



## Original Contribution

### Neighborhood Environment and Risk of Ischemic Stroke

#### The Brain Attack Surveillance in Corpus Christi (BASIC) Project

L. D. Lisabeth<sup>1,2</sup>, A. V. Diez Roux<sup>1</sup>, J. D. Escobar<sup>2</sup>, M. A. Smith<sup>2</sup>, and L. B. Morgenstern<sup>1,2</sup>

<sup>1</sup> Department of Epidemiology, University of Michigan School of Public Health, Ann Arbor, MI.

<sup>2</sup> Stroke Program, University of Michigan Medical School, Ann Arbor, MI.

Received for publication January 24, 2006; accepted for publication June 9, 2006.

The authors explored whether neighborhood-level characteristics are associated with ischemic stroke and whether the association differs by ethnicity, age, and gender. Using data from the Brain Attack Surveillance in Corpus Christi Project (January 2000–June 2003), they identified cases of ischemic stroke ( $n = 1,247$ ) from both hospital and out-of-hospital sources. Census tracts served as proxies for neighborhoods, and neighborhood socioeconomic status scores were constructed from census variables (higher scores represented less disadvantage). In Poisson regression analyses comparing the 90th percentile of neighborhood score with the 10th, the relative risk of stroke was 0.49 (95% confidence interval (CI): 0.41, 0.58). After adjustment for age, gender, and ethnicity, this association was attenuated (relative risk (RR) = 0.79, 95% CI: 0.63, 1.00). There was no ethnic difference in the association of score with stroke ( $p$  for interaction = 0.79). Significant effect modification was found for age ( $p$  for interaction < 0.001) and gender ( $p$  for interaction = 0.04), with increasing scores being protective against stroke in men and younger persons. Associations were attenuated after adjustment for education (men: RR = 0.77, 95% CI: 0.55, 1.07; persons aged <65 years: RR = 0.65, 95% CI: 0.41, 1.02). Neighborhood characteristics may influence stroke risk in certain gender and age groups. Mechanisms for these associations should be examined.

cerebrovascular accident; ethnic groups; residence characteristics; social class

Abbreviations: BASIC, Brain Attack Surveillance in Corpus Christi; CI, confidence interval; RR, relative risk; SD, standard deviation; SES, socioeconomic status.

Socioeconomic status (SES) is a risk factor for ischemic stroke, with a population attributable risk similar to that of hypertension (1, 2). Aside from individual-level measures of SES, such as income and education, there is growing interest in how neighborhoods or characteristics of neighborhoods affect health and health outcomes (3–5). Area-level indicators may provide information about living conditions potentially related to stroke risk that are not captured by individual-level variables, such as access to health care, healthy foods, and safe places to walk, social norms affecting health habits such as smoking or drinking, and area-based sources of psychosocial stress.

Low SES measured at the area level has been found to be associated with the onset of stroke in New Zealand (6), Sweden (7), and Australia (8) and with the prevalence of stroke risk factors, stroke severity, and recurrent vascular events in the United Kingdom (9). Little is known about the influence of residential environment on stroke risk or outcome in the United States (2). Neighborhood characteristics have been investigated for their role in cardiovascular disease in the United States. Residing in socioeconomically disadvantaged neighborhoods has been associated with coronary heart disease incidence (10) and prevalence (11, 12), the prevalence of subclinical

Reprint requests to Dr. Lewis B. Morgenstern, Stroke Program, University of Michigan Medical School, 1500 East Medical Center Drive, TC 1920/0316, Ann Arbor, MI 48109-0316 (e-mail: Lmorgens@umich.edu).

cardiovascular disease (13), and cardiovascular disease mortality (14, 15).

In the United States, ethnic minority groups, including Hispanic Americans, tend to live in more disadvantaged areas (16, 17) and experience a higher risk of ischemic stroke (18, 19). It is plausible that in addition to differences in risk factor profiles and individual-level SES, environmental conditions contribute to the ethnic disparity in stroke. Knowledge of neighborhood factors which may contribute to stroke risk could shed light on potential interventional strategies aimed at implementing environmental changes to prevent stroke and reduce ethnic disparities. Using data from a population-based stroke surveillance study carried out in a biethnic community, we investigated whether area-level SES is related to risk of ischemic stroke. Prior work has suggested that neighborhood health effects may differ by demographic characteristics (14, 15). Therefore, we also examined whether associations of neighborhood characteristics with stroke differed by ethnicity, age, and gender.

## MATERIALS AND METHODS

The methods used in the Brain Attack Surveillance in Corpus Christi (BASIC) Project have been previously reported (19, 20). Briefly, a combination of active and passive surveillance was used to capture cerebrovascular cases (ischemic stroke, intracerebral hemorrhage, and subarachnoid hemorrhage) among residents of Nueces County, Texas, older than 44 years of age between January 2000 and June 2003. Nueces County is a biethnic community (56 percent Mexican American, 38 percent non-Hispanic White) with a population of 313,645 (21). Over 95 percent of the population resides within the city of Corpus Christi, an urban area. Nueces County is over 150 miles (240 km) from San Antonio and Houston and serves as the regional referral area for the surrounding counties, which are sparsely populated. This distance affords complete case capture of initial medical contacts for acute stroke. Nueces County is not an immigrant community. The majority of Mexican Americans there are second- and third-generation US citizens.

Trained abstractors screened all cases identified from emergency department and inpatient sources ( $n = 7$  hospitals) in the county by manually searching emergency room visit and admission logs for a set of previously validated screening diagnostic terms. The abstractors routinely canvassed intensive care units and hospital floors searching for in-house strokes and cases not ascertained through screening logs. This "hot pursuit" surveillance method was supplemented by review of *International Classification of Diseases*, Ninth Revision, hospital discharge codes for stroke (430–438, excluding codes 433.x0, 434.x0 ( $x = 1-9$ ), 437.0, 437.2, 437.3, 437.4, 437.5, 437.7, 437.8, and 438).

Cases that screened positive for stroke were reviewed for eligibility criteria. For all eligible cases, data were abstracted and source documentation was copied. Study neurologists validated cases using the source documentation, blinded to the subject's age and ethnicity, on the basis of previously published criteria (22). Source documentation used for validation included admission history, physical examinations,

emergency room records, neurology consultations, and computed tomography/magnetic resonance imaging reports (19). Computed tomography/magnetic resonance imaging reports were available for all patients to distinguish hemorrhagic stroke from ischemic stroke. Only cases of ischemic stroke were included. Ischemic stroke was defined as the acute onset (minutes to hours) of a focal neurologic deficit specifically attributable to a cerebral artery distribution that persisted for more than 24 hours (except in cases of sudden death or if the development of symptoms was interrupted by a surgical or interventional procedure) and was not attributable to another disease process (seizure, brain tumor, hypoglycemia, metabolic encephalopathy, or hysteria). We distinguished ischemic stroke from hemorrhagic stroke, as the diagnosis of intracerebral hemorrhage or subarachnoid hemorrhage required the clinical symptoms mentioned above and neuroimaging that demonstrated a spontaneous focal collection of blood in either the parenchyma/ventricle (intracerebral hemorrhage) or subarachnoid space (subarachnoid hemorrhage). Alternatively, subarachnoid hemorrhage could be diagnosed by cerebrospinal fluid examination criteria.

An out-of-hospital sampling frame was used to identify nonhospitalized stroke cases. A random sample of 47 of 167 primary care physicians, four of 11 nursing homes, and all 11 neurologists in Nueces County was obtained, and out-of-hospital stroke cases were ascertained from this sample. Out-of-hospital stroke cases were validated using the same methods as those used for inpatient stroke cases.

## Interview methods

A random sample comprising two thirds of patients with validated cerebrovascular events was asked to participate in an in-person interview (response rate = 84 percent) (23). The interview contained questions regarding education. Patients unable to answer a series of orientation questions asked before the interview had a proxy interview conducted, in their presence whenever possible. Thirty-two percent (203 of 641) of the interviews were conducted with proxies. In 82 percent ( $n = 167$ ), the proxy subject was a spouse (28 percent) or child (54 percent) of the patient. Use of proxies did not differ by the gender or ethnicity of the cases. Interviews were performed in English or Spanish, depending on patient preference.

## Addresses and geocoding

Patients' home addresses were recorded from the medical record using either the face page or the hospital computer system. Addresses were sent to an external company for geocoding.

## Census data

Data from the 2000 US Census were the source for population counts and neighborhood-level variables. Population counts were obtained from Summary File 1 (which is derived from the Census short form administered to all persons in the United States) when available. Population counts

cross-classified by age, gender, ethnicity, and education were obtained from Summary File 4, and data on neighborhood socioeconomic status indicators were obtained from Summary File 3. Data in Summary Files 3 and 4 are derived from the US Census long form. Census long-form data are based on a randomly selected stratified sample (an average of one in six households) weighted to obtain estimates of actual figures that would have been obtained from a complete count (24). Because of confidentiality requirements, cross-classified population counts within census tracts are available only if there are at least 50 unweighted sample cases in the cell (24). Because of this restriction, population counts for several cross-classified cells of age, gender, ethnicity, and education (see “Statistical analysis” section) nested within census tracts were not available for analysis ( $n = 202$  cross-classified cells) and were excluded from models including education.

**Neighborhood scores**

Census tracts were used as proxies for neighborhoods. Census tracts tend to include approximately 4,000–7,000 persons and to represent homogeneous groups in terms of social and economic factors (25). We used six neighborhood-level census variables reflecting the domains of wealth/income: 1) median annual household income; 2) median value of occupied housing units; 3) percentage of households receiving interest, dividend, or net rental income; 4) percentage of adults who completed high school; 5) percentage of adults who completed college; and 6) percentage of persons in managerial or professional occupations. A summary score was constructed from these six variables on the basis of previously published methods (26). Data on median household income and the value of housing units were log-transformed because of skewed distributions prior to calculation of the summary score.

The summary score was constructed as follows. For each of the six census variables, a mean and standard deviation were calculated across census tracts. We estimated a Z score for each variable by subtracting the mean and dividing by the standard deviation for that variable. We then summed the Z scores for each variable to obtain a neighborhood score for the given census tract, with increasing scores representing increasing neighborhood advantage. Using the geocoded addresses for the stroke cases, we then linked neighborhood scores to stroke cases by census tract.

**Statistical analysis**

Because individual-level data were not available for non-stroke cases in the population, group-level data were analyzed using Poisson regression. Population counts from the census were used to represent the population at risk. We fitted a series of models to assess the association between neighborhood score and stroke risk. First, a Poisson regression model (model 1) was used to model the unadjusted association of neighborhood score and stroke risk at the census-tract level for a total of 64 census tracts in the study area. Neighborhood score was modeled continuously and as a series of dummy variables based on quartiles of the dis-

tribution, with the quartile representing the lowest score being used as the reference category.

Second, a Poisson regression model with a random intercept for each census tract (model 2) was used to model the association between neighborhood score (modeled continuously) and stroke risk, with adjustment for age, gender, and ethnicity. A multilevel model was fitted to cross-classified cells of age, gender, and ethnicity nested within census tracts (27). Numbers of stroke cases and population counts were calculated for the cross-classified cells of age (45–59, 60–74, and  $\geq 75$  years), gender (male and female), and ethnicity (Mexican American and non-Hispanic White) within each census tract (768 cells nested within 64 tracts). In the first stage (level 1—cells nested within tracts), a separate regression was defined for each census tract as follows.

$$\text{Log } Y_{ij} = b_{0j} + b_{1j} \text{ age1}_{ij} + b_{2j} \text{ age2}_{ij} + b_{3j} \text{ gender}_{ij} + b_{4j} \text{ ethnicity}_{ij} + \log n_{ij}, \tag{1}$$

where  $Y_{ij}$  is the number of stroke cases for the  $i$ th cross-classified cell in census tract  $j$ ;  $\text{age1}_{ij}$ ,  $\text{age2}_{ij}$ ,  $\text{gender}_{ij}$ , and  $\text{ethnicity}_{ij}$  are cell-level dummy variables for the  $i$ th cell in the  $j$ th census tract;  $b_{0j}$  is the census tract-specific intercept;  $b_{1j}$ ,  $b_{2j}$ ,  $b_{3j}$ , and  $b_{4j}$  are the census tract-specific effects of cell-level variables; and  $n_{ij}$  is the offset for each cell (the number of persons for the  $i$ th cross-classified cell in census tract  $j$ ).

In a second stage (level 2—census tracts), each of the regression coefficients defined in equation 1 was modeled as follows.

$$\begin{aligned} b_{0j} &= \gamma_{00} + \gamma_{01} \text{ score}_j + U_{0j} \\ b_{1j} &= \gamma_{10} + U_{1j} \\ &\dots \\ b_{4j} &= \gamma_{40} + U_{4j}, \end{aligned}$$

where  $\gamma_{00}$  is the common intercept across census tracts,  $\gamma_{01}$  is the fixed effect of neighborhood score,  $\gamma_{10}, \dots, \gamma_{40}$  are the common slopes associated with cell-level variables across census tracts, and  $U_{0j}, \dots, U_{4j}$  are the random effects associated with census tracts.

We expanded model 2 in order to investigate heterogeneity in the effect of neighborhood score by ethnicity, age, and gender. We did this by including a term for neighborhood score in each of the equations for the slopes ( $b_{1j}$ – $b_{4j}$ ) one at a time.

Third, a Poisson regression model with a random intercept for each census tract (model 3) was used to model the association between neighborhood score (modeled continuously) and stroke risk with adjustment for age, gender, ethnicity, and education. Similar to model 2, the number of stroke cases and population counts were calculated for cross-classified cells of age (45–64 and  $\geq 65$  years), gender (male and female), ethnicity (Mexican American and non-Hispanic White), and education (high school or more vs. less than high school) within each census tract ( $n = 1,024$ ). Population counts for this model were obtained from Summary File 4 because of the availability of education data

in this file as compared with the Summary File 1 data used for models 1 and 2. Age was modeled as a dummy variable representing persons aged 65 years or older, with persons aged 45–64 years used as the reference group, based on the availability of age and education data in Summary File 4. Models were fitted using the “proc nlmixed” function in SAS, version 9.1.3 (SAS Institute, Inc., Cary, North Carolina). For ease of interpretation, the effect of neighborhood score is summarized as the relative risk comparing the 90th percentile of neighborhood score with the 10th percentile.

### Imputation of education data

Educational attainment data were obtained from in-person interviews with a random sample of the stroke cases. For the remaining cases, years of education were imputed. We employed a multiple imputation technique using the sequential regression imputation method described by Raghunathan et al. (28), using software called IVEware (University of Michigan Survey Research Center, Ann Arbor, Michigan). Missing values for education were drawn from an appropriate distribution depending on the conditional relation of education to ethnicity, age, and gender. The procedure of drawing missing values for education was repeated five times. Because missing values were drawn from a distribution, there was a range of values imputed for each missing education data point, with the variation reflecting uncertainty about the education values. Each of the five complete data sets including imputed values for education was analyzed separately, and the results were combined.

This project was approved by the institutional review boards of the University of Michigan and all participating hospitals.

## RESULTS

There were 2,984 validated cerebrovascular events during the study period. Of the 2,984 events, 2,727 were geocoded to the census tract level. The remaining 8.6 percent were geocoded to the zip code level and were excluded. Of the 2,727 census-level events, 1,581 were ischemic strokes. Only non-Hispanic Whites and Mexican Americans were included in this study because of small numbers for other race-ethnicity groups; this resulted in 1,468 events. For persons with multiple ischemic strokes, only the first ischemic stroke event was considered. Therefore, the final population for analysis was limited to one stroke per individual among non-Hispanic Whites and Mexican Americans, for a total of 1,336 stroke cases. Of these 1,336 cases, 13 patients were nonresidents of Nueces County and were excluded. Because we were interested in measuring exposure to neighborhood SES over the time period relevant to development of stroke, cases who were currently residing in a nursing home or assisted living facility ( $n = 76$ ) were also excluded. This left 1,247 stroke cases for the final analysis.

The median age of the cases was 72.7 years (interquartile range, 63.5–80.6). Mexican Americans comprised 55.3 percent of the cases, while females comprised 51.7 percent. The

**TABLE 1. Distribution of neighborhood socioeconomic status scores and US Census variables across 64 census tracts, Nueces County, Texas, January 2000–June 2003**

	Median	Interquartile range
Neighborhood score*	0.33	–4.89, 4.30
Annual household income (dollars)	36,238	25,732, 44,274
Value of occupied housing units (dollars)	70,150	42,200, 93,450
% of households receiving interest, dividend, or net rental income	24	15, 35
% of adults who completed high school	78	58, 88
% of adults who completed college	15	7, 26
% of persons in managerial or professional occupations	27	16, 38

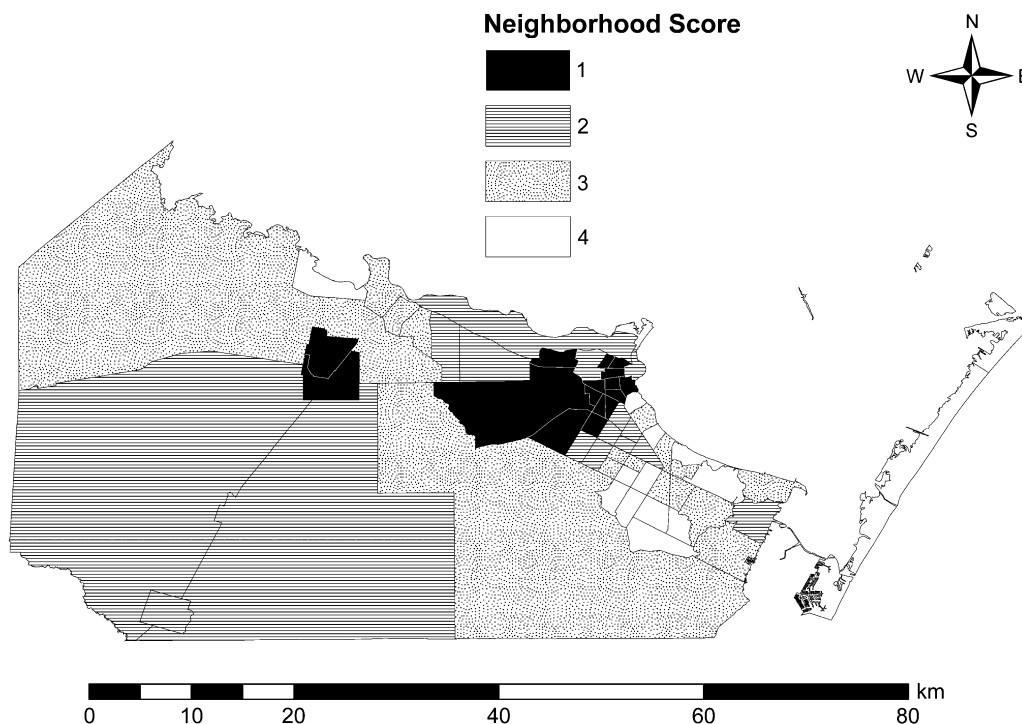
\* Neighborhood scores ranged from –10.49 to 11.89 across the 64 census tracts.

1,247 cases resided in 64 census tracts. Based on census data, the median percentage of persons per census tract who had lived in the same house in 1995 was 54 percent (interquartile range, 44–61). The median percentage of persons per census tract who had resided in the same county in 1995 was 85 percent (interquartile range, 77–89).

Table 1 shows the distribution of the neighborhood summary scores and neighborhood characteristics used to construct the summary score. Neighborhood scores ranged from –10.49 to 11.89 across the 64 census tracts. Figure 1 displays a map of Nueces County with shading based on quartiles of neighborhood score. Neighborhoods in the highest quartile as compared with the lowest quartile had a higher median household income (\$55,501 vs. \$19,622), a higher median percentage of residents with at least a high school education (92 percent vs. 49 percent), and a higher median percentage of residents in management, professional, or related occupations (46 percent vs. 13 percent). Among the cases, Mexican Americans had significantly lower neighborhood scores than non-Hispanic Whites, and age was significantly and positively associated with neighborhood score (table 2).

### Neighborhood score and stroke risk

Table 3 shows results from the unadjusted and adjusted models of neighborhood score. In the unadjusted model (model 1), neighborhood score was significantly associated with stroke risk. The relative risk of stroke comparing the 90th percentile (8.07) of neighborhood score with the 10th percentile (–7.95) was 0.49 (95 percent confidence interval (CI): 0.41, 0.58). In the model using quartiles of neighborhood score (data not shown), the relation between neighborhood score and stroke risk appeared to be graded, with no clear evidence of a threshold (quartile 4: relative risk (RR) = 0.50, 95 percent CI: 0.43, 0.59; quartile 3: RR = 0.66, 95



**FIGURE 1.** Distribution of neighborhood socioeconomic status scores across 64 census tracts, Nueces County, Texas, January 2000–June 2003. Quartiles of neighborhood score: 1 = score  $\leq -4.89$ ; 2 =  $-4.89 < \text{score} \leq 0.33$ ; 3 =  $0.33 < \text{score} \leq 4.30$ ; 4 = score  $> 4.30$ .

percent CI: 0.57, 0.76; quartile 2: RR = 0.71, 95 percent CI: 0.61, 0.82).

After adjustment for age, gender, and ethnicity (model 2), the association between neighborhood score and stroke risk was attenuated (RR = 0.79, 95 percent CI: 0.63, 1.00), but it remained borderline significant ( $p = 0.057$ ). In this model, Mexican Americans were more likely to experience stroke than non-Hispanic Whites (RR = 1.45, 95 percent CI: 1.27, 1.67). Women were less likely to experience stroke than

men (RR = 0.80, 95 percent CI: 0.71, 0.90), and stroke risk increased with advancing age (ages 60–74 years: RR = 3.54, 95 percent CI: 3.02, 4.15; ages  $\geq 75$  years: RR = 7.99, 95 percent CI: 6.82, 9.37).

When effect modification terms were added to model 2 (table 4), there was no evidence of an ethnic difference in the association of neighborhood score with stroke risk ( $p$  for interaction = 0.79). Significant effect modification was found for neighborhood score and age (for overall interaction,  $p < 0.001$ ). The protective effect of neighborhood score was driven by the two youngest age categories (45–59 and 60–74 years). In the oldest age category ( $\geq 75$  years), neighborhood score was positively related to stroke risk. The relative risks of stroke in the three age categories were 0.35 (95 percent CI: 0.23, 0.54) for ages 45–59 years, 0.55 (95 percent CI: 0.38, 0.81) for ages 60–74 years, and 1.60 (95 percent CI: 1.14, 2.26) for ages  $\geq 75$  years. Associations also differed by gender (for males, RR = 0.62, 95 percent CI: 0.46, 0.82; for females, RR = 1.01, 95 percent CI: 0.76, 1.35;  $p$  for interaction = 0.04).

**TABLE 2.** Distribution of neighborhood socioeconomic status scores among stroke cases by ethnicity, gender, and age ( $n = 1,247$ ), Nueces County, Texas, January 2000–June 2003

	Median	Interquartile range	$p$ value
Ethnicity			
Mexican American	-5.08	-6.55, -0.32	
Non-Hispanic White	2.18	-0.04, 5.31	<0.0001
Gender			
Male	-0.38	-5.32, 3.09	
Female	-0.32	-5.32, 3.09	0.3283
Age group (years)			
45–59	-2.62	-6.18, 2.19	
60–74	-0.38	-5.62, 2.73	
$\geq 75$	-0.04	-5.2, 4.17	0.0005

#### Individual-level SES, neighborhood score, and stroke risk

Of the 1,247 stroke cases, we had data on educational attainment for 631 (50.6 percent). In this subset, the mean number of years of education was 9.8 (standard deviation (SD), 4.3). After imputation, the distributions of years of

**TABLE 3. Association of neighborhood socioeconomic status score (90th percentile of neighborhood score vs. 10th percentile) with stroke risk, Nueces County, Texas, January 2000–June 2003**

	Model 1*		Model 2†		Model 3‡	
	RR§	95% CI§	RR	95% CI	RR	95% CI
Neighborhood score	0.49	0.41, 0.58	0.79	0.63, 1.00	1.06	0.81, 1.39
Ethnicity						
Mexican American			1.00		1.00	
Non-Hispanic White			1.45	1.27, 1.67	1.03	0.88, 1.20
Gender						
Male			1.00		1.00	
Female			0.80	0.71, 0.90	0.80	0.71, 0.89
Age group (years)						
45–59			1.00			
45–64					1.00	
60–74			3.54	3.02, 4.15		
≥65					4.05	3.53, 4.63
≥75			7.99	6.82, 9.37		
Education						
No high school					1.00	
High school					0.42	0.35, 0.50

\* Unadjusted.

† Adjusted for age (45–59, 60–74, and ≥75 years), gender, and ethnicity.

‡ Adjusted for age (45–64 and ≥65 years), gender, ethnicity, and education.

§ RR, relative risk; CI, confidence interval.

education were similar in the five imputed data sets (data set 1: mean = 9.7 (SD, 4.1); data set 2: mean = 9.6 (SD, 4.1); data set 3: mean = 9.5 (SD, 4.2); data set 4: mean = 9.7 (SD, 4.0); data set 5: mean = 9.7 (SD, 4.2)).

**TABLE 4. Association of neighborhood socioeconomic status score (90th percentile of neighborhood score vs. 10th percentile) with stroke risk, by gender and age group, Nueces County, Texas, January 2000–June 2003**

	Model 2*		Model 3†	
	RR‡	95% CI‡	RR	95% CI
Gender				
Male	0.62	0.46, 0.82	0.77	0.55, 1.07
Female	1.01	0.76, 1.35	1.43	1.04, 1.98
Age group (years)				
45–59	0.35	0.23, 0.54		
45–64			0.65	0.41, 1.02
60–74	0.55	0.38, 0.81		
≥65			1.28	0.95, 1.73
≥75	1.60	1.14, 2.26		

\* Adjusted for age (45–59, 60–74, and ≥75 years), gender, and ethnicity.

† Adjusted for age (45–64 and ≥65 years), gender, ethnicity, and education.

‡ RR, relative risk; CI, confidence interval.

Additional adjustment for education (model 3, table 3) resulted in no association of neighborhood score with stroke risk (RR = 1.06, 95 percent CI: 0.81, 1.39). In this model, persons with a high school education were 2.4 times less likely to have a stroke than persons without a high school education (RR = 0.42, 95 percent CI: 0.35, 0.50). Because we found significant effect modification for age and gender in model 2, we also stratified the data in model 3 by age and gender (table 4). Similar to the results found in model 2, increasing neighborhood score was associated with a reduced risk of stroke in men but not in women (RR = 0.77, 95 percent CI: 0.55, 1.07), although the association was not significant after adjustment for education. Also similar to model 2, the protective effect of neighborhood score on stroke risk was driven by the youngest age category. After adjustment for education, the relative risks of stroke comparing the 90th percentile of neighborhood score with the 10th percentile in the two age categories were 0.65 (95 percent CI: 0.41, 1.02) for ages 45–64 years and 1.28 (95 percent CI: 0.95, 1.73) for ages ≥65 years.

## DISCUSSION

This study suggests that low neighborhood SES may influence ischemic stroke risk in certain subgroups of the population, including men and persons under 75 years of age. Few studies have investigated the association between SES and stroke, and fewer have considered area-based

measures of deprivation (2). Residing in disadvantaged neighborhoods has been associated with coronary heart disease (10–12) and subclinical cardiovascular disease in the United States (13). The specific causal pathways that may explain the association between neighborhood-level SES and stroke are not clear but are probably similar to those hypothesized for cardiovascular disease (29). Neighborhood economic conditions influence the availability of healthy food and safe recreational areas for exercise, access to health care, attitudes towards healthy behaviors, etc.; all of these factors may contribute to an increased prevalence of stroke risk factors such as hypertension, smoking, high cholesterol, and obesity. Indeed, there is social patterning of risk factors for stroke (9, 30, 31). Thus, traditional risk factors may lie along a causal pathway between neighborhood socioeconomic characteristics and disease risk. This pathway may affect multiple health outcomes, including stroke and coronary heart disease.

Hispanic Americans tend to live in more disadvantaged areas within the United States (16, 17). In this population, neighborhood score was significantly lower among the Mexican American cases than among the non-Hispanic White cases. The effect of ethnicity on stroke risk was attenuated after adjustment for neighborhood score (RR = 1.54 vs. RR = 1.45), but it remained significant. This suggests that neighborhood disadvantage as measured in these analyses may contribute to the ethnic disparity in stroke risk but does not fully account for the difference. After adjustment for education, the effect of ethnicity was nullified, suggesting that individual-level SES may play a larger role in explaining ethnic differences in stroke. However, other risk factors, which may also differ by ethnicity, were not accounted for in the model. More research is needed to understand how neighborhood environment and SES affect ethnic differences in stroke risk.

In this analysis, we were not able to adjust for individual-level stroke risk factors such as hypertension, diabetes, coronary artery disease, and atrial fibrillation, since data on these factors were not available for persons without stroke. This lack of adjustment may have resulted in residual confounding. On the other hand, these factors may also lie within the causal pathway linking neighborhood characteristics to stroke risk; hence, adjusting for these factors in estimating neighborhood effects may not be appropriate. Furthermore, individual-level socioeconomic characteristics were not available for the entire population, making adjustment for these factors problematic. Because individual-level SES has been shown to be a strong predictor of stroke (1), we attempted to adjust for education as a marker of SES using imputation. Although we employed multiple imputation methods to minimize bias, approximately 50 percent of the educational data was imputed, so results based on these data should be interpreted with caution. After adjustment for education, the pattern of increasing stroke risk with decreasing neighborhood advantage did not persist, suggesting that the unadjusted results were confounded. However, in stratified analyses, there was some evidence to suggest that associations may be present in men and in younger age categories. We have no explanation as to why these associations were only present in some subgroups. The hypothesis

that low neighborhood SES may negatively influence stroke risk in some age and gender groups should be formally tested in analyses including data for individual-level SES factors, stroke risk factors, lifestyle factors, and area-level characteristics. If the relation of area characteristics to stroke risk in these groups holds true in further research, it will be important to investigate specific area attributes that contribute to the increased risk so that appropriate interventions can be developed to improve risk profiles.

The method employed in this analysis (multilevel analysis of cross-classified aggregate data) has not been used frequently in the study of neighborhood effects. The method is ideally suited to this area of research, because census data on neighborhood-level socioeconomic characteristics are readily available and are provided in aggregate to maintain confidentiality. The method offers the advantage of allowing investigators to model disease risk in relation to neighborhood characteristics using spatially aggregated data after adjusting for individual-level variables, even when detailed individual-level data on all study subjects are not available (27). Should an association be confirmed, this method could be extended to investigate features of a neighborhood which predispose some people to stroke. Results for the associations between age, gender, and ethnicity and stroke risk employing this method were similar to those of previous reports (19, 32) based on individual-level data, lending support to the findings.

Some limitations of this study warrant discussion. Census tracts are indirect proxies for neighborhoods and may not reflect meaningful neighborhood boundaries or define features of a neighborhood that are relevant to stroke, making it difficult to draw inferences regarding causal relations. An additional limitation is that census tracts in Nueces County vary widely in size. Although we used 2000 US Census data and case ascertainment began in 2000, a temporal relation between neighborhood characteristics and stroke is difficult to establish. We do know that this is a stable community, with a median of 84 percent of persons per census tract having resided in the same county in 1995. However, it is not clear whether current residence is a good marker for residence during the period relevant to development of stroke.

Differences in case ascertainment across neighborhoods are possible. While the majority of stroke patients visit a hospital and our study identified these cases, some stroke patients seek medical care at out-of-hospital locations. We attempted to account for this by including out-of-hospital strokes identified from a random sample. However, because of the small number of out-of-hospital cases identified ( $n = 29$ ), we were not able to account for the sampling procedure in our analysis. We did analyze the distribution of neighborhood scores by case source (out-of-hospital vs. hospital) and found no differences in median score ( $p = 0.40$ ). We have also previously demonstrated that out-of-hospital stroke cases account for less than 6 percent of ischemic stroke cases in this community (19). However, if out-of-hospital cases not identified in our sample varied by neighborhood SES, this would have introduced bias.

Nine percent ( $n = 257$ ) of addresses could not be geocoded to the census tract level and were excluded. If these

cases were different from the remaining cases with regard to neighborhood SES, bias would have been introduced. After manual review of these 257 addresses, we found that 100 cases were not cases of ischemic stroke, four patients experienced another event and were included in our data with a successfully geocoded address, and seven addresses were for nursing homes. Thus, the number of true exclusions for address-related reasons was 153, limiting the bias introduced by this source.

We did not measure environmental factors—the availability of healthy food, safe recreational areas, access to health care, etc.—which may explain the observed association. It would be interesting to determine whether these environmental factors also differ between census tracts and what role they may play in the relation between neighborhood and stroke risk.

In summary, we found that low neighborhood socioeconomic status may influence stroke risk, especially among men and persons under age 75 years. In future research, investigators should test the hypothesis that neighborhood environment influences stroke risk with adjustment for individual-level SES and medical risk factors.

## ACKNOWLEDGMENTS

This study was funded by National Institutes of Health grant RO1 NS38916.

Conflict of interest: none declared.

## REFERENCES

- Jakovljevic D, Sarti C, Sivenius J, et al. Socioeconomic status and ischemic stroke: The FINMONICA Stroke Register. *Stroke* 2001;32:1492–8.
- Cox AM, McKevitt C, Rudd AG, et al. Socioeconomic status and stroke. *Lancet Neurol* 2006;5:181–8.
- Diez-Roux AV. Bringing context back into epidemiology: variables and fallacies in multilevel analysis. *Am J Public Health* 1998;88:216–22.
- Diez Roux AV, Schwartz S, Susser E. Ecologic studies and ecological variables in public health research. In: Detels R, McEwen J, Beaglehole R, et al, eds. *The Oxford textbook of public health*. London, United Kingdom: Oxford University Press, 2002:493–507.
- Schwartz S. The fallacy of the ecological fallacy: the potential misuse of a concept and the consequences. *Am J Public Health* 1994;84:819–24.
- Brown P, Guy M, Broad J. Individual socio-economic status, community socio-economic status and stroke in New Zealand: a case control study. *Soc Sci Med* 2005;61:1174–88.
- Engstrom G, Jerntorp I, Pessah-Rasmussen H, et al. Geographic distribution of stroke incidence within an urban population: relations to socioeconomic circumstances and prevalence of cardiovascular risk factors. *Stroke* 2001;32:1098–103.
- Thrift AG, Dewey HM, Sturm JW, et al. Greater incidence of both fatal and nonfatal strokes in disadvantaged areas: The Northeast Melbourne Stroke Incidence Study. *Stroke* 2006;37:877–82.
- Aslanyan S, Weir CJ, Lees KR, et al. Effect of area-based deprivation on the severity, subtype, and outcome of ischemic stroke. *Stroke* 2003;34:2623–8.
- Diez Roux AV, Merkin SS, Arnett D, et al. Neighborhood of residence and incidence of coronary heart disease. *N Engl J Med* 2001;345:99–106.
- Stokols D, Allen J, Bellingham RL. The social ecology of health promotion: implications for research and practice. *Am J Health Promot* 1996;10:247–51.
- Diez-Roux AV, Nieto FJ, Muntaner C, et al. Neighborhood environments and coronary heart disease: a multilevel analysis. *Am J Epidemiol* 1997;146:48–63.
- Nordstrom CK, Diez Roux AV, Jackson SA, et al. The association of personal and neighborhood socioeconomic indicators with subclinical cardiovascular disease in an elderly cohort. The Cardiovascular Health Study. *Soc Sci Med* 2004;59:2139–47.
- Borrell LN, Diez Roux AV, Rose K, et al. Neighbourhood characteristics and mortality in the Atherosclerosis Risk in Communities Study. *Int J Epidemiol* 2004;33:398–407.
- Diez Roux AV, Borrell LN, Haan M, et al. Neighbourhood environments and mortality in an elderly cohort: results from the Cardiovascular Health Study. *J Epidemiol Community Health* 2004;58:917–23.
- Winkleby MA, Cubbin C. Influence of individual and neighbourhood socioeconomic status on mortality among black, Mexican-American, and white women and men in the United States. *J Epidemiol Community Health* 2003;57:444–52.
- Krieger N, Chen JT, Waterman PD, et al. Race/ethnicity, gender, and monitoring socioeconomic gradients in health: a comparison of area-based socioeconomic measures—The Public Health Disparities Geocoding Project. *Am J Public Health* 2003;93:1655–71.
- Sacco RL, Boden-Albala B, Gan R, et al. Stroke incidence among white, black, and Hispanic residents of an urban community: The Northern Manhattan Stroke Study. *Am J Epidemiol* 1998;147:259–68.
- Morgenstern LB, Smith MA, Lisabeth LD, et al. Excess stroke in Mexican Americans compared with non-Hispanic whites: The Brain Attack Surveillance in Corpus Christi Project. *Am J Epidemiol* 2004;160:376–83.
- Piriyawat P, Smajsova M, Smith MA, et al. Comparison of active and passive surveillance for cerebrovascular disease: The Brain Attack Surveillance in Corpus Christi (BASIC) Project. *Am J Epidemiol* 2002;156:1062–9.
- Bureau of the Census, US Department of Commerce. 2000 Census of Population and Housing. Summary File 1. Technical documentation. Washington, DC: US Department of Commerce, 2003. (<http://www.census.gov/prod/cen2000/doc/sf1.pdf>).
- Asplund K, Tuomilehto J, Stegmayr B, et al. Diagnostic criteria and quality control of the registration of stroke events in the MONICA Project. *Acta Med Scand Suppl* 1988;728:26–39.
- Smith MA, Risser JM, Lisabeth LD, et al. Access to care, acculturation, and risk factors for stroke in Mexican Americans: The Brain Attack Surveillance in Corpus Christi (BASIC) Project. *Stroke* 2003;34:2671–5.
- Bureau of the Census, US Department of Commerce. 2000 Census of Population and Housing. Summary File 4. Technical documentation. Washington, DC: US Department of



- Commerce, 2003. (<http://www.census.gov/prod/cen2000/doc/sf4.pdf>).
25. Krieger N. Overcoming the absence of socioeconomic data in medical records: validation and application of a census-based methodology. *Am J Public Health* 1992;82:703–10.
  26. Diez-Roux AV, Kiefe CI, Jacobs DR Jr, et al. Area characteristics and individual-level socioeconomic position indicators in three population-based epidemiologic studies. *Ann Epidemiol* 2001;11:395–405.
  27. Subramanian SV, Duncan C, Jones K. Multilevel perspectives on modeling census data. *Environ Planning* 2001;33:399–417.
  28. Raghunathan TE, Lepkowski JM, Van Hoewyk J, et al. A multivariate technique for multiply imputing missing values using a sequence of regression models. *Surv Methodol* 2001;27:85–95.
  29. Diez Roux AV. Residential environments and cardiovascular risk. *J Urban Health* 2003;80:569–89.
  30. Smith GD. Down at heart—the meaning and implications of social inequalities in cardiovascular disease. *J R Coll Phys Lond* 1997;31:414–24.
  31. Stern MP, Rosenthal M, Haffner SM, et al. Sex difference in the effects of sociocultural status on diabetes and cardiovascular risk factors in Mexican Americans. The San Antonio Heart Study. *Am J Epidemiol* 1984;120:834–51.
  32. Mohr J, Choi D, Grotta J, et al. Epidemiology of stroke. In: *Stroke: pathophysiology, diagnosis, and treatment*. Philadelphia, PA: Churchill Livingstone, 2004:13–15.