Built and Social Environments

Associations with Adolescent Overweight and Activity

Melissa C. Nelson, PhD, RD, Penny Gordon-Larsen, PhD, Yan Song, PhD, Barry M. Popkin, PhD

Background: Little is known about the patterning of neighborhood characteristics, beyond the basic urban, rural, suburban trichotomy, and its impact on physical activity (PA) and overweight.

Methods: Nationally representative data (National Longitudinal Study of Adolescent Health, 1994–1995, n =20,745) were collected. Weight, height, PA, and sedentary behavior were self-reported. Using diverse measures of the participants' residential neighborhoods (e.g., socioeconomic status, crime, road type, street connectivity, PA recreation facilities), cluster analyses identified homogeneous groups of adolescents sharing neighborhood characteristics. Poisson regression predicted relative risk (RR) of being physically active (five or more bouts/week of moderate to vigorous PA) and overweight (body mass index equal or greater than the 95th percentile, Centers for Disease Control and Prevention/National Center for Health Statistics growth curves).

Results: Six robust neighborhood patterns were identified: (1) rural working class; (2) exurban; (3) newer suburban; (4) upper-middle class, older suburban; (5) mixed-race urban; and (6) low-socioeconomic-status (SES) inner-city areas. Compared to adolescents living in newer suburbs, those in rural working-class (adjusted RR[ARR]=1.38, 95% confidence interval [CI]=1.13–1.69), exurban (ARR=1.30, CI=1.04–1.64), and mixed-race urban (ARR=1.31, CI=1.05–1.64) neighborhoods were more likely to be overweight, independent of individual SES, age, and race/ethnicity. Adolescents living in older suburban areas were more likely to be physically active than residents of newer suburbs (ARR=1.11, CI=1.04–1.18). Those living in low-SES inner-city neighborhoods were more likely to be active, though not significantly so, compared to mixed-race urban residents (ARR=1.09, CI=1.00–1.18).

Conclusions: These findings demonstrate disadvantageous associations between specific rural and urban environments and behavior, illustrating important effects of the neighborhood on health and the inherent complexity of assessing residential landscapes across the United States. Simple classical urban–suburban–rural measures mask these important complexities.
(Am J Prev Med 2006;31(2):109–117) © 2006 American Journal of Preventive Medicine

Introduction

Overweight and obesity have emerged as national public health concerns,[1,2] with adolescence as an important developmental period.[3] Built and social environments are important determinants of obesity-related health behavior (e.g., physical activity [PA]) and targets for intervention strategies.[4] Research studying neighborhood effects on health has relied largely on aggregate socioeconomic status (SES) measures to characterize neighborhood environments.[5–9] In an emerging literature exploring how environment facilitates or restricts health behavior, however, specific individual-level measures of neighborhood factors (e.g., crime/safety,[10–12] street connectivity, road type/traffic,[13–14] and activity-related recreation facility access[15–18]) have been independently associated with PA.

Neither aggregate indices of SES nor specific aspects of the built environment appear in isolation in neighborhoods. Clearly, factors such as SES, crime, lack of recreation facilities, and other community-level measures occur jointly.[15] In contrast to traditional risk factor approaches to data analysis that examine the independent effects of specific neighborhood characteristics, pattern analyses allow examination of the effects of multiple dimensions of the environment. Both independent risk factor analysis and pattern analysis may make important contributions to understanding how the environment affects behavior.

From the Division of Epidemiology and Community Health, University of Minnesota (Nelson), Minneapolis, Minnesota; and Department of Nutrition, University of North Carolina (Nelson, Gordon-Larsen, Popkin), Carolina Population Center (Nelson, Gordon-Larsen, Popkin), and Department of City and Regional Planning (Song), Chapel Hill, North Carolina

Address correspondence and reprint requests to: Melissa C. Nelson, PhD, RD, Division of Epidemiology and Community Health, University of Minnesota, 1300 S. 2nd Street, WBOB Suite 300, Minneapolis MN 55454-1015. E-mail: nelson@epi.umn.edu.

The full text of this article is available via AJPM Online at www.ajpm_online.net.

 doi:10.1016/j.amepre.2006.03.026

While the importance of covariance and joint effects of neighborhood features has gained recognition in this growing area of interest in environmental determinants of obesity, the study of patterning to date has been limited largely to index development as a means of measuring very specific aspects of the built environment.[13,19] There is little empirical evidence describing the diversity and covariance of community characteristics using data-driven techniques within nationally representative data sets, including a wide array of residential landscapes, most importantly, rural areas that are greatly understudied.

By using pattern analyses to measure the effects of multiple environmental characteristics on behavior, this research fills an important gap in the literature. Using data from a nationally representative, ethnically diverse sample of adolescents, the aims of this study were to (1) identify meaningful patterns of sociodemographic and built features in neighborhood environments that have been identified as potentially important determinants of PA, and (2) describe the cross-sectional associations between these neighborhood patterns and adolescent residents' PA and weight status.

Methods

Add Health

Add Health is a school-based longitudinal survey of youths, grades 7 through 12. A random sample of 80 high schools and 52 junior high feeder schools was selected. The Add Health sample was designed to be nationally representative of students in grades 7 through 12 in 1995 in the United States. Survey procedures[20] were previously approved by the Institutional Review Board at the University of North Carolina at Chapel Hill. The Wave-I in-home survey (1994–1995) included 20,745 adolescent participants. Analyses were conducted in 2005–2006.

Measuring Individual Neighborhood Environments

Residential location. Home street addresses of most (n =20,612, 99.4%) Wave-I participants were identified and geocoded, using primarily street-segment matches from commercial geographic information system (GIS) databases (n =17,119) or global positioning system (GPS) units (n =3242) (when a street-segment match was unavailable). When neither was available, residential location assignments used a ZIP+4/ZIP+2 or 5-digit ZIP centroid match, or the respondent's school location.[15] A relational database linked the location of a participant's home to (1) neighborhood attributes, based on buffers around each home; (2) block group, tract, and county attributes from U.S. Census and other federal sources; and (3) Add Health participants' survey responses. Add Health was designed for national representation of youth, not for geographic representation. Yet, the data include a wide array of geographic areas across the United States.

Buffers for respondent locations. To assess the variety of neighborhood characteristics, a 3-km buffer was drawn around each respondent's residential location using Euclidean distance. While there is some suggestion that 5-mile catchment areas may be relevant for adult PA,[15,17] it is likely that smaller areas influence adolescent PA, where travel is more limited. There is little empirical data to support appropriate buffer size for PA outcomes at a national level. Sensitivity analyses were conducted to determine the appropriate buffer size (3 km) for these analyses.

Physical activity facilities within 3 km. Commercially available, retrospective (1995), digitized Yellow Pages, using proprietary 4-digit extensions to the Standard Industrial Classification (SIC) codes, were obtained. These 8-digit codes correspond with those used by the Census, allowing for the detailed characterization of facility type. A comprehensive list representing PA facility/resource types (n =169) was compiled for these SIC codes.[15] SIC code counts were summed to measure all activity-related facilities, and subdivided to specifically assess parks. Park locations were verified using digital aerial photographs from the U.S. Geological Survey.

Walkability within 3 km. High street connectivity, or "walkability" (i.e., neighborhood street networks that are continuous, integrated, and maximize linkages between starting points and destinations, providing multiple route options) has been positively associated with residential activity patterns.[13,21] Indices of connectivity include (1) intersection density (three-way and four-way intersections), (2) alpha index (ratio of observed to maximum possible route alternatives [circuitry] between nodes, where the maximum possible circuits is the maximum number of links minus the number of links in a minimally connected network), (3) gamma index (ratio of observed node linkages to the maximum possible links in the network), and (4) cyclomatic index (number of route alternatives [circuits] between nodes).

Road type within 3 km. Road networks were mapped using retrospective U.S. Census TIGER (topologically integrated geographic encoding and referencing) line files (www.census.gov/geo/www/tiger/). Road types were assessed using Census feature class codes (www.census.gov/geo/www/tiger/appendxe.asc) (i.e., feature Class A categories). The presence of smaller, local roads (category A4x, which are more likely to have single traffic lanes, sidewalks, and lower speed limits) were of particular interest for this research, in comparison to the presence of larger roadways (A1x) on which walking/biking is more difficult (e.g., primary highways). Road types were assessed as the proportion of total roadways and the absolute total length.

Census measures. Census data reflecting individuals' residential block groups were extracted from the 1990 Census of Population and Housing summary tape file 3A (STF3A). A block group is a relatively small administrative unit in the U.S. Census (averaging 300 to 3000 residents). Variables used here were education (proportion of adults aged ≥25 years with a college degree), minority (proportion of nonwhites), poverty (proportion of people with incomes <185% of poverty level), housing units (proportion of housing units occupied by renters, proportion vacant, median housing unit age), and mobility (proportion of population living in same housing unit since 1985, proportion of population working in county

Table 1. Measured constructs of the neighborhood environment

Construct	Measure
Income/wealth	Income to poverty level (% less than 185%)
	Home age (year structure was built)
Race/ethnicity	Ethnicity (% minority)
Socioeconomic status and environment	Education (% with college degree)
	Occupancy status (% owner occupied, % renter occupied, % vacant)
	Mobility (% living in same house since 1985), median house age
	Proportion working in county of residence
Crime	Serious crimes (arrests) per 100,000 persons
Road type	Proportion of A1 and A4 roads in 3-km buffer
	Total length of A1 and A4 roads in 3-km buffer
Street connectivity (walkability)	Intersection density in 3-km buffer
	Gamma index in 3-km buffer
	Cyclomatic index in 3-km buffer (total route alternatives)
	Alpha index in 3-km buffer (observed total route alternatives)
Recreation facilities for physical activity	For each facility type: count in 3-km buffers

of residence). The metropolitan statistical area (MSA) of Add Health schools was also identified, and regions were broadly categorized as urban, suburban, or rural.

Crime. Reported crimes (per 100,000 population) were assessed using 1995 U.S. Federal Bureau of Investigation Uniform Crime Reporting county-level data from the National Archive of Criminal Justice Data (www.icpsr.umich.edu/NACJD/index.html), which have been shown to be associated with PA levels in this sample.[22] For Add Health respondents (*n* =366) in counties with no available 1995 crime data, crime rates were used from a previous year (1990 to 1994). For three counties (*n* =95 individuals), crime was imputed as average reported crime in surrounding counties.

Physical Activity/Sedentary Behaviors

Daily PA (e.g., housework, active play, sports, exercise) was self-reported using standard epidemiologic 7-day recall methodology.[23] Surveys (www.cpc.unc.edu/projects/addhealth/codebooks) employed questions similar to those used and validated in other large-scale studies.[23–26] Questions asked—variations on "During the past week, how many times did you . . ."—allowed estimation of activity frequency (bouts/week) by metabolic equivalent (MET) value. Moderate-to-vigorous PA (MVPA) was 5 to 8 METs.[27]

Adolescents also reported sedentary behaviors (e.g., watching/playing TV/videos, video or computer games [hours/week]), using recreation centers, and playing sports with parent(s). Overall activity frequency was summed to determine total weekly MVPA or sedentary behavior, as well as whether individuals met national recommendations for PA (i.e., recommendation is five or more bouts/week of MVPA)[28] and did not meet recommendations for sedentary behavior (i.e., recommendation is TO NOT exceed 14 hours/week "screen time").[29,30]

Weight Status

Self-reported height and weight were used to calculate body mass index (BMI) (kilograms/square meters).[31] The 95th percentile of nationally representative data (2000 Centers for Disease Control and Prevention/National Center for Health Statistics growth curves), was used to classify overweight.[32] Individuals aged 21 years (*n* =9) were considered as 20-year-olds for assessing overweight. Self-reported weight and height have been shown to correctly classify a majority of Add Health participants as overweight.[33]

Covariates

Adolescents self-reported race/ethnicity; reports were validated during in-home parent interviews. Parents reported highest level of achieved education, which was used to estimate SES. Income was reported in $1000 increments and imputed where missing, using parent occupation, family structure, and school community.

Identifying Patterns in Environment

Cluster analyses were used to identify patterns of environmental characteristics and to specify homogeneous, non-overlapping clusters (or patterns) of neighborhoods sharing various meaningful characteristics.[34] Multiple cluster analyses were conducted partitioning data into different numbers of clusters[4–10] by Euclidean distances between observations that were weighted for national representation, using SAS FASTCLUS, SAS version 9 (Research Triangle Institute, Research Triangle Park NC, 2004). Representing different constructs of the neighborhood, 19 variables were used (Table 1). Z-score transformations of variables were used to generate clusters, allowing for the appropriate weighting of variables with different scales.[34]

To identify initial cluster centers (i.e., seed values), 1000 iterations of each cluster procedure were conducted.[35] The initial group center for each iteration was randomly generated. The iteration with the largest overall r^2 value, which allowed for the evaluation of relative heterogeneity between clusters (vs homogeneity within clusters), was identified. Clusters best fitting the data maximized this inter- to intra-variability ratio, yielding a higher r^2. (For the six-cluster solution series—i.e., the final cluster solution—the maximum r^2 value identified through this iterative process was 0.41.) Results of these numerous analyses were assessed to identify common patterns appearing across various procedures. The final presented clusters were those representing the most robust data patterns.

Table 2. Major neighborhood types identified through cluster analysis procedures

Type	Description
Rural working class	Low SES, moderate-to-low minority population, little mobility Low connectivity (i.e., low intersection density and few possible routes between any two points) Low access to PA facilities, very low overall density of roadways
Exurban (urban/ suburban outgrowth)	Moderate SES, low minority population High % of relatively recently built housing units High % of population commuting to work outside of county of residence Low access to PA facilities, low street connectivity and low crime High proportion of large arterial roadways
New suburban developments (referent)	High SES, low minority population Relatively recently built housing units Low access to PA facilities, very poor street connectivity Few roadways overall (with high proportion of these roadways being of local)
Older, upper-middle class suburbia with highway access	High SES population with somewhat mixed racial/ethnic composition, little mobility High % of older housing units Moderate access to PA facilities, moderate street connectivity in linking intersections (and moderate-to-high number of alternative street routes between any two points) High density of roads overall (with a moderate proportion of large arterial roadways)
Mixed-race/ethnicity urban	Low SES, high poverty population Moderate access to PA facilities, moderate-to-high street connectivity and crime High density of local roadways
Low SES, inner city	Low SES, very high minority and very high poverty population Large proportion of older housing units Very high access to PA facilities, very high street connectivity and intersection density High crime, high density of local roadways

PA, physical activity; SES, socioeconomic status.

Demonstrating Cluster Variability

Cluster analytic procedures detect underlying data patterns, regardless of utility or substantive merit, but statistical methods for validating cluster analyses are limited.[34] To show that clusters fit the data in a meaningful way, clusters are often tested by predicting external variables not used to generate the patterns (although associated with the clusters in theory).

To demonstrate meaningful variability between patterns and to validate these findings, neighborhood clusters were assessed as independent variables in generalized linear models predicting adolescent PA, sedentary behavior, and overweight. As another tool for comparison, broad neighborhood characteristics were examined (e.g., broad urbanicity classifications of urban, suburban, and rural; median household income; percent college-educated population; percent minority population), which have been used extensively in previous literature.

All models controlled for important covariates (age, race/ethnicity, parent education/income). Observations with missing covariate or outcome data were excluded. Participants who were severely disabled ($n = 132$) and/or pregnant ($n = 379$) were also excluded. While logistic regression is commonly used in health research, the odds ratios yielded by these analyses substantially over-estimate risk ratios (RRs) that are >1.0 (and under-estimate those that are <1.0) when the outcome is not rare. The binomial outcomes included in these analyses were not rare (i.e., prevalence $>10\%$), so Poisson regression was used with robust variance estimates to generate valid estimates of the adjusted relative risk (ARR) (and relatively conservative confidence intervals [CIs]), instead of adjusted odds ratios.[36]

Descriptive analyses of individual adolescent characteristics used post-stratification sample weights, allowing results to be nationally representative. Survey design effects of multiple stages of cluster sampling were controlled using the survey (SVY) procedure series in STATA, version 9.0 (Stata Corp, College Station TX, 2004).

Results

Descriptive Statistics and Final Cluster Solution

The analysis sample ($n = 20{,}745$) generating the neighborhood clusters was composed of 50.1% males as well as 68.5% white, 15.2% black, 11.4% Hispanic and 4.0% Asian adolescents. Approximately 14.7% of participants' parents had less than a high school education, 32.5% had graduated from high school (or had a general equivalency diploma), 27.8% had some college, and 25.0% had a college degree or higher. Mean participant age was 15.4 (± 0.12) years.

Six robust neighborhood pattern types were identified by the final cluster solution and were observed across numerous iterations of analyses, representing non-overlapping groups of U.S. neighborhoods sharing various attributes. These clusters include (1) rural working class; (2) exurban (urban/suburban outgrowth); (3) new suburban developments (referent); (4) older, upper-middle class suburbia with highway access; (5) mixed-race urban; and (6) low-SES inner-city neighborhoods (Table 2). These neighborhood patterns are distinguished by important differences in the 19 neighborhood attributes used to generate the final cluster solution, including SES, race/ethnicity, socioenvironment, crime, road type, street connectivity, and recreation facilities (Table 3).

Table 3. Mean frequency of specific neighborhood characteristics by cluster[a]

	Cluster 1 Rural working class (*n* = 4725)	Cluster 2 Exurban (*n* = 2178)	Cluster 3 New suburban development (*n* = 3371)	Cluster 4 Older suburban developments (*n* = 4280)	Cluster 5 Mixed-race/ ethnicity urban (*n* = 3609)	Cluster 6 Low-SES inner city (*n* = 2582)
% population with a college education (in block group)	14.5 ± 0.1 (−0.56)	**21.2 ± 0.3 (−0.11)**	**34.5 ± 0.2 (0.77)**	**32.9 ± 0.3 (0.67)**	17.0 ± 0.2 (−0.39)	16.2 ± 0.2 (−0.45)
% nonwhite population (in block group)	20.0 ± 0.4 (−0.41)	15.7 ± 0.6 (−0.54)	15.0 ± 0.3 (−0.56)	36.5 ± 0.5 (0.06)	**50.8 ± 0.5 (0.47)**	**75.1 ± 0.6 (1.17)**
% population with income <185% poverty line (in block group)	**38.1 ± 0.2 (0.35)**	26.1 ± 0.4 (−0.23)	14.7 ± 0.2 (−0.79)	14.8 ± 0.1 (−0.78)	**45.8 ± 0.3 (0.73)**	**48.5 ± 0.4 (0.86)**
% renter-occupied housing units (in block group)	20.3 ± 0.2 (−0.44)	20.6 ± 0.3 (−0.43)	20.0 ± 0.3 (−0.45)	21.6 ± 0.3 (−0.38)	**47.9 ± 0.3 (0.87)**	**54.6 ± 0.4 (1.18)**
% vacant housing units (in block group)	11.6 ± 0.2 (0.44)	7.0 ± 0.2 (−0.10)	7.5 ± 0.2 (−0.04)	3.0 ± 0.1 (−0.57)	8.8 ± 0.1 (0.12)	8.9 ± 0.2 (0.12)
Median house age (years) (in block group)	23.9 ± 0.1 (−0.19)	21.5 ± 0.2 (−0.38)	**13.2 ± 0.1 (−1.03)**	31.2 ± 0.2 (0.38)	28.7 ± 0.2 (0.18)	41.1 ± 0.2 (1.15)
% population living in the same house for ≥5 years (in block group) (i.e., % low morbidity)	**62.6 ± 0.1 (0.47)**	58.6 ± 0.3 (0.20)	45.7 ± 0.3 (−0.65)	**62.6 ± 0.2 (0.47)**	45.8 ± 0.2 (−0.64)	54.5 ± 0.3 (−0.07)
% population working in county of residence (in block group)	70.5 ± 0.3 (−0.36)	**67.4 ± 0.5 (−0.52)**	75.1 ± 0.4 (−0.14)	81.9 ± 0.3 (0.19)	87.9 ± 0.2 (0.48)	83.9 ± 0.4 (0.29)
Number of facilities for physical activity (in 3 km)	3.3 ± 0.1 (−0.66)	5.7 ± 0.1 (−0.58)	8.1 ± 0.1 (−0.49)	**32.5 ± 0.2 (0.37)**	23.5 ± 0.2 (0.05)	**68.7 ± 0.9 (1.65)**
Number of parks (in 3 km)	<0.01 ± 0.001 (−0.21)	0.01 ± 0.002 (−0.20)	0.01 ± 0.001 (−0.20)	0.04 ± 0.003 (−0.11)	0.02 ± 0.002 (−0.16)	**0.55 ± 0.018 (1.21)**
Alpha index of road connectivity (in 3 km)	0.25 ± 0.001 (0.16)	0.19 ± 0.001 (−0.49)	0.16 ± 0.001 (−0.81)	0.22 ± 0.001 (−0.20)	0.26 ± 0.001 (0.22)	**0.35 ± 0.002 (1.19)**
Gamma index of road connectivity (in 3 km)	0.51 ± 0.001 (0.22)	0.46 ± 0.001 (−0.51)	0.44 ± 0.001 (−0.86)	0.48 ± 0.001 (−0.22)	0.51 ± 0.001 (0.21)	0.57 ± 0.001 (1.23)
Cyclomatic index of road connectivity (in 3 km)	84.4 ± 1.2 (−0.86)	120.9 ± 2.1 (−0.77)	185.8 ± 2.4 (−0.61)	548.8 ± 3.5 (0.32)	584.7 ± 3.2 (0.41)	**1172.3 ± 5.0 (1.91)**
Intersection density (in 3 km)	5.0 ± 0.1 (−1.00)	8.5 ± 0.1 (−0.81)	14.0 ± 0.2 (−0.51)	33.2 ± 0.2 (0.53)	32.3 ± 0.2 (0.48)	**53.9 ± 0.2 (1.65)**
Reported crime rate (in county)	3635.5 ± 29.3 (−0.83)	4483.8 ± 46.6 (−0.53)	6003.6 ± 37.5 (−0.004)	6195.1 ± 30.2 (0.06)	**6997.9 ± 35.0 (0.34)**	**9640.2 ± 61.6 (1.27)**
Total length (meters) of major arterial roadways (in 3 km)	114.6 ± 13.5 (−0.91)	**17,937.3 ± 139.6 (0.69)**	2910.7 ± 89.7 (−0.66)	**18,365.9 ± 180.9 (0.73)**	11,229.8 ± 159.4 (0.09)	**16,951.1 ± 197.7 (0.60)**
Total length (meters) of local roadways (in 3 km)	55,425.6 ± 381.8 (−1.07)	79,429.1 ± 810.7 (−0.81)	113,899.3 ± 971.7 (−0.44)	**206,137.1 ± 682.8 (0.57)**	**204,157.2 ± 845.8 (0.55)**	**291,885.8 ± 960.8 (1.51)**
% roadways that are major arterial roadways (in 3 km)	0.1 ± 0.02 (−0.78)	**16.1 ± 0.1 (1.99)**	1.6 ± 0.1 (−0.51)	6.6 ± 0.1 (0.34)	4.1 ± 0.1 (−0.09)	4.5 ± 0.1 (−0.03)
% roadways that are local roads (in 3 km)	**77.2 ± 0.2 (0.20)**	62.8 ± 0.2 (−1.26)	**80.2 ± 0.1 (0.51)**	73.7 ± 0.1 (−0.16)	**76.8 ± 0.1 (0.16)**	**76.0 ± 0.1 (0.08)**

[a]Mean ± standard error (mean z-score) of neighborhood characteristics (unweighted).

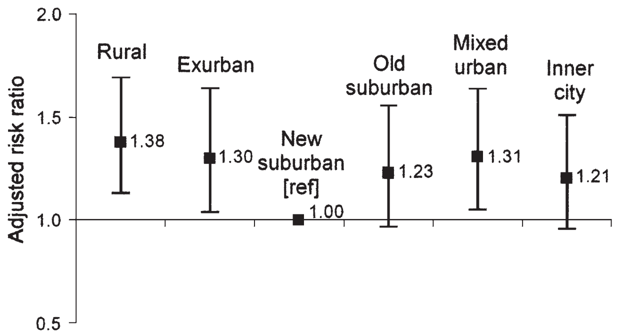

Figure 1. Adjusted risk ratios (95% confidence intervals) of overweight (BMI ≥95th percentile) by data-driven neighborhood cluster definitions (n = 19,029). Note: adjusted for household income, parental education, adolescent age, and race/ethnicity.

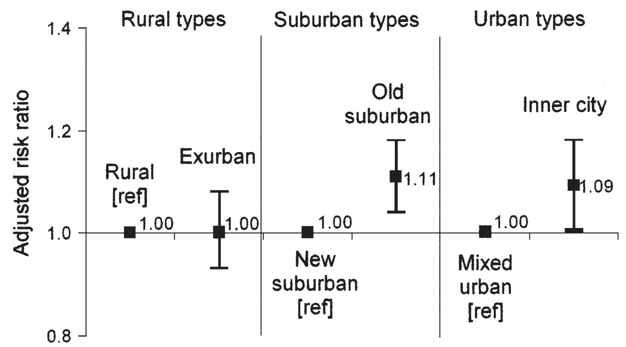

Figure 2. Adjusted risk ratios (and 95% confidence intervals) of achieving five or more bouts of moderate-to-vigorous physical activity (MVPA)/week by data-driven neighborhood cluster definition and type (rural, suburban and urban) (n = 19,531, across three models). Note: Adjusted for household income, parental education, adolescent age and race/ethnicity.

Neighborhood–Overweight Associations

Using the neighborhood clusters, important differences in the relative risk of overweight by community type were identified. Compared to adolescents living in newer suburban developments, those who lived in (1) rural working class, (2) exurban, and (3) mixed-race urban neighborhoods were 30% to 40% more likely to have a BMI≥95th percentile of age- and gender-specific national growth curves (Figure 1), independent of SES, adolescent age, and race/ethnicity.

Conversely, analyses using individual components or traditional measures of neighborhood characteristics show less-clear associations (Table 4). Using traditional analyses, no difference in the risk of overweight between low- and moderate-SES communities was found, although high-SES communities were less likely to be overweight, compared to moderate-SES communities. There was no difference in overweight status by the race/ethnicity of the community. The traditional urban–suburban–rural status (using MSAs) suggest a lower likelihood of overweight in adolescents residing in urban areas, but no differences in rural and suburban neighborhoods. A comparison of the traditional urban–suburban–rural breakdown with the approach presented in Table 3 highlights the differences in the two approaches.

Neighborhood–Activity Associations

Among adolescents living in older suburban neighborhoods, 39% reported engaging in five or more bouts of weekly MPVA, compared to 28% of adolescents living in mixed-race urban areas (data not shown). These findings indicate notable distinctions within suburban and urban community types; for example, independent of SES, race/ethnicity, and age, adolescents living in older suburban developments were 11% more likely to be physically active than those living in newer suburban areas (ARR=1.11, CI=1.04–1.18) (Figure 2). In addition, those living in low-SES inner-city areas were more likely to be active compared to those in mixed-race urban neighborhoods (RR=1.09, CI=1.00–1.18).

Table 4. Adjusted risk ratios (95% confidence intervals) of adolescent overweight (≥95th percentile body mass index) using broad, independent, and traditional measures of neighborhood characteristics

	Adjusted risk ratio of overweight by neighborhood features		
	Low tertile	**Moderate tertile**	**High tertile**
Median household income[a]	1.03 (0.91–1.17)	1.0 (ref)	0.83 (0.71–0.98)
% of population with ≥college education[b]	1.10 (0.99–1.22)	1.0 (ref)	0.74 (0.65–0.85)
% of minority population[c]	1.00 (0.85–1.15)	1.0 (ref)	1.13 (0.97–1.31)
MSA classification of school[d]	Rural: 1.9 (0.94–1.27)	Suburban: 1.0 (ref)	Urban: 0.85 (0.75–0.97)

Note: Adjusting for individual parental education, household income, race/ethnicity, and age of adolescent.
[a]Low-income tertile: median household income ≤\$23,775/year, moderate tertile: >\$23,775 to ≤\$36,440/year, high tertile: >\$36,440/year at the census block–group level (n = 19,029 in model).
[b]Low education tertile: ≤14.4% of population college educated; moderate tertile: >14.4% to ≤26.5%; high tertile: >26.5% at the census block–group level (n = 19,025 in model).
[c]Low minority tertile: ≤6.1% of population is minority, moderate tertile: >6.1% to ≤49.4%, high tertile: >49.4% at the census block–group level (n = 19,025 in model).
[d]MSA, metropolitan statistical area (n = 18,709 in model).

Independent of SES, race/ethnicity, and age, teens living in low-SES inner-city and older suburban areas were least likely to report playing a sport with a parent(s), and those living in the low-SES inner-city, mixed-race urban, and older suburban areas were the most likely to report using a neighborhood recreation facility. High levels of screen time (>14 hours/week, which is above recommended levels) were most likely among those living in low-SES inner-city neighborhoods (Table 5).

Discussion

Using cluster analysis, six robust patterns in residential neighborhood characteristics were identified, incorporating a range of sociodemographic and built environment characteristics: (1) rural working class, (2) exurban, (3) new suburban, (4) older suburban, (5) mixed-race urban, and (6) low-SES inner-city areas. To our knowledge, this is the first study using data-driven techniques to characterize neighborhoods by sociodemographic and built environment features in a nationally representative survey of adolescents. These results show differences in neighborhood patterns by adolescent PA and overweight (such as the disadvantageous associations between rural and urban environments and health). Rural populations have been particularly understudied; these findings indicate that these individuals have unique neighborhood characteristics that deserve further attention.

These findings show other important differences by activity and weight status. Contrary to recent reports supporting positive relationships between sprawl and adult overweight/obesity,[19,37,38] the current results indicate some beneficial associations between suburban living and activity/overweight. Cross-sectionally, these findings indicate that adolescents living in rural working class, exurban, and mixed-race urban areas were at the highest risk of overweight compared to those in the newer suburban cluster type. Teens living in older suburban communities were most likely to be physically active. In adolescence, there may be protective factors shared by those living in suburban communities (e.g., school-based sports/activity facilities, community organizations, low crime) overriding the deleterious effects of suburbia, which are believed to be at play in restricting activity of adult residents (e.g., low walkability, dependence on automobiles).

The natural patterns of neighborhood characteristics underlying these data represent diverse settings that span urban, suburban, and rural regions, and are shown to be valid through empirical evidence and theoretical frameworks. For example, these results are consistent with previous findings that low-income, racial/ethnic minority, and rural populations are less physically active and more overweight and obese.[3,39–41] In addition, the findings are supported by theoretical

Table 5. Adjusted risk ratios (and 95% confidence intervals) of adolescent physical activity and sedentary behavior by detailed data-driven neighborhood cluster definitions

	Cluster 1 Rural working class	**Cluster 2 Exurban**	**Cluster 3 New suburban development**	**Cluster 4 Older suburban developments**	**Cluster 5 Mixed-race/ethnicity urban**	**Cluster 6 Low-SES inner city**
≥5 bouts/week of any moderate-to-vigorous physical activity (n = 19,531)	1.03 (0.96–1.09)	1.04 (0.97–1.12)	1.0	1.11 (1.05–1.18)	0.96 (0.89–1.03)	1.05 (0.97–1.14)
Played a sport with a parent in the last month (n = 19,495)	0.95 (0.79–1.16)	1.01 (0.81–1.26)	1.0	0.78 (0.63–0.96)	0.81 (0.64–1.01)	0.63 (0.48–0.82)
Uses a neighborhood recreation center (n = 19,424)	1.11 (0.83–1.47)	0.97 (0.73–1.28)	1.0	1.41 (1.21–1.63)	1.40 (1.20–1.62)	1.51 (1.24–1.84)
>14 hours/week of TV/video viewing and video/computer gaming (n = 19,521)	1.03 (0.96–1.10)	1.06 (1.00–1.15)	1.0	1.06 (0.99–1.15)	1.06 (0.99–1.12)	1.12 (1.05–1.20)

Note: Poisson regression, adjusting for parental education, household income, race/ethnicity, and age of adolescent.
SES, socioeconomic status.

constructs of transect planning, an approach to urban planning emphasizing breadth and range in community design to recognize a broad range of human and environmental needs, rather than a "one size fits all" approach.[42] These neighborhood patterns overlap many of the eco-zones outlined in the transect planning literature, which to date has been based largely on theory and empirical evidence from smaller geographic samples with restrictive analytic approaches. Guided by these eco-zones in the transect theory literature, in conjunction with the empirically derived cluster definitions of this study (e.g., described by mean frequency of specific neighborhood characteristics by clusters), these neighborhood types may be re-created and explored in other data sets. In characterizing multidimensional neighborhood features, it is understood that they vary in structure and sociodemographic composition, and likely have differential impacts on population health behavior. By better understanding the complexities of today's American residential landscape, there is the opportunity to better conceptualize how to enhance built and social environments and tailor solutions that will promote healthy lifestyles.

While these data show important dimensions of U.S. communities and their association with health behavior, they are not without limitation. First, the data on PA and BMI are derived from self-reported measures, which are subject to error and bias. Second, these findings are cross-sectional, thus limiting causal inference, and may be influenced by residual confounding (due to unmeasured characteristics for which there is no control). Third, it is difficult to assess the true validity of the cluster method, given that cluster analysis will detect underlying patterns in data, regardless of meaning or utility. To address this issue, this research utilized several approaches, similar to those used in other previous applications of cluster methods, including (1) assessing numerous iterations of cluster solutions to ensure that the final solution represents robust patterns in the data, (2) comparing these findings to established theoretical frameworks, such as the eco-zones of transect planning, and (3) contrasting these clusters with external variables (that were not used in the cluster creation), such as PA and overweight status. The latter is critical for ensuring meaningful interpretability of the cluster results.

Finally, the strength in this study of using data characterizing multiple sites across the nation may also limit the ability to characterize neighborhood environments. While all of the neighborhood characteristics are relevant to individual participants, they represent varying degrees of "neighborhood," potentially carrying varying degrees of influence. Methods for defining one's neighborhood are highly debated; there is no consensus on appropriate buffer sizes to capture relevant exposure areas. The exposure areas assessed here may not be the most appropriate size (e.g., 1-mile buffers, county-level crime). Indeed, this is an important area for future research, including detailed analyses on methods of measuring environmental characteristics and important aspects of influential neighborhood features.

This is possibly the first research characterizing the national landscape in this way, understanding how neighborhoods function as a whole (rather than as individual components) and how they are associated with adolescent health behaviors and outcomes. Individuals live in neighborhoods rather than income brackets, and are simultaneously affected by factors such as crime, facility access, and street connectivity, which work in concert to affect health behavior. While traditional risk factor analysis provides important insights into the association between environment and behavior, these findings show that broad, traditional measures of neighborhood characteristics (e.g., median household income) may not capture the fine-grain detail and complexity needed to better understand how environments influence behavior. Thus, using pattern analysis with detailed environmental measures supplements the understanding of covariance in the environment and important environment–behavior relationships. Effective population-wide health promotion strategies and public policies need to address preexisting neighborhoods, which are composed of a variety of factors, many of which may be important determinants of activity patterns.

The patterns of neighborhood features identified here show meaningful variation, supported by urban planning theory and empirical evidence. Not only do these findings help illustrate the important effects of neighborhood on health, but they also demonstrate the inherent complexity of these relationships. This research highlights the extent to which neighborhoods vary, as well as how individuals may function differently in different environments, and points to the challenges of increasing population-wide PA through community design. Future research is needed to explore the specific mechanisms through which neighborhood form affects population health, as well as the interactive effects of a spectrum of co-varying community characteristics and the extent to which effective intervention and policy strategies can be tailored to promote healthy lifestyles.

Funding for this study and the development of the preliminary spatial measures comes from the National Institutes of Health, including the National Institute of Child Health and Human Development (NICHD) (R01-HD39183-01, R01 HD041375-01, and K01 HD044263-01), National Institute of Diabetes and Digestive and Kidney Diseases (DK56350), National Institute of Environmental Health Sciences (P30ES10126); a Cooperative Agreement with the Centers for Disease Control and Prevention (CDC SIP 5-00); and a New Investigator Dissertation Award from the Robert Wood Johnson Foundation's Active Living Research Program (050752).

We also thank the Spatial Analysis Unit at the University of North Carolina at Chapel Hill, particularly Phil Page, Jay Stewart, and Evan Hammer, for their assistance in data collection and processing.

This research used data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by the NICHD (P01-HD31921), with cooperative funding from 17 other agencies. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Anyone interested in obtaining data files from Add Health should contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill NC, 27516-2524 (www.cpc.unc.edu/addhealth/contract.html).

No financial conflict of interest was reported by the authors of this paper.

References

1. Flegal KM, Carroll MD, Ogden CL, Johnson CL. Prevalence and trends in obesity among U.S. adults, 1999–2000. JAMA 2002;288:1723–7.
2. Ogden CL, Flegal KM, Carroll MD, Johnson CL. Prevalence and trends in overweight among U.S. children and adolescents, 1999–2000. JAMA 2002;288:1728–32.
3. Gordon-Larsen P, Adair LS, Nelson MC, Popkin BM. Five-year obesity incidence in the transition period between adolescence and adulthood: the National Longitudinal Study of Adolescent Health. Am J Clin Nutr 2004;80:569–75.
4. Institute of Medicine. Does the built environment influence physical activity? Examining the evidence. Washington DC: National Academy of Sciences, 2005 (Transportation Research Board special report 282).
5. Diez Roux AV. Multilevel analysis in public health research. Annu Rev Public Health 2000;21:171–92.
6. Diez Roux AV. Investigating neighborhood and area effect on health. Am J Public Health 2001;91:1783–89.
7. 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.
8. Yen IH, Kaplan GA. Poverty area residence and changes in physical activity level: evidence from the Alameda County Study. Am J Public Health 1998;88:1709–12.
9. Sundquist J, Malmstrom M, Johansson S-E. Cardiovascular risk factors and the neighbourhood environment. Int J Epidemiol 1999;28:841–5.
10. Handy S, Clifton K. Local shopping as a strategy for reducing automobile travel. Transportation 2001;28:317–46.
11. Handy SL, Clifton KJ, Fisher J. The effectiveness of land use policies as a strategy for reducing automobile dependence: a study of Austin neighborhoods. Austin TX: Center for Transportation Research, Southwest Region University Transportation Center, 1998.
12. Booth ML, Owen N, Bauman A, Clavis O, Leslie E. Social-cognitive and perceived environment influences associated with physical activity in older Australians. Prev Med 2000;31:15–22.
13. Greenwald MJ, Boarnet MG. Built environment as determinant of walking behavior: analyzing nonwork pedestrian travel in Portland, Oregon. Transportation Res Rec 2001;1780:33–42.
14. Parsons Brinckerhoff Quade &Douglas. The pedestrian environment. Portland OR: 1000 Friends of Oregon, 1993.
15. Gordon-Larsen P, Nelson MC, Page P, Popkin BM. Inequality in the built environment underlies key health disparities. Pediatrics 2006;117:417–24. Available at: www.pediatrics.org/cgi/doi/10.1542/peds.2005-0058.
16. Giles-Corti B, Donovan RJ. Socioeconomic status differences in recreational physical activity levels and real and perceived access to a supportive physical environment. Prev Med 2002;35:601–11.
17. Sallis JF, Hovell MF, Hofstetter CR, et al. Distance between homes and exercise facilities related to frequency of exercise among San Diego residents. Public Health Rep 1990;105:179–85.
18. Brownson RC, Baker EA, Housemann RA, Brennan LK, Bacak SJ. Environmental and policy determinants of physical activity in the United States. Am J Public Health 2001;91:12.
19. Ewing R, Schmid T, Killingsworth R, Zlot A, Raudenbush S. Relationship between urban sprawl and physical activity, obesity, and morbidity. Am J Health Promot 2003;18:47–57.
20. Popkin BM, Udry JR. Adolescent obesity increases significantly in second and third generation U.S. immigrants. J Nutr 1998;128:701–6.
21. 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.
22. Gordon-Larsen P, McMurray RG, Popkin BM. Determinants of physical activity and inactivity patterns. Pediatrics 2000;105:e83.
23. Sallis JF, Buono MJ, Roby JJ, Micale FG, Nelson JA. Seven-day recall and other physical activity self-reports in children and adolescents. Med Sci Sport Exerc 1993;25:99–108.
24. Andersen RE, Crespo CH, Bartlett SJ, Cheskin LJ, Pratt M. Relationship of physical activity and television watching with body weight and level of fatness among children: results from the Third National Health and Nutrition Examination Survey. JAMA 1998;279:938–42.
25. Baranowski T. Validity and reliability of self-report measures of physical activity: an information processing perspective. Res Q Exerc Sport 1988;59:314–27.
26. Pate RR, Heath GW, Dowda M, Trost SG. Associations between physical activity and other health behaviors in a representative sample of U.S. adolescents. Am J Public Health 1996;86:1577–81.
27. Ainsworth B, Haskell WL, Leon AS. Compendium of physical activities: classification of energy costs of human physical activities. Med Sci Sport Exerc 1993;25:71–80.
28. Pate RR, Pratt M, Blair SN, et al. Physical activity and public health: a recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. JAMA 1995;273:402–7.
29. American Academy of Pediatrics Committee on Public Education. Children, adolescents, and television. Pediatrics 2001;107:423–6.
30. American Academy of Pediatrics Committee on Public Education. Media violence. Pediatrics 2001;108:1222–6.
31. WHO Expert Committee. Physical status: the use and interpretation of anthropometry. Geneva: World Heath Organization, 1995 (WHO Technical Report Series 854).
32. Centers for Disease Control and Prevention. Growth charts 2000: United States. National Center for Health Statistics. Available at: www.cdc.gov/growthcharts. Accessed April 21, 2003.
33. Goodman E, Hinden BR, Khandelwal S. Accuracy of teen and parental reports of obesity and body mass index. Pediatrics 2000;106:52–8.
34. Aldenderfer MS, Blashfield RK. Cluster analysis. Beverly Hills CA: Sage, 1984 (Sage University Paper Series on Quantitative Applications in Social Sciences, 44).
35. Nelson MC, Gordon-Larsen P, Adair LS, Popkin BM. Adolescent physical activity and sedentary behavior: patterning and long-term maintenance. Am J Prev Med 2005;28:259–66.
36. McNutt LA, Wu C, Xue X, Hafner JP. Estimating the relative risk in cohort studies and clinical trials of common outcomes. Am J Epidemiol 2003;157:940–3.
37. Lopez R. Urban sprawl and risk for being overweight or obese. Am J Public Health 2004;94:1574–9.
38. Vandegrift D, Yoked T. Obesity rates, income, and suburban sprawl: an analysis of U.S. states. Health Place 2004;10:221–9.
39. Gordon-Larsen P, Nelson MC, Popkin BM. Meeting national activity and inactivity recommendations: adolescence to adulthood. Am J Prev Med 2005;28:259–66.
40. Gordon-Larsen P, McMurray RG, Popkin BM. Adolescent physical activity and inactivity vary by ethnicity: the National Longitudinal Study of Adolescent Health. J Pediatr 1999;135:301–6.
41. Patterson PD, Moore CG, Probst JC, Shinogle JA. Obesity and physical inactivity in rural America. J Rural Health 2004;20:151–9.
42. Duany A, Talen E. Transect planning. J Am Planning Assoc 2002;68:245–66.