Methods Section

Use of Community Versus Individual Socioeconomic Data in Predicting Variation in Hospital Use

Timothy P. Hofer, Robert A. Wolfe, Philip J. Tedeschi, Laurence F. McMahon, and John R. Griffith

Objectives. (1) To examine the association of socioeconomic characteristics (SES) with hospitalization by age group, and when using measures of SES at the community as opposed to the individual level. (2) Thus, to support the inference that socioeconomic factors are important in the analysis of small area utilization data and address potential criticisms of this conclusion.
Data Sources. The 1989 Michigan Inpatient Database (MIDB), the 1990 U.S. Census, the 1989 Area Resource File (ARF), and the 1990 National Health Interview Survey (NHIS).
Study Design. A qualitative comparison of socioeconomic predictors of hospitalization in two cross-sectional analyses when using community as opposed to individual socioeconomic characteristics was done.
Data Extraction. Hospitalizations (excluding delivery) were extracted by county from the MIDB and by individual from the NHIS. SES variables were extracted from the U.S. Census for communities and from the NHIS for individuals. Measures of employment for communities were from the ARF and information on health insurance and health status of individuals from the NHIS.
Principal Findings. Both analyses show similar age-specific patterns for income and education. The effects were greatest in young adults, and diminished with increasing age. Accounting for multiple admissions did not change these conclusions. In the individual-level data the addition of variables representing health and insurance status substantially diminished the size of the coefficients for the socioeconomic variables.
Conclusions. By comparison to parallel individual-level analyses, small area analyses with community-level SES characteristics appear to represent the effect of individual-level characteristics. They are also not substantially affected by the inability to track individuals with multiple readmissions across hospitals. We conclude that the impact of SES characteristics on hospitalization rates is consistent when measured by individual- or community-level measures and varies substantially by age. These variables should be included in analyses of small area variation.

Key Words. Socioeconomic factors, hospitalization, small area analysis, health services misuse

The use of small area, or community, rates is increasing as healthcare providers attempt to meet price competition. Hospital admissions and procedures per capita are studied for potential savings under managed care. Low-use communities are frequently cited as benchmarks; comparisons are drawn and conclusions reached about reducing local hospitalization using these benchmarks. Thus, it is essential that the data used for such comparisons be as accurate as possible. This article explores the validity of using community-level estimates of socioeconomic characteristics (SES) as a proxy for individual-level socioeconomic characteristics in adjusting local hospitalization rates.

Small area studies have attracted attention since the initial work by Wennberg and Gittelsohn (1982), because these studies show that nearby, ostensibly similar communities have substantially and unpredictably different rates of hospitalization (Griffith, Restuccia, Tedeschi, et al. 1981; Paul-Shaheen, Clark, and Williams 1987; Wennberg and Gittelsohn 1982). Based on these studies, variation in hospitalization rates have often been assumed to be principally a function of provider practice style and capacity (Wennberg 1984; Wennberg and Cooper 1996), and the belief is widespread that educational and corrective policies directed at local provider groups can improve the quality and effectiveness of care (Iglehart 1984). A more general model of healthcare utilization suggests that these variations are a function of three different elements: (1) risk factors affecting the patient, (2) access to care through the supply of services or the impact of health insurance, and (3) provider practices (Aday and Andersen 1974).

An extensive literature describes the epidemiological evidence of associations between SES and the risk of morbidity and mortality. The causal

---

This research was supported in part by a grant from the Michigan Blue Cross and Blue Shield Foundation (formerly the Michigan Health Care Educational and Research Foundation). Dr. Hofer is supported by a career development grant from the Health Services Research and Development Service of the Department of Veterans Affairs.

Address correspondence and requests for reprints to Timothy P. Hofer, M.D., M.S., Research Scientist, VA Center for Practice Management and Outcomes Research, Ann Arbor VA HSR&D Field Program, P.O. Box 130170, Ann Arbor, MI 48103. E-mail thofer@umich.edu. Dr. Hofer is also an Assistant Professor of Medicine, Division of General Internal Medicine, University of Michigan. Robert A. Wolfe, Ph.D. is Professor of Biostatistics, Department of Biostatistics, School of Public Health; Philip J. Tedeschi, Ph.D. is an Assistant Research Scientist, Department of Health Management and Policy, School of Public Health; Laurence F. McMahon, M.D., M.P.H. is Professor of Medicine, Division of General Internal Medicine; and John R. Griffith, M.B.A. is Professor, Health Management and Policy, Department of Health Management and Policy, School of Public Health, all at the University of Michigan. This article submitted to *Health Services Research* on November 22, 1996, was revised and accepted for publication on July 29, 1997.

links for these observed relationships are not at all clear, but differences in the adoption of healthy behaviors or in the burden of disease are the most frequently proposed causal pathways (Pappas et al. 1993; Bunker, Gomby, and Kehrer 1989; Kaplan and Salonen 1990; Syme and Berkman 1976). While these relationships can have an impact on hospital utilization in many ways, one theory that has been studied recently is that the increased morbidity and decreased access to early treatment found in communities with lower SES levels result in higher hospitalization rates for those communities. There is some evidence in support of this theory for a few selected conditions (Billings, Zeitel, Lukomnik, et al. 1993; Bindman, Grumbach, Osmond, et al. 1995; Komaromy, Lurie, Osmond, et al. 1996).

Across broader ranges of diagnoses some small area researchers have failed to find a significant effect when using data sets that cover only a small number of communities (Wennberg and Gittelsohn 1982; Wennberg 1990) or health systems with universal healthcare insurance (Roos and Roos 1982). Wennberg has argued that provider capacity is a primary determinant of utilization and that SES are relatively unimportant in small area variation (Wennberg 1990; 1996). Socioeconomic adjustments are not included in the recently published *Dartmouth Atlas of Health Care* (Wennberg and Cooper 1996).

On the other hand, others have demonstrated that the variation in aggregate small area hospital utilization is associated with SES (Carlisle et al. 1995; Griffith, Restuccia, Tedeschi, et al. 1981; McLaughlin, Normolle, Wolfe, et al. 1989; McMahon et al. 1993; Wilson and Tedeschi (1984). At the DRG level, the amount of variability differs across diagnosis groups and is generally larger in nonprocedural groups (Griffith et al. 1985; McMahon, Tedeschi, Wolfe, et al. 1990; McLaughlin, Normolle, Wolfe, et al. 1989). McMahon et al showed that community measures of income and education are usually inversely associated with discharge rates and explain substantial portions of the variation in utilization (McMahon et al. 1993). A recent study in Maryland documented the complexity of income–discharge rate relationships, demonstrating positive relationships of income and use for more discretionary procedures and negative relationships for groups of diagnoses representing chronic diseases and medical conditions related to lifestyle (Gittelsohn and Powe 1995).

These conflicting conclusions about the role of SES may arise from differing effects across diagnoses or across age groups. They also may arise from using SES variables derived from community- versus individual-level measurements, a problem sometimes called the ecological fallacy. The ecological fallacy is said to occur if a community's SES profile is not representative of the status of the individuals actually using health services in that community. The age groups differ in important ways beyond their different diseases and conditions. For example, most people over age 65 have full Medicare coverage, and the absence of any health insurance is most common among young adults. Examining the interaction of community influences and age on hospital use is likely to improve our understanding of the role that socioeconomic factors play. But measures of SES are derived from census data for the communities, not from the individuals actually presenting for care. If the two differ significantly in their association with utilization, the results must be interpreted differently. While Geronimus et al. outline a statistical framework to describe the effect of using aggregate as opposed to individual socioeconomic characteristics as covariates in health outcomes studies, their work suggests that one cannot predict how the coefficients estimated from the two sources will be related when looking across different outcomes or measures of utilization (Geronimus, Bound, and Neidert 1996).

Thus, we set out to address two specific questions about the role of socioeconomic factors in contributing to small area variation: How does the importance of SES in small area analysis differ by age? And how are conclusions about the role of SES factors affected by the use of individual- as opposed to community-level data? We did a parallel analysis from two data sources, one representing a large number of geographic small areas and the other a national population–based survey of healthcare utilization. We also examined in the survey database whether conclusions about the importance of SES characteristics change when the hospitalization (as is common in small area studies) or the individual is the unit of analysis for the dependent utilization variable.

METHODS

Data Sources

The small area analysis used the 1989 Michigan Inpatient Data Base (MIDB). This database, which has been described in previous work, is a comprehensive discharge database for all Michigan hospitals and for hospitalizations of Michigan residents in hospitals in Ohio and Indiana (McMahon, Wolfe, and Tedeschi 1989; Tedeschi, Wolfe, and Griffith 1990). Socioeconomic variables were obtained from the 1990 Bureau of Census, Summary Tape File, STF3b, which organizes census data by zip code. An additional variable, an unemployment rate measure, was obtained from the 1989 Area Resource File (ARF) (Stambler 1988).

The data source for the individual-level analysis was the 1990 National Health Interview Survey (NHIS) Sample Person File, a population-based survey that collected detailed information on healthcare use, health status, health behaviors, and demographics from a sample of the civilian noninstitutionalized population (National Health Interview Survey 1990). The NHIS sampled 116,000 persons in 48,680 households using an in-person interview; the response rate was 97 percent.[1]

The NHIS records a total number of hospitalizations for each individual in a year, but does not try to code the diagnoses for each hospitalization beyond providing separate counts that include and exclude deliveries. Thus, parallel analyses could be constructed only for the aggregate variable of all hospitalizations excluding delivery, and no condition-specific analyses are presented. As previous work in the small area databases suggested that SES variables were highly significantly associated with hospital utilization even for aggregates of all hospitalizations, we anticipated that we would still be able to look for similar patterns in the individual-level database represented by the NHIS (McMahon et al. 1993).

Variables

In the small area analysis the dependent variable was the hospital discharge rate derived as a count of the number of hospitalizations over all DRGs excluding delivery and normal newborn divided by the population (in each of 12 age-sex groups). The unit of analysis was the county, and the hospitalization count may have contained multiple hospitalizations of a single individual. The hospitalization counts were indirectly age- and gender-adjusted. The independent variables included the SES and demographic variables shown in Table 1. The education variable was defined as the percentage of people over the age of 25 with at least a high school education; the poverty variable was defined as the percentage of the population below the poverty level; and the employment variable was defined as the percentage of the population unemployed in a small area. These variables were found in previous work to account for the largest amount of the variance (McMahon et al. 1993).

In the analysis of the National Health Interview Survey we constructed three dependent variables. In the first, which most closely parallels the small area analysis, we considered hospitalization as the unit of analysis and each hospitalization of an individual as an independent observation. In the second dependent variable we used the person as the unit of analysis and constructed

Table 1: Values of Socioeconomic Variables from the 1989 Area Resource File and 1990 Census for All Lower Peninsula Michigan Counties

Label	N	*Mean*	*s.d.*	*Minimum*	*Maximum*
% age > 25 and education ≥ high school	68	75.17	5.83	61.29	87.21
% below poverty	68	13.90	4.87	4.13	26.41
% urban	68	31.74	27.85	0	98.80
Income (thousand) per capita	68	11.768	2.384	8.19	21.12
Unemployment rate	68	8.3750	2.3644	4.20	16.20

a dichotomous dependent variable with 0 representing no hospitalizations in a year and 1 representing one or more hospitalizations in a year. The third dependent variable was a count of the number of hospitalizations in a single year, again using the person as the unit of analysis. The independent variables included age, gender, and urban-rural location, as well as individual-level dichotomous variables representing less than a high school education, income below the poverty level, and currently unemployed. In a single follow-up analysis, done only with the NHIS data, we used two variables to measure health status, one from a single-item scale rating overall health from poor to excellent, and the other a count of the number of days in the previous two weeks that the person had to reduce his or her activities due to a health problem. We also included in this analysis a variable asking whether the individual had health insurance coverage in the last month.

Analysis

Definition of Small Geographic Areas. In order to use the ARF variables, we chose to aggregate the data by county. In Michigan, counties have substantial overlap, with market or catchment areas defined on the basis of where the plurality of residents in the zip code area sought hospital care. Further, the effect of SES and provider variables have been shown to be similar for counties and market communities in Michigan (McLaughlin, Normolle, Wolfe, et al. 1989). Therefore, where zip codes crossed county borders the discharges were split randomly between the two counties in proportion to the census age-gender–specific population of the two counties.

Modeling. We estimated the SES coefficients in the small area analysis using a Poisson regression model with an extra-systematic component of variation (Wolfe et al. 1991). We chose this approach because almost half of Michigan counties are small, and this model makes a specific allowance for the random variability observed in small discharge counts. The Poisson regression equation is

$$\mathrm{Log}(m_{ij}) = \log(N_{ij}) + \beta z_i + \beta_0 \quad (1)$$

Where m_{ij} equals the expected count of discharges for the ith county and jth age-sex group, N_{ij} equals the county population by age and sex, and βz_i is the vector of socioeconomic adjusters.

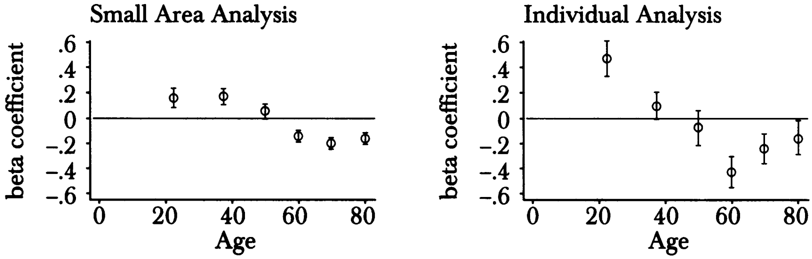
Another methodological problem of small area databases is the difficulty of identifying readmissions of the same person, which some have argued may be an important factor variation (Diehr et al. 1993; Diehr 1984). We used hospitalization as the dependent variable in both data sets, counting multiple admissions of the same person as independent cases. In the NHIS analysis we were also able to develop both a binary dependent variable representing whether an individual was hospitalized one or more times in a year and a count of the number of admissions in a single year by individual. We used logistic regression for the analyses of the first two dependent variables and Poisson regression in analyses of the count of number of admissions in one year. The significance and magnitude of the coefficients were similar in each of these analyses, so we proceeded, in the balance of the analysis, to use logistic regression with the hospitalization dependent variable in order to maximize its correspondence to the small area analysis. Separate models were run for each of the age groups examined (18–29, 30–44, 45–54, 55–64, 65–74, 75 and older) thus making each variable interact with age.

We display only qualitative comparisons of the beta-coefficients for the SES variables. It is not particularly meaningful to perform a quantitative comparison between, for example, the coefficient for a variable representing the percentage of people in a community with a high school or higher education level and a coefficient for a variable representing whether an individual person completed high school.

RESULTS

Results for the two parallel analyses are presented graphically in Figures 1 through 3 so that a qualitative comparison of the direction, magnitude, and significance of the socioeconomic variables can be made between the two analyses.

Figure 1: Comparison of the Effect of Female Gender on Hospitalization Rates Between a Small Area and Individual-Level Analysis



Legend. Each graph in this figure presents the beta coefficients for female gender from the parallel multiple regression models including age, gender, and socioeconomic factors in the small area and individual level (NHIS) databases. The bars represent the 95% confidence intervals for the estimated coefficient. The magnitude of the coefficients are not directly comparable, as the small area coefficient is a rate ratio from a Poisson regression estimating a community incidence rate and the individual-level analysis coefficient is an odds ratio from a logistic regression estimating the odds of a hospital discharge. The final models for the small area analysis were analyzed for each of the age groups shown with independent variables for gender, education (% of population > 25 with at least a high school education), poverty (% of population below the poverty level), employment (% of population unemployed), and % rural. The final models in the individual (NHIS) analysis were done for each of the age groups shown and included dichotomous independent variables for gender, less than high school education, income below the poverty level, currently unemployed, currently not in the labor force, and rural location.

Gender

Gender as a variable is ascertained at the individual level whether the analysis is done using small area discharge database data or survey data (as gender is coded on each discharge abstract). The coefficient patterns should be very similar for this variable as it is measured at the individual level (as opposed to the small area level) in both data sets. Reassuringly, the coefficients for the effect of female sex on hospitalization are similar for the small area and NHIS databases. The significance and direction are the same in every age group.

Socioeconomic Variables

For poverty level, the correspondence is close between the small area coefficients and the individual-level coefficients from the NHIS (Figure 2). Despite the differences in the units of the independent variable, the effect of poverty is substantially less among young adults (18–29), peaks in middle age, and declines somewhat thereafter. The significance and direction of the effect are the same for the community variable as for the individual variable in most age groups.

Both data sets show that higher education levels are generally associated with a reduced hospitalization rate. The general pattern, which shows that larger effects among the young population diminish with the age group, is also the same for both community and individual measures, although there is one somewhat anomalous coefficient (for those between 30 and 45 years old in the individual-level analysis). Unemployment is associated with higher hospitalization rates both in the small area and the individual analyses. The effects are generally the same in significance and direction for all age groups under 65. Unemployment as an individual characteristic is not particularly meaningful above the age of 65, and thus the coefficients for unemployment were not estimated for these older age groups in the NHIS data set.

Access and Need

Typically, variables representing access and need are not available in data sets used for small area analyses. Proxies for healthcare supply as one component of access, such as the number of hospital beds and physicians in a geographic area, have not been found to have significant coefficients in analyses of small area variation in hospital use (McMahon et al. 1993). Our supply variables—acute beds, long-term beds, and physician supply—were similarly not significant factors in the small area equations (analysis available from the authors).

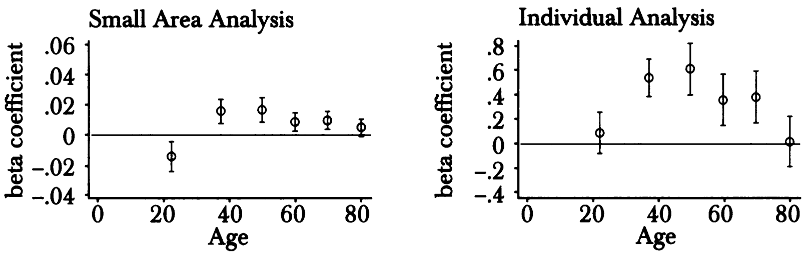
Although it is not possible to obtain variables representing access to and need for medical care in small area data sets, it is possible to examine what happens to the parallel effect of SES variables in the individual-level data from the National Health Interview Survey. We thus tested the hypothesis that SES variables should be less important if access and need variables (represented by insurance and health status) are entered in the individual-level analysis. In fact, the SES effects are always substantially diminished when insurance and health status are controlled for (Figure 3).

Unit of Analysis

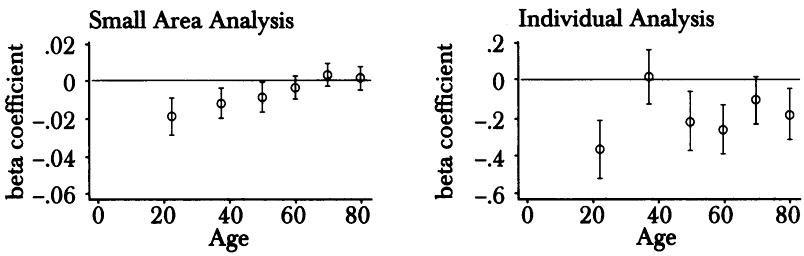
Table 2 illustrates the results of analyzing the data from the NHIS using three different dependent variables to represent a hospitalization. The definition

Figure 2: Comparison of the Effect of Socioeconomic Characteristics on Hospitalization Rates Between a Small Area and Individual-Level Analysis

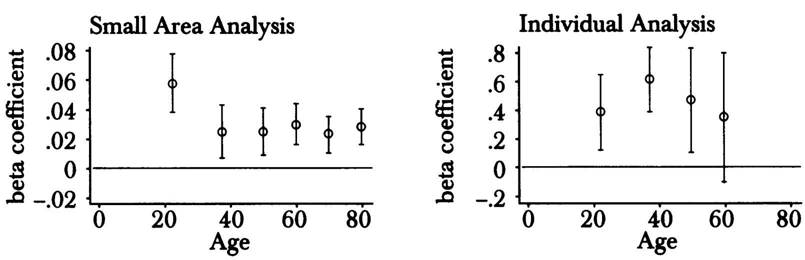
A. Poverty



B. Education



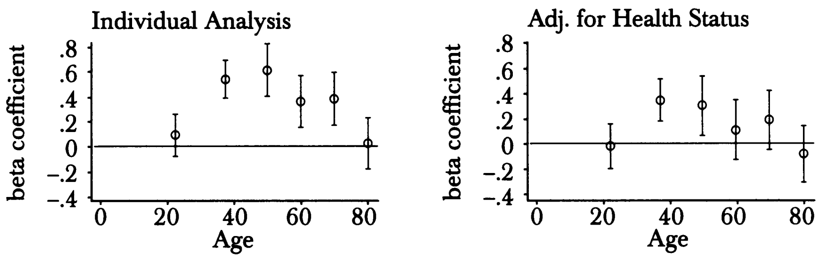
C. Unemployment



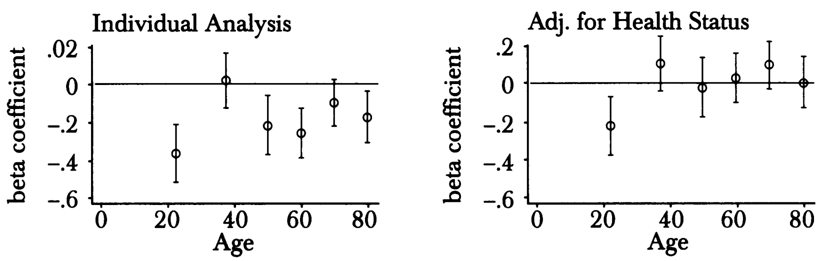
Legend. Each graph in this figure presents the beta coefficients for one socioeconomic variable from the parallel multiple regression models, including age, gender, and socioeconomic factors in the small area and individual-level (NHIS) databases. The bars represent the 95 percent confidence intervals for the estimated coefficient. The magnitudes of the coefficients are not directly comparable as the small area coefficient is a rate ratio from a Poisson regression estimating a community incidence rate, and the individual-level analysis coefficient is an odds ratio from a logistic regression estimating the odds of a hospital discharge. The variables in the model are described in the legend for Figure 1.

Figure 3: Adjusting for Insurance and Health Status in an Individual-Level Analysis

A. Poverty



B. Education



Legend. Each graph in this figure presents the beta coefficients of one of the socioeconomic variables from regression models, including age, gender, and socioeconomic factors in the individual-level (NHIS) databases, as described in the legend to Figure 1, with and without the inclusion of variables measuring overall health status and insurance status. The health status measures include two dichotomous variables for poor/fair overall health status and a dichotomous variable measuring if the person reduced his or her activity due to a health problem for one or more days in the previous two weeks. Insurance status is a dichotomous variable representing no health insurance in the previous month. The bars represent the 95 percent confidence intervals for the estimated coefficient.

that is most analogous to the small area analysis is to count every hospitalization as an independent observation. An individual with two hospitalizations in a year thus appears as two apparently independent observations, as they do in state discharge databases that lack individual identifiers. If one has an individual identifier it is possible to analyze hospitalization rates either as a dichotomous variable representing one or more hospitalizations or as a count of hospitalizations. The interpretation of each is, of course, slightly different. Table 2 illustrates that some small differences in the coefficients are

Table 2: Effect of Unit of Analysis on Beta Coefficients of Education and Income in the National Health Interview Survey

		Unit of Analysis		
Age	*Beta Coefficient (std. error)*	*Hospitalization (Without Individual Identifier)*	*Individual (Hospitalized One or More Times)*	*Individual (Count of Hospitalizations)*
18–29	Completed high school	−0.3718	−0.3166	−0.3553
	(standard error, H.S.)	(0.0777)	(0.0877)	(0.0744)
	Below poverty level	0.0884	0.1136	0.0858
	(standard error, poverty)	(0.0848)	(0.0952)	(0.0812)
30–44	Completed high school	0.0139	−0.0792	0.0240
	(standard error, H.S.)	(0.0739)	(0.0815)	(0.0703)
	Below poverty level	0.5338	0.4077	0.4926
	(standard error, poverty)	(0.0772)	(0.0893)	(0.0728)
45–54	Completed high school	−0.2232	−0.1986	−0.2055
	(standard error, H.S.)	(0.0790)	(0.0926)	(0.0729)
	Below poverty level	0.6085	0.4588	0.5443
	(standard error, poverty)	(0.1053)	(0.1288)	(0.0937)
55–64	Completed high school	−0.2688	−0.2499	−0.2400
	(standard error, H.S.)	(0.0660)	(0.0763)	(0.0597)
	Below poverty level	0.3563	0.2723	0.3091
	(standard error, poverty)	(0.1052)	(0.1246)	(0.0915)
65–74	Completed high school	−0.1084	−0.0879	−0.0955
	(standard error, H.S.)	(0.0626)	(0.0720)	(0.0558)
	Below poverty level	0.3748	0.3831	0.3092
	(standard error, poverty)	(0.1068)	(0.1210)	(0.0926)
above 75	Completed high school	−0.1871	−0.0803	−0.1716
	(standard error, H.S.)	(0.0691)	(0.0782)	(0.0598)
	Below poverty level	0.0155	0.0608	0.0057
	(standard error, poverty)	(0.1033)	(0.1173)	(0.0891)

Note: The NHIS data were analyzed in three ways. The first column refers to an analysis in which one observation exists for each individual not hospitalized and one observation was created *for each hospitalization* of an individual, analogous to small area analyses in databases that record hospitalizations but not individuals. The second column refers to an analysis in which the dependent variable is simply whether an individual had one or more hospitalizations in a year. The third column refers to an analysis of the number of hospitalizations for an individual in one year.

estimated for these three different dependent variables but that, on the whole, the patterns are quite similar.

CONCLUSIONS

It is important to consider SES characteristics in analyses of small area utilization data. We have found in parallel analyses that substantially similar conclusions are drawn about the relative direction by age group of socioeconomic variables (poverty, education, and unemployment), whether these predictors are individual or community characteristics. The association of community measures of education with utilization across age groups also has a correspondence, albeit somewhat less pronounced, with the associations found between individual measures of education and health utilization. The impact of these variables was far from small. For example, in the small area analysis an increase of ten points in the proportion of persons in a community with a high school education was associated with a 20 percent reduction in the hospitalization rate in the 18–29-year-old group and a 15 percent reduction for those 30–44 years old.

Obviously, these two sets of predictors—one at the community level and the other at the individual level—represent different things. The community characteristics are often described as compositional or contextual effects (Bryk and Raudenbush 1992). These characteristics may represent the average effect over the individuals in the community or may act as a proxy for omitted variables at the community level. "Ecological fallacy," as the term is used, reflects an incomplete analytical model of the hierarchical relationships between utilization and individual and community characteristics. Our data sources do not allow a hierarchical analysis, but by finding congruence between the SES effects at the individual and community level, we argue that the community characteristics in this case are acting as an appropriate summary of the individual-level characteristics.

In our analyses from the NHIS the inclusion of two health status measures attenuated the coefficients of education and income substantially, thus suggesting that SES variables are in part capturing unmeasured differences in the burden of disease in different populations. This might suggest that SES variables would not be needed in small area studies if health status were adequately captured. On the other hand, Bindman, Grumbach, Osmond, et al. (1995) show that—at least for the conditions examined in their study—SES factors remained significant even after controlling for community-level measures of the prevalence of disease and the propensity to seek care. Thus, it would clearly be too simple to attribute the relationships among income, education, and healthcare utilization completely to differences in health status.

By examining interactions between socioeconomic factors and age, we found reasons why studies in Medicare databases would not reveal significant relationships between small area variation and SES. The importance of these variables is much more prominent in younger age groups whether one looks at community characteristics or at individual-level characteristics. Billings et al., in looking at a number of ambulatory care–sensitive conditions, found strikingly similar patterns of interactions of income and age, peaking in the 25–44 year age groups, with the relationship between low income and higher hospitalization rates (Billings, Zeitel, Lukomnik, et al. 1993). This suggests that it is necessary to examine age interactions in assessing the importance of SES variables in small area studies, something that is not commonly done.

Why would the effect of SES variables decrease with increasing age? Several reasons are possible. First, access to outpatient care may improve through Medicare, or older users may, by virtue of being more frequent users of healthcare, have overcome some of the nonfinancial barriers to seeking outpatient care (Billings, Zeitel, Lukomnik, et al. 1993). Although better access to care has been shown to be associated with a reduced rate of hospital use for a few conditions in which hospitalization is thought to be preventable (Bindman, Grumbach, Osmond, et al. 1995; Billings, Zeitel, Lukomnik, et al. 1993), it seems unlikely that this effect would generalize to the aggregate hospitalization measure. Second, although previous small area studies found that SES variables had consistent directions in the 78–85 out of 114 modified diagnosis-related groups (MDRGs) (McMahon et al. 1993), many diseases and conditions affect a relatively narrow age distribution and SES factors may be more closely associated with utilization in conditions that occur in the younger age groups. Finally, within a particular condition, SES characteristics may be an important risk factor for the presence of that condition in younger age groups, while the influence of age-related risk factors may swamp this effect as people get older.

We also found that using counts of individuals hospitalized (rather than counts of hospitalizations) does not qualitatively change the interpretation of the socioeconomic predictors (Table 2). While some states such as California are finally moving to improve discharge databases to reflect readmissions, our analysis suggests that readmissions at least do not significantly change the conclusions drawn from analyses of large aggregates of diagnoses. Unique identifiers will almost certainly be important for diagnosis-specific analyses in diagnosis groups that have high readmission rates.

Our analysis has two important limitations. We lack a single database that would permit a hierarchical analysis that quantified the portion of total variance explained by the community and individual SES characteristics (Bryk and Raudenbush 1992). The disparity in the sources of data is both a limitation and a strength. When similar conclusions can be drawn from very different data sets it adds confidence in the robustness of those conclusions. We analyzed the difference only for the aggregate of all diagnoses. This is all that is possible for this analysis given that the NHIS does not include the DRG of hospitalizations. DRG-specific models have been reported for the small area database and the importance of SES indicators is remarkably consistent at the DRG and DRG aggregate level (McMahon et al. 1993). Furthermore, it seems likely that the impact of aggregating across diagnoses would tend to bias the coefficients of the SES variables toward the null hypothesis and that the relationships would be stronger only at the condition level. Finally, it is possible, although in our opinion unlikely, that small area hospitalization data from another state would show different results.

Despite these limitations, we believe that we have added further support to an argument that the small area variation in hospitalization rates depends significantly on SES effects, and that community-level measures of these variables appear to be a reasonable proxy for individual measures. Our work is consistent with the conclusions reached by Bindman, Grumbach, Osmond, et al. (1995). We believe that these adjustments should always be included in cross-sectional comparisons of small area rates. If residual variation in hospital utilization is ascribed to particular providers or groups of providers who care for a community, the impact that these SES covariates would have on provider rankings across communities is not predictable and would need to be assessed case by case. Communities and institutions seeking to reduce their use of hospital services safely should consider that SES characteristics may well capture population differences in health status or disease risk factors and could substantially change their interpretation of the residual variance.

NOTE

1. NCHS is responsible only for the initial data contained in the NHIS. Any analyses, interpretations, and conclusions based on these data are those of the authors alone.

REFERENCES

Aday, L. A., and R. Andersen. 1974. "A Framework for the Study of Access to Medical Care." *Health Services Research* 9, no. 3 (fall): 208–20.

Billings, J., L. Zeitel, J. Lukomnik, T. S. Carey, A. E. Blank, and L. Newman. 1993. "Impact of Socioeconomic Status on Hospital Use in New York City." *Health Affairs* (Millwood) 12, no. 1 (spring): 162–73.

Bindman, A. B., K. Grumbach, D. Osmond, M. Komaromy, K. Vranizan, N. Lurie, and others. 1995. "Preventable Hospitalizations and Access to Health Care." *Journal of the American Medical Association* 274, no. 4 (July): 305–11.

Bryk, A. S., and S. W. Raudenbush. 1992. *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, CA: Sage Publications.

Bunker, J. P., D. S. Gomby, and B. H. Kehrer, eds. 1989. *Pathways to Health: The Role of Social Factors*. Menlo Park, CA: The Henry J. Kaiser Family Foundation.

Carlisle, D. M., R. B. Valdez, M. F. Shapiro, and R. H. Brook. 1995. "Geographic Variation in Rates of Selected Surgical Procedures Within Los Angeles County." *Health Services Research* 30, no. 1 (April): 27–42.

Diehr, P. 1984. "Small Area Statistics: Large Statistical Problems." *American Journal of Public Health* 74, no. 4 (April): 313–14.

Diehr, P., K. Cain, Z. Ye, and F. Abdul-Salam. 1993. "Small Area Variation Analysis: Methods for Comparing Several Diagnosis-Related Groups." *Medical Care* 31, no. 5, Supplement (May): YS45–53.

Geronimus, A. T., J. Bound, and L. J. Neidert. 1996. "On the Validity of Using Census Geocode Characteristics to Proxy Individual Socioeconomic Characteristics." *Journal of the American Statistical Association* 91 (434): 529–37.

Gittelsohn, A., and N. R. Powe. 1995. "Small Area Variations in Health Care Delivery in Maryland." *Health Services Research* 30, no. 2 (June): 295–317.

Griffith, J. R., J. D. Restuccia, P. J. Tedeschi, P. A. Wilson, and H. S. Zuckerman. 1981. "Measuring Community Hospital Service in Michigan." *Health Services Research* 16, no. 2 (summer): 135–60.

Griffith, J. R., P. A. Wilson, R. A. Wolfe, and D. P. Bischak. 1985. "Clinical Profiles of Hospital Discharge Rates in Local Communities." *Health Services Research* 20, no. 2 (June): 131–51.

Iglehart, J. K. 1984. "Variations in Medical Practice." *Health Affairs* (Millwood) 3 (2): 1–160.

Kaplan, G. A., and J. T. Salonen. 1990. "Socioeconomic Conditions in Childhood and Ischaemic Heart Disease During Middle Age." *British Medical Journal* 301, no. 6761 (November): 1121–23.

Komaromy, M., N. Lurie, D. Osmond, K. Vranizan, D. Keane, and A. B. Bindman. 1996. "Physician Practice Style and Rates of Hospitalization for Chronic Medical Conditions." *Medical Care* 34, no. 6 (June): 594–609.

McLaughlin, C. G., D. P. Normolle, R. A. Wolfe, L. F. McMahon, Jr., and J. R. Griffith. 1989. "Small Area Variation in Hospital Discharge Rates: Do Socioeconomic Variables Matter?" *Medical Care* 27, no. 5 (May): 507–21.

McMahon, L. F., Jr., P. J. Tedeschi, R. A. Wolfe, J. R. Griffith, and C. G. McLaughlin. 1990. "Small-Area Analysis of Gastrointestinal Disease Hospital Discharge Variation: Are the Poor at Risk?" *Journal of Clinical Gastroenterology* 12, no. 2 (April): 132–39.

McMahon, L. F., Jr., R. A. Wolfe, J. R. Griffith, and D. Cuthbertson. 1993. "Socioeconomic Influence on Small Area Hospital Utilization." *Medical Care* 31, no. 5, Supplement (May): YS29–36.

McMahon, L. F., Jr., R. A. Wolfe, and P. J. Tedeschi. 1989. "Variation in Hospital Admissions Among Small Areas: A Comparison of Maine and Michigan." *Medical Care* 27, no. 6 (June): 623–31.

National Health Interview Survey. 1990. Health Promotion and Disease Prevention

(HPDP) Sample Person Supplement. [data tape]. National Center for Health Statistics, USDHHS, producer.

Pappas, G., S. Queen, W. Hadden, and G. Fisher. 1993. "The Increasing Disparity in Mortality Between Socioeconomic Groups in the United States, 1960 and 1986." *The New England Journal of Medicine* 329, no. 2 (July): 103–109.

Paul-Shaheen, P., J. D. Clark, and D. Williams. 1987. "Small Area Analysis: A Review and Analysis of the North American Literature." *Journal of Health Politics, Policy and Law* 12 (4): 741–809.

Roos, N. P., and L. L. Roos, Jr. 1982. "Surgical Rate Variations: Do They Reflect the Health or Socioeconomic Characteristics of the Population?" *Medical Care* 20, no. 9 (September): 945–58.

Stambler, H. V. 1988. "The Area Resource File: A Brief Look." *Public Health Reports* 103, no. 2 (March–April): 184–88.

Syme, S. L., and L. F. Berkman. 1976. "Social Class, Susceptibility and Sickness." *American Journal of Epidemiology* 104, no. 1 (July): 1–8.

Tedeschi, P. J., R. A. Wolfe, and J. R. Griffith. 1990. "Micro-Area Variation in Hospital Use." *Health Services Research* 24, no. 6 (February): 729–40.

Wennberg, J. E. 1984. "Dealing with Medical Practice Variations: A Proposal for Action." *Health Affairs* (Millwood) 3, no. 2 (summer): 6–32.

———. 1996. "On the Appropriateness of Small Area Analysis for Cost Containment." *Health Affairs* (Millwood) 15, no. 4 (summer): 164–67.

Wennberg, J., and A. Gittelsohn. 1982. "Variations in Medical Care Among Small Areas." *Scientific American* 246, no. 4 (April): 120–34.

Wennberg, J. E. 1990. "Small Area Analysis and the Medical Area Outcome Problem." In *AHCPR Conference Proceedings. Research Methodology: Strengthening Causal Interpretations of Nonexperimental Data, April 8–April 10, 1987, Tucson, AZ,* edited by L. Sechrest, E. Perrin, and J. Bunker. Department of Health and Human Services Publication No. 90-3454. Rockville, MD: DHHS.

Wennberg, J. E., and M. M. Cooper. 1996. *Dartmouth Atlas of Health Care Use in the United States.* Chicago: AHA Publishing.

Wilson, P., and P. Tedeschi. 1984. "Community Correlates of Hospital Use." *Health Services Research* 19, no. 3 (August): 333–55.

Wolfe, R. A., G. R. Petroni, C. G. McLaughlin, and L. F. McMahon, Jr. 1991. "Empirical Evaluation of Statistical Models for Counts or Rates." *Statistical Medicine* 10, no. 9 (September): 1405–16.