 4.$$

This final model is a random-effects model. The model includes the fixed effects of group-level variables (γ_{01}), individual-level variables (γ_{10}), and their interaction (γ_{11}) on the individual-level outcome Y_{ij} . It also includes a random intercept component (U_{0j}), and a random slope component (U_{1j}), which together with the individual-level errors (ε_{ij}) compose a complex error structure. As can be seen from the formula, the errors for observations within groups are correlated because U_{0j} and U_{1j} are common for observations within each group. In addition, the variance of the complex error is not constant because it depends on U_{0j} and U_{1j} , as well as on the value of I_{ij} . Thus the two assumptions of standard regression (independence and equal variance) are violated, and special estimation methods must be used. The parameters of the above equations (fixed effects, random group effects, variances of the random effects, and residual variance) are simultaneously estimated using iterative methods (4, 36, 43, 56).

Multilevel models allow investigation of a variety of interrelated research questions. The fixed-effects coefficients in Equation 4 can be used to estimate the independent effects of group-level variables (γ_{01}), individual-level variables (γ_{10}), and their interaction on individual-level outcomes. An alternative (and equivalent) interpretation of γ_{01} and γ_{11} derived from Equations 2 and 3 is that they represent the effects of the group-level variable on group-specific intercepts and slopes, respectively. In addition, the estimation of τ_{01} and τ_{11} and how they change as individual-level or group-level variables are added allows quantification of group-to-group variability and the degree to which it is statistically explained by characteristics of individuals and characteristics of groups. Thus, multilevel models allow separation of the effects of context (i.e. group characteristics) and of composition (characteristics of the individuals in groups): Do groups differ in average outcomes after controlling for the characteristics of individuals within them (e.g. does τ_{00} differ from 0)? Are group-level variables related to outcomes after controlling for individual-level variables? Multilevel models can also be used to examine whether the effects of individual-level variables differ across groups: Do individual-level associations vary from group to group, and is this partly a function of group-level variables (e.g. does τ_{11} differ from 0, and how does it change as group variables are added to Equation 3)? Do group-level variables modify the effects of individual-level variables? Multilevel models also allow quantification of variation at different levels: within group as summarized by σ^2 , and between group, as summarized by the variances of the random effects τ_{00} , τ_{11} .

The multilevel model described above admits many modifications. If, after accounting for the covariates in the model, there is little or no residual variability in intercepts or slopes across groups, the macro errors in Equations 2 and 3 will all be estimated as near zero, and the estimate of the variance of the random effects will approach 0 (43). Consequently, the random effects model in Equation 4 reduces to a standard regression model including both individual-level and group-level independent variables (a model with no random effects, in which all regression coefficients are modeled as fixed with no random component

at the group level). In this case individuals within groups can be considered independent, conditional on the individual-level and group-level variables in the model (i.e. there will be no residual correlation between individual-level outcomes within groups). On the other hand, the persistence of significant variation in intercepts or slopes after inclusion of group-level variables suggests that other group-level factors possibly responsible for this variation may need to be explored (assuming that all individual-level predictors that may result in differences between groups were already included in Equation 1). Hypotheses regarding whether macro-level error variances differ significantly from 0 (and consequently whether regression coefficients should be modeled as random) can be tested. It is also possible to model some coefficients as random and others as fixed (i.e. with no random component at the group level), by constraining the macro errors of some of the coefficients to be 0 (mixed-effects models). In addition, it is possible to model some coefficients as a function of group-level variables and others simply as random with no group-level predictors (no C_j in Equation 2 or 3). When the coefficient associated with an individual-level predictor (i.e. a slope) is modeled as a function of a group-level variable (as b_{1j} is above), an interaction term between the individual-level and the group-level variable (sometimes called a cross-level interaction) appears in the full model (as for γ_{11} above). Duncan et al 1998 (23) provide useful graphical summaries of different possibilities available in multilevel models in terms of the variation in intercepts and slopes, the covariance between both, and interactions between individual-level and group-level properties.

It is important to note that the power to estimate group-to-group variability and group-level effects is strongly dependent on the number of groups included in the analyses (56, 90). Therefore, the failure to observe significant group-to-group variability should not always be taken as an indication that groups can be ignored in the analyses, especially when dealing with small numbers of groups (or in cases involving binary dependent variables, for which current estimation methods may often underestimate random effects) (81). [Situations involving nested data structures with small numbers of "groups" are especially common, for example, in community intervention trials, where analytical strategies may need to take into account the fact that communities were the units randomized, even if no "statistically significant" community effects are detectable (54a).]

Although multilevel- or random-effects models were first developed for continuous dependent variables, analogous methods have been developed or are under development for binary outcomes, counts, multiple-category outcomes, and survival analysis (34, 36, 102). These models assume a specific distribution for the random part in the individual-level model (e.g. Equation 1 above is specified as a logistic model), while maintaining the normality assumption for the macro-level errors (23). In addition, the two-level model described above can be extended to allow three or more levels (e.g. multiple nested contexts) (4, 23, 36). Research on different statistical approaches to estimating the parameters of these models (particularly for the nonlinear case) is ongoing (20, 38, 73, 81).

ALTERNATIVES TO MULTILEVEL MODELS

Models in which individual-level dependent variables are explained in terms of both group-level and individual-level independent variables are not new. These types of models have also been called contextual models (2, 49). Although the terms contextual analysis and multilevel analysis have often been used synonymously (44, 95), today's multilevel models are more general than early contextual models. Early contextual models were simply models in which group-level predictors were included in standard regressions with individuals as the units of analysis. This is equivalent to a multilevel model in which no errors are allowed for in the group-level equations (e.g. all U_{0j} and U_{1j} are forced to be 0); hence, early contextual models were fixed-effects models. The inclusion of errors in the group-level equations has several potential advantages. By quantifying the variability in macro errors (which is summarized in τ_{00} and τ_{11}), multilevel models allow estimation of group-to-group variability in the group-specific regression coefficients and how it changes as individual- and group-level variables are added. In addition, the inclusion of these errors allows for the possibility that dependent variables for individuals within groups may be correlated even after accounting for the individual-level variables and group-level variables in the model. One reason for this correlation may have to do with the omission of important group-level variables that individuals within groups share. By taking into account this residual correlation, multilevel models correctly estimate standard errors associated with the regression coefficients. In addition, the allowance for errors in the group-level equations may be particularly appropriate if groups can be thought of as a sample of a larger population of groups (e.g. schools, neighborhoods, etc) about which inferences want to be made.

If interest centers in estimating the fixed effects of group-level and individual-level variables on an individual-level outcome (rather than group-to-group variability in coefficients), multilevel models may not always be necessary (6, 12). If the variances of the random effects are estimated as 0 (or, analogously, if there is no residual correlation between individuals within groups after accounting for the variables in the model), a fixed-effects model including relevant group-level and individual-level variables (as in early contextual models) may be an adequate and simpler formulation. Moreover, if residual correlation is indeed present, fixed effects contextual models can be modified to account for nonindependence of outcomes within groups. One option is to account for the correlation between individuals within groups by means of marginal models using estimation methods of the generalized-estimating equation (GEE) (107). Marginal models essentially involve the simultaneous estimation of two equations: one for the dependent variable, which includes all covariates thought to affect the outcome, and another for the correlations between outcomes. Whereas random-effects models model the dependent variable conditional on the random effects, marginal models (as their name indicates) model the marginal expectation of the dependent variables across the population (in a sense, averaged across the random effects). For this reason,

marginal models have also been termed population-average models, in contrast with unit-specific (or subject-specific for cases of repeat observations on individuals over time) random-effects models (6, 17, 107). For data structures involving individuals nested within groups, marginal models only describe the covariance among persons within a context or group. They model the population-averaged response as a function of covariates, without explicitly accounting for context-to-context heterogeneity (107). In contrast, random-effects models explain the source of context-to-context variation by modeling group-specific regression coefficients as a function of group-level variables as described in Equations 2 and 3 above (107). Differences between both types of models have consequences for the interpretation of regression coefficients. In the random effects model, the regression coefficient estimates how the response changes as a function of covariates conditional on the random effects; in the marginal model, the coefficient expresses how the response changes as a function of covariates averaged over the random effects (17, 107). For linear models (continuous dependent variables), these coefficients are mathematically equivalent, but in the nonlinear case (e.g. logistic models), the marginal parameter values will usually be smaller in absolute value than their random effects analogs (6, 107). In addition to the generalized-estimating equation approach, other methods routinely used in the analysis of clustered survey data to account for residual correlation can also be used to obtain correct standard errors. However, the generalized-estimating equation marginal models and other approaches mentioned above differ from multilevel modeling in that, although the correlation between outcomes within groups is appropriately accounted for, the source of this correlation is not directly investigated (the correlation and sometimes higher level effects themselves are viewed as nuisance parameters which must be taken into account but are not of direct interest). Therefore, these approaches do not allow examination of group-to-group variation, of the group-level or individual-level variables potentially related to it, or of the degree of variation present between and within groups, as multilevel models do.

OTHER APPLICATIONS OF MULTILEVEL MODELS

In addition to individuals nested within groups, multilevel models can be used in many other situations with nested sources of random variability. For example, multilevel models can be used for longitudinal data analysis (repeat observations nested within individuals over time) (33, 37, 36, 68, 82, 99), for multivariate responses (multiple outcomes nested within individuals) (22, 23), in the analysis of repeat cross-sectional surveys (multiple observations nested within time periods) (18), and in the examination of geographic variations in rates (rates for smaller areas nested within regions or larger areas) (10, 60, 61). Other applications of multilevel analysis include the examination of interviewer effects (respondents nested within interviewers) (45) and meta-analysis (individuals nested within studies) (4, 42).

Another important application of multilevel models is their use to obtain improved estimates of parameters for a given group (for example, estimates of within-group regression coefficients or rates for a particular group) by combining information from the group itself with information from all other groups investigated. This is particularly useful when estimating parameters for a group with few within-group observations. Multilevel methods allow researchers to obtain better estimates for a given group by borrowing information from other groups. For example, suppose researchers are interested in estimating the relationship between education and contraceptive use separately for each of a series of countries, but some countries have very few observations, making an estimate based exclusively on their own data unreliable. Three hypothetical possibilities are available to estimate the regression coefficient for a given country (19): (a) Use only data from that country; (b) use the macro model shown in Equations 2 and 3 by plugging in the appropriate values for the country variables (this option uses data from all countries, not just for the country for which estimates are desired); and (c) combine *a* and *b* into an optimally weighted average. These optimally weighted averages are termed empirical Bayes estimates. The weighted average shrinks the within-group estimate (*a*) toward the between-group estimate (*b*). The less precise the within-group estimate and the less the variability observed across groups, the greater the shrinkage. Thus, the estimate for a given group is based not only on its own data but also takes into account the effects for other groups and the characteristics groups share (78). This method [which is similar to a method proposed originally by Stein and others (reviewed in 27)] can be used, for example, to derive shrunken estimates of rates of death or diseases for small areas with few observations (7, 59) or to estimate rates of different health outcomes for individual providers (hospitals, physicians, etc) (94). The assumptions, strengths, and limitations of empirical Bayes estimates are reviewed elsewhere (4, 19, 39, 56, 70). In other applications (which do not involve the structure of individuals within groups described here, although they are directly analogous to it), empirical Bayes estimates of regression coefficients have been used to obtain improved estimates of associations in studies investigating the role of multiple exposures (101).

EXAMPLES OF EMPIRICAL APPLICATIONS IN PUBLIC HEALTH INVOLVING INDIVIDUALS NESTED WITHIN GROUPS OR CONTEXTS

Over the past few years, multilevel models have been used in public health to examine the independent and interacting effects of group-level and individual-level factors on health outcomes. The groups or contexts investigated using multilevel analysis have included countries, states, regions, neighborhoods or communities, schools, families, workplaces, and health care providers (see for example 24, 25, 29, 32, 43, 87, 92, 100). Multilevel analysis has been used in demography (e.g. 29, 44), health services research (e.g. 32, 51, 64, 65, 79, 87,), evaluation of

interventions (31, 43), and the study of addictions (98, 100). Multilevel models have also been used increasingly in the investigation of the social determinants of health. Within this field, one of the main research areas in which multilevel models have been applied is the investigation of the effects of neighborhood social environments on health outcomes. [Recently, fixed-effects contextual models have also been used to examine the relation between income inequality and health (e.g. 11, 30, 53), but this literature is not reviewed here.] This research has been stimulated by theoretical arguments and empirical studies suggesting that neighborhoods may differ in many aspects potentially related to health (52, 67). A key issue in investigating neighborhood effects on health is separating out the effects of neighborhood characteristics (context) from the effects of individual-level attributes that persons living in certain types of areas may share (composition). Because neighborhoods can be thought of as groups or contexts with individuals nested within them, multilevel models have been used to investigate how neighborhood factors, individual-level factors, and their interactions influence health. Using examples drawn from this field, this section illustrates some of the capabilities of multilevel analysis, as well as the challenges raised by its use.

One objective of the use of multilevel analysis in the investigation of neighborhood effects has been to simultaneously examine between-neighborhood and within-neighborhood variability in outcomes and the degree to which between-neighborhood variability is accounted for by neighborhood-level and individual-level variables. As shown in Equations 2 and 3, multilevel models can allow both neighborhood-specific (equivalent to group-specific) intercepts and neighborhood-specific slopes to vary across neighborhoods. In the simplest case, researchers can define a model at the individual level (as in Equation 1), including all relevant individual-level predictors, and then allow the intercept to vary randomly across neighborhoods, as in Equation 2 (while modeling the slopes as fixed). The presence of significant variability in intercepts across neighborhoods (as evidenced by the value of τ_{00}) suggests that neighborhood-level factors (or, alternatively, omitted individual-level factors closely associated with neighborhoods) may be related to average outcomes for neighborhoods. For linear models (continuous dependent variables) with a random intercept, it is possible to estimate the percent of the total variation in the outcomes that is between groups (i.e. the intraclass correlation coefficient) by estimating the ratio of τ_{00} to $(\tau_{00} + \sigma^2)$. [The estimation of the intraclass correlation coefficient is not straightforward in linear models with random slopes or in the nonlinear case (4, 91).] Several studies have documented statistically significant variability across neighborhoods (or areas), which persists after accounting for differences in the social class of residents (3, 21, 24, 26, 40, 41, 48, 50, 86), although the percent of total variability between neighborhoods has generally been small (3, 26, 41, 50). By examining changes in the value of τ_{00} as additional individual-level variables are added to Equation 1 or neighborhood variables are added to Equation 2, researchers can examine the degree to which neighborhood-to-neighborhood variability is accounted for by omitted individual-level or neighborhood-level variables. For

example, some studies have found that neighborhood characteristics (such as deprivation) partly explain the neighborhood-to-neighborhood variations observed, as evidenced by the reduction in τ_{00} when neighborhood deprivation is added to the equation for the neighborhood-specific intercepts (although statistically significant interneighborhood variation sometimes persists) (24, 26, 48, 86). On the other hand, the presence of significant reductions in τ_{00} as additional individual-level variables are added to Equation 1 would suggest that between-neighborhood differences may be at least partly attributable to differences in the characteristics of individuals (i.e. differences in neighborhood composition) (see e.g. reference 3). In more complex models, the effects of individual-level predictors (slopes) can also be allowed to vary across neighborhoods (as in Equation 3), and the contribution of neighborhood factors to this variability can be examined. For example, does the effect of individual-level education on the outcome differ by neighborhoods (i.e. if b_1 is the slope associated with individual-level education, does τ_{11} differ significantly from 0)? Is this variability partly a function of neighborhood deprivation (how much does τ_{11} change when deprivation is added to Equation 3)? This ability to simultaneously examine and model within-neighborhood and between-neighborhood variability is a unique characteristic of the multilevel-modeling approach.

Another related objective of the use of multilevel analysis in the investigation of neighborhood effects has been to estimate associations of neighborhood characteristics with individual-level outcomes after adjustment for individual-level confounders. Thus, for example, neighborhood characteristics such as deprivation or other indicators of socioeconomic context have been found to be associated with adverse health outcomes after accounting for individual-level indicators of social class (16, 15, 24, 26, 48, 50, 54, 72, 75, 86, 106) (although associations with outcomes are generally stronger for individual-level indicators of social position than for characteristics of neighborhood environments). The estimates of interest (neighborhood effect, individual effect, and neighborhood-individual interaction) can be obtained from the corresponding coefficients shown in Equation 3 above (γ_{01} , γ_{10} , γ_{11}). In deriving these estimates, multilevel models are used chiefly as a way to account for residual correlation between outcomes within neighborhoods, an objective that can also be achieved using other statistical approaches, as previously discussed (e.g. 1, 74, 80, 97). [The effects of neighborhood context have also been investigated using fixed-effects contextual models, by including characteristics of neighborhoods in standard regressions (e.g. 63, 88, 89, 105). In the absence of residual correlation between individuals within neighborhoods, these methods yield correct estimates of standard errors.]

The advent of multilevel analysis as a statistical tool has undoubtedly stimulated research into the neighborhood determinants of health. However, its use also highlights some of the challenges faced by researchers interested in multilevel analysis. Results of research on neighborhood effects and health using multilevel models have been rather mixed. Although researchers have sometimes found significant variation in outcomes across neighborhoods, the percent of total

variation between neighborhoods (in cases where this can be assessed) has been small (3, 26, 41, 50). In addition, neighborhood-level factors have sometimes been found to be significantly associated with outcomes even in the absence of significant between-neighborhood variability as assessed by multilevel models. These results raise methodological questions related to whether the studies performed so far have the power to detect between-neighborhood variability, given the number of neighborhoods, the average number of persons per neighborhood, and the estimation methods commonly used (particularly in the nonlinear case) (81). In addition, data limitations have forced researchers to focus on existing data sources to define and characterize neighborhoods, resulting in the use of very crude and indirect proxies for neighborhoods and neighborhood variables. But perhaps most importantly, results so far highlight the need to develop theories that articulate how neighborhood factors may affect health. This implies specifying how neighborhoods should be defined, what things about neighborhoods may be relevant for specific health outcomes, and the mechanisms through which these effects may operate. The development of such models will allow researchers to test more specific hypotheses, using more appropriate study designs. Theories on neighborhood effects on health may also need to place the process of residential differentiation within a broader theory of social organization and social stratification. For example, how do neighborhood differences reinforce and how are they in turn reinforced by individual social class differences? A better understanding of neighborhood effects may require examination of how neighborhood-level and individual-level factors are reciprocally interrelated, and thinking of processes and mediators in addition to the estimation of independent effects. This implies moving beyond the explanation of residual variance across neighborhoods, once individual-level social class indicators are accounted for (which has been the focus of much of our research on neighborhood effects to date). These challenges, which are related more broadly to our conceptual thinking of how groups and individuals are interrelated, influence each other, and may jointly influence health, are common to multilevel analysis generally. They will be discussed in more detail below.

CHALLENGES TO MULTILEVEL ANALYSIS

Multilevel Theories of Disease Causation

In his 1984 review, Blalock (2) described many of the theoretical and methodological challenges facing contextual analysis. Despite the methodological sophistication of multilevel models, many of these challenges are still valid today. Perhaps chief among these is the need to develop theories that specify how group-level and individual-level factors may jointly shape the distribution of health and disease, theories that can be operationalized and tested. An example of the use of multilevel analysis in the context of a theoretical model that specifies how neighborhood attributes may be related to violent crime is provided by Sampson et al (83). Based on their underlying model, Sampson and collaborators conceptualized the relevant

neighborhood-level attributes and developed operational measures of them. Multilevel analysis was then used as a statistical tool to examine aspects of the model in different ways. An important challenge to public health researchers is to develop substantive explanations and move beyond the use of multilevel analysis simply to document and statistically explain residual variability across groups after accounting for individual-level variables. In the absence of this, multilevel analysis runs the risk of being reduced to a method that examines variation across meaningless groups or associations with meaningless group-level variables and of either not finding much or finding patterns that are difficult to understand.

Groups and Group Properties

A crucial component of multilevel analysis closely related to the theoretical model being tested is the definition of the relevant group and of the relevant group-level variables. The groups that are (or should be) investigated in multilevel analysis are not arbitrary or convenient groupings of individuals, but rather groups that are hypothesized to be meaningful in explaining the outcome. The increasing sophistication of multilevel models now allows them to accommodate multiple or overlapping contexts (e.g. 35), but the more substantial issues of defining relevant contexts, specifying the relevant group-level variables, and collecting the necessary data remain a challenge.

A key component of the rationale for multilevel analysis is the notion of emergent group properties, the idea that group-level variables may provide information that is not captured by individual level data. Several different types of group-level variables [including both derived and integral variables, as reviewed elsewhere (14, 62, 95, 96)] may be used in multilevel analysis. The key issue is that group-level variables are used as measures of relevant group-level constructs (rather than as proxies for unavailable individual-level data). For example, the construct of neighborhood unemployment is distinct from individual-level unemployment, and both may be important to health. Similarly, inequality in the distribution of income within a group measures a different construct than individual-level income. In formulating the conceptual model for a research project it is particularly important to identify the constructs of interest and the level at which they are defined and measured. Sometimes the distinction between group-level and individual-level constructs (or variables) is clear-cut, but other times it may be complex. For example, individual-level variables can be used to categorize people into groups, such as age groups; however, age itself remains an individual-level attribute. Of course, it is possible that age groups themselves may have emergent group-level properties (related for example, to the types and patterns of interactions between individuals), which may be related to the outcome being studied. Another perhaps more subtle issue is that many variables measured at the individual-level (such as individual social class or race/ethnicity) may only be meaningfully understood in the context of how individuals are related to each other in groups or societies. Their meaning and implications for health are tightly intertwined with (and dependent on) group-level attributes. The fact that the appropriate unit for measuring a given

characteristic is the individual does not imply that it is individually-determined (as opposed to determined by characteristics of the social organization of the group to which that individual belongs). In fact, many hypotheses regarding the social determinants of disease can and should be tested with individual-level data. In some cases it may still be relevant to add another level to the analysis, for example, to examine the effects of race/ethnicity (a socially defined construct measured at the individual level) across groups with different types of social organizations, social norms, or public policies related to race/ethnic discrimination (group-level variables) (see e.g. 103). Another example in which group-level and individual-level attributes are tightly intertwined is provided by patterns of interactions between individuals within groups. Patterns of relationships or interactions between individuals may be important in understanding both group-level and individual-level health outcomes (55, 55a). In some cases these patterns of interactions can be summarized in the form of a group-level attribute (for example, network size or structure), which may affect all individuals within a group, in which case they can be thought of as true group-level properties (62, 95). In other cases, they may pertain to smaller groups within a larger group, or may even vary from individual to individual depending on individual contact patterns (in which case they can be thought of as individual-level variables that depend on how individuals are related to each other in groups). The specification of relevant constructs and the levels at which they are defined and measured is a key challenge to multilevel analysis.

Separating Out Independent Effects

As do all models, multilevel models necessarily simplify complex processes. One limitation in this regard, which multilevel analysis shares with other regression methods, is its focus on teasing apart independent effects. The extent to which the group- and individual-level effects can be meaningfully separated depends on the model that hypothetically links them. One of the critiques leveled at multilevel analysis has been that group-level effects may simply reflect unaccounted for (or mismeasured) individual-level predictors or, more generally, misspecification of the individual-level model (Equation 1) (14). However, the degree to which it makes sense to control for individual-level attributes in examining group effects depends on whether the individual-level variable is conceptualized as a true confounder or a mediator. In addition, group effects do not operate magically: If group attributes affect health they must get into the body and therefore are necessarily mediated through individual-level processes. Strictly speaking, therefore, group-level attributes cannot affect individuals independently of all individual-level attributes, but this does not imply that group-level variables are reducible to individual-level variables.

In addition, multilevel models generally do not allow examination of the full range of complex and reciprocal interrelationships between variables (2). They still posit a relatively simple regression structure in which a single variable depends on a number of other variables (4). For example, multilevel models do not model the possibility that individual-level properties (or individual-level relations

between variables) may influence group characteristics (19, 69, 76) and, vice versa, that group characteristics may shape individual-level independent variables. This may be more or less relevant depending on the problem investigated. Entwisle (28) provides an example from the demography and fertility literature. It has been argued that the implementation of family planning programs is often stimulated by the initiation of fertility decline in some groups of the population [leading, for example, to increased differentials at the individual level by socioeconomic status (SES)]. If this is so, then the individual-level coefficient for differential SES fertility behavior within a group at time 1 could be used to predict a group variable such as the existence of a family planning program at time 2. Modifications to multilevel models to account for these types of relationships have been proposed (28, 69). An additional complexity arises from the fact that individuals may self-select themselves into certain groups based on unmeasured individual characteristics, making the interpretation of any contextual effects that are observed problematic. The degree to which self-selection operates may differ for different types of contexts (e.g. groups of friends vs communities) and may also differ from setting to setting (2).

Model Complexity

Many of the advantages of multilevel models over traditional methods are at the expense of greater model complexity. More complicated models may be closer to reality but testing model fit and examination of model assumptions is more difficult (19, 73). If the model is true, multilevel estimates are less biased and more efficient than those obtained using other methods; however, models are less parsimonious and need larger data sets, and estimation becomes complicated (12, 56). Sample size and power calculations for multilevel hypotheses testing are particularly complex (8, 56, 90). Power, for example, depends both on the number of groups and on the number of individuals per group (56, 90). The centering of explanatory variables also raises more complicated issues than it does in traditional regression models (4, 57, 56), as does the estimation of variance explained at different levels and by different variables (56, 91), particularly for models with many random coefficients and for nonlinear models. Several authors have warned against the rapid incorporation of complex multilevel models before their performance is adequately understood and evaluated and especially when it is done with little regard to the adequacy of the data and the inferences that can be drawn from it (12, 20, 73).

CONCLUSION

Multilevel analysis has many features that may be of use in public health research. By explicitly acknowledging the existence of groups, modeling group-to-group variation simultaneously with individual-to-individual variation, and including group-level properties with individual-level variables in the analyses, multilevel

models allow for the importance of both groups and individuals in understanding health outcomes. Multilevel models can be used to draw individual-level inferences, but inferences can also be made regarding group-to-group variation, including whether it exists in the data, and the extent to which it is accounted for by group- and individual-level characteristics. In a sense, multilevel analysis provides one way to link the traditionally distinct ecological- and individual-level studies and to overcome the limitations inherent in focusing only at one level. However, it is also true that, for many research questions, multilevel analysis may not be necessary. Relevant questions may often be formulated and answered within a single level.

As the term “multilevel analysis” appears more and more often in the public health literature, it is important to distinguish the statistical method of multilevel analysis from the more general issue of thinking of multilevel conceptual models or hypotheses, which may be addressed using many different types of methods. The term “multilevel,” broadly used, refers to more than the hierarchical data structures of observations nested within groups, addressed by the multilevel models described here. More broadly, the term multilevel has been used to refer to the multiplicity of qualitatively different levels (e.g. society, groups, individuals, organ systems, cells, and genes) that are important in understanding health and disease (85). The multilevel-analysis approach described above may be useful in understanding some aspects of the multilevel continuum (the aspects that fit the data structures that multilevel analysis can handle) but not others, which may need to be addressed using other types of methods. Like other statistical methods, multilevel analysis will help describe, summarize, and quantify patterns present in the data. But it will not explain these patterns; explanation will emerge from the reciprocal interplay between theory formulation and empirical testing. If the advent of multilevel analysis as a statistical tool helps stimulate our thinking on how factors at multiple levels are important to understanding health and disease (as well as contributing to the empirical examination of aspects of these theories), then it will have accomplished very much indeed.

ACKNOWLEDGMENTS

The author thanks Bruce Link and Sharon Schwartz for helpful comments on an earlier version. This work was supported by R29 HL59386-01 (National Heart, Lung and Blood Institute of the National Institutes of Health).

Visit the Annual Reviews home page at www.AnnualReviews.org

LITERATURE CITED

1. Anderson R. 1997. Mortality effects of community socioeconomic status. *Epidemiology* 8:42–47
2. Blalock HM. 1984. Contextual-effects models: theoretical and methodological issues. *Annu. Rev. Sociol.* 10:353–72
3. Boyle MH, Willms JD. 1999. Place effects for areas defined by administrative

- boundaries. *Am. J. Epidemiol.* 149:577–85
4. Bryk AS, Raudenbush SW. 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, CA: Sage. 265 pp.
 5. Bullen N, Jones K, Duncan C. 1997. Modelling complexity: analysing between individual and between-place variation—a multilevel tutorial. *Environ. Plan.* 29:585–609
 6. Burton P, Gurrin L, Sly P. 1998. Extending the simple linear regression model for correlated responses: an introduction to generalized estimating equations and multi-level mixed modeling. *Stat. Med.* 17: 1261–91
 7. Clayton D, Kaldor J. 1987. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* 43:671–81
 8. Cohen MP. 1998. Sample size in hierarchical linear models. *J. Off. Stat.* 14:267–76
 9. Coleman J. 1991. Social theory, social research and a theory of action. *Am. J. Sociol.* 91:1309–35
 10. Congdon P. 1995. The impact of area context on long term illness and premature mortality: an illustration of multi-level analysis. *Reg. Stud.* 29:327–44
 11. Daly M, Duncan G, Kaplan G, Lynch J. 1998. Macro to micro links in the relation between income inequality and mortality. *Milbank Q.* 76:315–39
 12. deLeeuw J, Kreft IGG. 1995. Questioning multilevel models. *J. Educ. Behav. Stat.* 20:171–89
 13. Dempster AP, Rubin DB, Tsutakawa RK. 1981. Estimation in covariance components models. *J. Am. Stat. Assoc.* 76:341–56
 14. Diez-Roux AV. 1998. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *Am. J. Public Health* 88:216–22
 15. Diez Roux AV, Nieto FJ, Caulfield L, Tyroler HA, Watson RA, Szklo M. 1999. Neighbourhood differences in diet: the Atherosclerosis Risk in Communities (ARIC) study. *J. Epidemiol. Community Health* 53:55–63
 16. Diez-Roux A, Nieto F, Muntaner C, Tyroler HA, Comstock GW, et al. 1997. Neighborhood environments and coronary heart disease: a multilevel analysis. *Am. J. Epidemiol.* 146:48–63
 17. Diggle PJ, Liang KY, Zeger SL. 1994. *Analysis of Longitudinal Data*. New York: Oxford Univ. Press. 253 pp.
 18. DiPrete T, Grusky D. 1990. The multilevel analysis of trends with repeated cross-sectional data. *Sociol. Methodol.* 20:337–68
 19. DiPrete TA, Forristal JD. 1994. Multilevel models: methods and substance. *Annu. Rev. Sociol.* 20:331–57
 20. Draper D. 1995. Inference and hierarchical modeling in the social sciences. *J. Educ. Behav. Stat.* 20:115–48
 21. Duncan C, Jones K, Moon G. 1993. Do places matter? A multi-level analysis of regional variation in health-related behavior in Britain. *Soc. Sci. Med.* 37:725–33
 22. Duncan C, Jones K, Moon G. 1996. Health-related behaviour in context: a multi-level modelling approach. *Soc. Sci. Med.* 42:817–30
 23. Duncan C, Jones K, Moon G. 1998. Context, composition, and heterogeneity: using multilevel models in health research. *Soc. Sci. Med.* 46:97–117
 24. Duncan C, Jones K, Moon G. 1999. Smoking and deprivation: Are there neighbourhood effects? *Soc. Sci. Med.* 48:497–505
 25. Duncan T, Duncan S, Hops H, Alpert A. 1997. Multi-level covariance structure analysis of intra-familial substance use. *Drug Alcohol Depend.* 46:167–80
 26. Ecob R. 1996. A multilevel modelling approach to examining the effects of area of residence on health. *J. R. Stat. Soc. A.* 159:61–75
 27. Efron B, Morris C. 1977. Stein's paradox in statistics. *Sci. Am.* 236:119–27
 28. Entwisle B. 1991. Micro-macro theoretical linkages in social demography: a

- commentary. In *Macro-micro Linkages in Sociology*, ed. J Huber, pp. 280–86. Newbury Park, CA: Sage
29. Entwisle B, Mason W, Hermalin H. 1986. The multilevel dependence of contraceptive use on socioeconomic development and family planning program strength. *Demography* 23:199–216
 30. Fiscella K, Franks P. 1997. Poverty or income inequality as predictor of mortality: a longitudinal cohort study. *Br. J. Med.* 314:1724–28
 31. Forster J, Murray D, Wolfson M, Blaine T, Wagenaar A, Hennrikus D. 1998. The effects of community policies to reduce youth access to tobacco. *Am. J. Public Health* 88:1193–98
 32. Gatsonis C, Normand S, Liu C, Morris C. 1993. Geographic variation of procedure utilization: a hierarchical model approach. *Med. Care* 31:YS54–59
 33. Gibbons R, Hedeker D, Waternaux C, Davis J. 1988. Random regression models: a comprehensive approach to the analysis of longitudinal psychiatric data. *Psychopharmacol. Bull.* 24:438–43
 34. Goldstein H. 1991. Nonlinear multilevel models, with an application to discrete response data. *Biometrika* 78:45–51
 35. Goldstein H. 1994. Multilevel cross-classified models. *Sociol. Methods Res.* 22:364–75
 36. Goldstein H. 1995. *Multilevel Statistical Models*. New York: Halsted. 178 pp.
 37. Goldstein H, Healy M, Rasbash J. 1994. Multilevel time series models with applications to repeated measures data. *Stat. Med.* 13:1643–55
 38. Goldstein H, Rasbash J. 1996. Improved approximations for multilevel models with binary responses. *J. R. Stat. Soc. A.* 159: 505–13
 39. Goldstein H, Spiegelhalter D. 1996. League tables and their limitations: statistical issues in comparisons of institutional performance. *J. R. Stat. Soc. A.* 159:385–443
 40. Gould M, Jones K. 1996. Analyzing perceived limiting long-term illness using U.K. census microdata. *Soc. Sci. Med.* 42:857–69
 41. Hart C, Ecob R, Davey Smith G. 1997. People, places, and coronary heart disease risk factors: a multilevel analysis of the Scottish heart health study archive. *Soc. Sci. Med.* 45:893–902
 42. Hedeker D, Gibbons R, Davis J. 1991. Random regression models for multicenter clinical trials data. *Psychopharmacol. Bull.* 27:73–77
 43. Hedeker D, McMahan S, Jason L, Salina D. 1994. Analysis of clustered data in community psychology: with an example from a worksite smoking cessation project. *Am. J. Community Psychol.* 22:595–615
 44. Hermalin A. 1986. The multilevel approach: theory and concepts. The methodology for measuring the impact of family planning programs on fertility. Addendum Manual IX *Popul. Stud.* 66:15–31
 45. Hox JP, de Leeuw ED, Kreft IGG. 1991. The effect of interviewer and respondent characteristics on the quality of survey data: a multilevel model. In *Measurement Errors in Surveys*, ed. PP Biemer, LE Lyberg, NA Mathiowetz, S Sudman. New York: Wiley & Sons
 46. Hox JP, Kreft IGG. 1994. Multilevel analysis methods. *Sociol. Methods Res.* 22:283–99
 47. Huber J. 1991. Introduction. In *Macro-micro Linkages in Sociology*, ed. J Huber, pp. 6–9. Newbury Park, CA: Sage
 48. Humphreys K, Carr-Hill R. 1991. Area variations in health outcomes: artefact or ecology. *Int. J. Epidemiol.* 20:251–58
 49. Iversen G. 1991. *Contextual Analysis*. Newbury Park, CA: Sage. 84 pp.
 50. Jones K, Duncan C. 1995. Individuals and their ecologies: analysing the geography of chronic illness within a multilevel modelling framework. *Health Place.* 1:27–40
 51. Jones K, Moon G. 1991. Multilevel assessment of immunisation uptake as a

- performance measure in general practice. *Br. Med. J.* 303:28–31
52. Kaplan G. 1996. People and places: contrasting perspectives on the association between social class and health. *Int. J. Health Serv.* 26:507–19
 53. Kennedy B, Kawach I, Glass R, Prothrow-Stith D. 1998. Income distribution, socioeconomic status and self rated health in the United States: a multilevel analysis. *Br. Med. J.* 31:917–21
 54. Kleinschmidt I, Hills M, Elliott P. 1995. Smoking behavior can be predicted by neighborhood deprivation measures. *J. Epidemiol. Community Health* 49:S72–77
 - 54a. Koepsell TD, Diehr PH, Cheadle A, Kristal A. 1995. Invited commentary: symposium on community intervention trials. *Am. J. Epidemiol.* 142:594–99
 55. Koopman J, Longini I, Jacquez J, et al. 1991. Assessing risk factors for transmission of infection. *Am. J. Epidemiol.* 133:1199–209
 - 55a. Koopman JS, Lynch J. 1999. Individual causal models and population system models in epidemiology. *Am. J. Public Health* 89:1170–74
 56. Kreft I, deLeeuw J. 1998. *Introducing Multilevel Modeling*. London: Sage. 149 pp.
 57. Kreft I, deLeeuw J, Aiken LS. 1995. The effect of different forms of centering in hierarchical linear models. *Multivariate Behav. Res.* 30:1–20
 58. Laird NM, Ware H. 1982. Random effects models for longitudinal data. *Biometrics* 38:963–74
 59. Langford IH. 1994. Using empirical Bayes estimates in the geographical analysis of disease risk. *Area* 26:142–49
 60. Langford IH, Bentham G. 1996. regional variations in mortality rates in England and Wales: an analysis using multilevel modelling. *Soc. Sci. Med.* 42:897–908
 61. Langford I, Bentham G, McDonald A. 1998. Multi-level modelling of geographically aggregated health data: a case study of malignant melanoma mortality and uv exposure in the European community. *Stat. Med.* 17:41–57
 62. Lazarsfeld P, Menzel H. 1971. On the relation between individual and collective properties. In *A Sociological Reader on Complex Organizations*, ed. A Etzioni, pp. 499–516. New York: Holt, Rinehart, & Winston
 63. Leclere F, Rogers R, Peters K. 1998. Neighborhood social context and racial differences in women's heart disease mortality. *J. Health Soc. Behav.* 39:91–107
 64. Leung K, Elashoff R, Rees K, Hasan M, Legorreta A. 1998. Hospital- and patient-related characteristics determining maternity length of stay: a hierarchical linear model approach. *Am. J. Public Health* 88:377–81
 65. Leyland A, Boddy F. 1998. League tables and acute myocardial infarction. *Lancet* 351:555–58
 66. Longford NT. 1993. *Random Coefficient Models*. Oxford, UK: Clarendon. 270 pp.
 67. Macintyre S, McIver S, Sodman A. 1993. Area, class, and health: should we be focusing on places or people? *J. Soc. Pol.* 22:213–34
 68. Manor O, Kark J. 1996. A comparative study of four methods for analyzing repeated measures data. *Stat. Med.* 15:1143–59
 69. Mason W. 1991. Problems in quantitative comparative analysis: ugly ducklings are to swans as ugly scatter plots are to...? In *Macro-micro Linkages in Sociology*, ed. J Huber, pp. 231–43. Newbury Park, CA: Sage
 70. Mason W. 1995. Comment. *J. Educ. Behav. Stat.* 20:221–27
 71. Mason W, Wong G, Entwisle B. 1983. Contextual analysis through the multilevel linear model. In *Sociological Methodology 1983–1984*, ed. S Leinhardt, pp. 72–103. San Francisco: Jossey-Bass

72. Matteson D, Burr J, Marshall J. 1998. Infant mortality: a multi-level analysis of individual and community risk factors. *Soc. Sci. Med.* 47:1841–54
73. Morris C. 1995. Hierarchical models in educational data: an overview. *J. Educ. Behav. Stat.* 20:190–200
74. O'Campo P, Gielen A, Faden R, Xue X, Kass N, Wang M. 1995. Violence by male partners against women during the child-bearing year: a contextual analysis. *Am. J. Public Health* 1092–97
75. O'Campo P, Xue X, Wang MC, McCaughy M. 1997. Neighborhood risk factors for low birthweight in Baltimore: a multilevel analysis. *Am. J. Public Health* 87:1113–18
76. Oppenheimer MK. 1991. Multilevel analysis in the study of social institutions and demographic change. In *Macro-micro Linkages in Sociology*, ed. J Huber, pp. 223–30. Newbury Park, CA: Sage
77. Paterson L, Goldstein H. 1991. New statistical methods for analysing social structures: an introduction to multilevel models. *Br. Educ. Res. J.* 17:387–93
78. Rice N, Jones A. 1997. Multilevel models and health economics. *Health Econ.* 6:561–75
79. Rice N, Leyland A. 1996. Multilevel models: applications to health data. *J. Health Serv. Res. Policy* 1:154–64
80. Robert S. 1998. Community-level socioeconomic status effects on adult health. *J. Health Soc. Behav.* 39:18–37
81. Rodriguez G, Goldman N. 1995. An assessment of estimation procedures for multilevel models with binary responses. *J. R. Stat. Soc. A.* 158:73–89
82. Rutter C, Elashoff R. 1994. Analysis of longitudinal data: random coefficient regression modeling. *Stat. Med.* 13:1211–31
83. Sampson R, Raudenbush S, Earls F. 1997. Neighborhoods and violent crime: a multilevel study of collective efficacy. *Science* 277:918–24
84. Schwartz S. 1994. The fallacy of the ecological fallacy: the potential misuse of a concept and its consequences. *Am. J. Public Health* 84:819–24
85. Schwartz S, Susser E, Susser M. 1999. A future for epidemiology? *Annu. Rev. Public Health* 20:1–19
86. Shouls S, Congdon P, Curtis S. 1996. Modeling inequality in reported long term illness in the UK: combining individual and area characteristics. *J. Epidemiol. Community Health* 50:366–76
87. Sixma HJ, Spreeuwenberg PM, van der Pasch MA. 1998. Patient satisfaction with the general practitioner: a two-level analysis. *Med. Care* 36:212–29
88. Sloggett A, Joshi H. 1994. Higher mortality in deprived areas: community or personal disadvantage. *Br. Med. J.* 309:1470–74
89. Sloggett A, Joshi H. 1998. Deprivation indicators as predictors of life events 1981–92 based on the UK ONS longitudinal study. *J. Epidemiol. Commun. Health* 52:228–33
90. Snijders TA, Bosker RJ. 1993. Standard errors and sample sizes in two-level research. *J. Educ. Stat.* 19:237–59
91. Snijders TA, Bosker RJ. 1994. Modeled variance in two-level models. *Sociol. Methods Res.* 22:342–63
92. Soderfeldt B, Soderfeldt M, Jones K, et al. 1997. Does organization matter? A multilevel analysis of the demand-control model applied to human services. *Soc. Sci. Med.* 44:527–34
93. Susser M. 1994. The logic in ecological: I. The logic of analysis. *Am. J. Public Health* 84:825–29
94. Thomas N, Lonford N, Rolph J. 1994. Empirical Bayes methods for estimating hospital-specific mortality rates. *Stat. Med.* 13:889–903
95. van den Eeden P, Huttner HJ. 1982. Multilevel research. *Curr. Sociol.* 30:1–182
96. Von Korff M, Koepsell T, Curry S, Diehr P. 1992. Multi-level research in epidemiologic research on health behaviors and outcomes. *Am. J. Epidemiol.* 135:1077–82
97. Waitzman N, Smith K. 1998. Phantom of

- the area: poverty-area residence and mortality in the United States. *Am. J. Public Health* 88:973–76
98. Wang J, Siegal HA, Falck RS, Carlson RG. 1998. Needle transfer among injection drug users: a multilevel analysis. *Am. J. Drug Alcohol Abuse* 24:225–37
99. Ware J. 1985. Linear models for the analysis of longitudinal studies. *Am. Stat.* 39:95–101
100. Wilcox Rountree P, Clayton R. 1999. A contextual model of adolescent alcohol use across the rural-urban continuum. *Subst. Use Misuse* 34:495–519
101. Witte JS, Greenland S, Haile RW, Bird CL. 1994. Hierarchical regression analysis applied to a study of multiple dietary exposures and breast cancer. *Epidemiology* 5:612–21
102. Wong G, Mason W. 1985. The hierarchical logistic regression model for multilevel analysis. *J. Am. Stat. Assoc.* 80:513–24
103. Wong G, Mason W. 1991. Contextually specific effects and other generalizations of the hierarchical linear model for comparative analysis. *J. Am. Stat. Assoc.* 86:487–503
104. Wu Y-W. 1995. Hierarchical linear models: a multilevel data analysis technique. *Nurs. Res.* 44:123–26
105. Yen I, Kaplan G. 1998. Poverty area residence and changes in physical activity level. *Am. J. Public Health* 88:1709–12
106. Yen I, Kaplan G. 1999. Neighborhood social environment and risk of death: multilevel evidence from the Alameda County study. *Am. J. Epidemiol.* 149:898–907
107. Zeger S, Liang K, Albert P. 1988. Models for longitudinal data: a generalized estimating equation approach. *Biometrics* 44:1049–60



CONTENTS

PUBLIC HEALTH GENETICS: An Emerging Interdisciplinary Field for the Post-Genomic Era, <i>Gilbert S. Omenn</i>	1
HOST-PATHOGEN INTERACTIONS IN EMERGING AND RE-EMERGING INFECTIOUS DISEASES: A Genomic Perspective of Tuberculosis, Malaria, Human Immunodeficiency Virus Infection, Hepatitis B, and Cholera, <i>Janet M. McNicholl, Marie V. Downer, Venkatachalam Udhayakumar, Chester A. Alper, David L. Swerdlow</i>	15
NUTRITION, GENETICS, AND RISKS OF CANCER, <i>Cheryl L. Rock, Johanna W. Lampe, Ruth E. Patterson</i>	47
POPULATION SCREENING IN HEREDITARY HEMOCHROMATOSIS, <i>Arno G. Motulsky, Ernest Beutler</i>	65
THE INTERFACE OF GENETICS AND PUBLIC HEALTH: Research and Educational Challenges, <i>Melissa A. Austin, Patricia A. Peyser, Muin J. Khoury</i>	81
LOOKING BACK ON ""CAUSAL THINKING IN THE HEALTH SCIENCES, <i>J. S. Kaufman, C. Poole</i>	101
CAUSAL EFFECTS IN CLINICAL AND EPIDEMIOLOGICAL STUDIES VIA POTENTIAL OUTCOMES: Concepts and Analytical Approaches, <i>Roderick J. Little, Donald B. Rubin</i>	121
BUILDING BRIDGES BETWEEN POPULATIONS AND SAMPLES IN EPIDEMIOLOGICAL STUDIES, <i>W. Kalsbeek, G. Heiss</i>	147
MULTILEVEL ANALYSIS IN PUBLIC HEALTH RESEARCH, <i>Ana V. Diez-Roux</i>	171
SHOULD WE USE A CASE-CROSSOVER DESIGN, <i>M. Maclure, and M. A. Mittleman</i>	193
WATER RECLAMATION AND UNRESTRICTED NONPOTABLE REUSE: A New Tool in Urban Water Management, <i>Daniel A. Okun</i>	223
EPIDEMIOLOGY AND PREVENTION OF INJURIES AMONG ADOLESCENT WORKERS IN THE UNITED STATES, <i>Carol W. Runyan, Ronda C. Zakocs</i>	247
THE EFFECTS OF CHANGING WEATHER ON PUBLIC HEALTH, <i>Jonathan A. Patz, David Engelberg, John Last</i>	271
TOXICOLOGICAL BASES FOR THE SETTING OF HEALTH-RELATED AIR POLLUTION STANDARDS, <i>M. Lippmann, R. B. Schlesinger</i>	309
RELIGION AND HEALTH: Public Health Research and Practice, <i>Linda M. Chatters</i>	335
A REVIEW OF COLLABORATIVE PARTNERSHIPS AS A STRATEGY FOR IMPROVING COMMUNITY HEALTH, <i>Stergios Tsai Roussos, Stephen B. Fawcett</i>	369
ORAL HEALTH IN THE UNITED STATES: The Post-Fluoride Generation, <i>P. Milgrom, S. Reisine</i>	403
THE NEW PUBLIC HEALTH LITIGATION, <i>W. E. Parmet, R. A. Daynard</i>	437
BABY AND THE BRAIN: Advances in Child Development, <i>Janet A. DiPietro</i>	455
HEALTH PROMOTION IN THE CITY: A Review of Current Practice and Future Prospects in the United States, <i>N. Freudenberg</i>	473

THE RISE AND DECLINE OF HOMICIDE- AND WHY, <i>Alfred Blumstein, Frederick P. Rivara, Richard Rosenfeld</i>	505
INCOME INEQUALITY AND HEALTH: What Does the Literature Tell Us?, <i>Adam Wagstaff, Eddy van Doorslaer</i>	543
EVALUATING THE STATE CHILDREN'S HEALTH INSURANCE PROGRAM: Critical Considerations, <i>Barbara Starfield</i>	569
PREFERENCE-BASED MEASURES IN ECONOMIC EVALUATION IN HEALTH CARE, <i>Peter J. Neumann, Sue J. Goldie, Milton C. Weinstein</i>	587
TELEMEDICINE: A New Health Care Delivery System, <i>Rashid L. Bashshur, Timothy G. Reardon, Gary W. Shannon</i>	613
THE CHANGING NATURE OF RURAL HEALTH CARE, <i>Thomas C. Ricketts</i>	639
ASSESSMENT IN LONG-TERM CARE, <i>R. L. Kane, R. A. Kane</i>	659