

Building the Evidence—U.S. Approaches

Relationship Between Urban Sprawl and Physical Activity, Obesity, and Morbidity

Reid Ewing, Tom Schmid, Richard Killingsworth, Amy Zlot, Stephen Raudenbush

Abstract

Purpose. To determine the relationship between urban sprawl, health, and health-related behaviors.

Design. Cross-sectional analysis using hierarchical modeling to relate characteristics of individuals and places to levels of physical activity, obesity, body mass index (BMI), hypertension, diabetes, and coronary heart disease.

Setting. U.S. counties (448) and metropolitan areas (83).

Subjects. Adults ($n = 206,992$) from pooled 1998, 1999, and 2000 Behavioral Risk Factor Surveillance System (BRFSS).

Measures. Sprawl indices, derived with principal components analysis from census and other data, served as independent variables. Self-reported behavior and health status from BRFSS served as dependent variables.

Results. After controlling for demographic and behavioral covariates, the county sprawl index had small but significant associations with minutes walked ($p = .004$), obesity ($p < .001$), BMI ($p = .005$), and hypertension ($p = .018$). Residents of sprawling counties were likely to walk less during leisure time, weigh more, and have greater prevalence of hypertension than residents of compact counties. At the metropolitan level, sprawl was similarly associated with minutes walked ($p = .04$) but not with the other variables.

Conclusion. This ecologic study reveals that urban form could be significantly associated with some forms of physical activity and some health outcomes. More research is needed to refine measures of urban form, improve measures of physical activity, and control for other individual and environmental influences on physical activity, obesity, and related health outcomes. (*Am J Health Promot* 2003;18[1]:47–57.)

Key Words: Physical Activity, Urban Design, Sprawl, Obesity, Prevention Research

INTRODUCTION

The links between physical activity and health outcomes are well established. At the time of the Surgeon General's *Report on Physical Activity and Health* in 1996, hundreds of research studies were amassed providing evidence of these links.¹ Physical inactivity contributes to increased risk of many chronic diseases and conditions, including obesity, hypertension, non-insulin-dependent diabetes, colon cancer, osteoarthritis, osteoporosis, and coronary heart disease. Despite the health benefits of physical activity, 74% of U.S. adults do not get enough physical activity to meet public health recommendations and about one in four U.S. adults remains completely inactive during their leisure time.^{2,3}

One consequence of physical inactivity—obesity—has reached epidemic proportions across age, race/ethnic, and socioeconomic groups.^{4,5} Recent data from the National Health and Nutrition Examination Survey (NHANES) found that 64.5% of the U.S. adult population is overweight and almost one in three is obese (30.5%).⁶ Excess weight and physical inactivity are reported to account for over 300,000 premature deaths each year, second only to tobacco-related deaths among preventable causes of death.^{7,8}

There is growing interest in how physical inactivity, obesity, and related chronic health problems are affected by environmental factors. Public health researchers are expanding their horizons, moving beyond individual models of behavior to more inclusive ecologic models that recognize the importance of both physical

Reid Ewing completed this work while with the Bloustein School of Planning and Public Policy, Rutgers University, New Brunswick, New Jersey. He is currently with the National Center for Smart Growth, University of Maryland, College Park, Maryland. Tom Schmid is from the Centers for Disease Control and Prevention, NCCDPHP, DNPA, Physical Activity and Health Branch. Richard Killingsworth is with Active Living By Design, University of North Carolina, Chapel Hill, North Carolina. Amy Zlot is with the Centers for Disease Control and Prevention, NCCDPHP, OIIRM, Atlanta, Georgia. Stephen Raudenbush is with the Department of Education, Department of Statistics, and Survey Research Center, University of Michigan, Ann Arbor.

Send reprint requests to Dr. Reid Ewing, University of Maryland, National Center for Smart Growth, Preinkert Field House, College Park, MD 20742; (301)405-6788 (tel).

This manuscript was submitted November 15, 2002; revisions were requested January 9 and March 19, 2003; the manuscript was accepted for publication June 3, 2003.

Copyright © 2003 by American Journal of Health Promotion, Inc.
0890-1171/03/\$5.00 + 0

and social environments as determinants of health.⁹⁻¹⁴ For physical activity researchers, this interest is relatively new. A review published in 1998 found only seven such studies.¹⁴ Since then, several studies have researched environmental determinants of physical activity.¹⁵⁻¹⁹ One such study found that urban and suburban residents living in homes built before 1946 (a proxy for older neighborhoods) were more likely to walk long distances with some frequency than those living in newer homes.¹⁷ This result was attributed to the greater likelihood of sidewalks, denser interconnected streets, and a mix of business and residential uses in older neighborhoods.

Urban planning and transportation researchers are also expanding their horizons, giving increased attention to how their fields affect human behavior and health.²⁰ In the past decade, more than 50 studies have related aspects of the built environment to travel for utilitarian purposes.²¹ Utilitarian travel is travel not for its own sake but, rather, to engage in activities at the trip end, such as going to work, shopping, or school. It is distinct from leisure time physical activity such as walking for exercise, an end in itself. Several recent studies have focused on the relationship between the built environment and the choice of travel mode (e.g., driving a car, taking a bus, or walking).²²⁻²⁹ Walking for utilitarian purposes is consistently found to be more prevalent in dense, mixed-use neighborhoods when compared to lower density, exclusively residential neighborhoods.³⁰ For example, a study of two pairs of neighborhoods in the San Francisco Bay Area concluded that walking trips to commercial areas were more frequent in the older neighborhoods with nearby stores and grid-like street networks than in the newer more homogeneous neighborhoods.²⁹

The quantity and quality of such studies, although based on cross-sectional and case study designs, are increasing and some of these studies are now being reviewed as part of the evidence base for the *Guide for Community Preventive Services*. Developed by public health experts, this

guide will recognize the importance of community design in promoting leisure time physical activity.³¹

The study reported in this paper measured urban form at the county and metropolitan levels. Urban form at these levels is often characterized as more or less "sprawling." *Poor accessibility* is the common denominator of urban sprawl—nothing is within easy walking distance of anything else.³² Although variously defined by others, we consider sprawl to be any environment characterized by (1) a population widely dispersed in low-density residential development; (2) rigid separation of homes, shops, and workplaces; (3) a lack of distinct, thriving activity centers, such as strong downtowns or suburban town centers; and (4) a network of roads marked by large block size and poor access from one place to another. Compact development is the antithesis of sprawl, keeping complementary uses close to one another.

Our working hypotheses, based on the planning and public health literature, were that residents of sprawling places would (1) walk less, (2) weigh more, and (3) have higher prevalence of health problems linked to physical inactivity than those living in more compact places. These hypotheses were tested using data from the Behavioral Risk Factor Surveillance System (BRFSS) for 1998 to 2000.

METHODS

Design

The research design in this study was cross-sectional and ecologic. The degree of sprawl within counties or metropolitan areas was related to levels of physical activity, obesity, body mass index (BMI), hypertension, diabetes, and coronary heart disease (CHD) for BRFSS respondents from these particular counties or metropolitan areas. Hierarchical linear and nonlinear modeling (HLM) methods were used to control for covariates, such as age, race/ethnicity, and education, at the individual level while examining the effects of sprawl at the population level.

Behavioral and health status variables extracted from BRFSS are listed

in Table 1. All data are self-reported. A condition was assumed to exist if a health care practitioner had told the respondent that he or she had the condition.

Three leisure time physical activity variables served as dependent or outcome variables: *any physical activity*, reporting any amount of leisure time physical activity over the past month; *recommended physical activity*, getting the recommended levels of physical activity in the past month; and *minutes walked*, total minutes of walking as leisure time physical activity in the past month. A person was considered to have met the physical activity recommendations if she or he reported ≥ 30 minutes of moderate-level physical activity on ≥ 5 days of the week or if he or she reported ≥ 20 minutes of vigorous activity on ≥ 3 days of the week. Walking was emphasized because of its documented relationship to urban form and its dominance as a leisure time activity (reported with almost six times the frequency of the next most common leisure time activity, gardening).

Two weight-related measures were included as outcome variables: *BMI* and *obesity*. BMI was defined as weight in kilograms divided by height in meters squared (kg/m^2) and obesity was defined as a BMI of ≥ 30.0 .

Three health status variables were also modeled: *hypertension*, *diabetes*, and *coronary heart disease (CHD)*. These three were selected for their known relationships to inactivity and obesity.

Unless otherwise noted, gender, race/ethnicity, education, age, smoking status, and fruit and vegetable consumption were included in models as individual-level covariates. The reference groups for sociodemographic variables were females, white non-Hispanics, college graduates, and persons aged 18 to 30 years. Race/ethnicity, for example, was represented by three dummy variables (1 if yes, 0 otherwise): black non-Hispanic, Hispanic, and Asian or other race. In this case, white non-Hispanics were the reference group (for other covariates, see Table 2).

A *metropolitan sprawl index*, developed for Smart Growth America (SGA), was used in this study to mea-

Table 1

Sample Sizes (n), Means, and Standard Deviations (SD) for Health Behavior and Health Status Variables, 1998 to 2000*

	n for County Models With All Covariates (N = 206,992)†	n for Metropolitan Models With All Covariates (N = 175,609)	Means (SD) for County Models	Means (SD) for Metropolitan Models
Any physical activity‡	149,835	126,893	0.730 (0.444)	0.733 (0.442)
Recommended physical activity§	135,344	115,006	0.268 (0.443)	0.273 (0.445)
Minutes walked	147,305	124,764	247.8 (493.3)	251.2 (499.6)
Body mass index (BMI, kg/m ²)	137,263‡‡	116,779‡‡	26.06 (5.15)	26.03 (5.15)
Obesity¶	137,409‡‡	116,913‡‡	0.181 (0.385)	0.181 (0.385)
Hypertension#	85,465	68,927	0.239 (0.426)	0.235 (0.424)
Diabetes**	142,685‡‡	121,292‡‡	0.056 (0.230)	0.055 (0.228)
Coronary heart disease††	40,651	31,563	0.042 (0.201)	0.041 (0.197)

* For exact wording of Behavioral Risk Factor Surveillance System (BRFSS) questions and to see how calculated variables were determined, go to <http://www.cdc.gov/brfss/calcvvars.htm>.

† N, initial sample before any BRFSS variables entered.

‡ Reported any leisure time physical activity in the last month.

§ Met recommended level of physical activity in the last month: Recommended amount is 30 minutes of moderately intense physical activity at least 5 days per week and/or 20 minutes of vigorously intense physical activity at least 3 days per week.

|| Minutes walked for leisure during last month.

¶ BMI \geq 30.

Ever been told had hypertension.

** Ever been told had diabetes.

†† Ever been told had coronary heart disease.

‡‡ Includes fruit and vegetable consumption as a covariate, which reduced sample size.

Table 2

Sociodemographic and Behavioral Covariates From BRFSS Surveys*

Gender	Male (dichotomous)
Age	Ages 18 to 29, 30 to 44, 45 to 64, 65 to 74, 75+ (categorical)
Race/ethnicity	White non-Hispanic, black non-Hispanic, Hispanic, other race (categorical)
Education	College graduate, some college, high school graduate, less than high school (categorical)
Smoking	Currently smoke (dichotomous)
Diet	Fruit or vegetable consumption three or more times per day (dichotomous)

* For exact wording of Behavioral Risk Factor Surveillance System (BRFSS) questions and to see how calculated variables were determined, go to <http://www.cdc.gov/brfss/calcvvars.htm>.

sure urban form at the metropolitan level. The metropolitan sprawl index is a linear combination of 22 land use and street network variables. A simpler *county sprawl index* was used to measure urban form at the county level. It is a linear combination of six variables from the larger set, these six being available for counties, whereas many of the larger set are available only for metropolitan areas. The derivation of these indices is described in the "Measures" section.

Sample

BRFSS surveys for 1998, 1999, and 2000 provided data on leisure time physical activity levels, BMI and obesity, hypertension, diabetes, and CHD.³³ Our samples consisted of 206,992 respondents from 448 counties and 175,609 respondents from 83 metropolitan areas for the pooled 1998, 1999, and 2000 BRFSS surveys. These respondents were selected from the larger BRFSS samples because they had known places of resi-

dence for which urban sprawl indices were available. Hence, it was possible to link urban sprawl indices directly to health data for all respondents. Data for 3 years were pooled to increase the statistical power of the analysis.

Metropolitan areas, as defined by the U.S. Office of Management and Budget, consist of one or more counties having a high degree of economic and social integration with one another. Our sample of respondents was smaller at the metropolitan than county level because metropolitan sprawl indices were available only for the largest metropolitan areas (500,000 population or more) with complete urban form datasets (all 22 variables that make up the metropolitan sprawl index). The sample of respondents at the county level included residents of counties that are part of smaller metropolitan areas, metropolitan areas with only partial datasets (although always with all six variables that make up the county sprawl index), or both.

As illustrated in Table 1, actual sample sizes varied among BRFSS

outcome measures because of missing responses and exclusion of certain questions in certain years. For instance, physical activity data were collected by all states in 1998 and 2000 but only by certain states in 1999. Although diabetes data were gathered by all states for all 3 years, many cases were lost because fruit and vegetable consumption was included as an explanatory variable in the diabetes analysis. Fruit and vegetable consumption data were collected by all states in 1998 and 2000 but by only a small subset of states in 1999.

Sample sizes for individual counties ranged from 6 to 6429, with 353 counties having samples of 50 or more. Sample sizes were more than adequate to support stable and powerful statistical analysis. HLM uses the method of maximum likelihood to optimally combine information from different samples. In this study, counties with small samples contributed less information to the estimation of parameters than counties with large samples. Because maximum likelihood took into account the information from each county and because the number of counties in this study was large ($n = 448$), counties with small samples were not problematic from a statistical standpoint.^{34–36}

Measures

BRFSS is a population-based, random digit-dialed telephone survey administered to U.S. civilian noninstitutionalized adults aged ≥ 18 years. For the years under study, BRFSS collected data from 150,000 to 185,000 respondents in the 50 states and the District of Columbia. Surveys consisted of a core module of questions asked annually, a rotating core asked every other year, and optional modules asked at states' discretion. A recent review found high reliability and validity for demographic questions (e.g., age, sex, race) and moderate to high reliability and validity for behavioral and health status questions (e.g., hypertension, diabetes, level of physical activity, weight, BMI, fruit and vegetable consumption).³⁷ Further information on specific questions and how variables were calculated

can be found at <http://www.cdc.gov/brfss/index.htm>.

Smart Growth America's *metropolitan sprawl index*, used in this study to represent urban form at the metropolitan level, is the most comprehensive representation of sprawl for metropolitan areas yet developed. Technical details, including operational definitions, are available in the full technical report at the SGA web site (www.smartgrowthamerica.org).

To construct the index, 83 metropolitan areas in the United States with a total population of more than 150 million people in 2000, over half the U.S. population, were rated in four urban form dimensions. For each dimension, a composite factor was extracted from several observed variables via principal components analysis.

- *Residential density* was defined in terms of gross and net densities and proportions of population living at different densities; seven variables made up the metropolitan density factor.
- *Land use mix* was defined in terms of the degree to which land uses are mixed and balanced within subareas of the region; six variables made up this factor.
- *Degree of centering* was defined as the extent to which development is focused on the region's core and regional subcenters; six variables made up this factor.
- *Street accessibility* was defined in terms of the length and size of blocks; three variables made up this factor.

The four factors were combined into an overall index by summing them and then adjusting for the size of the metropolitan area. The four were given equal weight in the overall index. Scores were then converted to a scale with a mean of 100 and standard deviation of 25. The bigger the value of the index, the more compact the metropolitan region. The smaller the value, the more sprawling the metropolitan region. A few metropolitan regions are compact in all dimensions; New York, New York; San Francisco, California; Boston, Massachusetts; and Portland, Oregon rank near the top in overall

score. Others near the top, despite one factor score below average, include Jersey City, New Jersey; Providence, Rhode Island; Honolulu, Hawaii; and Omaha, Nebraska. A few regions sprawl badly in all dimensions. These include Atlanta, Georgia; Raleigh-Durham and Greensboro-Winston-Salem-High Point, North Carolina; and Riverside-San Bernardino, California. They rank at or near the bottom in overall score.

In an earlier study, the metropolitan sprawl index was found to have good explanatory power. The index explained a significant proportion of the variance across metropolitan areas in percent walking or taking transit to work, average vehicle ownership, vehicle miles traveled per capita, traffic fatality rates, and ground-level ozone concentration.³⁸

In order to examine the effects of urban form at a finer geographic scale, we developed a *county sprawl index* using a process similar to that used to develop the *metropolitan sprawl index*. The county is the smallest geographic unit that can be matched to BRFSS data. The index was estimated for 448 metropolitan counties or statistically equivalent entities (e.g., independent towns and cities). These counties comprised the 101 most populous metropolitan statistical areas, consolidated metropolitan statistical areas, and New England county metropolitan areas in the United States as of the 1990 census, the latest year for which metropolitan boundaries were defined as this study began. Nonmetropolitan counties and metropolitan counties in smaller metropolitan areas were excluded from the sample. More than 183 million Americans, nearly two thirds of the U.S. population, lived in these 448 counties in 2000.

Although sprawl has the four characteristics noted above, only two of these could be measured at the county level: low residential density and poor street accessibility. Six variables became part of the county sprawl index (as shown in Table 3). We used U.S. Census data³⁹ to derive three population density measures for each county: (1) gross population density (persons per square mile); (2) percentage of the county popula-

Table 3
County Sprawl Index Variables and Factor Loadings

Observed Variable	Factor Loading*
Gross population density in persons per square mile	0.846
% of population living at densities <1500 persons per square mile	-0.698
% of population living at densities >12,500 persons per square mile	0.846
County population divided by the amount of urban land in square miles	0.849
Average block size in square miles	-0.698
% of blocks 1/100 of a square mile or less in size (about 500 feet on a side, a traditional urban block)	0.821

* Correlation with county sprawl index.

tion living at low suburban densities, specifically, densities between 101 and 1499 persons per square mile, corresponding to less than one housing unit per acre; and (3) percentage of the county population living at moderate to high urban densities, specifically, more than 12,500 persons per square mile, corresponding to about eight housing units per acre, the lower limit of density needed to support mass transit. When deriving these county population density measures, we excluded census tracts with fewer than 100 inhabitants per square mile (corresponding to rural areas, desert tracts, and other undeveloped lands) located within the county, because we were only concerned about sprawl in developed areas where the vast majority of residents live. A fourth density variable, the net density in urban areas, was derived from estimated urban land area for each county from the Natural Resources Inventory of the U.S. Department of Agriculture.⁴⁰

Data reflecting street accessibility for each county were obtained from the U.S. Census, based on information concerning block size.⁴¹ A census block is defined as a statistical area bounded on all sides by streets, roads, streams, railroad tracks, or geopolitical boundary lines, in most cases. A traditional urban neighborhood is composed of intersecting bounding roads that form a grid, with houses built on the four sides of the block, facing these roads. Therefore, the length of each side of that block, and therefore its block size, is relatively small. By contrast, a contemporary suburban neighborhood

does not make connections between adjacent cul-de-sacs or loop roads. Instead, local streets only connect with the road at the subdivision entrance, which is on one side of the block boundary. Thus, the length of a side of this block is quite large, and the block itself often encloses multiple subdivisions to form a superblock a half mile or more on a side. Large block sizes indicate a relative paucity of street connections and alternate routes.

For each county, we calculated (1) average block size and (2) percentage of blocks with areas less than 1/100 square mile, the size of a typical traditional urban block bounded by sides just over 500 feet in length. Tracts with blocks larger than 1 square mile were excluded from these calculations because they were likely to be in rural or other undeveloped areas.

The six variables were combined into one factor representing degree of sprawl within the county. This was accomplished via principal components analysis. The principal component selected to represent sprawl was the one capturing the largest share of common variance among the six variables (i.e., the one on which the observed variables loaded most heavily; see Table 3). This one component explained almost two thirds of the variance in the dataset (63.4%). Because this component captured the majority of the combined variance of these variables, no subsequent components were considered.

To derive a county sprawl index, we transformed the principal component, which had a mean of 0 and

standard deviation of 1, to a scale with a mean of 100 and standard deviation of 25. This transformation produced a more familiar metric (like an IQ scale) and ensured that all values would be positive, thereby enhancing our ability to test nonlinear relationships.

The bigger the value of the index, the more compact the county. The smaller the value, the more sprawling the county. Scores ranged from a high of 352 to a low of 63. At the most compact end of the scale were four New York City boroughs—Manhattan, Brooklyn, The Bronx, and Queens; San Francisco County, California; Hudson County (Jersey City), New Jersey; Philadelphia County, Pennsylvania; and Suffolk County (Boston), Massachusetts. At the most sprawling end of the scale were outlying counties of metropolitan areas in the Southeast and Midwest, such as Goochland County in the Richmond, Virginia, metropolitan area and Geauga County in the Cleveland, Ohio, metropolitan area. The county sprawl index is positively skewed. Most counties clustered around intermediate levels of sprawl. In the United States, few counties approach the densities of New York or San Francisco counties. (A complete list of counties and their respective “sprawl” scores is available on request.)

Analysis

In this cross-sectional, ecologic study, relationships between urban sprawl and leisure time physical activity levels, BMI and obesity, hypertension, diabetes, and CHD were estimated with HLM 5 (Hierarchical Linear and Nonlinear Modeling) software.⁴² Many BRFSS respondents share characteristics of a given place, which tends to produce a dependence among respondents, violating the independence assumption of ordinary least squares (OLS) regression. Standard errors of regression coefficients associated with place characteristics based on OLS will consequently be underestimated. Moreover, OLS regression coefficient estimates will be inefficient. Hierarchical (multilevel) modeling overcomes these limitations, accounting for the dependence among respon-

dents residing in a given place and producing more accurate standard error estimates.⁴³

The hierarchical models estimated in this study can be characterized as pairs of linked statistical models. At the first level, respondent health status or behavior were modeled within each place as a function of respondent characteristics plus a random error. Thus, each place had a place-specific regression equation that described the association between respondent characteristics and respondent health status or behavior within that place. At the second level, the place-specific intercept and coefficients were conceived as outcomes and were modeled in terms of place characteristics plus random effects.

In some HLM models, only the place-specific intercepts vary across places, and all of the place-specific regression coefficients are invariant across places. These are often termed “random intercept” models to denote that only the intercept randomly varies. In other HLM models, the place-specific regression coefficients randomly vary as well. These are often termed “random coefficient” models.

In this study, all models initially assumed the random intercept form. Only the intercept term in the place-specific model was allowed to vary, and all place-specific coefficients were taken as fixed. Then this assumption was relaxed, and coefficients were allowed to vary as a function of place characteristics, effectively permitting interactions between place and respondent characteristics.

Interactions between place and respondent characteristics were seldom significant and never sufficiently large to appreciably affect the relationships between place characteristics and outcome variables. Hence, the only results reported are for random intercept models.

Linear models were estimated for continuous outcome variables such as minutes of walking per month. Non-linear models were estimated for the binary outcomes such as meeting or failing to meet recommended physical activity levels; specifically, the log-odds of the outcome was equated to

a linear function of the explanatory variables.

Using HLM software, we were able to apply BRFSS final weights to observations, thereby partially accounting for different probabilities of sample selection and survey response. However, we were unable to account for the complex cluster and stratified sample survey designs used by state health departments when conducting the health surveys on which this study relies. This capability lies beyond the current HLM software.^a

RESULTS

County-level Analysis

Physical Activity Outcomes. As has been found in previous research, the likelihood of engaging in any leisure time physical activity in the past month was greater for males than females and for white non-Hispanics than other races/ethnicities. The likelihood declined with age and increased with educational attainment (see Table 4 for regression coefficients, *t*-ratios, and significance levels).

The likelihood of engaging in recommended levels of physical activity in the past month followed a similar pattern, with one exception. Those age 65 or older were more likely to meet recommended levels than were younger adults because of the greater amount of leisure time walking they do.

The amount of leisure time walking was greater for females than males and increased with age up to 75 years. Education was positively associated with minutes of walking and being physically active in general.

Controlling for these covariates, the likelihood of reporting any leisure time physical activity was not significantly related to the county index ($t = 1.01, p = .313$). The likelihood of getting recommended levels of physical activity was related to the county index, but just short of the traditional .05 probability level ($t = 1.94, p = .052$). The number of min-

utes walked varied directly with the county index, with residents of more compact places reporting more leisure time walking than residents of more sprawling places. The difference was not large but was statistically significant ($t = 2.95, p = .004$).

All else being equal, residents of a county one standard deviation (25 units) above the mean county index would be expected to walk for leisure 14 minutes more each month compared to residents of a county one standard deviation below the mean (i.e., 50 units \times 0.275 minutes per unit). Comparing the extremes (New York County with an index of 352 and Geauga County with an index of 63), New York residents would be expected to walk for leisure 79 minutes more each month.

Weight-related Outcomes. BMI was higher for males than females; increased with age up to middle age (45 to 64 years), and then declined; was higher for blacks and Hispanics than for whites and lower for other races (primarily Asian); was higher for the less educated relative to the college educated; was lower for smokers than nonsmokers; and was lower for those who consume three or more servings of fruits and vegetables daily (Table 5).

After controlling for these covariates, the county index was related to BMI in the expected direction and at a highly significant level ($t = -2.84, p = .005$). Residents of a more compact county, one standard deviation above the mean county index, would be expected to have BMIs 0.17 kg/m² lower than residents of a more sprawling county, one standard deviation below the mean (i.e., 50 \times -.00344). Again, comparing the extremes, New York residents would have BMIs almost 1 kg/m² less than their counterparts in Geauga County. For the BRFSS sample mean BMI (26.1 kg/m²), this translates into 6.3 fewer pounds of body weight.

The binary variable obesity was also modeled and had a highly significant relationship to the county index ($t = 4.24, p < .001$). The odds of being obese in a more compact county, one standard deviation above the mean county index, were 0.90

^aTo account for the complexities of the BRFSS sample survey design, the package of choice is SUDAAN. However, SUDAAN software is not capable of multilevel modeling, a significant shortcoming in a study of this sort.

Table 4

Relationship Between Individual Characteristics, County Sprawl Index, and Leisure Time Physical Activity, 1998 to 2000 (With Coefficients, *t*-ratios, and Significance Levels)

	Any Physical Activity			Recommended Physical Activity			Minutes Walked		
	Coefficient	<i>t</i>	<i>p</i>	Coefficient	<i>t</i>	<i>p</i>	Coefficient	<i>t</i>	<i>p</i>
Male	0.246	12.1	<0.001	0.087	4.44	<0.001	-82.5	-22.1	<0.001
Age 30 to 44	-0.396	-14.7	<0.001	-0.228	-8.17	<0.001	39.4	7.95	<0.001
Age 45 to 64	-0.596	-17.5	<0.001	-0.159	-5.68	<0.001	102.2	14.9	<0.001
Age 65 to 74	-0.639	-13.6	<0.001	0.054	1.38	0.167	139.7	16.4	<0.001
Age 75+	-1.067	-26.7	<0.001	0.187	4.78	<0.001	74.1	6.65	<0.001
Black non-Hispanic	-0.322	-10.9	<0.001	-0.176	-4.96	<0.001	4.24	0.62	0.537
Hispanic	-0.625	-14.7	<0.001	-0.217	-6.15	<0.001	-27.6	-3.58	0.001
Other race	-0.553	-9.43	<0.001	-0.276	-4.49	<0.001	-37.8	-3.26	0.001
Some college	-0.417	-13.3	<0.001	-0.226	-10.3	<0.001	-8.33	-1.66	0.097
High school graduate	-0.854	-31.8	<0.001	-0.525	-21.1	<0.001	-19.8	-3.74	<0.001
Less than high school	-1.353	-39.6	<0.001	-0.946	-20.9	<0.001	-65.3	-9.24	<0.001
Currently smoke	-0.357	-15.7	<0.001	-0.273	-11.0	<0.001	-5.65	-1.16	0.245
County sprawl index	0.000552	1.01	0.313	0.000872	1.94	0.052	0.275	2.95	0.004

Table 5

Relationship Between Individual Characteristics, County Sprawl Index, and Weight, 1998 to 2000 (With Coefficients, *t*-ratios, and Significance Levels)

	Body Mass Index			Obesity		
	Coefficient	<i>t</i>	<i>p</i>	Coefficient	<i>t</i>	<i>p</i>
Male	1.190	22.4	<0.001	0.0535	2.07	0.038
Age 30 to 44	1.696	27.7	<0.001	0.578	16.0	<0.001
Age 45 to 64	2.547	43.0	<0.001	0.852	24.2	<0.001
Age 65 to 74	1.995	23.5	<0.001	0.574	12.3	<0.001
Age 75+	0.517	6.29	<0.001	0.0542	0.98	0.327
Black non-Hispanic	1.604	20.1	<0.001	0.563	17.5	<0.001
Hispanic	0.744	8.71	<0.001	0.308	6.45	<0.001
Other race	-1.075	-10.2	<0.001	-0.448	-7.32	<0.001
Some college	0.818	14.7	<0.001	0.397	13.7	<0.001
High school graduate	1.102	17.9	<0.001	0.520	17.0	<0.001
Less than high school	1.693	19.7	<0.001	0.758	17.4	<0.001
Currently smoke	-0.985	-16.6	<0.001	-0.381	-11.4	<0.001
Fruit/vegetable consumption	-0.327	-7.54	<0.001	-0.154	-5.94	<0.001
County sprawl index	-0.00344	-2.84	0.005	-0.00212	-4.24	<0.001

times the odds in a more sprawling county, one standard deviation below the mean index (95% CI, 0.86 to 0.95). Table 6 reports odds ratios and confidence intervals for all binary outcome variables.

Morbidity Outcomes. Males were more likely to report having diabetes and coronary heart disease than were females. The probability of having these conditions, as well as hypertension, generally increased with age. The probability of having hyperten-

sion and diabetes generally decreased with educational attainment. Probabilities varied with race in more complex ways (Table 7).

The only morbidity outcome statistically linked to sprawling places was hypertension ($t = -2.37, p = .018$). The odds of suffering from hypertension in a more compact county, one standard deviation above the mean sprawl index, was 0.94 times the odds in a more sprawling county, one standard deviation below the mean index (95% CI, 0.90 to 0.99). As for diabe-

tes and coronary heart disease, the county index had the expected sign in both equations, but the relationships were not statistically significant.

Direct and Indirect Effects on BMI and Obesity. To explore the mechanisms by which sprawl affects BMI and obesity, additional analyses were conducted that included *minutes walked* as an independent variable in the level-1 equations for both BMI and obesity. We wanted to see whether living in compact counties was independently

Table 6

Odds of Leisure Time Physical Activity, Obesity, and Morbidity One Standard Deviation Above the Mean County Sprawl Index Compared to One Standard Deviation Below, 1998 to 2000

	Odds Ratio (95% Confidence Interval)
Any physical activity	1.028 (0.974–1.084)
Recommended physical activity	1.045 (0.996–1.092)
Obesity	0.899 (0.856–0.945)
Hypertension	0.942 (0.897–0.990)
Diabetes	0.971 (0.930–1.014)
Coronary heart disease	0.994 (0.988–1.000)

related to weight, after controlling for the amount of reported leisure time walked. Results are presented in Table 8. Both variables—minutes walked and county index—were significantly (and independently) associated with BMI. BMI declined as leisure time walking increased at the individual level, and BMI declined as the county index increased at the population level. The same pattern applied to the binary variable obesity.

Thus, sprawl appears to have direct relationships to BMI and obesity, plus indirect relationships through the number of minutes walked, which varies with the county sprawl index. A portion of the overall sprawl-weight relationship is mediated through the amount of time people spend walking

for leisure. The direct effect is much stronger. A 25-unit increase in the county index (1 SD) is associated directly with a .085 kg/m² (25 × .00338) decrease in BMI. The same 25-unit increase is associated indirectly with only a .001 kg/m² (25 × 0.275 × .000128) decrease in BMI through its effect on leisure time walking.

Metropolitan-level Analysis

We also examined relationships between sprawl at the metropolitan level and health and health-related behaviors (see Table 9). The metropolitan sprawl index proved significantly related to only one outcome variable, *minutes walked* as a leisure time activity ($t = 2.09, p = .04$). Model coefficients for the county and metropolitan

sprawl indices can be compared because they were standardized on the same basis, with means of 100 and standard deviations of 25. In most cases, the county index was more strongly associated with outcomes than was the metropolitan index.

DISCUSSION

This ecologic study reveals that urban form could be significantly associated with some forms of physical activity and with some health outcomes. After controlling for demographic and behavioral covariates, the county sprawl index had small but significant associations with minutes walked ($p = .004$), obesity ($p < .001$), BMI ($p = .005$), and hypertension ($p = .018$). Those living in sprawling counties were likely to walk less, weigh more, and have greater prevalence of hypertension than those living in compact counties. At the metropolitan level, sprawl was similarly associated with minutes walked ($p = .04$) but not with the other variables.

Although the magnitude of the effects observed in this study are small, they do provide added support for the hypothesis that urban form affects health and health-related behaviors. Furthermore, as Geoffrey Rose has pointed out, even a small shift in

Table 7

Relationship Between Individual Characteristics, County Sprawl Index, and Morbidity, 1998 to 2000 (With Coefficients, *t*-ratios, and Significance Levels)

	Hypertension			Diabetes			Coronary Heart Disease		
	Coefficient	<i>t</i>	<i>p</i>	Coefficient	<i>t</i>	<i>p</i>	Coefficient	<i>t</i>	<i>p</i>
Male	0.0191	0.74	0.46	0.221	7.32	<0.001	0.0207	13.1	<0.001
Age 30 to 44	0.689	16.8	<0.001	1.064	12.8	<0.001	0.0164	10.4	<0.001
Age 45 to 64	1.778	44.5	<0.001	2.435	31.7	<0.001	0.0594	31.8	<0.001
Age 65 to 74	2.435	52.2	<0.001	2.958	37.2	<0.001	0.0949	18.3	<0.001
Age 75+	2.456	48.9	<0.001	2.736	34.8	<0.001	0.123	19.8	<0.001
Black non-Hispanic	0.597	15.5	<0.001	0.731	18.5	<0.001	0.0167	-4.19	<0.001
Hispanic	-0.101	-1.10	0.27	0.413	7.38	<0.001	-0.0304	-5.40	<0.001
Other race	0.0203	0.24	0.81	0.284	3.06	0.003	-0.0168	-3.07	0.003
Some college	0.253	7.41	<0.001	0.361	8.33	<0.001	0.0162	7.47	<0.001
High school graduate	0.287	7.73	<0.001	0.383	9.66	<0.001	0.0128	5.93	<0.001
Less than high school	0.427	11.4	<0.001	0.869	18.2	<0.001	0.0680	8.62	<0.001
Currently smoke	-0.0454	-1.51	0.13	-0.232	-5.68	<0.001	-0.00087	-0.43	0.67
Fruit/vegetable consumption	—	—	—	0.0909	2.63	0.009	—	—	—
County sprawl index	-0.00119	-2.37	0.018	-0.00059	-1.32	0.19	-0.00011	-1.82	0.069

the distribution at the population level can have important public health implications.⁴⁴

Heretofore, BRFSS data have not generally been used to examine county- or metropolitan-level relationships. In this study, the consistency of findings with those generally found in previous research on associations between health outcomes and covariates, such as gender, age, and race/ethnicity, provides some assurance that our observations on health and urban form also have validity.

Our finding that relationships are stronger for the county index than for the larger scale metropolitan index is not surprising. Most metropolitan areas consist of multiple counties whose built environments vary significantly between central and outlying counties. The county environment might be more representative of what is actually experienced on a day-to-day basis by residents than is the overall metropolitan environment. By implication, as research shifts from the macroscale (metropolitan and county) to the meso- and microscales (community and neighborhood), we might expect that the explanatory power of environmental variables to predict outcomes will improve.

This study is exploratory and subject to important limitations that call for additional research.

- Because this study is ecologic and cross-sectional in nature, it is premature to imply that sprawl causes obesity, hypertension, or any other health condition. Our study simply indicates that sprawl is associated with certain outcomes. Future research using quasi-experimental designs is needed to tackle the more difficult job of testing for causality.
- As shown in Figure 1, the presumptive relationships between environment (urban form), physical activity, and health are multiple and complex. In particular, leisure time physical activity constitutes only one of four major sources of physical activity, the others being related to occupation, household, and transportation. Greater precision in characterizing physical activity will help disentangle the effects of urban form on health. Recognizing the need to

Table 8
Relationship of County Sprawl Index and Leisure Time Walking to Body Mass Index (BMI) and Obesity, 1998 to 2000*

	County Index			Minutes Walked		
	Coefficient	t	p	Coefficient	t	p
BMI	-0.00338	-2.87	0.005	-0.000128	-2.93	0.004
Obesity	-0.00216	-4.35	<0.001	-0.000061	-2.30	0.022

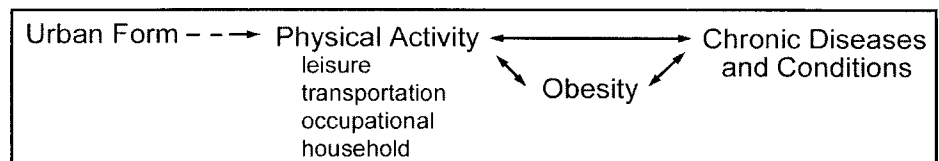
* Models included gender, age, race, education, smoking status, fruit and vegetable consumption, and minutes of walking for leisure as level-1 covariates.

Table 9
Comparison of Relationships of County and Metropolitan Sprawl Indices to Leisure Time Physical Activity, Obesity, and Morbidity Outcomes, 1998 to 2000*

	County Index			Metropolitan Index		
	Coefficient	t	p	Coefficient	t	p
Any physical activity	0.000552	1.01	0.313	0.000760	0.83	0.411
Recommended physical activity	0.000872	1.94	0.052	0.00141	1.49	0.139
Minutes Walked	0.275	2.95	0.004	0.338	2.09	0.040
BMI	-0.00344	-2.84	0.005	-0.00142	-1.03	0.307
Obesity	-0.00212	-4.24	<0.001	-0.000800	-1.02	0.312
Hypertension	-0.00119	-2.37	0.018	-0.000325	-0.49	0.626
Diabetes	-0.000586	-1.32	0.187	-0.000400	-0.60	0.548
Coronary heart disease	-0.000113	-1.82	0.069	na		

* Models included gender, age, race, education, and smoking status as level-1 covariates. Models for body mass index (BMI), obesity, and diabetes also included fruit and vegetable consumption.

Figure 1
Established (Solid) and Speculative (Dashed) Relationships



- monitor more than just leisure time physical activity, the 2001 BRFSS questions were modified to include transportation-, household-, and work-related physical activity.
- In this study, we were not able to account for the complex nature of the BRFSS sampling design, reinforcing the need for cautious interpretation of these early findings. There is growing interest in using BRFSS at the local level, and CDC

is in the process of developing methods to adjust the state-based weights for use at the local level.

- Better measures of walking are needed to improve our ability to trace potential differences that are attributable to urban form. The variable *minutes walked* is based on people who reported walking as one of their top two forms of leisure time activity. It excludes walking as a less frequent form of leisure time activity

or walking for other purposes. The new BRFSS questions should help produce more comprehensive measures of walking.

- We recognize that the relationships between sprawl and behavior or weight are probably not completely linear. It might be that certain thresholds or critical levels of “compactness” are needed before community design begins to have a palpable influence on physical activity—increasing density from one or two houses per acre to three or four might not meet the threshold needed for change. Subsequent research will have to explore threshold effects.
- This study relates physical activity and health to the built environment at the county and metropolitan levels, which are large areas compared to the living and working environments of most residents. If environmental effects are felt most strongly at the community or neighborhood level, at least for walking, this study needs to be supplemented with research at a finer geographic scale. Future research will need to use geographic information system (GIS) data to hone in on the specific living and working environments of individuals.
- Because they are not directly measured in either of the sprawl indices, many other environmental variables that might act directly or interact to influence physical activity, such as availability and quality of parks, sidewalks, and bike trails, are not accounted for in this study. Also missing from this analysis are potentially important environmental variables such as climate, topography, and crime. Future research will have to fill this void by specifying more complete outcome models.
- By focusing on physical activity, this study largely ignores the other side of the energy equation—calories consumed as opposed to calories expended. In this study, leisure time walking accounts for only a small portion of the relationship between urban form and BMI. Although we expect other forms of physical activity to fill some of this gap, differing patterns of food consumption must also be explored. Only our fruit and

vegetable consumption variable begins to get at that dimension of the problem. Caloric intake could have a spatial component. Future research could, for example, relate the density of fast food restaurants and availability of food choices to diet and obesity.

The growing interest in how policies and the environment serve to encourage or discourage health-related behaviors is attested to by the new focus on these issues in journals such as this one and by new initiatives of governmental and nongovernmental organizations such as the CDC, with its Active Community Environments (ACES) research group, and Robert Wood Johnson Foundation, with its commitment of more than 70 million dollars to promote active living. Over the past several decades, we have engineered much of the physical activity out of our daily lives. Now our task is to understand how opportunities for physical activity can be revived.

SO WHAT: Implications for Health Promotion Practitioners and Researchers

This exploratory study seems to indicate that, after controlling for individual differences, those living in sprawling counties are likely to walk less in their leisure time, weigh more, and have greater prevalence of hypertension than those living in more compact places. Combined with other research from public health and urban planning, there is moderate support for the assertion that urban form can have significant (positive or negative) influences on health and health-related behaviors.

If this assertion holds true, health practitioners can improve public health by advocating for more compact development patterns. Public health researchers can refine their understanding of physical activity, obesity, and morbidity by including urban form variables in their analyses.

References

1. US Dept of Health and Human Services. *Physical Activity and Health: A Report of the Sur-*

1. *geon General*. Atlanta, Ga: Centers for Disease Control and Prevention; 1996.
2. Physical activity levels for U.S. overall. Available at: <http://apps.nccd.cdc.gov/dnpa/piRec.asp?piState=us&PiStateSubmit=Get+Stats>. Accessed October 31, 2002.
3. Pratt M, Macera CA, Blanton C. Levels of physical activity and inactivity in children and adults in the United States: current evidence and research issues. *Med Sci Sports Exerc*. 1999; 31(suppl 11):S526–S533.
4. Mokdad AH, Serdula MK, Dietz WH, et al. The spread of the obesity epidemic in the United States, 1991–1998. *JAMA*. 1999;282: 1519–1522.
5. Mokdad AH, Bowman BA, Ford ES, et al. The continuing epidemics of obesity and diabetes in the United States. *JAMA* 2001;286:1195–1200.
6. Flegal K, Carroll M, Ogden C, Johnson C. Prevalence and trends in obesity among US adults, 1999–2000. *JAMA*. 2002;288:1723–1727.
7. Allison DB, Fontaine KR, Manson JE, et al. Annual deaths attributable to obesity in the United States. *JAMA*. 1999;282:1530–1538.
8. McGinnis JM, Foege WH. Actual causes of death in the United States. *JAMA*. 1993;270: 2207–2212.
9. Sallis JF, Owen N. *Physical Activity and Behavioral Medicine*. Thousand Oaks, Calif: Sage Publications; 1999.
10. Humpel N, Owen N, Leslie E. Environmental factors associated with adults’ participation in physical activity. *Am J Prev Med*. 2002;22:188–199.
11. King AC, Jeffery RW, Fridinger F, et al. Environmental and policy approaches to cardiovascular disease prevention through physical activity: issues and opportunities. *Health Educ Q*. 1995;22:499–511.
12. Schmid TL, Pratt M, Howze E. Policy as intervention: environmental and policy approaches to the prevention of cardiovascular disease. *Am J Public Health*. 1995;85:1207–1211.
13. Sallis JF, Owen N. *Ecological Models*. In: Glanz K, Lewis FM, Rimer BK, eds. *Health Behavior and Health Education: Theory, Research, and Practice*. 2nd ed. San Francisco, Calif: Jossey-Bass; 1997:403–424.
14. Sallis JF, Bauman A, Pratt M. Environmental and policy interventions to promote physical activity. *Am J Prev Med*. 1998;15:379–397.
15. Bauman A, Smith B, Stoker L, et al. Geographical influences upon physical activity participation: evidence of a ‘coastal effect.’ *Aust N Z J Public Health*. 1999;23:322–324.
16. Craig CL, Brownson RC, Craig SE, Dunn AL. Exploring the effect of the environment on physical activity: a study examining walking to work. *Am J Prev Med*. 2002;23(2S):36–43.
17. Berrigan D, Troiano RP. The association between urban form and physical activity in U.S. adults. *Am J Prev Med*. 2002;23(2S):74–79.
18. King AC, Castro C, Eyster AA, et al. Personal and environmental factors associated with physical inactivity among different racial-ethnic groups of U.S. middle-aged and older-aged women. *Health Psychol*. 1999;19:354–364.
19. Brownson RC, Baker EA, Houseman RA, et al. Environmental and policy determinants of physical activity in the United States. *Am J Public Health*. 2001;91:1995–2003.
20. Handy SL, Boarnet MG, Ewing R, Killingsworth RE. How the built environment affects physical activity: views from urban planning. *Am J Prev Med*. 2002;23:64–73.
21. Ewing R, Cervero R. Travel and the built environment. *Transp Res Rec*. 2001;1780:87–114.
22. Greenwald M, Boarnet MG. The built environment as a determinant of walking behavior: analyzing non-work pedestrian travel in

- Portland, Oregon. *Transp Res Record*. 2001; 1780:33–42.
23. Handy SL. Urban form and pedestrian choices: a study of Austin neighborhoods. *Transp Res Record*. 1996;1552:135–144.
 24. Moudon AV, Hess P, Snyder MC, Stanilov K. Effects of site design on pedestrian travel in mixed-use, medium-density environments. *Transp Res Record*. 1997;1578:48–55.
 25. Frank LD, Pivo G. Impacts of mixed use and density on utilization of three modes of travel: single-occupant vehicle, transit, and walking. *Transp Res Rec*. 1994;1466:44–52.
 26. Shriver K. Influence of environmental design on pedestrian travel behavior in four Austin neighborhoods. *Transp Res Rec*. 1997;1578:64–75.
 27. Cervero R, Gorham R. Commuting in transit versus automobile neighborhoods. *J. Am Plann Assoc*. 1995;61:210–225.
 28. Hess PM, Moudon AV, Snyder MC, Stanilov K. Site design and pedestrian travel. *Transp Res Rec*. 1999;1674:9–19.
 29. Handy SL. Understanding the link between urban form and nonwork travel behavior. *J. Plann Educ Res*. 1996;15:183–198.
 30. Saelens BE, Sallis JF, Frank LD. Environmental correlates of walking and cycling: findings from the transportation, urban design, and planning literatures. *Ann Behav Med*. 2003;25:80–91.
 31. Centers for Disease Control and Prevention. Increasing physical activity: a report on recommendations of the Task Force on Community Preventive Services. *Morb Mortal Wkly Rep*. 2001;rr-18:50.
 32. Ewing R. Is Los Angeles-style sprawl desirable? *J Am Planning Assoc*. 1997;63:107–126.
 33. BRFSS. Available at: <http://www.cdc.gov/brfss/>. Accessed November 5, 2002.
 34. Raudenbush SW. Statistical analysis and optimal design for cluster randomized trials. *Psychol Methods*. 1997;2:173–185.
 35. Raudenbush SW, Liu, X. Statistical analysis and optimal design for multisite randomized trials. *Psychol Methods*. 2000;5:199–213.
 36. Raudenbush SW. Many small groups. To appear in the Handbook of Multilevel Analysis. DeLeeuw, J, Kreft, I (eds.). Chapter 6: Kluwer. In press.
 37. Nelson DE, Holtzman D, Bolen J, et al. Reliability and validity of measures from the Behavioral Risk Factor Surveillance System (BRFSS). *Soc Prev Med*. 2001;46(suppl 1):S34.
 38. Ewing R, Pendall R, Chen D. Measuring sprawl and its transportation impacts. *Transp Res Rec*. 2003;1831.
 39. US Census Bureau. *U.S. Census of Population and Housing. Census 2000 Summary File 1*. Washington, DC: US Census Bureau; 2001.
 40. US Dept of Agriculture. *1997 National Resources Inventory*. Machine-readable data [CD-ROM], revised data, alpha release. USDA; March 2001.
 41. US Census Bureau. *UA Census 2000. TIGER/LineR files* [machine-readable data files]. Washington, DC: Dept of Commerce; 2002.
 42. Raudenbush SW, Bryk A, Cheong YF, Congdon R. *HLMs: Hierarchical Linear and Nonlinear Modeling*. Chicago, Ill: Scientific Software International; 2000.
 43. Raudenbush SW, Byrk AS. *Hierarchical Linear Models: Applications and Data Analysis Methods*. 2nd ed. Thousand Oaks, Calif: Sage Publications; 2002.
 44. Rose G. Sick individuals and sick populations. *Int J Epidemiol* 1985;14:32–38.