An exploratory spatial data analysis of low birth weight prevalence in Georgia
Wei Tu a,*, Stuart Tedders b, Jie Tian a

a Department of Geology and Geography, Georgia Southern University, Statesboro, GA 30460-8149, USA
b Jiann-Ping Hsu College of Public Health, Georgia Southern University, Statesboro, GA 30460-8149, USA

A B S T R A C T

Low birth weight (LBW), defined as a live birth weighing less than 2500 g, is a significant public health problem in this country, and this condition is a significant predictor of infant mortality (Collier & Hogue, 2006; Merves, Decoufle, Murphy, & Yergin-Allsopp, 1995; Mssall & Tremont, 2002; Rantakallio & von Wendt, 1985; Robinson & Robinson, 1965). In the United States, the LBW rate was reported to be 8.2% in 2007 representing a 16.0% increase over the 1990 rate (CDC, 2010). In the State of Georgia, the unadjusted LBW rate in 2006 was 9.6%, and this, too, represented an increase in prevalence over the 1990 rate (GDCH, 2010). At both the national and state level, reported prevalence rates of LBW are well above the Healthy People 2010 intended target of 5.0% (USDHHS, 2000). These statistics suggest that LBW remains a challenging public health issue in this country and further research is necessary to understand the epidemiology of this problem.

Significant disparities in the rates of LBW among certain racial and ethnic groups are also reported. After adjusting for factors known to be associated with LBW, observed rates among non-Hispanic blacks are still two times greater than rates observed for non-Hispanic white mothers (CDC, 2010). LBW rates also differ markedly over a given study area, and, in the southeastern United States, clusters of significantly high and low LBW rates have been documented in both Florida (Hunt et al., 2002; Reader, 2001) and Georgia (GDCH, 2010). Even after controlling for known maternal and area risk factors, clusters of LBW rate are also reported to vary considerably among black and non-black subgroups across neonatal service regions in the United States (Thompson, Goodman, Chang, & Stukel, 2005).

The importance of location, spatial interaction, spatial structure, and spatial processes has been well established in the public health literature (Elliott, Wakefield, Best, & Briggs, 2000; Gatrell & Rigby, 2004; Moore & Carpenter, 1999). The application of Exploratory Spatial Data Analysis (ESDA) is one of the more commonly used spatial techniques in initial studies involving LBW, as well as other health-related issues (Anselin, 1994; Crichton, Elliott, Moinnедин, Kanaroglu, & Upshur, 2007; Ganapati, Ganapati, De La Rosa, & Rojas, 2010). The utility of ESDA tools allows researchers to map spatial patterns, identify local variability in health data, and assess efficacy of spatial models (Bertazzon, Olson, & Knudtson, 2010; Kieffer, Alexander, Lewis, & Mor, 1993; Pouliou & Elliott, 2009).

Introduction

Low birth weight (LBW), defined as a live birth weighing less than 2500 g, is a prevalent health problem in this country, and this condition is a significant predictor of infant mortality (Collier & Hogue, 2006; Merves, Decoufle, Murphy, & Yergin-Allsopp, 1995; Mssall & Tremont, 2002; Rantakallio & von Wendt, 1985; Robinson & Robinson, 1965). In the United States, the LBW rate was reported to be 8.2% in 2007 representing a 16.0% increase over the 1990 rate (CDC, 2010). In the State of Georgia, the unadjusted LBW rate in 2006 was 9.6%, and this, too, represented an increase in prevalence over the 1990 rate (GDCH, 2010). At both the national and state level, reported prevalence rates of LBW are well above the Healthy People 2010 intended target of 5.0% (USDHHS, 2000). These statistics suggest that LBW remains a challenging public health issue in this country and further research is necessary to understand the epidemiology of this problem.

Significant disparities in the rates of LBW among certain racial and ethnic groups are also reported. After adjusting for factors known to be associated with LBW, observed rates among non-Hispanic blacks are still two times greater than rates observed for non-Hispanic white mothers (CDC, 2010). LBW rates also differ markedly over a given study area, and, in the southeastern United States, clusters of significantly high and low LBW rates have been documented in both Florida (Hunt et al., 2002; Reader, 2001) and Georgia (GDCH, 2010). Even after controlling for known maternal and area risk factors, clusters of LBW rate are also reported to vary considerably among black and non-black subgroups across neonatal service regions in the United States (Thompson, Goodman, Chang, & Stukel, 2005).

The importance of location, spatial interaction, spatial structure, and spatial processes has been well established in the public health literature (Elliott, Wakefield, Best, & Briggs, 2000; Gatrell & Rigby, 2004; Moore & Carpenter, 1999). The application of Exploratory Spatial Data Analysis (ESDA) is one of the more commonly used spatial techniques in initial studies involving LBW, as well as other health-related issues (Anselin, 1994; Crichton, Elliott, Moinnедин, Kanaroglu, & Upshur, 2007; Ganapati, Ganapati, De La Rosa, & Rojas, 2010). The utility of ESDA tools allows researchers to map spatial patterns, identify local variability in health data, and assess efficacy of spatial models (Bertazzon, Olson, & Knudtson, 2010; Kieffer, Alexander, Lewis, & Mor, 1993; Pouliou & Elliott, 2009).
In particular, the ESDA toolkit includes cluster analysis which is frequently used for disease surveillance activities. The cluster analysis option enables researchers to identify geographically bounded group occurrences that are unlikely to have occurred by chance (Knox, 1989). Although the identification of clusters does not provide a causal explanation for the spatial pattern and spatial clustering of specific health-related events, working hypotheses are often more easily generated and can be subsequently tested in future research efforts (Pickle, 2002). Published literature regarding LBW in Georgia is sparse, particularly in rural areas of the state. Given the paucity of information related to this health problem, the intent of this research was to investigate the baseline prevalence of LBW in Georgia. In addition, this study sought to examine the spatial patterns of LBW in Georgia using an ESDA approach because, to date, this methodology has been
commonly employed in Georgia to investigate the spatial pattern and trends of LBW. It is anticipated that outcomes gleaned from this research will provide public health professionals with baseline measures of the geographic distribution of LBW in Georgia that is fundamental for establishing a best practice framework for planning health promotion program. Armed with this information, public health practitioners will be able to facilitate a targeted approach to prevention and intervention activities designed to impact high-risk populations as well as geographic regions, thereby potentially narrowing LBW disparities in Georgia.

Although geographic concepts, themes, and tools have been increasingly applied in various real world problems by professionals both inside and outside academia in the past three decades (Frazier, 1978; Montz & Tobin, 2007; Warf, Janelle, & Hansen, 2004), there is a continuing need to "...bring new and innovative methods and information to the solution of society's most vexing problem" (Frazier, 2004, p. 205). This paper contributes to the medical/health theme in applied geography literature by addressing a significant public health issue using geospatial technology.

**Methods**

**Data sources and study design**

The research presented in this paper can be classified as a cross-sectional, ecologically-based study involving birth outcomes in Georgia. Prevalence of LBW, as well as associated maternal risk factors, for this research was derived from electronic birth certificate data (BCD) obtained from the state Vital Records Office in Atlanta, Georgia. Analysis of these data was limited to the year 2000 so outcomes could be more conveniently correlated with decennial census-based demographic profiles for following studies. A detailed description of the birth data is available elsewhere (GDCH, 2010).

In the year 2000, a total of 132,287 births were recorded in Georgia. Given that multiple births (i.e., more than one fetus carried during a single pregnancy event) are known to significantly reduce birth weight, these birth records \( n = 3968 \) were excluded prior to the analysis and only singleton and live births were investigated \( n = 127,319 \). Birth records were first geo-coded and aggregated by geographical units. Unadjusted LBW rates were then calculated. These data were examined for three separate group aggregates. These aggregates included births to all women, births to black women, and births to white women. Births to other racial and ethnic groups were not examined as an independent cohort due to the small numbers occurring in a given year in Georgia, as well as statistical variability observed with small numbers of events that often translates into the instability of the rate calculations. In addition, all three group aggregates were examined at both the county \( n = 159 \) and census tract \( n = 1618 \) level.

The U.S. Census cartographic boundary files were acquired and used as geographical boundaries (county and census tract) for this study (U.S. Census Bureau, 2010). Other supplemental data such as city boundaries were collected from various state-level sources such as Georgia County Guide (GCC, 2011) and Georgia GIS Clearinghouse (GCC, 2011). Fig. 1 illustrates the boundaries and names of the Georgia counties, major cities, and state highways. Fig. 2 illustrates the 18 public health district boundaries and major cities in the state.

**Exploratory spatial data analyses**

Statistical treatment of these data followed a two-step analytical process. The first step involved mapping the spatial distribution of unadjusted and adjusted LBW rates for all births, births to black mothers, and births to white mothers at both county and census tract levels. The second step involved the global and local spatial clustering analyses of LBW prevalence. The Moran’s I statistic was used as a measure of the global clustering and was assessed by testing the null hypothesis that the spatial pattern of these data were random. Therefore, rejection of the null hypothesis implies a non-random spatial pattern that is also referred to as spatial autocorrelation. More specifically, positive spatial autocorrelation indicates similar values occur at adjacent locations; whereas negative autocorrelation implies that high values appear next to...
low values. Moran’s I ranges approximately from +1 (for positive spatial autocorrelation) to −1 (negative autocorrelation), and its expected value in the absence of autocorrelation approximates zero. In addition, Queen’s case adjacency was used to define neighbor relationships, which considers that all counties/census tracts sharing at least a corner as neighbors (Anselin, 1995).

Anselin’s local indicator of spatial autocorrelation (LISA) was also used in the analysis of these data. Specifically, LISA is an indicator of local spatial association that measures whether or not the LBW rate for a particular spatial unit (e.g., a county) is closer to the values of a neighboring unit or to the average of the study area (Anselin, 1995). To test for significance of these associations, a Monte Carlo permutation approach was used. This permutation approach assumes, under a randomization assumption, that the rate under study is equally as likely to be observed at any location, so the observed values are randomly shuffled over the given spatial units and the LISA is re-calculated for each permutation. The significance of the LISA is then determined by generating a reference distribution using a sufficiently large number of permutations. For this study, the value 999 was used as the reference distribution using a sufficiently large number of the permutations. Lastly, cluster maps were produced using the results of statistical testing to illustrate the local spatial patterns.

The results of this analysis yielded five categories of spatial units (at either the county or census tract level). These categories were defined as “high-high (HH),” “low-low (LL),” “high-low (HL),” “low-high (LH),” or “not significant (NS).” A HH category indicates clustering of high values of unadjusted LBW rates, while a LL category indicates clustering of low values of unadjusted LBW rates. These outcomes are equivalent to a positive spatial autocorrelation. In addition, a HL category indicates that high values are adjacent to low values of the unadjusted LBW rates, while a LH category indicates that low values are adjacent to high values of unadjusted LBW rates. These outcomes are equivalent to a negative spatial autocorrelation. Lastly, the NS category indicates that there is no statistically significant spatial autocorrelation.

One common issue in mapping and analyzing public health data is the varying size of populations (Diehr, 1984). To lessen the impact of population heterogeneity, several smoothing methods have been developed to stabilize local estimates. These methods help stabilize rates using various techniques such as borrowing strength from either neighboring areas or a prior distribution (Clayton & Kaldor, 1987; Gelman, Carlin, Stern, & Rubin, 2004; Kafadar, 1996; Talbot, Kolfzijff, Forand, & Haley, 2000). Considerably, adjustments on Moran’s I statistic were proposed to account for the varying population size (Assuncao & Reis, 1999; Oden, 1995; WaldhÖR, 1996). Assuncao and Reis (1999) demonstrated that their empirical Bayes-based approach is superior over tests such as the ones proposed by WaldhÖR (1996) and Oden (1995).

All data analyses, including data processing, mapping, and statistical testing, were conducted using the ArcGIS Desktop 9.3.1 software package (ESRI, 2008) and OpenGeoDa 0.9.8.15 (Anselin, Syabri, & Kho, 2006).

## Results

County-level unadjusted LBW rates for the three cohorts under investigation are reported in Table 1. According to these data, the mean [M] LBW rate for Georgia was 7.09% (Standard Deviation [SD] = 0.023). The LBW rate among black mothers (M = 10.23%, SD = 0.057) represented a nearly 2 fold increase over the observed rate among white mothers (M = 5.23%, SD = 0.052).

Fig. 3 illustrates the spatial distribution of unadjusted LBW rates for the three cohorts under investigation. Relative to the state, the total LBW rate is higher as indicated by more than two standard deviations above the mean was observed in four Georgia counties in the central and southwestern part of the state (Wilkinson, Talbot, Clay, and Brooks counties). Moreover, Talbot, Brooks, and Twiggs counties had observed rates of 14.1%, 12.8%, and 12.2%, respectively, and were ranked as the highest LBW rates among all 159 counties in Georgia. Two low-rate clusters were observed in the northern (the Towns County and the White County) and the southeastern (the Pierce County) areas of Georgia. In fact, there were 85 births in the Towns County and none were classified as LBW.

Among births to black mothers (Fig. 3), only 36 counties (22.6%) in Georgia experienced LBW rates lower than the state average. LBW rates at or above one standard deviation from the state average were widely distributed throughout Georgia. Overall, considerable variation in LBW rates among births to black mothers was observed throughout the state and ranged from 2.63% in the Long County in southeast Georgia to 40.00% in Pickens County in north Georgia. Two “clusters” of counties with lower than expected LBW rates among black mothers were observed in the northern and southeastern areas of the state.

Unadjusted LBW to white mothers in 137 counties (86.2%) in Georgia were lower than the state average as indicated by standard deviations of the mean (Fig. 3). Compared to that of the black births, the LBW rates among whites tended to vary much less substantially across counties of the state. The range of observed rates was 1.54% in the Wilcox County to 7.73% in the Burke County. There seems to be a cluster of high-rate counties in the central part of the state. Among the births to white mothers, two larger “clusters” of counties with low rates were observed. One cluster was found in north Georgia in the Atlanta Area while the other cluster centers at Pierce County located in southeast Georgia.

County-level empirical Bayes smoothed LBW rates for the three cohorts under investigation are reported in Table 2. According to these data, the adjusted mean [M] LBW rate for Georgia was 6.75% (standard deviation [SD] = 0.010). The LBW rate among black mothers (M = 10.23%, SD = 0.005) represented more than 2 fold increase over the adjusted rate among white mothers (M = 4.83%, SD = 0.010). Overall, the unadjusted rates were adjusted toward prior means. Standard deviations of adjusted rates were also decreased significantly. Because the base value must be nonzero to calculate empirical Bayes (EB) rates and because three counties (Dawson, Fannin and Towns Counties) in our data had zero black births in 2000, these three counties were merged into one of their neighboring counties with which they share the longest boundary with prior to calculating the empirical Bayes smoothed LBW rates. As such, the Dawson County was merged into the Lumpkin County, the Fannin County was merged into the Gilmer County, and the Towns County was merged into the Union County. Therefore, the total mapping and analytical units was reduced from 179 to 156. For simplicity, “county” will still be used throughout the remainder of the paper.

Fig. 4 illustrates the spatial distribution of empirical Bayes smoothed LBW rates for the three cohorts under investigation. Relative to the state, the total LBW rate is higher as indicated by more than two standard deviations above the mean was observed in two Georgia counties in the central and southwestern part of the state (the Bibb County and the Brooks County). Moreover, the Bibb County, the Brooks County and the Tift County had adjusted rates of 9.13%, 8.91%, and 8.67%, respectively, and were ranked as the

### Table 1

<table>
<thead>
<tr>
<th>Demographic group</th>
<th>Mean (%)</th>
<th>Range (%)</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>10.23</td>
<td>0—40</td>
<td>0.057</td>
</tr>
<tr>
<td>White</td>
<td>5.23</td>
<td>0—16.67</td>
<td>0.052</td>
</tr>
<tr>
<td>All</td>
<td>7.09</td>
<td>0—14.13</td>
<td>0.023</td>
</tr>
</tbody>
</table>

Fig. 3 illustrates the spatial distribution of empirical Bayes smoothed LBW rates for the three cohorts under investigation. Relative to the state, the total LBW rate is higher as indicated by more than two standard deviations above the mean was observed in two Georgia counties in the central and southwestern part of the state (the Bibb County and the Brooks County). Moreover, the Bibb County, the Brooks County and the Tift County had adjusted rates of 9.13%, 8.91%, and 8.67%, respectively, and were ranked as the
highest LBW rates among all 156 counties in Georgia. Low-rate clusters were observed in the northern (the Paulding County, the Cherokee County, the Forsyth County, the Hall County, and the White County), eastern (the Columbia County), and the southeastern (the Pierce County) areas of Georgia.

Considerable variation in LBW rates among births to black mothers was observed throughout the state (Fig. 4). The adjusted rate ranged from 8.86% in the Gwinnett County in the Atlanta area to 12.13% in the Spalding County in central west Georgia. LBW rates at or above one standard deviation from the state mean were widely distributed throughout Georgia. Two “clusters” of counties with lower than two standard deviations from the state mean among black mothers were observed in the Atlanta area (the Gwinnett County, the Cobb County, the Clayton County, and Dekalb County) and in the Camden County in southeastern Georgia.

Compared to that of the black births, the adjusted LBW rates among whites varied less substantially across counties of the state.

### Table 2

<table>
<thead>
<tr>
<th>Demographic group</th>
<th>Mean (%)</th>
<th>Range (%)</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>10.23</td>
<td>8.86–12.13</td>
<td>0.005</td>
</tr>
<tr>
<td>White</td>
<td>4.85</td>
<td>3.90–6.27</td>
<td>0.003</td>
</tr>
<tr>
<td>All</td>
<td>6.75</td>
<td>3.99–10.54</td>
<td>0.010</td>
</tr>
</tbody>
</table>
The range of adjusted rates was 3.90% in the Forsyth County to 6.26% in the Walker County. There seems to be a cluster of high-rate counties in the central part of the state. Among the births to white mothers, two “clusters” of counties with low rates were observed. One cluster was found in the Atlanta area (the Gwinnett County, the Fulton County, the Forsyth County, and the Hall County) and the Columbia County in eastern Georgia.

Table 3 illustrates spatial correlation statistics and corresponding p-values as estimated by the Moran’s I statistic. According to these data, the level of spatial autocorrelation is 0.3766 (p < 0.001) for all the births occurring in Georgia during the study period. Among births to black mothers, a Moran’s I statistic of 0.2677 (p < 0.001) was observed while a value of 0.4365 (p < 0.001) was observed among births to white mothers. For total births, births to black mothers, and births to white mothers, the Moran’s I statistic and corresponding p-values all suggest non-randomness in the overall spatial pattern of LBW in Georgia.

Table 4 illustrates spatial correlation statistics and corresponding p-values as estimated by the Moran’s I statistic based on empirical Bayes smoothed LBW rates (Assuncao & Reis, 1999).
According to these data, the level of spatial autocorrelation is 0.2250 (p < 0.001) for all the births occurring in Georgia during the study period. Among births to white mothers, a Moran’s I statistic of 0.1120 (p < 0.05) was observed while a value of 0.0404 (p = 0.158) was observed among births to black mothers. For total births, as well as births to white mothers, the Moran’s I statistic and corresponding p-values suggest non-randomness in the overall spatial pattern of LBW in Georgia. This trend was not observed for births to black mothers, thereby suggesting that the spatial pattern of LBW prevalence among this cohort appears to be random. In addition, comparing with the results in Table 2, the values of Moran’s I statistic and corresponding p-values decreased in all the three Moran’s I measures.

County-level LISA (based on unadjusted LBW rates) significant clusters (HH, LL, HL, and LH) were observed for all the three cohorts and the illustration of these clusters can be found in Fig. 4. For all the births in Georgia, one larger hot spot (HH) was detected in the central portion of the state centering at the Washington County and one smaller hot spot (HH) was found in the Clay County in the southwest. Two cool spots (LL) were observed in the northern and southeastern regions of Georgia. For all the black births, there was a one-county hot spot located in the Tift County. There were also two cool spots, one found in the northeastern and the other in the southeastern part of the state. For all white births, one hot spot was detected in the central part of the state including the Glascock County, the Washington County, the Wilkinson County, and the Twiggs County. There were also two cool spots, one being the Pierce County and the other comprising the Stewart County, the Webster County, and the Marion County. In addition, several south central and southwestern counties showed a dispersed pattern of cluster (HL or LH) as displayed in Fig. 5.

County-level LISA (based on empirical Bayes LBW rates) significant clusters (HH, LL, HL, and LH) were observed for all the three cohorts and the illustration of these clusters can be found in Fig. 6. Note that the total analytical units are 156 due to above-mentioned reason. For all the births in Georgia, one larger hot spot (HH) was detected in the central portion of the state centering at the Washington County and two smaller hot spots (HH) were found in the southwest. Three one-county cool spots (LL) were observed in, or near, the Atlanta area. For all the black births, there were two hot spots located in central west and southwest Georgia. There was one cool spot in the Bacon County in the southeastern part of the state. For all the white births, one large hot spot (HH) was detected in the northwest corner of the state including the Dade County, the Walker County, the Catoosa County, the Gordon County, the Chattooga County, and the Floyd County. A one-county hot spot (HH) was found in the Thomas County in the southwest. There were also two one-county cool spots (LL) in the Baker County and the Clayton County. In addition, several counties in northern, southeastern and southwestern areas of the state showed a dispersed pattern of cluster (HL or LH) for all the births.

The illustration of the census-level LISA significant clusters for the three cohorts can be found in Figs. 7 through 10. For all births, hot spots were found in Atlanta area, central and southern regions of the state while cold spots were seen in both northeastern and southeastern regions of the state. Four types of clusters were identified in finer geographic scales. A spatial pattern with mixed high-rate and low-rate census tracts was observed in downtown areas of Georgia’s four most populous cities, Atlanta, Augusta, Columbus, and Savannah.

**Table 3**

<table>
<thead>
<tr>
<th>Demographic group</th>
<th>Moran’s I</th>
<th>Pseudo p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>0.2677</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>White</td>
<td>0.4365</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>All</td>
<td>0.3766</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

**Table 4**

<table>
<thead>
<tr>
<th>Demographic group</th>
<th>Moran’s I</th>
<th>Pseudo p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Black</td>
<td>0.0404</td>
<td>0.158</td>
</tr>
<tr>
<td>White</td>
<td>0.1120</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>All</td>
<td>0.2250</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

**Discussion**

This study illustrates the marked variability of LBW prevalence in the State of Georgia at both the county and census tract level. In addition, the significant positive spatial autocorrelation identified by Moran’s I statistic in all three birth cohorts, as well as significant local clusters identified by LISA, confirm the presence of spatial heterogeneity in unadjusted LBW rates (Anselin, 1994). The identified spatial pattern and clustering of events provide important information for the development and refinement of geographically and population specific prevention programs to reduce LBW risk. In addition, this information is useful to health planners because many current policies or health initiatives are principally based on the assumptions of spatial homogeneity (Katzmarzyk, 2002; Vanasse, Demers, Hemjari, & Courteau, 2006).

The difference in unadjusted LBW rates between black and white births is consistent with results from previous studies in the United States (Dunlop, Salihu, Freymann, Smith, & Brann, 2010; English et al., 2003; O’Campo, Xue, Wang, & Caughy, 1997). The elevated rates in rural Georgia are a multi-dimensional problem and observed associations are related to many different factors that deserve further research. However, one possible explanation for these patterns is the unfavorable socio-economic conditions unique to many of Georgia’s rural counties (Levernier, 1996; Levernier & White, 1998). Moreover, the low rates in the northern portion of the state, the Blue Ridge region, may suggest the physical environment as a contributing factor to reducing LBW risks. However, it can more likely be explained by the uniqueness of population demographics in that part of Georgia. Overall, extra caution should be used when interpreting these results due to the heterogeneity of total birth across Georgia counties. The results based on empirical Bayes smoothed LBW rates significantly lessen the spatial variation of LBW. The results, on the other hand, also confirmed marked variation between overall low rates in the Atlanta area and neighboring counties, as well as the observed high rates in more rural regions of the state.

Similar to the situation in many other states, health districts in Georgia are administrative units designed to assess and prioritize the health needs of residents, as well as the health goals of the district. In addition, health districts work to reconcile the needs of the residents and the availability of resources at hand to develop an integrated approach toward population health and health promotion (GDCH, 2010). Therefore, detection of LBW clusters using both global and local measures could be an effective method of public health surveillance, thus emerging clusters can possibly be quickly identified. A timely response to any adverse health outcome, including LBW, could position public health practitioners to more effectively and efficiently mount a public health response in order to minimize morbidity and mortality of the problem. Moreover, information provided through a technique similar to the one described in this research may allow public health decision makers to allocate more human and capital resources to communities that are identified as hot spots. In addition, health initiatives including,
but not limited to, educational programs could also be more effectively developed to target high-risk population groups and geographic regions. While it is important to focus resources on high-risk areas, it may also be advantageous from the perspective of health promotion and program planning to further explore areas of low prevalence to ascertain factors that might be protective.

This study provides an important basis for further research to better understand behavioral, demographic, and socio-economic risk factors of LBW that, in turn, may yield plausible explanations for causality. The mixed high/low clustering pattern in Georgia’s metropolitan areas, for example, may be related to the segregated residential patterns associated with urban areas in Georgia. However, much more in-depth research is necessary to infer a cause and effect relationship in both rural and urban areas alike.

There are four major limitations of this study that need to be addressed. First, the addresses used to geo-code births are mother’s self-reported mailing addresses, which are the best available data that could be utilized when this study was conducted. The problem is that the information regarding length of residence is not included in the data. This is a significant issue because the length of

![Fig. 5. County-level LISA cluster map for unadjusted LBW rates in 2000.](image-url)
residence is directly related to the degree of exposure to the physical and built environment of study subjects. This factor could be accounted for in the future studies by incorporating survey data with birth certificate data (O’Campo, 2003). This additional information may allow a researcher to exclude subjects from any given residential area and reassign that subject to their previous residential area to more accurately estimate the relationship between residential area and time-at-risk (Caughy, O’Campo, & Muntaner, 2003).

The second limitation is due to the fact that spatial patterns of LBW rates may change depending on the spatial scales and units used in the analyses, an issue which is commonly known as the Modifiable Areal Unit Problem (MAUP) or Ecological Fallacy (Openshaw, Chailton, Wymer, & Craft, 1987; Robinson, 1950). To partly address this problem, spatial analyses were conducted at both the county and the census tract levels in this study, but statistical testing for the influence of the MAUP effect on spatial patterns and spatial clustering is beyond the scope of this study. Additionally, although a widely accepted solution to this issue has yet to be published, a number of methods for mitigating the MAUP problem have been proposed and these methodologies can certainly be tested in the future studies (Amrhein, 1995; Andrew et al., 2008; Hayward, 2009; King, 1979; Openshaw et al., 1987; Swift, Liu, & Uber, 2008). Given the MAUP effect was not fully addressed, the results reported in this study should be interpreted with caution.

**Fig. 6.** County-level LISA cluster map for empirical Bayes smoothed LBW rates in 2000.
The third limitation is due to the use of unadjusted LBW rates in this study. Rates in areas with small total number of births are subject to the “small number problem”, and estimated rates are likely to be artificially elevated (Diehr, 1984). In this study, this problem is more obvious at the census tract level. To illustrate this problem, the unadjusted LBW rate in the census tract 13121006810 was 100%, which was an obvious overestimation of LBW risk. In this instance, there was only one birth in this census tract for the year and it was classified as a LBW birth. Given this a well-documented issue in public health research, many solutions to overcome the small number problem have been proposed (Crosse, Alder, Østbye, & Campbell, 1997; Grady & Enander, 2009; Haining, Wises, & Blake, 1994). For example, an emerging approach is to construct a geographic area with a sufficiently large base population and with attribute homogeneity and spatial contiguity (Murray & Shyy, 2000). To conceptually illustrate this approach, two counties (the

Fig. 7. Census tract-level LISA cluster map for unadjusted LBW rates in 2000 – black births.

Fig. 8. Census tract-level LISA cluster map for unadjusted LBW rates in 2000 – white births.
Fig. 9. Census tract-level LISA cluster map for unadjusted LBW rates in 2000 — all births.

Fig. 10. Census tract-level LISA cluster map for unadjusted LBW rates in four cities in 2000 — all births.
Glascok County and the Taliaferro County) with a total birth of less than 30 were selected and merged to one of their neighboring counties prior to the analyses. The rule was set to merge a source county into a neighboring county with which it shares the longest boundary. The Glascok County was then merged into the Warren County and the Taliaferro County was merged into the Jefferson County. As a result, each analytic unit now had at least 30 total births and the total analytic units were reduced from 159 to 157.

Table 5 illustrates spatial correlation statistics and corresponding p-values as estimated by the Moran’s I statistic after reconstructing the geographic units. Overall, the results are consistent with the data presented in Table 2, but the Moran’s I values and p-value are slightly different. In future studies, more sophisticated zoning algorithms could be applied to formally addressed the “small number” problem using this method (Grady & Enander, 2009; Mu & Wang, 2008).

The last noted limitation is due to several inherent problems regarding the Moran’s I statistic. First, both the normality and the randomization assumptions are criticized as inappropriate to many real world phenomenon; second, the LISA statistic bears several theoretical issues such as its analytic distributional properties as well as the well-known multiple testing problem (Wall & Getway, 2004); third, the values of Moran’s I statistic are sensitive to factors such as the choice of the null hypothesis, the definition of neighborhood, and the analytical scale.

However, these limitations should not discourage the application of global and local Moran’s I measures in particular and ESDA methods in general in applied public health studies. Indeed, the major objectives of this type of research are to help generate working hypotheses and design more sophisticated research protocol for future research efforts. In addition, this research is ideal for providing public health professionals and policy makers with information about areas of elevated risk so effective intervention programs can be designed. The data presented in this study represents the first comprehensive examination of LBW prevalence in the State of Georgia using a GIS-based ESDA approach. This study also serves to strengthen the results of previous studies reported in the literature by confirming marked race-specific regional disparities in the LBW prevalence in the United States. The identification of spatial clusters provides additional information on the local geographical variation of LBW prevalence in Georgia. In the long run, we hope that information gleaned from this research and more in-depth studies based on this research will enable a regionally focused strategy of prevention and health promotion, as well as the effective allocation of scarce resources and services based on community and population specific demands.

Acknowledgments
We appreciate the suggestions of Drs. Fahui Wang at Louisiana State University and Xingyou Zhang at the US Center for Disease Control and Prevention. We would like to thank Joseph Lowry for providing research assistance for the manuscript. We would also like to thank the editor of Applied Geography and the two anonymous referees for their insightful comments and constructive criticism. The analyses and opinions in this article are entirely the responsibility of the authors.

References


