The Logic in Ecological: I. The Logic of Analysis

Mervyn Susser, MB, BCh, FRCP(E), DPH

A B S T R A C T

This paper addresses ecological studies in public health research in terms of the logic of their analysis. It makes several distinctions between studies based on ecological and individual units. First, it identifies the variables common to both types of study and those particular to ecological studies. Second, it shows how ecological and individual units combine in two classes: unmixed (purely ecological, purely individual) and mixed. Third, it details how the relationships among and between individual and grouped units (expressed in terms of regression coefficients between independent and dependent variables) yield four coefficients: for all individual members; for all groups; for all individuals within each group; and for all individuals within groups (a weighted average). Equipped with an understanding of the dimensions involved at ecological and individual levels and of the relationships between them, researchers are in a position to exploit the public health potential of the ecological approach. (*Am J Public Health.* 1994;84:825–829)

Introduction

The prime justification for the ecological approach in epidemiology is to study health in an environmental context. The aim is ambitious: to understand how context affects the health of persons and groups through selection, distribution, interaction, adaptation, and other responses. Measures of individual attributes cannot account for these processes; pairings, families, peer groups, schools, communities, cultures, and laws are all contexts that alter outcomes in ways not explicable by studies that focus solely on individuals. With these contexts unmeasured, neither patterns of mortality and morbidity, nor epidemic spread, nor sexual transmission can be explained.

As E. L. Thorndike warned in 1939[1] and W. S. Robinson in 1950,[2] problems arise in making inferences about individuals from studies of groups. H.C. Selvin labeled this the "ecological fallacy" in 1958.[3] The naming of names often influences attitudes and thought. In epidemiology particularly, the fallacy has brought the ecological approach into disrepute.

This work will in some ways be one of rehabilitation. Its concern is with the utility of the ecological approach in its own right and not with its use as a substitute for more respectable approaches. In epidemiology, much attention has been given to abuses,[4–6] the barest minimum to uses.[7,8] The ecological fallacy is an issue of analysis and inference. Hence, one needs to clarify the logic of analysis and inference, the concern of this paper, before going on to the logic of design and practice.[9]

Analytic Distinctions

Analytically, the ecological approach reduces to the technical if not simple matter of taking groups and not individuals as the unit of study. Its essence lies in the distinctions between levels of organization. Individual units at one level are assembled into groups; these groups become the units of the next level. Each level acquires collective properties that are more than the sum of the properties of its individual members. It follows that the properties of neither level (individual or group) are wholly predictable from those of the other. The problems special to ecological analysis thus arise when one extrapolates upward or downward from any level to another[10–13]— either in the ascent from gene to molecule, to cell, to tissue or organ, to person, to group, or in the descent in reverse.

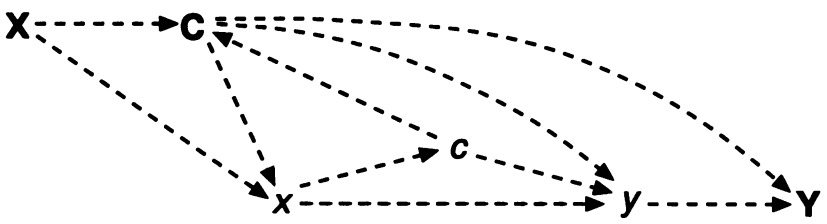
The task is to unravel the phenomena that emerge when persons, the individual units of the usual analytic currency of epidemiology, are assembled into larger units. These phenomena require a study of the effects of the group dimension itself on the manifestations under observation. Here, two or more persons make a group. Individual-level analysis cannot capture the effects of this dimension, including the interactions between one person and others in the transmission of infection, or behavior, or values. Hence, two broad kinds of variables enter the calculation: everyday variables common to individual and grouped units, and special variables peculiar to groups alone.

The author is the Editor of the Journal.

Requests for reprints should be sent to the Journal office.

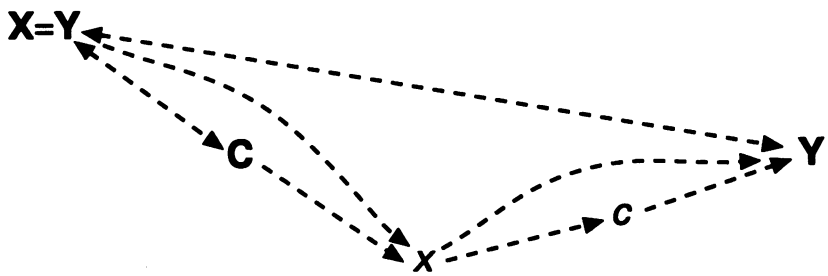
This paper was accepted July 15, 1993.

Editor's Note. Nicole Schupf was the Editor in charge of the blind peer review for this paper. See related articles by Susser (p 830), Koopman and Longini (p 836), and Schwartz (p 819) and editorial by Poole (p 715) in this issue.



Note. X = population exposure; *x* = individual exposure; C = probability of contact with infection; *c* = individual susceptibility; *y* = individual infection; Y = prevalence.

FIGURE 1—Causal paths in the spread of communicable disease.



Note. Y = prevalence of population exposure; C = probability of contact with infection; *x* = individual exposure; *c* = individual susceptibility.

FIGURE 2—Contagion as exposure: modified causal paths in the spread of communicable disease.

Variables in Common

Three general kinds of variable are shared by individual and group analysis:

Independent: individual, *x*; group, *X*.
Dependent: individual, *y*; group, *Y*.
Associated: individual, *a*; group, *A*.

Association is with either the independent variable, the dependent variable, or both. Thus, *A* along with *a* may be antecedent conditions, intervening variables, moderator variables or effect modifiers, potential confounders, etc.

Variables Peculiar to Groups

Two analytic dimensions (here termed *integral* and *contextual*) are unique to group analysis, as explained below. In the analysis of group phenomena, both integral and contextual variables can serve as independent, as associated or, if mutable, as dependent variables.

Integral variable. An integral variable (*I*) affects all or virtually all members of a group.[14] It reflects an antecedent condition that varies between groups but not appreciably within groups. *I* may be discrete and dichotomous (e.g., an intervention or a disaster), scaled and polychotomous (e.g., social disorganization, intensity of newborn care), or continuous (e.g., altitude, latitude). When groups, as the analytic unit, are assembled in a manner that renders the members within each group uniform with regard to *I*, meaningful individual unit analysis is precluded.

Contextual variable. A contextual variable (*C* = median, mean, or proportion of an attribute (*c*) of individual group members), derived from a measured attribute of individuals within each group, characterizes the group and not the individual.[15–17] Thus, in the case of disease, the knowledge that individuals are or are not infected collectively affords group prevalence rates as a measure of context. The grouped variable *C* will always be the appropriate measure of group effects (adjusted as necessary when groups are compared); the association of *C* with the grouped dependent variable *Y* incorporates the individual as well as the group effects of *C* on *y*. For individuals, *C* will be effective when it contributes variation to the individual dependent variable *y* over and above that found with the ungrouped individual-level variable *c*.

For instance, as a communicable disease spreads (as in Figure 1), the independent variable *x* characterizes individual exposure to microorganisms through contact, and the independent variable *X* characterizes population exposure. An individual also has an immune state (*c*) that governs susceptibility and, hence, the likelihood of becoming infected (*y*). A group will have a threshold value for herd immunity, which is a function of the proportion of susceptibles (*C*, itself a result of the spread of previous infection but also of other factors). The contextual variable *C* is a main determinant of the probability of individual contact with others infected (*x*) and thus modifies the effects of the individual immune state (*c*) and the likelihood of infection (*y*) (a likelihood that, of course, alters incidence and prevalence *Y* as well as subsequent spread in the population).

For convenience, we exemplify the mathematical relations of individual and group variables with a linear regression model. (In many instances, categorical multivariable models better reflect the data.[5]) Here, *c* is an individual predictor of *y* and *C* is a contextual predictor of *y*; both are measured by the position of the individual on the slope $x \rightarrow y$. c/C (*c* in the presence of *C*) allows for simultaneous effects at the individual and group levels—that is, the interaction of individual and contextual predictors of the position of the individual on the slope $x \rightarrow y$. All the above are incorporated in the group relations $X \rightarrow Y$.

Contagion. Another relation that can obtrude on measured associations is an analogous contextual effect of grouped dependent variables on outcome, aptly described as "contagion." This is a group effect, not of independent or associated variables but of the dependent variable *Y*. In Ronald Ross' Theory of Happenings,[18] the "dependent happenings" recently resurrected by Halloran and Struchiner[19] serve well to describe this fundamental dynamic of communicable disease. For instance, with any transmissible infection in a defined group, the prevalence at any moment modifies the likelihood of an individual becoming infected. The sequential path of infection set out above needs to be amended because in this case the population exposure *X* is, in fact, *Y* (see Figure 2).

Koopman's synthetic model of the transmission of human immunodeficiency virus (HIV) infection provides an ex-

ample.[20] It is a cogent reminder that, in the face of group effects, individual measures of risk (odds ratios) may serve poorly. Where prevalence is low, the odds ratios of infection with high-risk behavior rise as prevalence rises; where prevalence is high, the risks with low-risk behavior increase, and the odds ratio of high- to low-risk behavior declines. The dependent happenings of contagion render the individual measure unstable since it changes as contagion (i.e., context) changes.

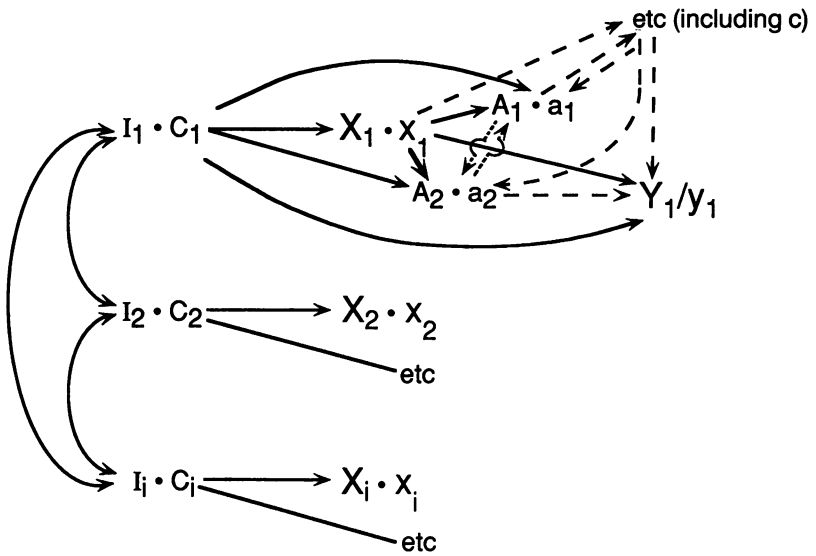
The moral is that analysis at the individual level, which clearly cannot capture epidemic spread at the group level, does not capture the entirety of spread through individuals either. To gauge risks of transmission, the supraindividual context is a key element. This includes the transactions between individuals with different modifying attributes and, in turn, the social and cultural dynamics that rule those transactions (about which more is said below under "Unmixed Studies").

Figure 3 is a simple path model that illustrates the relations engendered by the variables described. Within groups, the presence of associated variables multiplies the number of pathways (x_i, a_i, y_i). The presence of A at group level further exaggerates the number of relationships among individuals within and across groups. The more variables, the more complex the analysis and, for that reason alone, the more liability for mismeasurement, bias, and confounding. Interactions are likely to be more intrusive if not more detectable, covariation or colinearity compounded and difficult to disentangle, and confounding ever present and inapparent. These are problems of scale and are not unique to group variables.

Between groups, however, pathways special to groups are added; these multiply relations within groups still further. Integral and contextual dimensions, when in the form of associated variables, can be bearers of bias, and confounding remediable only by group analysis. With HIV infection, for instance, integral variables like geography (say, Africa vs the United States) and contextual variables (say, the prevalence of parenteral drug use or, as noted, of HIV infection) act on risks of transmission. The possibilities of imprudent inference from extrapolation between levels are patent but avoidable.

Combinations of Individual and Group Units

We now turn to the arrangement of variables in ecological studies. A fourfold table of independent and dependent variables at individual and group levels separates studies that do and do not mix levels (Table 1). Where extrapolation must occur is made evident.



Note. Solid lines represent definite relationships between variables that must be accounted for. Broken lines represent potentially influential relationships within each group. Arrows indicate causal direction. Letters designate variables: I = integral, C = contextual; X = grouped independent, *x* = individual independent; Y = grouped dependent, y = individual dependent; A = grouped associated, a = individual associated.

FIGURE 3—Scheme of paths between individual and grouped variables in *I* groups.

TABLE 1—Unmixed and Mixed Levels of Variables

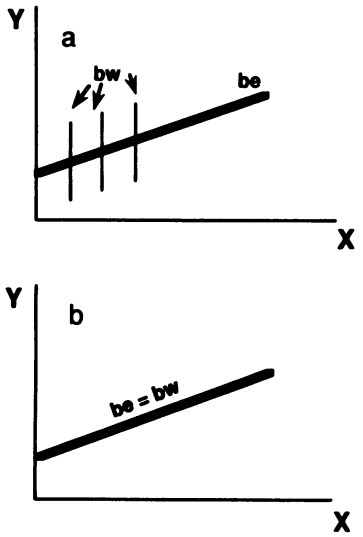
Dependent Variable	Independent Variable: Ecological	Independent Variable: Individual
Ecological	a. *XY*	b. *xY*
Individual	c. *Xy*	d. *xy*

Unmixed Studies

Unmixed studies (a and d in Table 1) involve variables on the same level. Extrapolation can arise only if one level is used to infer to the other in the absence of requisite data (i.e., *XY* to *xy* or vice versa). Unmixed group studies present few special problems except for those of scale and a resulting complexity of confounding, control, and interpretation. Unmixed individual studies present the familiar problems of epidemiology except for the neglect of group effects, as when subjects are assembled from more than one group.

The problems of complexity in themselves are neither insignificant nor simple. They arise from the multifarious ways in which individuals can be assembled into groups.[7-10] The nature of the problem is illustrated in Figure 4. In this model of associations at the individual and group level, the same regression slope between groups (*XY*) for each of two groups (coefficient B_e) can conceal quite different slopes for individuals (*xy*) within groups (coefficient B_w).[21]

By no means does the advantage always lie with the simpler individual analysis, and this is notably so with dependent happenings and contagion. The instance of epidemic infection and the relation $X \rightarrow Y$, with X (as temperature) representing an integral ecological effect, have been well illustrated for a severe dengue fever epidemic in Mexico. The mosquito *Aedes Egypti* is the vector for dengue and hence an obligatory condition for virus transmission. The integral condition, temperature during the rainy season, has by far the most powerful association (*XY*) with prevalence.[22] At cool temperatures, the incubation of the virus within the *Aedes Egypti* mosquito requires a period longer than many of the mosquitoes can survive, and this governs its prevalence.



Source. Based on an idea from Lincoln and Zeitz.[21]

FIGURE 4—Example of possible within group (bw) and between group (be) relations (expressed as regression coefficients) in the form of regression slopes between two variables at individual (xy) and group (XY) levels.

Surprisingly, no association between counts of *Aedes* larvae and dengue fever could be demonstrated at the individual level (*xy*). Risks of infection were raised for households where *Aedes* larvae were found and for communities where the prevalence of the mosquito was high. Here, context in the form of contagion played its part. Presumably, the dynamic between the prevalence of the mosquito, the ratio of infected to uninfected persons, and transmission between them led to the same risk of infection for entire households or communities, without distinction among individuals. Only ecological analysis could detect these contextual effects.

Mixed Studies

Mixed studies (b and c in Table 1) do present special problems. These stem from the choices to be made about levels of variables for analysis and from the potential intrusion of integral and contextual effects. In mixed studies, some degree of extrapolation is inherent. The relation *Xy*, with *X* a contextual effect, appeared above in the influence of the proportion vaccinated (*X*) on the probability of individual infection (*y*). *xY*, with *Y* a contextual effect, is illustrated by vaccinated persons (*x*) who help contain both the spread of infection and the prevalence (*Y*).

The distinction between levels is not always sharp. Aside from the hybrid forms in Table 1, blurring can be a matter of degree. Only nice distinctions between measures of independent or dependent variables may exist. For instance, a true individual risk can be derived from observing the timing of an outcome (single or repeated) in relation to repetitive exposures in one person, as with exercise and cardiac ischemia or life stress and epileptic seizures. To measure individual risk using the convention of cumulative incidence in a cohort, however, is, for a purist, a step toward grouping; to do so using period incidence or incidence density as a measure is a still larger step.

Relationships among Variables

Four basic relationships between independent (*x* or *X*) and dependent (*y* or *Y*) variables inhere in an assembly of data that contains measures of the character both of groups and of the individual members who comprise them (following Piantodosi et al.[23]). Previous writers have expressed these relationships in terms of correlations and linear regression coefficients (although, as already noted, categorical variables and logistic regression often serve better). For convenience, we shall use only regression coefficients as the more robust.[24,25] Correlations add more problems of instability when, as is common, they vary across groups according to the distributions of the variables within each group.

The coefficients refer to the following: B_i = Between individuals within each or any given group, $x \leftrightarrow y$. B_t = Between the total assembly of individuals and ignoring groups, i.e., $(x \leftrightarrow y)_i$. B_e = Between groups (ecological, with groups the unit), i.e., $X \leftrightarrow Y$. B_w = Weighted average of the group-specific coefficients for individuals within groups.*

B_w thus differs from B_e in that it adjusts for the known differences between groups in the distribution of variables that might affect the observed relations between *x* and *y*. If contextual and integral effects special to groups or grouping are absent, then $B_t = B_w = B_e$. Conversely, if contextual and integral variables contribute to the individual-level coefficients, they produce an "aggregation effect" and cause B_e to differ from B_t and B_i (improperly termed an *aggregation bias* unless individuals are the sole concern). If grouping distributes individual characteristics differently among groups and produces specification effects, B_w will differ from B_e and B_t (improperly termed *specification bias* unless estimates of associations between both groups and individuals are distorted owing to misspecification, confounding, and interaction).

A logical estimate of relations among the whole assembly of individuals is the weighted average group-specific coefficient, B_w. When available, B_w is also likely to be a logical estimate for groups; it facilitates adjustment for the different distributions of the characteristics of group members. Although in fact B_t is a weighted average of B_w and B_e and must lie between them, in theory (given random assembly of both groups and group membership) the order of magnitude of the two group coefficients cannot be predicted.[23] In practice, B_w and B_e are commonly larger than B_t or B_i for two reasons: naturally formed human groups tend to have a greater degree of homogeneity and covariance than occurs randomly; and, by design, competent research eliminates as much extraneous variation as possible and allows a larger proportion to be explained by the study variables.

The Fallacy

At this juncture, the bugbear of the ecological fallacy can be defined. The nub is simply the assumption that an association at one level of organization can be inferred from that at another level. In regression analysis terms, the fallacy is to assume that a between-group coefficient (B_e) will be equal either to the individual-level coefficients or, more elaborately, to the weighted average within-group coefficient—i.e., $B_e = B_i$, B_t, or B_w.

The fallacy stems from cross-level bias.[26,27] This bias occurs in either direction, according to initial perspective. From an individual perspective, it stems either from aggregation effects of contextual and integral variables or from specifi-

*See National Auxiliary Publication Service (NAPS) document 051103 for 4 pages of appendix material, by Dr Bruce Levin, that gives mathematical expression to the relations between individual and group variables in terms of linear regression coefficients. Please order directly from NAPS c/o Microfiche Publications, PO Box 3513, Grand Central Station, New York, NY 10163-3513. Enclose with your order $7.75 for paper copy or $4.00 for microfiche (US only funds from a US bank). For orders from outside the United States and Canada, add postage of $4.50 for paper copy or $1.50 for microfiche. There is a $15.00 invoicing charge for all orders filled before payment.

cation errors stemming from group composition, or both. Aggregation bias is special to groups; specification bias is not, although it is often aggravated by grouping.[4,16] Equally, from a group perspective, cross-level bias stems either from the "atomistic fallacy" inherent in individual observations[28] that ignore group effects, or from specification bias to which individual analyses are also prone.

Besides the sources of bias discussed above in the form of special group effects and macrolevel confounding or interactions, special problems reside in measurement.[5] We have illustrated how a variable measured as an attribute of individuals—say, infection, immunity, social class, or education—takes on added meaning as an attribute of groups when it signifies context. But other results can follow. Suppression of associations may result from the transfer of variables across levels *in either direction.* Loss of specificity results when a mean or proportion masks the variation between individuals. Conversely, an inadequate individual measure can suppress associations present at group level. In individuals, measures of blood pressure or urinary sodium or diet intake are notoriously unreliable; in groups, they are distinctly less so.[29,30]

Conclusion

Equipped with an understanding of the dimensions involved at ecological and individual levels and of the relationships between them, one is in a position to exploit the public health potential of the ecological approach. Effective researchers do not despair in the face of confounding and error; they guard against them and search them out. The means to do so reside in imaginative research design, measurement of potential confounders, informed analysis, and forceful inference. □

Acknowledgments

This work was supported by National Institutes of Health Grant 5-P50-MH43520. This paper was first undertaken as the opening for a special symposium convened by Nancy Padian, King Holmes, and Sevgi Aral at the meeting of the International Society for the Study of Sexually Transmitted Diseases in Banff, British Columbia, Canada, November 1991; it was further developed for the meeting of the American Epidemiological Society in Ann Arbor, Mich, March 1992, and for a meeting of the National Research Council of Italy on Models and Epidemiologic Methods of Research on HIV Infection, convened by Alfredo Nicolosi in Capri, September 4–7, 1992.

I owe thanks for helpful comments to Sander Greenland, David Jacobs, James Koopman, Bruce Levin, Nicole Schupf, and Zena Stein.

References

1. Thorndike EL. On the fallacy of imputing the correlations found for groups to the individuals or smaller groups composing them. *Am J Psychol.* 1939;52:122–124.
2. Robinson WS. Ecological correlations and the behavior of individuals. *Am Sociol Rev.* 1950;15:351–357.
3. Selvin H. Durkheim's suicide and problems of empirical research. *Am J Sociol.* 1958;63:607–619.
4. Brenner H, Savitz DA, Jockel K-H, Greenland S. Effects of nondifferential exposure misclassification in ecologic studies. *Am J Epidemiol.* 1992;135:85–95.
5. Greenland S, Morgenstern H. Ecological bias, confounding, and effect modification. *Int J Epidemiol.* 1989;18:269–274.
6. Greenland S. Divergent biases in ecological and individual-level studies. *Stat Med.* 1992;11:1209–1223.
7. Susser M. *Causal Thinking in the Health Sciences: Concepts and Strategies of Epidemiology.* New York, NY: Oxford University Press; 1973.
8. Morgenstern H. Uses of ecologic analysis in epidemiologic research. *Am J Public Health.* 1982;72:1336–1344.
9. Susser M. The logic in ecological: II. the logic of design. *Am J Public Health.* 1994;84:830–835.
10. Menzel H. Comment on Robinson's ecological correlations and the behavior of individuals. *Am Sociol Rev.* 1950;15:674.
11. Goodman L. Ecological regressions and behavior of individuals. *Am Sociol Rev.* 1953;18:663–664.
12. Blalock HM. *Causal Inferences in Nonexperimental Research.* Chapel Hill, NC: University of North Carolina Press; 1964.
13. Lazarsfeld PF, Menzel H. On the relation between individual and collective properties. In: Etzioni A, ed. *Complex Organizations: A Sociological Reader.* Glencoe, Ill: Free Press; 1961:422–440.
14. Selvin HC, Hagstrom WO. The empirical classification of formal groups. *Am Sociol Rev.* 1963;28:399–411.
15. Barton AH. Comments on Hauser's "context and consex." *Am J Sociol.* 1970;76:514.
16. Hammond JL. Two sources of error in ecological correlations. *Am Sociol Rev.* 1973;38:764–777.
17. Farkas G. Specification, residuals, and contextual effects. *Sociol Methods Res.* 1974;2:333–363.
18. Ross R. *The Prevention of Malaria.* 2nd ed. London, England: John Murray; 1911.
19. Halloran EM, Struchiner CJ. Study designs for dependent happenings. *Epidemiol.* 1991; 2:331–338.
20. Koopman JS, Longini IM, Jacquez JA, et al. Assessing risk factors for transmission of infection. *Am J Epidemiol.* 1991;133:1199–1209.
21. Lincoln JR, Zeitz G. Organizational properties from aggregate data: separating individual and structural effects. *Am Sociol Rev.* 1980;45:391–408.
22. Koopman JS, Prevots DR, Vaca Marin MA, et al. Determinants and predictors of dengue infection in Mexico. *Am J Epidemiol.* 1991;133:1168–1178.
23. Piantadosi S, Byar DP, Green SB. The ecological fallacy. *Am J Epidemiol.* 1988;127: 893–903.
24. Duncan OD, Davis B. An alternative to ecological correlation. *Am Sociol Rev.* 1953;18:665–666.
25. Goodman LA. Some alternatives to ecological correlation. *Am J Sociol.* 1959;64:610–625.
26. Hannan MT, Burstein L. Estimation from grouped observations. *Am Sociol Rev.* 1974;39:374–392.
27. Firebaugh G. A rule for inferring individual-level relationships from aggregate data. *Am Sociol Rev.* 1978;43:557–572.
28. Riley MW. *Sociological Research.* New York, NY: Harcourt, Brace, Jovanovich; 1963;1:700–718.
29. Frost CD, Law MR, Wald NJ. By how much does dietary salt reduction lower blood pressure? II. Analysis of observational data within populations. *Br Med J.* 1991;302:815–818.
30. Rush D, Kristal AR. Methodologic studies during pregnancy: the reliability of the 24-hour dietary recall. *Am J Clin Nutr.* 1982;35(suppl 5):1259–1268.