



HHS Public Access

Author manuscript

Appl Geogr. Author manuscript; available in PMC 2021 December 17.

Published in final edited form as:

Appl Geogr. 2018 October ; 99: 12–21. doi:10.1016/j.apgeog.2018.07.016.

Spatial clustering patterns of child weight status in a southeastern US county

S. Morgan Hughey^{a,*}, Andrew T. Kaczynski^{b,c}, Dwayne E. Porter^d, James Hibbert^e, Gabrielle Turner-McGrievy^b, Jihong Liu^f

^aDepartment of Health and Human Performance, College of Charleston, USA

^bDepartment of Health Promotion, Education, and Behavior, Arnold School of Public Health, University of South Carolina, USA

^cPrevention Research Center, Arnold School of Public Health, University of South Carolina, USA

^dDepartment of Environmental Health Sciences, Arnold School of Public Health, University of South Carolina, USA

^eYork County Government, SC, USA

^fDepartment of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, USA

Abstract

Youth obesity is a major public health concern due to associated physical, social, and psychological health consequences. While rates and disparities of youth obesity levels are known, less research has explored spatial clustering patterns, associated correlates of spatial clustering, comparing patterns in urban and rural areas. Therefore, this study 1) examined spatial clustering of youth weight status, 2) investigated sociodemographic correlates of spatial clustering patterns, and 3) explored spatial patterns by level of urbanization. This study occurred in a southeastern US county (pop:474,266) in 2013. Trained physical education teachers collected height and weight for all 3rd-5th grade youth ($n = 13,469$) and schools provided youth demographic attributes. BMI z-scores were calculated using standard procedures. Global Moran's Index and Anselin's Local Moran's I (LISA) were used to detect global and local spatial clustering, respectively. To examine correlates of spatial clustering, BMI z-score residuals from a series of four linear regression models were spatially analyzed, mapped, and compared. SAS 9.4 and GeoDA were used for analyses; ArcGIS was used for mapping. Significant, positive global clustering (Index = 0.04, $p < 0.001$) was detected. LISA results showed that about 4.7% ($n = 635$) and 7.9% ($n = 1058$) of the sample were identified as high and low obesity localized spatial clusters ($p < 0.01$), respectively. Individual and neighborhood sociodemographic characteristics accounted for the majority of spatial clustering and differential patterns were observed by level of urbanization. Identifying

*Corresponding author. Department of Health and Human Performance, College of Charleston, Silcox Center, Room 336, 30 George Street, Charleston, SC, 29401, USA. hugheysm@cofc.edu (S.M. Hughey).

Declarations of interest

None.

geographic areas that contain significant spatial clusters is a powerful tool for understanding the location of and exploring contributing factors to youth obesity.

Keywords

Childhood; Overweight; Obesity; Spatial clustering; GIS; Urban and rural

1. Introduction

Childhood obesity has been recognized as a major public health problem of the 21st century due to the wide array of physical, social, and emotional health consequences that often accompany overweight and obesity in youth (Ogden, Carroll, Kit, & Flegal, 2014; Reilly et al., 2003; Strong et al., 2005). Studies have also documented that overweight and obese youth have a higher risk for increased weight status into adolescence and adulthood (Cunningham, Kramer, & Narayan, 2014; Wang & Beydoun, 2007). and persistent adult obesity is related to decreased quality of life, increased rates of chronic disease and healthcare costs, as well as increased morbidity and mortality (Sallis, Floyd, Rodríguez, & Saelens, 2012; Wang, McPherson, Marsh, Gortmaker, & Brown, 2011). Disparities in childhood obesity rates also exist in most developed countries; in the U.S. youth who are low-income, racial/ethnic minorities, and live in the southeastern U.S. have higher rates of overweight and obesity (Singh et al., 2007, 2008).

Researchers and practitioners have recognized complex causes of youth obesity, with many individual, interpersonal, community, environmental, and societal factors contributing to weight status (Han, Lawlor, & Kimm, 2010; Pereira, Nogueira, & Padez, 2018; Xu, Wen, & Wang, 2015). As a result, multidisciplinary theoretical models are frequently employed frameworks to understand the childhood obesity determinants and patterns at a population level (McLeroy, Bibeau, Steckler, & Glanz, 1988; Pereira et al., 2018; Sallis et al., 2006, 2012; Story, Kaphingst, Robinson-O'Brien, & Glanz, 2008; Xu et al., 2015). Integrating advanced spatial analytical tools and analyses with childhood obesity research is one example of applying a multidisciplinary approach to a widespread public health problem in order to advance this area of research.

Spatial variables and analyses are essential elements when exploring geographic patterns of health outcomes, which relies on computer-based geographic information systems (GIS) software and technology to visualize, measure, and conduct analyses (Auchincloss, Gebreab, Mair, & Roux, 2012; Casey et al., 2014; Jerrett, Gale, & Kontgis, 2010). Although broad public health literature has seen an increase in the use of GIS applications, obesity-related research could benefit from continued application of spatial tools and analyses when examining patterns and determinants (Auchincloss et al., 2012). Many studies documenting the prevalence of obesity distribution by geographic areas have aggregated data at administratively-defined units (e.g., census tracts, ZIP codes, cities, states) to analyze and describe the prevalence or rates of overweight or obesity (Auchincloss et al., 2012; Ford, Mokdad, Giles, Galuska, & Serdula, 2005; Koh, Grady, & Vojnovic, 2015; Singh et al., 2007, 2010; Wang & Beydoun, 2007); these methodologies have served as a

critical foundation for understanding obesity rates locally, nationally, and internationally. Utilizing global and local spatial clustering analyses to explore localized patterns that are not necessarily constrained by administrative units can offer unique insight into individual-level geographic obesity patterns (Huang, Moudon, Cook, & Drewnowski, 2015;Laraia et al., 2014; Penney, Rainham, Dummer, & Kirk, 2014; Poulou & Elliott, 2009).

To date, two of the primary spatial tools that have been most used in public health research are spatial proximity (i.e., measuring the distance between two points) and spatial aggregation methods. Spatial clustering is a useful spatial analysis that can conceptually, technically, and practically advance obesity research. First, spatial clustering analyses offers a tool to measure the nature and strength of geographical interdependence between data, which can conceptually show researchers where patterns of health outcomes may or may not be closely related (Auchincloss et al., 2012; Penney et al., 2014; Rushton, 2003). Indeed, other researchers have applied spatial clustering tools and analyses to understanding distributions of important public health problems, such as human anthrax (Barro et al., 2015) and cancer (Lin et al., 2015), yet obesity remains less explored with these powerful tools. The second reason to conduct these analyses is if significant spatial autocorrelation is present, the statistical assumption of independent observations for many additional statistical analyses may be violated (Rezaeian, Dunn, Leger, & Appleby, 2007). Consequently, assessing spatial autocorrelation is recommended as a first step in place-focused obesity research to minimize overstating significance between exposures and outcomes; this study demonstrates how this methodology can be applied in research (Rezaeian et al., 2007). Third, mapping patterns that are identified with clustering analyses can result in powerful visualizations, which may be used to pinpoint communities uniquely impacted by chronic disease outcomes, like obesity (Huang et al., 2015; Penney et al., 2014). These types of maps, combined with maps showing obesity prevalence statistics could be particularly impactful for practitioners in highlighting priority areas for intervention. Last, exploring social and economic determinants of particular spatial clustering patterns, not only the obesity outcome itself, is a critical step towards understanding how geographic patterns emerge.

To date, some studies have employed spatial clustering analyses to examine unique geographic patterns of obesity; however, there are some advancements that warrant further attention in this area of literature (Fritz, Schuurman, Robertson, & Lear, 2013). First, among studies that have examined obesity clustering, the vast majority have focused on adults (Curtis & Lee, 2010; Gartner, Taber, Hirsch, & Robinson, 2016; Huang et al., 2015; Laraia et al., 2014; Mobley, Finkelstein, Khavjou, & Will, 2004; Poulou & Elliott, 2009; Schuurman, Peters, & Oliver, 2009); to our knowledge, only a few studies have investigated spatial clustering of child or adolescent obesity (Hernández-Vásquez et al., 2016; Jin & Lu, 2017; Penney et al., 2014). Second, many studies that have explored spatial clustering of obesity conducted analyses with large-scale administrative units, such as census tracts, zip codes, or states (Curtis & Lee, 2010; Gartner et al., 2016; Hernández-Vásquez et al., 2016; Mobley et al., 2004; Poulou & Elliott, 2009; Schuurman et al., 2009). Conducting spatial clustering analyses at an individual level (i.e., point data) may provide added information on smaller scale, or localized, patterns in the study area that would not be identified (Huang et al., 2015; Laraia et al., 2014). Finally, some studies have also

examined whether demographic (e.g., socioeconomic status) and community-level factors (e.g., physical activity and nutrition environments) are related to the geographic patterning of obesity. Among those studies, economic indicators have emerged as some of the main explanatory variables of observed spatial patterns (Chalkias et al., 2013; Huang et al., 2015; Jin & Lu, 2017; Laraia et al., 2014). For example, Chalkias and colleagues found that education level was the most significant predictor of childhood obesity in Greece (Chalkias et al., 2013), whereas Chen and Troung found that socioeconomic disadvantage was only significantly related to obesity in specific geographic townships in Taiwan (Chen & Truong, 2012). Examining how identified spatial clusters change as economic and other demographic variables are included in spatial analytical models is essential to better understand location-specific patterns of childhood obesity.

In addition to these aforementioned gaps, few studies have explored patterns of obesity by varying levels of urbanization. Indeed, in large metropolitan areas may only contain urban areas and not warrant more nuanced analyses. However, many cities across the globe contain substantial diversity in levels of urbanization within the boundaries, including suburban and rural areas proximal to urban city centers. The contextual differences between urban, suburban, and rural may substantially influence youth health behaviors and weight status, and subsequently, result in varying types and degrees of spatial clustering of youth obesity (Hennessy et al., 2010; Liu et al., 2008, 2012). Furthermore, rural areas have been acknowledged as another focus of youth obesity disparities in the U.S. because children in these areas demonstrate higher rates of overweight and obesity (Lutfiyya, Lipsky, Wisdom-Behounek, & Inpanbutr-Martinkus, 2007). More spatially specific analyses are needed to compare patterns by urban, suburban, and rural areas.

To address these gaps in the literature and apply infrequently used spatial analyses to advance childhood obesity research, the objectives of this study were to 1) analyze spatial clustering patterns of childhood obesity in a southeastern US County, 2) examine whether sociodemographic characteristics were associated with spatial clustering patterns of youth obesity, and 3) explore differential spatial clustering patterns of obesity by levels of urbanization.

2. Methods

2.1. Study setting

This study occurred in 2013 in a large county in the southeastern United States, which had a total population of 474,266, of which 77.1% was Non-Hispanic White, 18.5% was African American, and 8.5% was Hispanic or Latino (United States Census Bureau, 2013a). In 2013, the median household income of the county was \$48,886 and approximately 15.0% of residents lived below the federal poverty line (United States Census Bureau, 2013a). The county encompassed approximately 750 square miles of land area.

2.2. Measures and data collection

2.2.1. Youth obesity and demographic characteristics—As part of regular protocol, trained physical education teachers measured and recorded the height and weight

for all children in 3rd through 5th grade ($n = 14,232$) enrolled in all 51 public elementary schools in the county school district. Height, weight, date of birth, and date of testing were used to calculate body mass index (BMI) percentiles and BMI z-scores using standardized protocols for youth from the Centers for Disease Control and Prevention (Center for Disease Control, 2014). Demographic information and address were also obtained for each individual. Three demographic variables were categorized for all youth. First, sex was classified as male or female. Second, individual, or child, level socioeconomic status (SES) was indicated by whether students were eligible for free or reduced lunch (low SES) or not (high SES). Finally, race/ethnicity was classified as African American, Hispanic, White, or Other. Data were acquired through a data sharing agreement between the research team, a local health-focused coalition, and the school district in the study setting. This study was approved by the Institutional Review Board from the University of South Carolina.

2.2.2. Block group characteristics—Several variables were collected for all census block groups (BGs, $n = 255$) in the study county. The following key BG variables were included in this study. First, racial/ethnic composition was measured by calculating the total percentage of racial and ethnic minorities (i.e., all persons other than those identifying as non-Hispanic White) (Hughey et al., 2016). Second, a SES index at the BG level was calculated by summing the standardized scores for the following area-level SES indicators: percent unemployed, percent of the population under 125% of the federal poverty threshold, percent less than high school education, and percent of renter occupied housing (Hughey et al., 2016; Kirby & Kaneda, 2005). The third BG variable was population density per square mile, which was calculated by dividing the total population of each BG by the land area (sq. miles) of the BG (Hall, Kaufman, & Ricketts, 2006). The final BG variable included in this study indicated the level of urbanization. The Census Bureau identifies two types of urban areas – urbanized areas (50,000 people or more) and urban clusters (at least 2500 and less than 50,000 people); rural areas are classified as those not defined as urban (Hall et al., 2006; United States Census Bureau, 2013b). Urbanized areas and urban clusters are represented with TIGER/Line Topological Faces (polygons with geocodes); this file was spatially overlaid with the BGs (the unit of analysis used to calculate all other area-level variables). If the BG contained only urban areas, it was classified as ‘urban’, whereas BGs with both urban and rural areas were classified as ‘mixed’; rural areas were defined as BGs that had no urban topological faces present (Hall et al., 2006; United States Census Bureau, 2013b). BG shapefiles with the aforementioned variables were joined to all individual data points using ArcMap 10.2.2 to assign these area-level characteristics to each participant.

2.3. Geospatial approach

2.3.1. Geocoding—Youth addresses ($n = 14,232$) were geocoded at the point address level in ArcGIS 10.2.2 using ESRI’s 2013 StreetMap data file, concurrent with youth obesity data collection. A total of 98.5% of the addresses were geocoded at either the street address ($n = 269$) or point address levels ($n = 13,835$), the two most accurate means of geocoding; 128 observations were removed from the dataset because they were geocoded at less precise levels (e.g., postal codes, municipality), which would negatively impact localized spatial analyses. Additional observations were removed due to the address residing

outside of the study county boundary ($n = 632$) and extreme BMI values ($n = 3$), resulting in a final sample of 13,469 youth.

2.3.2. Spatial clustering—Two of the overarching spatial analytic techniques used to assess spatial clustering patterns are global and local; global clustering measures overall patterns in a specified area without pinpointing the exact locations, whereas local clustering measures test for small-scale patterns across the study area (Auchincloss et al., 2012; Mitchell, 2005). Both levels of clustering have different substantive meaning and interpretations that were used in this study in distinct, yet complimentary ways.

One of the most robust and commonly-used approaches to detect global spatial clustering across multiple fields is Global Moran's Index (I), where the pattern of a set of features (i.e., study participants) and an associated attribute (i.e., youth weight status) are evaluated for clustering, dispersion (i.e., checkerboard), or random distribution (Laraia et al., 2014; Moran, 1948, 1950; Rushton, 2003). Global Moran's I was also chosen for this dataset because it is intended for data where high and low value clusters are assumed to exist, which has been supported in previous studies (Laraia et al., 2014).

Despite the ability to detect broad clustering in the study area, the Global Moran's I test does not indicate the physical location of clustering (Auchincloss et al., 2012; Mitchell, 2005). Therefore, a local cluster detection measure, Anselin's Local Moran's I (LISA), was used to assess more fine-grained patterns within the study County (Anselin, 1995; Ord & Getis, 2001). Often referred to as a hot-spot analysis, this clustering test provided an indication of the degree that localized areas represent unexpectedly high or low BMI z-score values compared to the overall, or global BMI z-score average across the sample (Anselin, 1995). Furthermore, this test can identify five categories of various spatial patterns that may be present in the data: Not Significant, High-High, High-Low, Low-High, and Low-Low (Pouliou & Elliott, 2009). This study was particularly interested in the High-High and Low-Low patterns of clustering that represent areas where youth with high and low BMI values, respectively, are surrounded by youth with similar values, indicating areas of geographic concentrations of high or low youth obesity. Results from the LISA analysis were mapped to show the location of identified clusters.

2.3.3. Geospatial analyses—For the Moran's I global spatial measure, the null hypothesis tested was that there is no spatial clustering across the study area and the analysis produces an overall estimate of clustering for the entire county (Laraia et al., 2014; Mitchell, 2005; Moran, 1950). The Global Moran's I statistic was calculated as follows, where I = Global Moran's I statistic for spatial autocorrelation in the study county, n = sample size ($n = 13,469$), i = individual observation and j = observations at other locations, $w_{i,j}$ is spatial weight matrix (distance threshold of one-half mile, or 805 m), x_i = individual BMI z-score value, \bar{x} = mean BMI z-score value, and S_0 = aggregate of all spatial weights (Moran, 1950):

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^n \sum_{j=1}^n w_{i,j} (x_i - \bar{x})(x_j - \bar{x})}{\sum_{j=1}^n (x_j - \bar{x})^2}$$

Similarly, the null hypothesis for the LISA measure was that all spatial patterns across the study were random. However, this analysis produces analytical output for each individual in the dataset. The LISA statistic for each individual was calculated as follows (Anselin, 1995), where I_i = Local Moran's I statistic for localized spatial autocorrelation, n = sample size ($n = 13,469$), i = individual observation and j = observations in another location, x_i = individual BMI z-score value, \bar{x} = mean BMI z-score value, σ^2 = variance of BMI z-score, $w_{i,j}$ is spatial weighting (distance threshold of one-half mile, or 805 m) (Anselin, 1995):

$$I_i = \frac{(x_i - \bar{x})}{\sigma^2} \sum_{j=1}^n w_{i,j}(x_j - \bar{x})$$

When conducting Global Moran's I and LISA (GeoDA 1.8.14), several specifications were selected, including distance options, number of permutations, and significance values. First, the distance selected to conduct the spatial clustering analyses has varied across studies and context, yet it is critical to the test and results (James et al., 2014). Both conceptual and empirical rationale determined the distance for analyses in this study. The Optimized Hot Spot Analysis tool in ArcGIS 10.2.2 was first employed to compare spatial autocorrelation values at a series of distances to determine the highest, or peak, spatial clustering value in the study area (ArcS for Desktop. Optim, 2016). The peak spatial clustering distance for the entire sample was 826.5 m, or 0.51 miles. A half-mile distance has been recognized as a critical threshold for measuring access to and use of health-promoting community features for youth (e.g., schools, parks, food outlets) (Carroll-Scott et al., 2013; Kligerman, Sallis, Ryan, Frank, & Nader, 2007; National Recreation and P, 2016; The Trust for Public Land, 2016). Given the close approximation of the peak spatial clustering distance and the empirical foundation, a half-mile, non-weighted distance band was used for Global Moran's I and LISA analyses.

In addition, the GeoDA software provides researchers several options for conducting permutations and examining clusters at various levels of significance. Permutations are a numerical approach that uses data-driven processes to determine statistical significance (Anselin, 2005). For these analyses, permutations determined how likely it would be to observe the Moran's I value of an actual distribution under conditions of spatial randomness. For each analysis, we used 99 permutations to examine the test statistic and a more conservative p-value of 0.01 (compared to 0.05) to adjust for multiple comparisons of testing all potential clusters.

The first research objective was the analyze spatial clustering patterns of childhood obesity. Using BMI z-score as the dependent variable (Must & Anderson, 2006; Himes, 2009) and the described specifications above, both Global and Local Moran's I analyses were conducted.

The second research objective aim was to examine how spatial clustering of weight status changed after including the described socio-demographic variables. In order to be able to conduct the Global and Local Moran's I analyses and compare patterns with different variables added, we first conducted a series of multivariate linear regression models. In total,

four linear regression models with BMI z-score as the dependent variable were estimated consecutively. The first model was an unconditional model with no covariates. Then, model 2 included all youth characteristics (i.e., age, gender, race/ethnicity, SES), while model 3 included all BG characteristics (i.e., percent racial/ethnic minority, SES index, population density, level of urbanization). All youth and BG characteristics were both included in model 4. In the final model, variance inflation factor (VIF) values indicated no evidence of multicollinearity bias.

For each of these regression models, we outputted and used the residuals as the dependent variables for the Global and Local Moran's I spatial analyses (second research objective). Researchers have used residuals as dependent variables when adjusting for covariates in spatial clustering analyses because the value (i.e., distance from predicted value) and direction (i.e., positive or negative) of the residual can be used to indicate high and low clusters (Cook, Li, Arterburn, & Tiwari, 2010; Huang et al., 2015; Jerrett et al., 2010; Laraia et al., 2014). In this study, high positive residuals indicated higher than expected BMI z-scores whereas large negative residuals indicated lower than expected BMI z-scores (Huang et al., 2015). With the same analytical parameters (e.g., distance, permutations, p-value), we conducted four Global and Local Moran's I analyses that correspond with the described regression models, with the residuals from those regression models used as the dependent variable. Statistical output for Global Moran's I as well as the number and location of localized clusters for the Local Moran's I were compared and mapped for each set of analyses, as the covariates were added to each model (maps described further in results section).

Lastly, to achieve the third research objective and explore the differential patterns of spatial clustering of youth obesity by level of urbanization, we first separated the geocoded file of youth addresses by three types of areas identified: urban ($n = 6788$), mixed urban-rural ($n = 6040$), and rural ($n = 641$) (United States Census Bureau, 2013b). Then, we conducted the same series of spatial analyses on these three sets of data points, with the exception of the urban/rural classification variable included in the analyses. However, different distances were used based on the results from the Optimized Hot Spot Analysis for each area type given the differences in average distances between participants: 500 m for participants in urban areas, 730 m for participants in mixed urban-rural areas, and 3186 m for participants in rural areas.

3. Results

3.1. Sample characteristics

All youth and block group sample characteristics are presented in Table 1. The majority of the sample was white (62%) while 45% were low SES (eligible for free or reduced lunch). The average BMI z-score was 0.5 ($SD = 1.1$), with 15.7% and 18.8% classified as overweight and obese, respectively. On average, youth were about 10 years old. Approximately half of youth lived in urban BGs, while 44.8% lived in urban-rural mixed BGs and 4.8% lived in rural BGs.

3.2. Global spatial clustering

As shown in Table 2, Global Moran's I results showed that there was a small, but significant, positive global spatial autocorrelation for BMI z-score across the study area (Index = 0.04, Z-value = 14.3, $p < 0.001$), indicating a general tendency for BMI z-score values to be located near other similar BMI z-score values (Model 1). The subsequent models that included different sets of covariates resulted in changes in the Index values, z-values, and significance values. In Model 2, which adjusted for youth characteristics, the Index value was substantially reduced but still statistically significant (Index = 0.007, Z-value = 2.6, $p = 0.005$). Model 3 included all BG characteristics and no youth variables; the results showed a slight decrease in the Index value from the unconditional model (Model 1) but adjusting for these variables still resulted in significant, positive global spatial autocorrelation (Index = 0.014, Z-value = 5.1, $p < 0.001$). In the final model that included both youth and BG variables, global spatial clustering was no longer significant (Index = 0.003, Z-value = 0.9, $p < 0.169$).

3.3. Local spatial clustering

Results from the Local Anselin Moran's I tests showed significant local clustering patterns; the number of high-high and low-low cluster observations for Models 1 through 4 are presented in Table 2. High-high cluster observations represent youth that have elevated BMI z-scores compared to the overall population and are also surrounded by other youth that have similarly high BMI z-scores and vice-versa for low-low clusters. In Model 1 with no covariates, there were a total of 635 high-high and 1058 low-low statistically significant spatial cluster observations, representing 4.7% and 7.9% of the sample, respectively.

Similar patterns were observed in the LISA results as Global Moran's I throughout the model building process. Specifically, after adjusting for youth characteristics (Model 2), the number of high-high and low-low cluster observations were reduced by over half, with only 1.9% and 2.5% of participants now located in high-high and low-low statistically significant local clusters, respectively. In Model 3 that adjusted for BG characteristics, fewer significant clustered observations were noted compared to Model 1 but more than were identified in Model 2 (Table 2). Finally, Model 4 showed the lowest proportion of local spatial cluster observations.

A series of maps were developed to visually represent the changes in the local clustering patterns throughout the model building process. Maps representing the concentrations of each type of spatial cluster were developed using the point density function (Laraia et al., 2014). Each statistically significant point was smoothed a half-mile, concurrent with the distance used to conduct the clustering analyses. One map was created for each of the four models with shades of red areas showing the concentration of high-high clusters while shades of blue representing the concentration of low-low obesity clusters (Laraia et al., 2014). As shown in Fig. 1, Model 1, the western areas of the study county had substantial areas of high-high obesity clusters, whereas the eastern of the county showed high concentrations of low-low obesity clusters. Interestingly, several high-high and low-low clusters are observed adjacent to one another, particularly near the middle of the map,

corresponding with areas surrounding the city center. The images for Models 2, 3, and 4 show the changes in the concentration of local clusters as correlate variables were examined.

3.4. Spatial clustering by level of urbanization

Using the same model building process described for the overall sample, global and local spatial clustering was also explored by level of urbanization by analyzing the same set of statistical models separately for youth living in urban, urban-rural mixed, and rural areas (Table 3). For Global Moran's I, significant global clustering was observed in the urban (Index = 0.030, Z-score = 6.6, $p < 0.01$), urban-rural mixed (Index = 0.027, Z-score = 5.4, $p < 0.01$), and rural areas (Model 1; Index = 0.029, Z-score = 2.7, $p = 0.001$). Differential patterns were observed based on level of urbanization as correlate variables were included. For youth that lived in urban areas, global spatial clustering was attenuated in Models 2 and 4 after including individual socio-demographic characteristics and both individual and neighborhood socio-demographic characteristics, respectively (Table 3). However, in urban-rural mixed areas, global clustering was only attenuated by including the combination of individual and BG characteristics (Model 4), whereas significant global clustering was present in all models for youth that lived in rural areas. The local clustering results showed similar patterns across all four models as the global patterns described (lower half of Table 3). However, in Model 1, urban areas had a higher proportion of high-high cluster observations (5.6%) compared to urban-rural mixed (2.9%) and rural (2.3%) areas.

4. Discussion

This study applied spatial clustering analyses for youth weight status at the individual data point level (Hernández-Vásquez et al., 2016; Huang et al., 2015; Laraia et al., 2014), and explored spatial clustering by levels of urbanization. Further, these spatial analyses contributed to literature on obesity by examining geographic-specific patterns and determinants of observed spatial distribution, differing from previous research investigating determinants of the obesity outcome itself (Auchincloss et al., 2012; Huang et al., 2015; Laraia et al., 2014). Results showed statistically significant global clustering across the study area and local spatial clustering in specific regions of the county. Global clustering was attenuated and the number of individual local clusters was greatly reduced after adjusting for both youth and neighborhood socio-demographic characteristics, though variations were discovered by level of urbanization.

Significant, positive spatial global autocorrelation was found, indicating that BMI z-score values were not randomly distributed within the study boundaries and that high and low values were more proximal to other high and low values, respectively (Anselin, 1995). The global spatial clustering results showed a low to moderate level of global autocorrelation (Laraia et al., 2014). As discussed in the methods, Moran's I values can range from -1 (completely dispersed) to $+1$ (completely clustered), and the Global Moran's I value found in this study was 0.04, which was both lower and higher than previously reported values. For example, Laraia and colleagues found a Global Moran's I value of 0.05 (Laraia et al., 2014), while Penney and colleagues found a Global Moran's I value of -0.017 (Penney et al., 2014). However, caution should be used when interpreting and comparing Global

Moran's I statistics within and across studies. As described in the methods, this spatial analysis is sensitive to the specifications chosen by the researcher, particularly the spatial weights (Anselin, 2005; Anselin, Syabri, & Kho, 2010). Unlike a standardized Pearson correlation, spatial weights often differ within and across studies. Importantly, the Moran's I value in this study was statistically significant with a conservative p-value, indicating researchers can reject the null hypothesis of random spatial patterning. The geographical interdependence of these data suggest that one or more underlying characteristics could be contributing to similar patterns in the same areas. Other researchers have also found positive, global spatial clustering patterns (Laraia et al., 2014), while some studies have reported no global autocorrelation of obesity (Penney et al., 2014; Schuurman et al., 2009). These observed inconsistencies may be attributable to varied unit of analyses (e.g., individual points, census tracts) used in spatial clustering analyses for obesity and contextual differences between study cities, resulting in varying spatial dependency in health outcomes, like obesity (Michimi & Wimberly, 2010).

In addition to global autocorrelation, localized spatial patterns of obesity were detected using Local Anselin Moran's I (LISA) (Ord & Getis, 2001). Overall, about 13% of the sample was located in either high or low local spatial clusters. A large concentration of low weight status clusters was found in eastern areas of the county, whereas high weight status clusters were more prominent in the western region. Researchers using the same Local Moran's I analysis for BMI among adults in Seattle, WA reported similar low and high patterns in distinct regions (i.e., northern and southern) of their study area (Huang et al., 2015). Such gradients highlight divergent spatial patterns of health indicators within the same cities, showing distinct geographic differences across neighborhoods.

Knowing the overall prevalence of youth obesity for a particular neighborhood or city can certainly highlight areas where a high proportion of children are affected and subsequently, be used by public health researchers, practitioners, and community advocates to intervene to address the health problem. Spatial clustering analyses and results can complement these statistics by exploring a more nuanced role of place or geography; that is, do children in the same neighborhood tend to have more similar weight status and what might be driving such patterns? Significant clustering of obesity could also be indicative of underlying social factors that create an environment where the determinants and behaviors contributing to youth weight status are more concentrated. Combining overall prevalence and spatial clustering results may highlight places where obesity prevention strategies and intervention are needed most (Huang et al., 2015; Kirby, Delmelle, & Eberth, 2016). Furthermore, future research can take these analyses a step further by critically examining similarities and differences between individuals or communities that are identified in opposite spatial clusters (i.e., high or low). For example, important multilevel factors, such as local policies, community resources, and built environment features, could be compared for individuals that are in high-high spatial clusters compared to those that are in low-low spatial clusters. A plethora of research has examined how various policies and environments are related to obesity, yet less work has examined whether these factors contribute to specific concentrations of spatial patterns among individuals and communities. Better understanding the processes and determinants that have led to specific spatial patterns of obesity will be critical for determining the most effective strategies to combat childhood obesity.

The second main analysis of this study examined the correlates of spatial clustering of youth obesity by including numerous demographic and economic indicators in subsequent analytical models. Overall, individual-level and neighborhood-level variables accounted for a portion of global clustering, separately, but the combination of individual and neighborhood variables attenuated global autocorrelation and substantially reduced the concentration of significant local cluster points. Similarly, one study in northern California showed that combined individual and neighborhood-level characteristics accounted for the majority of global and local spatial clustering of adult obesity (Laraia et al., 2014), while the previously described study in Seattle reported that property values, an area-level SES indicator, was the primary variable accounting for the local spatial clustering of adult obesity (Huang et al., 2015).

The combination of individual and block group characteristics explaining the observed global and local spatial clustering closely resonates with the multifaceted social ecological model of health (Bronfenbrenner, 1997; McLeroy et al., 1988). This theoretical framework posits that health outcomes are impacted by multiple levels of influence (e.g., intrapersonal, interpersonal, institutional, community, broad policy). This study highlights how both individual and neighborhood-level socio-demographic factors contribute to manifestations of varying spatial patterns of youth weight status at a local level. While this study shows important correlates of spatial clustering, the socio-demographic variables included in the analytical models (e.g., BG SES index, percentage of racial/ethnic minority residents) are likely reflective of complex processes that contribute to these spatial patterns (Schulz & Northridge, 2004). For example, the BG SES index was comprised of multiple elements of SES, including education, employment, housing, and poverty (Kirby & Kaneda, 2005). These economic indicators are reciprocally related to important social and public policies and conditions (e.g., access to quality educational opportunities) and health-promoting environments (e.g., access to quality, nutritious foods) (Schulz & Northridge, 2004). Future research should seek innovative ways to incorporate more contextual variables in studies that intersect health and geography.

Importantly, this study found nuanced clustering patterns and correlates of spatial clustering based on the level of urbanization where youth lived. Previous research has detected high and low significant local spatial clusters of obesity in areas characterized by varying levels of urbanization (Huang et al., 2015; Laraia et al., 2014). For example, two studies found that low obesity clustering was found in more urban areas for adults, whereas high obesity clustering was more prominent in less densely populated areas (Huang et al., 2015; Laraia et al., 2014). In this study, the overall proportion of overweight and obese youth was highest among youth living in rural areas (42.5%) compared to urban (33.4%) and urban-rural mixed (35.2%) areas. However, the spatial clustering analyses revealed specific spatial patterns that somewhat differed from the overall prevalence, further highlighting the differences between a non-spatial statistical model and spatial model. Local spatial clustering results showed that a higher proportion of the significant high-high clusters were identified in urban areas whereas a higher proportion of low-low clusters were in the urban-rural mixed areas. This indicates that while the overall proportion of overweight and obesity was highest among children living in rural areas, a higher concentration of spatially-linked clusters, where children tended to have similar weight status, were located in urban and urban-rural mixed

areas. Additional research should investigate whether these patterns are similar across cities and how specific conditions of urban and suburban areas may generate these clustering patterns.

For example, while this phenomenon was not observed in all urban and suburban areas of the County, this pattern may partially be explained by urban sprawl (Frumkin, 2002). Urban sprawl reflects patterns of expansion outside of centralized urban areas, often marked by more affluent residents shifting residence, leaving high concentrations of low-income residents in urban areas (Frumkin, 2002). Future work could explore whether local spatial patterns of high and low obesity are found in historically disadvantaged or advantaged areas, respectively, and whether these patterns persist over time. Several studies have documented differences in the prevalence of youth overweight and obesity by urban and rural areas, but to our knowledge, this is one of the first studies to conduct localized spatial analyses by level of urbanization.

4.1. Limitations

Several study limitations should be noted. First, this study is focused on one county in the United States, which limits the generalizability of the findings (Shadish, Cook, & Campbell, 2002). However, to our knowledge, this is the first study to examine point-level spatial clustering of child weight status in the southeastern US area, a region with notably high rates of obesity (Singh et al., 2007). Furthermore, this study demonstrates the utility of spatial clustering analyses and results for one county, which could be scaled to larger regions. At the same time, many spatial analyses conducted within a defined space illustrates the boundary, or border effects, problem/issue (Griffith, 1983). Often, administratively-defined units, including city or county boundaries, are used to define the area and population of interest. Indeed, the data sharing agreement was focused in on one specific county; however, we acknowledge the limitation the spatial relationships could have continued beyond the boundaries. Decisions regarding this boundary issue are likely specific to the scope of the research and research objectives. For example, the data sharing agreement played a large role in defining the area and data were not removed to maintain the entire population and sample size. However, for other situations, removing data points could be viable solutions to address the boundary issue (Griffith, 1983). Similar to the discussion of geographic scope, though this study included a large sample of elementary-aged youth (all 3rd-5th grade students in public schools), this limited the scope of ages analyzed. Comparing spatial patterns for populations across the lifespan could help researchers and practitioners understand what geographic factors influence health outcomes such that interventions, including policy and environment changes, can best meet the needs of a diversity of populations.

Additionally, this study was cross-sectional and no causality can be attributed to the findings (Shadish et al., 2002). This has been a limitation for many spatial clustering studies focused on health, particularly as this field is rapidly growing (Auchincloss et al., 2012). However, there is a need for longitudinal studies analyzing obesity patterns in terms of space and time to better understand whether spatial patterns of obesity persist over time and from a community standpoint, whether clustering patterns of obesity remain in the same areas over

time (Kirby et al., 2016). Such information could indicate priority areas for community health needs assessments, in order to understand localized factors that contribute and detract from health. For example, with multiple years of data, we could observe whether the high-high and low-low spatial clusters persists, are exacerbated, or dissipate as the same children age. The capabilities of spatial software programs are advancing in ways that can handle innovative and longitudinal spatiotemporal analyses that can provide important insight into obesity patterns over time (Anselin et al., 2010; Kirby et al., 2016). Finally, obesity is a complex health condition that is influenced by many factors, many of which were not measured in this specific population. For example, data on behavioral patterns like nutrition and physical activity were not available for this large sample, but likely play a substantial role in impacting obesity and the spatial patterns observed.

4.2. Conclusions

In summary, the results of this study showed global and local spatial patterning of youth obesity in a southeastern U.S. county, which reinforces the importance of spatial relationships among health conditions, including obesity. Individual-level socio-demographic characteristics were identified as a primary correlate of the spatial patterns identified, though more work is needed to explicate the mechanisms driving these associations. Overall, identifying geographic areas that contain significant spatial clusters is a powerful tool for understanding the location of and exploring contributing factors to childhood obesity.

Acknowledgments

Funding

We would like to acknowledge the Piedmont Health Foundation, University of South Carolina Office for the Vice President of Research (SPARC), and University of South Carolina Provost for their funding support of this project.

References

- Anselin L. (1995). Local indicators of spatial Association—LISA. *Geographical Analysis*, 27(2), 93–115. 10.1111/j.1538-4632.1995.tb00338.x.
- Anselin L. (2005). *Exploring spatial data with GeoDa: A workbook*. Urbana-Champaign: University of Illinois.
- Anselin L, Syabri I, & Kho Y. (2010). *GeoDa: An introduction to spatial data analysis*. In M. M. Fischer, & A. Getis (Eds.). *Handbook of applied spatial analysis* (pp. 73–89). Springer Berlin Heidelberg. 10.1007/978-3-642-03647-7_5.
- ArcGIS for Desktop. Optimized Hot Spot Analysis. <http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-statistics-toolbox/optimized-hot-spot-analysis.htm>. Accessed May 28, 2016.
- Auchincloss AH, Gebreab SY, Mair C, & Roux AVD (2012). A review of spatial methods in epidemiology, 2000–2010. *Annual Review of Public Health*, 33, 107–122. 10.1146/annurev-publhealth-031811-124655.
- Barro AS, Kracalik IT, Malania L, Tsertsvadze N, Manvelyan J, Imnadze P, et al. (2015). Identifying hotspots of human anthrax transmission using three local clustering techniques. *Applied Geography*, 60, 29–36. 10.1016/j.apgeog.2015.02.014.
- Bronfenbrenner U. (1997). Ecological models of human development. *Read Dev Child*, 5, Accessed date: 16 September 2014.
- Carroll-Scott A, Gilstad-Hayden K, Rosenthal L, Peters L, McCaslin SM, Joyce R, et al. (2013). Disentangling neighborhood contextual associations with child body mass index, diet, and physical

- activity: The role of built, socioeconomic, and social environments. *Social Science & Medicine*, 95, 106–114. 10.1016/j.socscimed.2013.04.003. [PubMed: 23642646]
- Casey R, Oppert J-M, Weber C, Charreire H, Salze P, Badariotti D, et al. (2014). Determinants of childhood obesity: What can we learn from built environment studies? *Food Quality and Preference*, 31, 164–172. 10.1016/j.foodqual.2011.06.003.
- Center for Disease Control (2014). Healthy weight: About BMI for children and teens. Published http://www.cdc.gov/healthyweight/assessing/bmi/childrens_bmi/about_childrens_bmi.html, Accessed date: 18 September 2014.
- Chalkias C, Papadopoulos AG, Kalogeropoulos K, Tambalis K, Psarra G, & Sidossis L. (2013). Geographical heterogeneity of the relationship between childhood obesity and socio-environmental status: Empirical evidence from Athens, Greece. *Applied Geography*, 37, 34–43. 10.1016/j.apgeog.2012.10.007.
- Chen D-R, & Truong K. (2012). Using multilevel modeling and geographically weighted regression to identify spatial variations in the relationship between place-level disadvantages and obesity in Taiwan. *Applied Geography*, 32(2), 737–745. 10.1016/j.apgeog.2011.07.018.
- Cook AJ, Li Y, Arterburn D, & Tiwari RC (2010). Spatial cluster detection for weighted outcomes using cumulative geographic residuals. *Biometrics*, 66(3), 783–792. 10.1111/j.1541-0420.2009.01323.x. [PubMed: 19751250]
- Cunningham SA, Kramer MR, & Narayan KMV (2014). Incidence of childhood obesity in the United States. *New England Journal of Medicine*, 370(5), 403–411. 10.1056/NEJMoa1309753.
- Curtis AJ, & Lee W-AA (2010). Spatial patterns of diabetes related health problems for vulnerable populations in Los Angeles. *International Journal of Health Geographics*, 9, 43. 10.1186/1476-072X-9-43. [PubMed: 20796322]
- Ford ES, Mokdad AH, Giles WH, Galuska DA, & Serdula MK (2005). Geographic variation in the prevalence of obesity, diabetes, and obesity-related behaviors. *Obesity Research*, 13(1), 118–122. 10.1038/oby.2005.15. [PubMed: 15761170]
- Fritz CE, Schuurman N, Robertson C, & Lear S. (2013). A scoping review of spatial cluster analysis techniques for point-event data. *Geospatial Health*, 7(2), 183–198. 10.4081/gh.2013.79. [PubMed: 23733283]
- Frumkin H. (2002). Urban sprawl and public health. *Public Health Reports*, 117, 201–217. [PubMed: 12432132]
- Gartner DR, Taber DR, Hirsch JA, & Robinson WR (2016). The spatial distribution of gender differences in obesity prevalence differs from overall obesity prevalence among US adults. *Annals of Epidemiology*, 26(4), 293–298. 10.1016/j.annepidem.2016.02.010. [PubMed: 27039046]
- Griffith DA (1983). The boundary value problem in spatial statistical analysis. *Journal of Regional Science*, 23(3), 377–387. [PubMed: 12279616]
- Hall SA, Kaufman JS, & Ricketts TC (2006). Defining urban and rural areas in U.S. epidemiologic studies. *Journal of Urban Health*, 83(2), 162–175. 10.1007/s11524-005-9016-3. [PubMed: 16736366]
- Han JC, Lawlor DA, & Kimm SY (2010). Childhood obesity. *The Lancet*, 375(9727), 1737–1748. 10.1016/S0140-6736(10)60171-7.
- Hennessy E, Kraak VI, Hyatt RR, Bloom J, Fenton M, Wagoner C, et al. (2010). Active living for rural children: Community perspectives using PhotoVOICE. *American Journal of Preventive Medicine*, 39(6), 537–545. 10.1016/j.amepre.2010.09.013. [PubMed: 21084074]
- Hernández-Vásquez A, Bendezú-Quispe G, Díaz-Seijas D, Santero M, Minckas N, Azañedo D, et al. (2016). Spatial analysis of childhood obesity and overweight in Peru, 2014. *Revista Peruana de Medicina Experimental y Salud Pública*, 33(3), 489–497. 10.17843/rpmesp.2016.333.2298. [PubMed: 27831612]
- Himes JH (2009). Challenges of accurately measuring and using BMI and other indicators of obesity in children. *Pediatrics*, 124(Supplement 1), S3–S22. 10.1542/peds.2008-3586D. [PubMed: 19720665]
- Huang R, Moudon AV, Cook AJ, & Drewnowski A. (2015). The spatial clustering of obesity: Does the built environment matter? *Journal of Human Nutrition and Dietetics*, 28(6), 604–612. 10.1111/jhn.12279. [PubMed: 25280252]

- Hughey SM, Walsemann KM, Child S, Powers A, Reed JA, & Kaczynski AT (2016). Using an environmental justice approach to examine the relationships between park availability and quality indicators, neighborhood disadvantage, and racial/ethnic composition. *Landscape and Urban Planning*, 148, 159–169. 10.1016/j.landurbplan.2015.12.016.
- James P, Berrigan D, Hart JE, Hipp JA, Hoehner CM, Kerr J, et al. (2014). Effects of buffer size and shape on associations between the built environment and energy balance. *Health & Place*, 27, 162–170. 10.1016/j.healthplace.2014.02.003. [PubMed: 24607875]
- Jerrett M, Gale S, & Kontgis C. (2010). Spatial modeling in environmental and public health research. *International Journal of Environmental Research and Public Health*, 7(4), 1302–1329. 10.3390/ijerph7041302. [PubMed: 20617032]
- Jin H, & Lu Y. (2017). The relationship between obesity and socioeconomic status among Texas school children and its spatial variation. *Applied Geography*, 79, 143–152. 10.1016/j.apgeog.2016.12.008.
- Kirby RS, Delmelle E, & Eberth JM (2016). Advances in spatial epidemiology and geographic information systems. *Annals of Epidemiology*, 0(0), 10.1016/j.annepidem.2016.12.001.
- Kirby JB, & Kaneda T. (2005). Neighborhood socioeconomic disadvantage and access to health care. *Journal of Health and Social Behavior*, 46(1), 15–31. [PubMed: 15869118]
- Kligerman M, Sallis JF, Ryan S, Frank LD, & Nader PR (2007). Association of neighborhood design and recreation environment variables with physical activity and body mass index in adolescents. *American Journal of Health Promotion*, 21(4), 274–277. [PubMed: 17375494]
- Koh K, Grady SC, & Vojnovic I. (2015). Using simulated data to investigate the spatial patterns of obesity prevalence at the census tract level in metropolitan Detroit. *Applied Geography*, 62, 19–28. 10.1016/j.apgeog.2015.03.016.
- Laraia BA, Blanchard SD, Karter AJ, Jones-Smith JC, Warton M, Kersten E, et al. (2014). Spatial pattern of body mass index among adults in the diabetes study of Northern California (DISTANCE). *International Journal of Health Geographics*, 13(1), 48. 10.1186/1476-072X-13-48. [PubMed: 25471753]
- Lin H, Ning B, Li J, Ho SC, Huss A, Vermeulen R, et al. (2015). Lung cancer mortality among women in Xuan Wei, China: A comparison of spatial clustering detection methods. *Asia Pacific Journal of Public Health*, 27(2), NP392–NP401. 10.1177/1010539512444778.
- Liu J, Bennett KJ, Harun N, & Probst JC (2008). Urban-rural differences in overweight status and physical inactivity among US children aged 10–17 years. *The Journal of Rural Health*, 24(4), 407–415. 10.1111/j.1748-0361.2008.00188.x. [PubMed: 19007396]
- Liu J-H, Jones SJ, Sun H, Probst JC, Merchant AT, & Cavicchia P. (2012). Diet, physical activity, and sedentary behaviors as risk factors for childhood obesity: An urban and rural comparison. *Childhood Obesity*, 8(5), 440–448. 10.1089/chi.2012.0090. [PubMed: 23061499]
- Lutfiyya MN, Lipsky MS, Wisdom-Behounek J, & Inpanbutr-Martinkus M. (2007). Is rural residency a risk factor for overweight and obesity for U.S. children? *Obes Silver Spring Md*, 15(9), 2348–2356. 10.1038/oby.2007.278.
- McLeroy KR, Bibeau D, Steckler A, & Glanz K. (1988). An ecological perspective on health promotion programs. *Health Education & Behavior*, 15(4), 351–377.
- Michimi A, & Wimberly MC (2010). Spatial patterns of obesity and associated risk factors in the conterminous U.S. *American Journal of Preventive Medicine*, 39(2), e1–e12. 10.1016/j.amepre.2010.04.008. [PubMed: 20621256]
- Mitchell A. (2005). *The ESRI guide to GIS Analysis: Spatial measurements and statistics*, Vol. 2. ESRI Press.
- Mobley LR, Finkelstein EA, Khavjou OA, & Will JC (2004). Spatial analysis of body mass index and smoking behavior among WISEWOMAN participants. *Journal of women's health*, 13(5), 519–528.
- Moran PAP (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society*, 10(2), 243–251.
- Moran PAP (1950). Notes on continuous stochastic phenomena. *Biometrika*, 37(1–2), 17–23. [PubMed: 15420245]

- Must A, & Anderson SE (2006). Body mass index in children and adolescents: Considerations for population-based applications. *International Journal of Obesity*, 30(4), 590–594. 10.1038/sj.ijo.0803300. [PubMed: 16570087]
- National Recreation and Park Association. Safe Routes to Parks, Equitable Park Access. <http://www.nrpa.org/Safe-Routes-To-Parks/>. Accessed November 29, 2016.
- Ogden CL, Carroll MD, Kit BK, & Flegal KM (2014). Prevalence of childhood and adult obesity in the United States, 2011–2012. *Journal of the American Medical Association*, 311(8), 806–814. 10.1001/jama.2014.732. [PubMed: 24570244]
- Ord JK, & Getis A. (2001). Testing for local spatial autocorrelation in the presence of global autocorrelation. *Journal of Regional Science*, 41(3), 411–432. 10.1111/0022-4146.00224.
- Penney TL, Rainham DGC, Dummer TJB, & Kirk SFL (2014). A spatial analysis of community level overweight and obesity. *Journal of Human Nutrition and Dietetics*, 27(Suppl 2), 65–74. 10.1111/jhn.12055. [PubMed: 23701323]
- Pereira M, Nogueira H, & Padez C. (2018). Association between childhood obesity and environmental characteristics: Testing a multidimensional environment index using census data. *Applied Geography*, 92, 104–111. 10.1016/j.apgeog.2018.01.005.
- Pouliou T, & Elliott SJ (2009). An exploratory spatial analysis of overweight and obesity in Canada. *Preventive Medicine*, 48(4), 362–367. 10.1016/j.ypmed.2009.01.017. [PubMed: 19463485]
- Reilly JJ, Methven E, McDowell ZC, Hacking B, Alexander D, Stewart L, et al. (2003). Health consequences of obesity. *Archives of Disease in Childhood*, 88(9), 748–752. [PubMed: 12937090]
- Rezaeian M, Dunn G, Leger SS, & Appleby L. (2007). Geographical epidemiology, spatial analysis and geographical information systems: A multidisciplinary glossary. *Journal of Epidemiology & Community Health*, 61(2), 98–102. 10.1136/jech.2005.043117. [PubMed: 17234866]
- Rushton G. (2003). Public health, GIS, and spatial analytic tools. *Annual Review of Public Health*, 24(1), 43–56. 10.1146/annurev.publhealth.24.012902.140843.
- Sallis JF, Cervero RB, Ascher W, Henderson KA, Kraft MK, & Kerr J. (2006). An ecological approach to creating active living communities. *Annual Review of Public Health*, 27(1), 297–322. 10.1146/annurev.publhealth.27.021405.102100.
- Sallis JF, Floyd MF, Rodríguez DA, & Saelens BE (2012). Role of built environments in physical activity, obesity, and cardiovascular disease. *Circulation*, 125(5), 729–737. 10.1161/CIRCULATIONAHA.110.969022. [PubMed: 22311885]
- Schulz A, & Northridge ME (2004). Social determinants of health: Implications for environmental health promotion. *Health Education & Behavior*, 31(4), 455–471. 10.1177/1090198104265598. [PubMed: 15296629]
- Schuurman N, Peters PA, & Oliver LN (2009). Are obesity and physical activity clustered? A spatial analysis linked to residential density. *Obesity*, 17(12), 2202–2209. 10.1038/oby.2009.119. [PubMed: 19390521]
- Shadish WR, Cook TD, & Campbell DT (2002). Experimental and quasi-experimental designs for generalized causal inference. Wadsworth Cengage learning <http://impact.cgiar.org/pdf/147.pdf>, Accessed date: 18 February 2016.
- Singh GK, Kogan MD, & Dyck PC van(2007). A multilevel analysis of state and regional disparities in childhood and adolescent obesity in the United States. *Journal of Community Health*, 33(2), 90–102. 10.1007/s10900-007-9071-7.
- Singh GK, Kogan MD, & Dyck PC van (2010). Changes in state-specific childhood obesity and overweight prevalence in the United States from 2003 to 2007. *Archives of Pediatrics and Adolescent Medicine*, 164(7), 598–607. 10.1001/archpediatrics.2010.84. [PubMed: 20603458]
- Singh GK, Kogan MD, Van Dyck PC, & Siahpush M. (2008). Racial/Ethnic, socioeconomic, and behavioral determinants of childhood and adolescent obesity in the United States: Analyzing independent and joint associations. *Annals of Epidemiology*, 18(9), 682–695. 10.1016/j.annepidem.2008.05.001. [PubMed: 18794009]
- Story M, Kaphingst KM, Robinson-O'Brien R, & Glanz K. (2008). Creating healthy food and eating environments: Policy and environmental approaches. *Annual Review of Public Health*, 29, 253–272.

- Strong WB, Malina RM, Blimkie CJR, Daniels SR, Dishman RK, Gutin B, et al. (2005). Evidence based physical activity for school-age youth. *The Journal of Pediatrics*, 146(6), 732–737. 10.1016/j.jpeds.2005.01.055. [PubMed: 15973308]
- The Trust for Public Land. 10-minute walk. <https://www.tpl.org/slides/10-minute-walk>. Accessed November 29, 2016.
- United States Census Bureau. State and county quick facts. Accessed June 2, 2013.
- United States Census Bureau. Urban and Rural Classification. <https://www.census.gov/geo/reference/ua/uafaq.html>. Accessed December 5, 2013.
- Wang Y, & Beydoun MA (2007). The obesity epidemic in the United States—gender, age, socioeconomic, racial/ethnic, and geographic characteristics: A systematic review and meta-regression analysis. *Epidemiologic Reviews*, 29, 6–28. 10.1093/epirev/mxm007. [PubMed: 17510091]
- Wang YC, McPherson K, Marsh T, Gortmaker SL, & Brown M. (2011). Health and economic burden of the projected obesity trends in the USA and the UK. *Lancet Lond Engl*, 378(9793), 815–825. 10.1016/S0140-6736(11)60814-3.
- Xu Y, Wen M, & Wang F. (2015). Multilevel built environment features and individual odds of overweight and obesity in Utah. *Applied Geography*, 60, 197–203. 10.1016/j.apgeog.2014.10.006. [PubMed: 26251559]

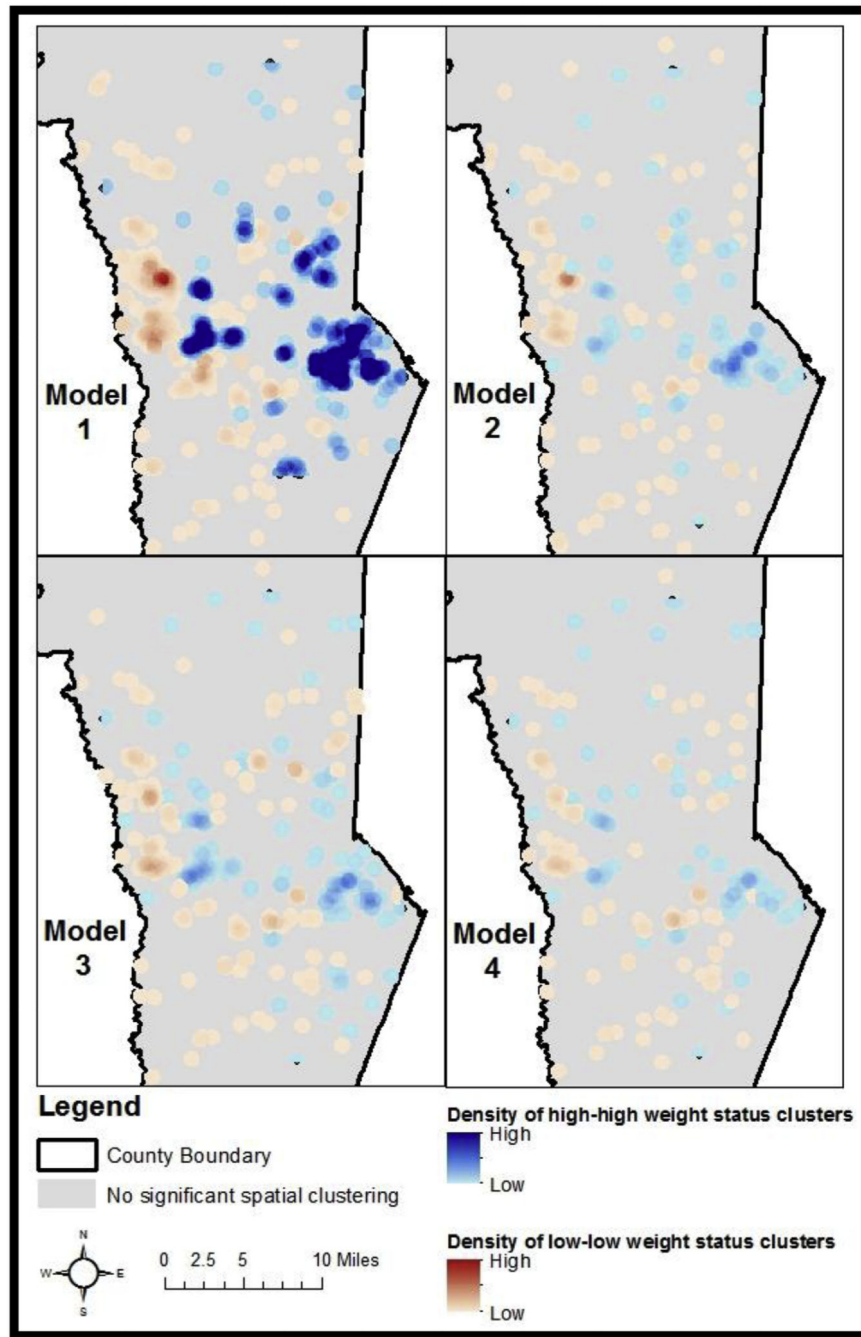


Fig. 1. Density of Significant High-high and Low-low Local Spatial Clusters of Youth Weight Status in a Southeastern US County. Notes: Model 1 includes no covariates, Model 2 adjusts for youth socio-demographics, Model 3 adjusts for block group socio-demographics, and Model 4 adjusts for both youth and block group characteristics.

Table 1

Sample characteristics.

Youth Characteristics (n = 13,469)	Mean or %	SD	Range
Body Mass Index (BMI) z-score	0.5	1.1	(-8.1, 3.0)
BMI percentile	64.0	1.0	(0, 99.9)
BMI categories			
Underweight	3.1%		
Normal weight	62.3%		
Overweight	15.7%		
Obese	18.8%		
Age (years)	9.7	1.0	(7, 13)
Gender			
Male	50.8%		
Female	49.2%		
Socioeconomic Status (SES)			
High (Full priced lunch)	54.7%		
Low (Eligible for free or reduced price lunch)	45.3%		
Race/ethnicity			
White	62.2%		
African American	18.9%		
Hispanic	11.5%		
Other	7.4%		
Block Group Characteristics			
BG Percent racial/ethnic minority (%)	28.0	20.4	(0, 98.6)
BG Socioeconomic status (SES) index	-0.8	2.7	(-5.5, 9.9)
BG Population density (persons per sq. mile)	1554.64	1073.94	(15.3, 11555.5)
BG Level of urbanization			
Urban	50.4%		
Urban-Rural Mixed	44.8%		
Rural	4.8%		

Table 2

Global Moran's I and local Moran's I results.

Model #	Global Moran's I			Local Moran's I	
	Index value	Z-score	p-value	High-High Cluster Observations (#, %)*	Low-Low Cluster Observations (#, %)*
Model 1 ^a	0.039	14.3	0.001	635 (4.7%)	1058 (7.9%)
Model 2 ^b	0.007	2.6	0.005	260 (1.9%)	339 (2.5%)
Model 3 ^c	0.014	5.1	0.001	335 (2.5%)	411 (3.1%)
Model 4 ^{b & c}	0.003	0.9	0.169	205 (1.5%)	185 (1.4%)

p < 0.001, Distance tested: 0.5 mile.

^aNo covariates included, unconditional model.

^bAdjusted for youth characteristics (age, gender, SES, race/ethnicity).

^cAdjusted for block group characteristics (percent racial/ethnic minority, SES index, population density).

Table 3

Global Moran's I and local Moran's I results by level of urbanization.

Global Moran's I						
	Urban (n = 6788)		Urban-Rural Mixed (n = 6040)		Rural (n = 641)	
Model #	Index value	Z-score	Index value	Z-score	Index value	Z-score
Model 1 ^a	0.052	11.6 ^{***}	0.031	6.4 ^{***}	0.029	2.7 [*]
Model 2 ^b	0.005	1.0	0.007	1.6 [*]	0.021	1.9 [*]
Model 3 ^c	0.019	4.5 ^{***}	0.015	3.2 ^{**}	0.033	2.9 ^{**}
Model 4 ^{b,c}	0.002	0.4	0.006	1.3	0.024	2.1 [*]

Local Moran's I						
	Urban (n = 6788)		Urban-Rural Mixed (n = 6040)		Rural (n = 641)	
	High-High Cluster Observations (#, %) [*]	Low-Low Cluster Observations (#, %) [*]	High-High Cluster Observations (#, %) [*]	Low-Low Cluster Observations (#, %) [*]	High-High Cluster Observations (#, %) [*]	Low-Low Cluster Observations (#, %) [*]
Model 1 ^a	381 (5.6%)	381 (5.6%)	174 (2.9%)	377 (6.2%)	15 (2.3%)	15 (2.3%)
Model 2 ^b	184 (2.7%)	100 (1.5%)	100 (1.7%)	146 (2.4%)	11 (1.7%)	9 (1.4%)
Model 3 ^c	195 (2.9%)	155 (2.3%)	123 (2.0%)	175 (2.9%)	14 (2.2%)	17 (2.7%)
Model 4 ^{b&c}	151 (2.2%)	85 (1.3%)	97 (1.6%)	100 (1.7%)	15 (2.3%)	12 (1.9%)

* p < 0.05

** p < 0.01

*** p < 0.001.

Distances tested: Urban (500 m), Urban-Rural Mixed (730 m), Rural (3000 m).

^aNo covariates included, unconditional model.

^bAdjusted for youth characteristics (age, gender, SES, race/ethnicity).

^cAdjusted for block group characteristics (percent racial/ethnic minority, SES index, and population density).