Health & Place 26 (2014) 118–126

Contents lists available at ScienceDirect

Health & Place

journal homepage: www.elsevier.com/locate/healthplace

Relationship between urban sprawl and physical activity, obesity, and morbidity – Update and refinement

Reid Ewing *, Gail Meakins [1], Shima Hamidi [2], Arthur C. Nelson [3]

Department of City and Metropolitan Planning, College of Architecture and Planning, 375 S 1530 E RM 235, Salt Lake City, UT 84112, USA

ARTICLE INFO

Article history:
Received 11 April 2013
Received in revised form
10 December 2013
Accepted 15 December 2013
Available online 21 December 2013

Keywords:
Obesity
Compactness

ABSTRACT

Aims: This study aims to model multiple health outcomes and behaviors in terms of the updated, refined, and validated county compactness/sprawl measures.
Methods: Multiple health outcomes and behaviors are modeled using multi-level analysis.
Results: After controlling for observed confounding influences, both original and new compactness measures are negatively related to BMI, obesity, heart disease, high blood pressure, and diabetes. Indices are not significantly related to physical activity, perhaps because physical activity is not defined broadly to include active travel to work, shopping, and other destinations.
Conclusions: Developing urban and suburban areas in a more compact manner may have some salutary effect on obesity and chronic disease trends.

1. Introduction

The prevalence of adult obesity and overweight in the United States has risen significantly in the last 30 years (Khan et al., 2009). Data from the 2009–2010 National Health and Nutrition Examination Survey (NHANES) indicate that 36% of adults and 17% of youth are obese (Ogden et al., 2012). If these trends continue, more than 44% of people in the United States could be obese by the year 2030 (Levi et al., 2012). The rising prevalence of obesity presents serious long term challenges including the increased prevalence of chronic diseases resulting in decreased life expectancy, the potential for negative impacts on an individual′s quality of life, the availability and cost of future health care, and the viability and productivity of future generations (Trogdon et al., 2008).

The fundamental cause of obesity and overweight is an imbalance between calories consumed and calories expended. While there are many influences impacting both weight and health, including genetics, socioeconomic status, race/ethnicity, and gender, two modifiable risk factors are unhealthy diets and physical inactivity, both of which have a spatial component (Black and Macinko, 2008; Trost et al., 2001, 2002). Physical inactivity has been identified as the fourth leading risk factor for global mortality causing an estimated 3.2 million deaths annually (WHO, 2012). It is commonly recognized that even a moderate amount of physical activity can result in significant health benefits (Centers for Disease Control and Prevention, 2009). Yet current research indicates that physical activity levels have declined, with many adults in the United States (43%) failing to meet the recommended physical activity requirements (CDC, 2009). In the last fifty years, activity levels have dropped for a variety of reasons including new technologies and automation that make our lives easier, television and computer use, and changes in the built environment that have led to sedentary life styles (Transportation Research Board and Institute of Medicine Committee on Physical Activity, Health, Transportation, and Land Use, 2005). Automobile use has substituted for active travel, and urban sprawl, the dominant development pattern in the United States, all but guarantees automobile dependence (Committee on Physical Activity, 2005).

In this article, we update a “sprawl index” first associated with obesity in 2003 (see Ewing et al., 2003b). The update is to 2010, using recent census and other data. The reason for updating is provide researchers and policy professionals with current data on sprawl, the earlier metrics now being more than decade old. Additionally, we develop a refined version of the same index that incorporates additional built environmental variables. The earlier metrics only covered two dimensions of sprawl, development density and street accessibility, while sprawl is often defined in terms of land use diversity (or lack thereof) and population and employment centering

* Corresponding author. Tel.: +1 801 581 8255.
E-mail addresses: ewing@arch.utah.edu (R. Ewing), g.meakins@gmail.com (G. Meakins), shima.hamidi@gmail.com (S. Hamidi), acnelson@utah.edu (A.C. Nelson).
[1] Tel.: +1 801 891 5179.
[2] Tel.: +1 801 581 8254.
[3] Tel.: +1 801 581 8253.

1353-8292
http://dx.doi.org/10.1016/j.healthplace.2013.12.008

(or lack thereof). The refined metrics cover all four dimensions. Principal component analysis is used to derive a density factor from five density variables, a mix factor from three variables, a centering factor from four variables, and a street accessibility factor from four variables. Finally, we apply the resulting indices to health data from the Behavioral Risk Factor Surveillance System (BRFSS) to see if reported relationships have changed over the decade since the first index was published. Checking for stability in relationships over time is the best way to check the validity and reliability of the 2003 results.

1.1. Literature

In 2003, Ewing et al. (2003b) first established a relationship between health behaviors, health outcomes, and a "county sprawl index," which became the most widely cited academic article in the Social Sciences as of late 2005, according to *Essential Science Indicators* (Reuters). After controlling for age, education, fruit and vegetable consumption, and other sociodemographic and behavioral covariates, they found that adults living in sprawling counties have higher body mass indices (BMIs) and are more likely to be obese (BMI > 30) than are their counterparts living in compact counties.

In the years since the original study, there has been a plethora of research studies in both planning and public health investigating the relationship between the built environment and health outcomes (Galvez et al., 2010, p. 202; Casey et al., 2011; Dunton et al., 2009; Finkelstein et al., 2005; Feng et al., 2010; Lachowycz and Jones, 2011; Papas et al., 2007; Withrow and Alter, 2010). Research has established statistically significant links between elements of the built environment and the risk of obesity (Booth et al., 2005; Papas et al., 2007; Feng et al., 2010), suggesting that some built environments may be more "obesogenic" than others (Black and Macinko, 2008).

Also since the original study, there have been numerous applications of the original county sprawl index (which has also been referred to as a compactness index, since compactness and sprawl anchor opposite ends of the same scale). The original sprawl index was made available to researchers who wished to explore the various costs and benefits of sprawl. Sprawl has now been linked, in one or another study, to physical inactivity, obesity, traffic fatalities, poor air quality, residential energy use, emergency response times, teenage driving, lack of social capital, private-vehicle commute distances and times, and coronary heart disease (Ewing et al., 2003a, 2003b, 2003c; Kelly-Schwartz et al., 2004; Sturm and Cohen, 2004; Cho et al., 2006; Doyle et al., 2006; Ewing et al., 2006; Kahn, 2006; Kim et al., 2006; Plantinga and Bernell, 2007; Ewing and Rong, 2008; Joshua et al., 2008; Stone, 2008; Trowbridge and McDonald, 2008; Fan and Song, 2009; McDonald and Trowbridge, 2009; Trowbridge et al., 2009; Lee et al., 2009; Nguyen, 2010; Stone et al., 2010; Schweitzer and Zhou, 2010; Gregson, 2011; Kostova, 2011; Zolnik, 2011; Holcombe and Williams, 2012; Griffin et al., 2012; James et al., 2013; Bereitschaft and Debbage, 2013).

1.2. Geographic scale

Since the 2003 study, most investigators have chosen to characterize the built environment of individuals at the neighborhood scale, whether in terms of census tracts, block groups, or small buffers around individuals' homes (starting with Frank et al., 2004). There has been an implicit assumption that walking distance from home is the operative scale at which the built environment affects physical activity, food availability, and ultimately weight. This is just an assumption. Adults spend most of their waking hours away from home. An estimated 30–40% of all trips are non-home-based. A sprawling metropolitan area produces long commutes, which cut into leisure time and hence physical activity. Access to healthy foods may be more difficult in sprawling environments. To our knowledge, only one study has compared the power of neighborhood and county environments as predictors of obesity (Joshua et al., 2008). While this study found that perceived neighborhood characteristics were more important than objectively measured county characteristics, it is likely that environmental factors at both scales are relevant for understanding obesity and physical activity. Better measures of macro-scale characteristics such as sprawl are needed to represent the broad settings that shape people's health-related activities.

2. Methods

2.1. Data and Measures

This study represents the built environment at the county scale rather than the smaller neighborhood scale. The main reason is expediency, since the health database used in this study, for reasons of confidentiality, only supplies geocodes for respondents by county, and then only for larger counties. However, the preceding discussion suggests that the county may be an appropriate scale for health research in an auto-oriented society like our own.

Health-related data come from the Behavioral Risk Factor Surveillance System (BRFSS), a telephone survey conducted by state health departments and managed by the Centers for Disease Control and Prevention (CDC). Over 350,000 adults are interviewed nationally each year to collect detailed information on health risk behaviors, preventive health practices, and health care access primarily related to chronic disease and injury.

We use a subsample of individuals for which county geocodes of residence are available for public use. Our data come from the Selected Metropolitan/Micropolitan Area Risk Trends (SMART) project which is populated with BRFSS data for metropolitan and micropolitan statistical areas with 500 or more respondents. We have included data for survey years 2007 through 2010. Different questions are asked in different survey years (all four years for most variables but only two years for some). This accounts for the different sample sizes for different variables in Table 1.

Our health outcome variables fall into three categories: weight status, physical activity, and chronic diseases. Weight status variables are calculated from self-reported height and weight. Body mass index (BMI) is a continuous variable defined as weight in kilograms divided by height in meters squared (kg/m^2). Obesity status is dichotomous, defined as having a BMI greater than or equal to 30.0. Weight status data are available for all four survey years.

One physical activity outcome is dichotomous: whether a respondent reported "any physical activity." The question reads: "During the past month, other than your regular job, did you participate in any physical activities or exercises such as running, calisthenics, golf, gardening, or walking for exercise?" This question is included in all four years of the BRFSS survey. The phrasing, particularly the reference to exercise, likely means the kind of active travel that occurs in compact areas will not be reported by respondents.

A second physical activity variable is continuous: minutes of moderate physical activity per week, which presumably includes the kind of walking we expect to see in compact areas. The 2003 study found that minutes of leisure-time walking were positively related to county compactness (Ewing et al., 2003b). This was the only physical activity variable with a significant relationship to compactness. More recent surveys have not asked about specific physical activities such as walking and bicycling, but instead have asked about moderate and vigorous physical activity generally. If any relationship is likely to show up between compactness and physical activity, it will be in minutes of moderate activity.[4] This is

[4] In BRFSS 2007 and 2009, respondents were asked if they engaged in moderate physical activities outside work for at least 10 min at a time. The specific question included in the 2007 and 2009 surveys was as follows:

Table 1
BRFSS variables and sample sizes.

	Scale	n
Level 1 dependent variables		
Body mass index	Continuous	675,784
Obese status	Dichotomous	675,784
Any physical exercise	Dichotomous	709,099
Minutes of moderate physical activity per week	Continuous	159,965
Diagnosed high blood pressure	Dichotomous	354,826
Diagnosed heart disease	Dichotomous	703,942
Diagnosed diabetes	Dichotomous	709,234
Level 1 independent variables		
Male	Dichotomous	709,889
Age 30–44	Dichotomous	709,889
Age 45–64	Dichotomous	709,889
Age 65–74	Dichotomous	709,889
Age 75+	Dichotomous	709,889
Black non-Hispanic	Dichotomous	701,572
Other non-Hispanic	Dichotomous	701,394
Hispanic	Dichotomous	701,572
Less than high school	Dichotomous	709,889
High school graduate	Dichotomous	709,889
Some college	Dichotomous	709,889
Income < $25,000	Dichotomous	709,889
Income $25,000–$50,000	Dichotomous	615,801
Income $50,000–$75,000	Dichotomous	615,801
Current smoker	Dichotomous	705,549
Recommended servings fruits/vegetables	Dichotomous	343,695
Level 2 independent variables		
Violent crime rate per 100,000 population	Continuous	316
Annual precipitation	Continuous	316
Annual heating degree days	Continuous	316
Annual cooling degree days	Continuous	316
Percentage park land (relative to total land area)	Continuous	316
Original county compactness index for 2010 (using the same 2000 index variables)	Continuous	316
New county compactness index for 2010 (including additional variables compared to 2000 index)	Continuous	316

the case because walking, repeatedly shown to have a relation to the built environment, involves moderate activity.

Chronic disease outcomes are the status conditions of hypertension, diabetes, and coronary heart disease as diagnosed by a health care professional and reported by the respondent. These variables are also dichotomous.

Gender, age, race/ethnicity, income, and educational attainment are included as control variables representing individual-level sociodemographics. The reference groups for these variables are females, white non-Hispanics, college graduates, persons aged 18–30 years, and income $75,000 or greater. Smoking status and fruit and vegetable consumption are also included as control variables representing individual health behaviors.

2.2. Analysis method

Individual respondents are nested within counties and hence share the characteristics of the county's built, social, and natural environments. Having individuals nested within counties, this data set violates the independence assumption of Ordinary Least Squares (OLS). So instead of OLS, models were estimated using multi-level modeling and the statistical package HLM 6.08. Hierarchical linear modeling was applied to the continuous outcomes (BMI and minutes of moderate physical activity per week), while hierarchical nonlinear modeling (logistic modeling) was applied to the dichotomous outcomes (all other outcome variables). Two different sets of models were estimated, differing only in that the original county sprawl index was controlled in one set of models (Original Models) and a new refined index was controlled in the other set (Refined Models). Selection of models was based on significance levels and log-likelihood ratios.

3. Results

3.1. Original sprawl index

The first county compactness index developed for 2010 is almost identical to the index for 2000 used in the 2003 study (Ewing et al., 2003b). Using principal component analysis, six variables were reduced to one, that being the principal component that accounted for the greatest variance in the dataset. Factor loadings (that is, correlations of these variables with the first principal component) are shown in Table 2. The eigenvalue of the first principal component is 3.56, which means that this one variable accounts for more of the variance in the original dataset than three of the original variables combined.

As expected, four of the variables load positively on the first principal component: gross population density of urban and suburban census tracts; percentage of the population living at gross densities of more than 12,500 persons per square mile, a transit-supportive density; net population density of lands

(*footnote continued*)

We are interested in two types of physical activity: vigorous and moderate. Vigorous activities cause large increases in breathing or heart rate while moderate activities cause small increases in breathing or heart rate. Now, thinking about the moderate activities you do when you are not working (if employed) in a usual week, do you do moderate activities for at least 10 min at a time, such as brisk walking, bicycling, vacuuming, gardening, or anything else that causes some increase in breathing or heart rate?

If the answer was "yes," follow-up questions were asked about amounts of physical activity. Some of the values provided were unrealistically high. Therefore, values were truncated at 1260 min a week, which represents 3 h per day 7 days a week and included 99% of all respondents.

Table 2
Original county sprawl index variables and factor loadings in 2010 (data sources).

Observed variable	Factor loading[a]	
Gross population density, excluding rural census tracts with fewer than 100 persons per square mile	0.858	2010 Census
Percentage of the population living at less than 1500 persons per square mile, a low suburban density	−0.658	2010 Census
Percentage of the population living at more than 12,500 persons per square mile, a transit-supportive urban density	0.821	2010 Census
Net population density of developed lands	0.876	2010 Census and national land cover database
Average block size, excluding rural blocks of greater than one square mile.	−0.664	2010 Census
Percentage of small urban blocks less than 0.01 square miles	0.711	2010 Census
Eigenvalue	3.56	
Explained variance	59.3%	

[a] Correlation with county sprawl index.

classified as developed; and percentage of census blocks of less than 0.01 square miles, or about 500 feet on a side, an urban block. Also, as expected, two of the variables load negatively on the first principal component: the percentage of population living at less than 1500 persons per square mile, a low suburban or exurban density; and average block size, which is inversely related to street accessibility. Thus, for all component variables, better accessibility translates into higher values of the first principal component.[5]

We transformed the overall compactness score into an index with a mean of 100 and a standard deviation of 25, similar to an IQ score. This was done for the sake of consistency with the 2003 study, and ease of understanding. With this transformation, the more compact counties have index values above 100, while the more sprawling have values below 100.

The most compact counties are as expected, central counties of large, older metropolitan areas. The most sprawling counties are outlying counties of large metropolitan areas, or component counties of smaller metropolitan areas.

3.2. Refined sprawl indices

The original county sprawl index operationalized only two dimensions of urban form – residential density and street accessibility. Grants from NIH and the Brookings Institution provided for the development of refined measures of county compactness. These measures are modeled after the more complete metropolitan sprawl indices developed by Ewing et al. (2002). The refined indices operationalize four dimensions, thereby characterizing county sprawl in all its complexity. The four are development density, land use mix, population and employment centering, and street accessibility. The dimensions of the new county indices parallel the metropolitan indices, basically representing the relative accessibility from one land use to another provided by the county.

Data sources for the new index, and the variables derived from each sources, are described in detail elsewhere (Ewing and Hamidi, 2014). Principal components were extracted from each set of variables, and the principal component that accounted for the greatest variance became our density, mix, centering, and street accessibility factors (Table 3).[6] While correlated, as one might expect, the four compactness factors seem to represent distinct constructs based on their bivariate correlations (Table 4).

The next issue we had to wrestle with was how to combine the four factors into a single sprawl index. A priori, there is no “right” way to do so, only ways that have more or less face validity. Should the four factors be weighted equally, or should one or another be given more weight than the others? Density has certainly received more attention as an aspect of sprawl than has, say, street accessibility. However, beyond play in the literature, we could think of no rationale for differential weights. The factors all contribute to the accessibility or inaccessibility of different development patterns, none presumptively more than the others. Depending on their values, all move a county along the continuum from sprawl to compact development. Thus they were simply summed, in effect giving each dimension of sprawl equal weight in the overall index.

As with the individual sprawl factors, we transformed the overall compactness score into an index with a mean of 100 and a standard deviation of 25. This was done for the sake of consistency with original compactness index. With this transformation, the more compact counties have index values above 100, while the more sprawling have values below 100.

The ten most compact and ten most sprawling counties are shown in Tables 5 and 6. The most compact counties are central counties of large, older metropolitan areas. The most sprawling counties are outlying counties of large metropolitan areas, or component counties of smaller metropolitan areas. Values range from 42 for Oglethorpe County, GA outside Athens, the most sprawling county in 2010, to 332 for New York County (Manhattan), the most compact county in 2010 (see Figs. 1 and 2). County compactness (sprawl) scores for 2010, both original and refined, are posted on an NIH website (http://gis.cancer.gov).

It would seem that the original and new compactness indices are measuring the same construct, but that is not quite true. The original compactness index is dominated by density variables (four of six variables in the index) and only slightly diluted by street variables (two of the six), which correlate strongly with density. The new compactness index dilutes the role of density by adding two new factors (mix and centering). The simple correlation coefficient between original and new indices is 0.865, which means that about 25% of the variance in each index is unexplained by the other. We would expect that they have similar but not identical relationships to outcome variables, and similar but not identical predictive power.

[5] To derive the county compactness index, the first principal component, which has a mean of 0 and standard deviation of 1, was transformed to a scale with a mean of 100 and standard deviation of 25. This transformation produces a more familiar metric (like an IQ scale) and ensures that all values will be positive, thereby enhancing our ability to test for nonlinear relationships. With this transformation, the more compact counties have scores above 100, while the more sprawling have scores below 100.

[6] One principal component represents density, another mix, a third centering, and a fourth street accessibility. County principal component values, standardized such that the mean value of each is 100 and the standard deviation is 25, were summed to create one overall compactness index, which was also on a scale with a mean of 100 and a standard deviation of 25. The simple structure of the original county sprawl index became more complex, but also more nuanced and comprehensive, in line with definitions of sprawl in the technical literature.

Table 3
New county sprawl index variables and factor loadings in 2010.

Observed variable	Factor loading	Data source
County density factor		
Gross population density of urban and suburban census tracts	0.983	2010 Census
Percentage of the population living at low suburban densities	0.848	2011 Census
Percentage of the population living at medium to high urban densities	−0.440	2012 Census
Net population density of urban lands	0.850	2006 NLCD
Gross employment density of urban and suburban census tracts	0.977	2010 LED
Eigenvalue	3.56	
Explained variance	71.1%	
County mix factor		
Job-population balance which measures the countywide average degree of balance between jobs and residents	0.891	2010 Census 2010 LED
Degree of job mixing which measures the countywide average degree of job mixing using an entropy formula	0.942	2011 LED
Walk score which measure the countywide average walk score for census tracts in the county	0.784	Walk Score, Inc.
Eigenvalue	2.30	
Explained variance	76.6%	
County centering factor		
Coefficient of variation in census block group population densities, defined as the standard deviation of block group densities divided by the average density of block groups.	0.085	2010 Census
Coefficient of variation in census block group employment densities, defined as the standard deviation of block group densities divided by the average density of block groups.	0.642	2010 LED
Percentage of county population in CBD or sub-centers	0.820	2010 Census
Percentage of county employment in CBD or sub-centers	0.932	2010 LED
Eigenvalue	1.43	
Explained variance	49.1%	
County street factor		
Average block size excluding rural blocks of more than one square mile	−0.764	2010 Census
Percentage of small urban blocks of less than one hundredth of a square mile	0.901	2011 Census
Intersection density for urban and suburban census tracts within the county, excluding rural tracts with gross densities of less than 100 persons per square mile	0.836	ESRI (TomTom)
Percentage of 4-or-more-way intersections, again excluding rural tracts	0.545	ESRI (TomTom)
Eigenvalue	2.39	
Explained variance	59.8%	

Table 4
Simple correlations among compactness/sprawl factors.

	Density factor	Mix factor	Centering factor	Street factor
Density factor				
Pearson correlation	1	0.399**	0.523**	0.583**
N	981	973	977	980
Mix factor				
Pearson correlation	0.399**	1	0.421**	0.647**
N	973	980	969	979
Centering factor				
Pearson correlation	0.523**	0.421**	1	0.438**
N	977	969	977	977
Street factor				
Pearson correlation	0.583**	0.647**	0.438**	1
N	980	979	977	992

** Correlation is significant at the 0.01 level (2-tailed).

Table 5
10 most compact counties in 2010 according to the four-factor index (excluding Massachusetts counties).

		Metropolitan area	Index
1	New York County, NY	New York–Northern New Jersey–Long Island, NY–NJ–PA	425.2
2	Kings County, NY	New York–Northern New Jersey–Long Island, NY–NJ–PA	265.2
3	San Francisco County, CA	San Francisco–Oakland–Fremont, CA	251.3
4	Bronx County, NY	New York–Northern New Jersey–Long Island, NY–NJ–PA	224.0
5	Philadelphia County, PA	Philadelphia–Camden–Wilmington, PA–NJ–DE–MD	207.2
6	District of Columbia, DC	Washington–Arlington–Alexandria, DC–VA–MD–WV	206.4
7	Queens County, NY	New York–Northern New Jersey–Long Island, NY–NJ–PA	204.2
8	Baltimore city, MD	Baltimore–Towson, MD	190.9
9	Norfolk city, VA	Virginia Beach–Norfolk–Newport News, VA–NC	179.6
10	Hudson County, NJ	New York–Northern New Jersey–Long Island, NY–NJ–PA	178.7

Table 6
10 most sprawling counties in 2010 according to the four-factor index (excluding Massachusetts counties).

		Metropolitan area	Index
958	Spencer County, KY	Louisville/Jefferson County, KY–IN	60.4
959	Morrow County, OH	Columbus, OH	58.8
960	Brown County, IN	Indianapolis–Carmel, IN	58.5
961	Blount County, AL	Birmingham–Hoover, AL	56.6
962	Greene County, NC	Greenville, NC	56.6
963	Harris County, GA	Columbus, GA–AL	55.1
964	Elbert County, CO	Denver–Aurora–Broomfield, CO	54.3
965	Macon County, TN	Nashville–Davidson–Murfreesboro–Franklin, TN	54.3
966	Grant Parish, LA	Alexandria, LA	53.8
967	Oglethorpe County, GA	Athens–Clarke County, GA	45.5

Fig. 1. Most compact county according to the four-factor index (New York County, NY).

Fig. 2. Most sprawling county according to the four-factor index (Oglethorpe County, GA).

Compared to the original county compactness index, the new four-factor index has greater construct and face validity. It has greater construct validity because it captures four different dimensions of the construct “compactness” (density, mix, centering, and street accessibility), whereas the original index captures only two dimensions (density and street accessibility).

The greater face validity of the new four-factor index requires some explanation. The very first county compactness indices were derived for only 448 counties in the largest 101 metropolitan areas (Ewing et al., 2003b). The most sprawling counties, such as Geauga County outside Cleveland, have “classic sprawl patterns” of low-density suburban development, segregated land uses, commercial strips, and curvy streets ending in cul-de-sacs.

Expanding to 994 counties and adding smaller metropolitan areas, the picture becomes more complicated. The ten most compact counties based on the original index largely overlap with the top ten based on the new index (with the notable exception of Suffolk County (Boston), for which we do not have all required variables). New York County (Manhattan) is the most compact according to both indices. Kings County (Brooklyn) is the second most compact according to both indices.

However, the ten most sprawling counties are entirely different when measured by different indices. Which index has greater face validity? We reviewed satellite imagery for the ten most sprawling counties, according to both indices, and found that the development patterns for the new index are much more representative of classic suburban sprawl. While all 20 counties are part of metropolitan areas, many of the counties rated as most sprawling according to the original index have different development patterns than expected. They would best be described as exurban counties with small towns surrounded by farmlands. The small towns have moderate densities and gridded streets. The fact they are part of larger census tracts, our units of analysis, depresses their densities and compactness scores. They are not examples of classic suburban or exurban sprawl. On the other hand, the counties rated as most sprawling according to the new four-factor index have census tracts with very low residential densities, commercial strips, and cul-de-sac street networks.

3.3. Commuting

The relationships between sprawl and travel outcomes can be used to validate our county sprawl measures, and also to see if one measure has more predictive power than the other. If sprawl has any consistently recognized outcome, it is automobile dependence. In the final report of our NIH project, we validated our compactness/sprawl measures against vehicle ownership and commuting data from the 2010 American Community Survey, 5-year estimates (Ewing and Hamidi, 2014). We would expect to find, and found, that after controlling for other relevant influences, compact counties have relatively low vehicle ownership, high transit and walking mode shares on work trips, and short drive times to work. The “other relevant influences” were socioeconomics, climate, fuel price, and metropolitan area size. Both compactness indices were significant in the expected directions. The original county compactness index was more strongly related to average household vehicle ownership and transit mode share, while the new index was more strongly related to walk mode share and average drive time to work. Results for the four individual compactness factors, presented in the final report of the NIH project, are generally supportive of the hypothesis that

Table 7
Relationships between county compactness and public health outcomes (controlling for individual and other county-scale variables).

	Original compactness index			Refined compactness index		
	Coeff.	T-ratio	PseudoR2	Coeff.	T-ratio	PseudoR2
BMI	−0.00897	−4.25***	0.079	−0.00910	−4.50***	0.079
Obesity	−0.0035	−4.75***	0.50	−0.00401	−6.11***	0.57
Any physical activity	0.000012	0.021	0.67	0.00079	1.41	0.67
Moderate physical activity	−0.14	−2.43**	0.16	−0.154	−2.46**	0.16
High blood pressure	−0.0018	−3.15**	0.54	−0.0021	−3.62***	0.55
Coronary heart disease	−0.0024	−1.79	0.51	−0.0028	−2.12*	0.58
Diabetes	−0.0015	−2.22*	0.57	−0.0016	−2.27*	0.58

* $p=0.05$.
** $p=0.01$.
*** $p=0.001$.

compact development reduces automobile dependence (Ewing and Hamidi, 2014).

3.4. Weight status

BMI is higher for males than females and for black non-Hispanics and Hispanics than white non-Hispanics (see Table 7). BMI is lower for other non-Hispanics than whites, the former being mostly Asian. BMI is higher for those with less education, though the pattern is complicated. BMI increases with age until age 65, when it begins to decline. BMI declines with income following a smooth curve. BMI is lower for those who smoke and those who meet the recommended servings of five fruit or vegetables a day. BMI is lower in counties with more land devoted to parks, and higher in counties with more heating degree days. In both cases, there is a logical connection to physical activity. Parks are sites of physical activity, and also generate access trips by walking and bicycle.

As for the variables of greatest interest, those related to the built environment, BMI is negatively related to the original and new compactness indices. The new compactness index has a slightly stronger relationship to BMI and a slightly higher significance level. By either measure, controlling for sociodemographic and behavioral covariates, residents of more compact counties have a lower BMIs.

A one standard deviation increase in the compactness index (that is, a 25 point increase) translates into almost a two pound drop in weight for the average American male (5 ft 10 in. tall). This is about half the effect size of eating recommended servings of fruits and vegetables and a third the effect size of smoking, both of which reduce weight.

Comparatively, compactness measures have a larger effect and higher significance level in this study than in the 2003 study (Ewing et al., 2003b). This is not due to the sample size. There are actually fewer counties represented in the 2010 sample. Perhaps the reason for the added significance is the limitation of this sample to larger and more urban counties with a minimum of 500 BRFSS respondents. The earlier sample included some counties that are much smaller and less urban.

As with BMI, males, black non-Hispanics, and Hispanics have higher probabilities of being obese than their reference categories (see Table 7). The likelihood of obesity declines with income and generally with education. Obesity is less prevalent among current smokers and those who consume the recommended servings of fruits and vegetables. In both models, the likelihood of obesity is higher for residents of counties with more heating degree days. In two models, prevalence of obesity is higher in counties with more violent crime. In another, the prevalence of obesity is lower in counties with more park space. There are logical connections to physical activity. Controlling for covariates, the original and new compactness indices are negatively related to obesity, with the new index having a slightly stronger relationship. Results for the four individual compactness factors, presented in the final report of the NIH project, are generally supportive of the hypothesis that compact development reduces overweight and obesity (Ewing and Hamidi, 2014).

3.5. Physical activity

The first PA variable indicates whether a respondent engaged in any physical activity. Males are more likely to be physically active than females (Table 7). The likelihood of any physical activity generally increases with education and income, and declines with age. The two compactness indices are not significantly related to engaging in any physical activity. This result parallels that of the 2003 study (Ewing et al., 2003b).

The second physical activity variable is a calculated value for minutes of moderate physical activity per week (see Table 7). Results are similar to those for the other PA variable, except that respondents with less education get more moderate physical activity than those with college degrees. As one might expect, moderate physical activity is negatively related to precipitation, heating degree days, and cooling degree days, these environmental conditions apparently discouraging outside activities. Surprisingly, after controlling for other variables, moderate physical activity appears to be negatively related to county compactness.

3.6. Chronic diseases

Three health outcomes were modeled in this study: the diagnosed conditions of high blood pressure, heart disease, and diabetes (Table 7). These are known to be related to obesity and physical inactivity, and the former at least is related to sprawl. However, these three chronic conditions are “downstream” outcomes of obesity and physical inactivity, and highly dependent on diet and heredity. Thus, a priori, we cannot say whether they will have a relationship to the urban sprawl.

Largely tracking the findings for obesity, the prevalence of diabetes, heart disease, and high blood pressure increases with age, declines with income, and declines with education. All three conditions are more prevalent in males than females. The effect of race is mixed.

More interestingly, the two overall compactness indices are negatively associated with all three chronic diseases. The new compactness index has slightly stronger relationships to these conditions than does the original compactness index.

4. Discussion

The prevalence of adult obesity and overweight in the United States has risen significantly in the last 30 years. It is widely agreed that environmental factors, broadly defined, have played a key role in this trend. Many studies have now linked obesity and overweight to characteristics of the built environment (Booth et al., 2005; Papas et al., 2007; Feng et al., 2010; Black and Macinko, 2008). Urban sprawl has been implicated in rising obesity (Kelly-Schwartz et al., 2004; Sturm and Cohen, 2004; Lopez, 2004; Doyle et al., 2006; Ewing et al., 2006; Plantinga and Bernell, 2007; Joshua et al., 2008). The causal mechanism is debatable, but low levels of active travel may be partially responsible. Rates of chronic diseases have also increased over the past 30 years. Urban sprawl has been implicated in the prevalence of high blood pressure and diabetes (Ewing et al., 2003b).

This study replicates the findings of the widely cited 2003 study (Ewing et al., 2003b) showing that, after controlling for confounding influences, residents of more compact counties have lower BMIs and lower probabilities of obesity and chronic diseases. There is ample evidence that active travel is more prevalent in compact areas, which provides a causal mechanism (Ewing and Cervero, 2010).

Why, then, are compactness indices unrelated to measures of physical activity derived from BRFSS? First, the measure of "any physical activity" appears to relate to leisure time physical activity. The term "exercise" appears in the BRFSS question. We do not postulate that this type of physical activity is associated with the built environment, only that active travel for work, shopping, and errands may be. Second, the continuous measure of physical activity would be unlikely to capture active travel, as the question is specifically worded to capture physical activity that causes an "increase in breathing or heart rate" (see footnote). Individuals may not perceive the walk to a bus stop or a store as doing so.

We believe that the new county sprawl indices are superior to the original index. Compared to the original county compactness index, the new four-factor index has greater construct validity, internal validity, and face validity. It has greater construct validity because it captures four different dimensions of the construct "compactness" (density, mix, centering, and street accessibility), whereas the original index captures only two dimensions (density and street accessibility).

The new index has greater internal validity in the sense that it is a stronger predictor of outcomes associated with sprawl. For virtually every outcome, the new index has a larger regression coefficient and higher significance level.

The greater face validity of the new four-factor index has been discussed at length above. Basically, the new index assigns the lowest values to counties with classic patterns of suburban sprawl, whereas the original index assigns the lowest values to counties consisting of small towns surrounded by farmland.

4.1. Limitations

This analysis is subject to important limitations that call for additional research. Because this study is ecological and cross sectional in nature, we cannot say that sprawl causes obesity, high blood pressure, or any other health condition. Our study simply indicates that sprawl is associated with these conditions. Future research using longitudinal data is needed to tackle the more difficult job of testing for causation (Transportation Research Board and Institute of Medicine Committee on Physical Activity, Health, Transportation, and Land Use, 2005).

The presumptive relationships between sprawl and health are multiple and complex (Booth et al., 2005; Frank et al., 2006; Kelly-Schwartz et al., 2004; Rodriguez et al., 2006). In particular, leisure time physical activity constitutes only one of four major sources of physical activity. Greater precision in characterizing physical activity will help disentangle the effects of sprawl on health (Trost et al., 2002; Westerterp, 2009).

We recognize that the relationships between sprawl and behavior or weight are probably not completely linear (Ewing et al., 2003b). It may 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 1 or 2 houses per acre to 3 or 4 may not meet the threshold needed for change.

This study relates health to the built environment at the county scale, which is large compared to the living and working environments of most residents (Black and Macinko, 2008; Feng et al., 2010). Geocodes are only available from BRFSS down to the county level. If environmental effects are felt most strongly at the community or neighborhood scale, these results may understate the effects of the built environment on health (Booth et al., 2005).

Omitted variables could bias our results. In particular, because they are not directly measured in any of the compactness/sprawl measures, this study does not account for many other environmental variables such as sidewalks and topography that may act directly or interact to influence physical activity and hence health. Attitudes toward walking and biking are also omitted variables. Residential self-selection, the tendency of those who want to be more physically active and physically fit to choose to locate in compact counties, may account for some the relationship between compactness and health outcomes (Cao et al., 2009; Frank et al., 2007; Plantinga and Bernell, 2007). Many travel studies show that attitudes about walking and transit use attenuate the effects of the built environment on travel (Cao et al., 2009), and the same could be true of physical activity.

By focusing on physical activity, this study largely ignores the other side of the energy equation, calories consumed as opposed to calories expended (Trost, et al., 2002; Biro & Wien, 2010; Black and Macinko, 2008; Feng et al., 2010). Only our fruit and vegetable consumption variable begins to get at that dimension of the problem. Caloric intake may 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 (Kelly et al., 2011; Larson et al., 2009).

This study is subject to the "multiple comparison fallacy." In inductive reasoning, there is always some chance that the conclusion will be false even if the evidence is true. When the confidence level is set at 95%, there is a probability of 1 in 20 – that is, 5% – that a misleading result will occur simply by chance. We estimated many models. As the number of models estimated increases, it becomes more likely that the compactness variables will appear significant in at least one model. Our confidence that a result will generalize to independent data will generally be weaker if it is observed as part of an analysis that involves multiple comparisons, rather than an analysis that involves only a single comparison.

Finally, the pseudo-R2s are low for models with continuous dependent variables. We have shown pseudo-R2s because urban planners, one of our target audiences, are used to dealing with R2s and may want this information. Pseudo-R2s in multi-level modeling are not equivalent to R2s in ordinary least squares regression, and should not be interpreted the same way. The pseudo-R2 bears some resemblance to the statistic used to test the hypothesis that all coefficients in the model are zero, but there is no construction by which it is a measure of how well the model predicts the outcome variable in the way that R2 does in conventional regression analysis.

Notwithstanding these limitations, this study adds to the weight of evidence that the built environment has a relationship to obesity. In the now slightly dated literature review by Papas et al. (2007), most articles (84%) reported a statistically significant association between some aspect of the built environment and obesity. Future research will have to delve further into the exact causal mechanism by which one affects the other (it does not appear to be by increasing or decreasing leisure time physical activity) and will need to strengthen causal inference by adding a longitudinal element to studies of sprawl vs. obesity (studying movers, in particular). One such study by the authors in currently in progress.

References

Biro, F.M., Wien, M., 2010. Childhood obesity and adult morbidities. Am. J. Clin. Nutr. 91 (5), 1499S–1505S.
Black, J.L., Macinko, J., 2008. Neighborhoods and obesity. Nutr. Rev. 66 (1), 2–20.
Booth, K.M., Pinkston, M.M., Poston, W.S., 2005. Obesity and the built environment. J. Am. Diet. Assoc. 105 (5 Suppl 1), S110–7.
Cao, X., Mokhtarian, P.L., Handy, S.L., 2009. Examining the impacts of residential self-selection on travel behaviour: a focus on empirical findings. Transp. Rev. 29 (3), 359–395.
Casey, R., Oppert, J.M., Weber, C., Charreire, H., Salze, P., Badariotti, D., Simon, C., 2011. Determinants of childhood obesity: what can we learn from built environment studies? Food Qual. Prefer. , http://dx.doi.org/10.1016/j.foodqual.2011.06.003.
Centers for Disease Control and Prevention, 2009. Fact Sheet for health professionals on physical activity guidelines for adults. In: CDC editors, Centers for Disease Control and Prevention, Division of Nutrition, Physical Activity, and Obesity. 〈http://www.health.gov/paguidelines/guidelines/default.aspx#toc〉. (accessed January 2012).
Cho, S.H., Chen, Z., Yen, S.T., Eastwood, D.B., 2006. The effects of urban sprawl on body mass index: where people live does matter. Consum. Interests Annu., 52.
Doyle, S., Kelly-Schwartz, A., Schlossberg, M., Stockard, J., 2006. Active community environments and health: the relationship of walkable and safe communities to individual health. J. Am. Plan. Assoc. 72 (1), 19–31.
Dunton, G.F., Kaplan, J., Wolch, J., Jerrett, M., Reynolds, K.D., 2009. Physical environmental correlates of childhood obesity: a systematic review. Obes. Rev. 10, 393–402.
Ewing, R., Pendall, R., Chen, D. 2002. Measuring Sprawl and Its Impacts. Smart Growth America, Washington, DC.
Ewing, R., Pendall, R., Chen, D., 2003a. Measuring Sprawl and Its Transportation Impacts. Transp. Res. Rec. 1832, 175–183.
Ewing, R., Schmid, T., Killingsworth, R., Zlot, A., Raudenbush, S., 2003b. Relationship between urban sprawl and physical activity, obesity, and morbidity. Am. J. Health Promot. AJHP 18 (1), 47–57.
Ewing, R., Schieber, R.A., Zegeer, C.V., 2003c. Urban sprawl as a risk factor in motor vehicle occupant and pedestrian fatalities. J. Inf. 93, 9.
Ewing, R., Brownson, R.C., Berrigan, D., 2006. Relationship between urban sprawl and weight of United States youth. Am. J. Prev. Med. 31 (6), 464–474.
Ewing, R., Rong, F., 2008. The impact of urban form on US residential energy use. Hous. Policy Debate 19 (1), 1–30.
Ewing, R., Cervero, R., 2010. Travel and the built environment: a meta-analysis. J. Am. Plan. Assoc. 76 (3), 265–294.
Ewing, R., Hamidi, S., 2014. Measuring Urban Sprawl and Validating Sprawl Measures. National Institutes of Health and Smart Growth America, Washington, DC.
Fan, Y., Song, Y., 2009. Is sprawl associated with a widening urban–suburban mortality gap? J. Urban Health 86 (5), 708–728.
Feng, J., Glass, T., Curriero, F., Stewart, W., Schwartz, B., 2010. The built environment and obesity: a systematic review of the epidemiologic evidence. Health Place 16, 175–190.
Finkelstein, E., Fiebelkorn, I., Wang, G., 2005. The costs of obesity among full-time employees. Am. J. Health Promot. 20 (1), 45–53.
Frank, L.D., Andresen, M.A., Schmid, T.L., 2004. Obesity relationships with community design, physical activity, and time spent in cars. Am. J. Prev. Med. 27 (2), 87–96.
Frank, L.D., Sallis, J.F., Conway, T.L., Chapman, J.E., Saelens, B.E., Bachman, W., 2006. Many pathways from land use to health: associations between neighborhood walkability and active transportation, body mass index, and air quality. J. Am. Plan. Assoc. 72 (1), 75–87.
Frank, L., Saelens, B., Powell, K.E., Chapman, J.E., 2007. Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity? Soc. Sci. Med. 65, 1898–1914.
Galvez, M.P., Meghan, P., Yen, I., 2010. Childhood obesity and the built environment: a review of the literature from 2008–2009. Curr. Opin. Pediatr. 22 (2), 202–207.
Gregson, J., 2011. Poverty, sprawl, and restaurant types influence body mass index of residents in California counties. Public Health Rep. 126 (Suppl 1), 141–149 (May–June).
Griffin, B.A., Eibner, C., Bird, C.E., Jewell, A., Margolis, K., Shih, R., Escarce, J.J., 2012. The relationship between urban sprawl and coronary heart disease in women. Health Place 20, 51–61.
Holcombe, R.G., Williams, D.W., 2012. Urban sprawl and transportation externalities. Rev. Reg. Stud. 40 (3), 257–272.
James, P., Troped, P.J., Hart, J.E., Joshu, C.E., Colditz, G.A., Brownson, R.C., Ewing, R., Laden, F., 2013. Urban sprawl, physical activity, and body mass index: nurses' health study and nurses' health study II. Am. J. Public Health 103 (2), 369–375.
Joshua, C.E., Boehmer, T.K., Brownson, R.C., Ewing, R., 2008. Personal, neighbourhood and urban factors associated with obesity in the United States. J. Epidemiol. Community Health 62, 202–208.
Kahn, M.E., 2006. The quality of life in sprawled versus compact cities. Prepared for the OECD ECMT Regional Round, Berkeley, California, pp. 27–28.
Kelly, B., Flood, V., Yeatman, H., 2011. Measuring local food environments: an overview of available methods and measures. Health Place 17, 1284–1293.
Kelly-Schwartz, A.C., Stockard, J., Doyle, S., Schlossberg, M., 2004. Is sprawl unhealthy? a multilevel analysis of the relationship of metropolitan sprawl to the health of individuals. J. Plan. Educ. Res. 24, 184–196.
Khan, L.K., Sobush, K., Keener, D., Goodman, K., Lowry, A., Kakietek, J., Zaro, S., 2009. Centers for disease control and prevention. recommended community strategies and measurements to prevent obesity in the United States. MMWR 58 (RR-7), 1–29.
Kim, D., Subramanian, S.V., Gortmaker, S.L., Kawachi, I., 2006. US state-and county-level social capital in relation to obesity and physical inactivity: a multilevel, multivariable analysis. Soc. Sci. Med. 63 (4), 1045–1059.
Kostova, D., 2011. Can the built environment reduce obesity quest; the impact of residential sprawl and neighborhood parks on obesity and physical activity. East Econ. J. 37 (3), 390–402.
Lachowycz, K., Jones, A.P., 2011. Green space and obesity: a systematic review of the evidence. Obes. Rev. 12, 183–189.
Larson, N., Story, M., Nelson, M., 2009. Neighborhood environments disparities in access to healthy foods in the USAm. J. Prev. Med. 36 (1), 74–81.
Lee, I., Ewing, R., Sesso, H.D., 2009. The built environment and physical activity levels: the Harvard alumni health study. Am. J. Prev. Med. 37 (4), 293–298.
Levi, J., Vinter, S., St Laurent, R., Segal, L. 2012. F as in Fat. Trust for America's Health. Available from: 〈http://healthyamericans.org/reports/obesity2010/〉. (accessed 15.02.13).
Lopez, R., 2004. Urban sprawl and risk for being overweight or obese. Am. J. Public Health 94 (9), 1574–1579.
McDonald, N., Trowbridge, M., 2009. Does the built environment affect when American teens become drivers? Evidence from the 2001 national household travel survey. J. Saf. Res. 40 (3), 177–183.
Nguyen, D., 2010. Evidence of the impacts of urban sprawl on social capital. Environ. Plan. B: Plan. Des. 37 (4), 610–627.
Plantinga, A.J., Bernell, S., 2007. The association between urban sprawl and obesity: is it a two way street? J. Reg. Sci. 47 (5), 857–879.
Ogden C.L., Carroll M.D., Kit B.K., Flegal K.M., 2012. National Center for Health Statistics. Prevalence of obesity in the United States, 2009–2010. NCHS data brief, no. 82.
Papas, M.A., Alberg, A.J., Ewing, R., Helzlsouer, K.J., Gary, T.L., Klassen, A., C., 2007. The built environment and obesity. Epidemiol. Rev. 29 (1), 129–143.
Rodriguez, D.A., Khatak, A.J., Evenson, K.R., 2006. Can new urbanism encourage physical activity? J. Am. Plan. Assoc. 72, 1.
Schweitzer, L., Zhou, J., 2010. Neighborhood air quality outcomes in compact and sprawled regions. J. Am. Plan. Assoc. 76 (3), 363–371.
Stone, B., 2008. Urban sprawl and air quality in large US cities. J. Environ. Manage. 86 (4), 688–698.
Stone, B., Hess, J.J., Frumkin, H., 2010. Urban form and extreme heat events: are sprawling cities more vulnerable to climate change than compact cities? Environ. Health Perspect. 118 (10), 1425.
Sturm, R., Cohen, D.A., 2004. Suburban sprawl and physical and mental health. Public health 118 (7), 488–496.
Transportation Research Board & Institute of Medicine Committee on Physical Activity, Health, Transportation, and Land Use, 2005. Does The Built Environment Influence Physical Activity? Examining the Evidence. National Academy of Sciences, Washington, DC. (Special Report 282).
Trogdon, J.G., Finkelstein, E.A., Hylands, T., Dellea, P.S., Kamal-Bahl, A.S.J., 2008. Indirect costs of obesity: a review of the current literature. Obes. Rev. 9, 489–500.
Trost, S., Kerr, L., Ward, D., Pate, R., 2001. Physical activity and determinants of physical activity in obese and non-obese children. Int. J. Obes. 25, 822–829.
Trost, S.G., Owen, N., Bauman, A.E., Sallis, J.F., Brown, W., 2002. Correlates of adults' participation in physical activity: review and update. Med. Sci. Sports. Exerc. 34 (12), 1996–2001.
Trowbridge, M.J., McDonald, N.C., 2008. Urban sprawl and miles driven daily by teenagers in the United States. Am. J. Prev. Med. 34 (3), 202–206.
Trowbridge, M.J., Gurka, M.J., O'connor, R.E., 2009. Urban sprawl and delayed ambulance arrival in the US. Am. J. Prev. Med. 37 (5), 428.
World Health Organization (WHO), 〈http://www.who.int/dietphysicalactivity/pa/en/index.html〉 (accessed October 2012).
Westerterp, K., 2009. Assessment of physical activity: a critical appraisal. Eur. J. Appl. Physiol. 105, 823–828.
Withrow, D., Alter, D.A., 2010. The economic burden of obesity worldwide: a systematic review of the direct costs of obesity. Obes. Rev. 12, 131–141.
Zolnik, E.J., 2011. The effect of sprawl on private-vehicle commuting outcomes. Environ. Plan. – Part A 43 (8), 1875.