한국지역지리학회지 제16권 제5호(2010) 610-622

Spatial Dependency and Heterogeneity of Adult Diseases : In the Cases of Obesity, Diabetes and High Blood Pressure in the U.S.A.

Yang, Byungyun* · Hwang, Chulsue.**

성인병의 공간적 의존성과 이질성 : 미국의 비만, 당뇨, 고혈압을 사례로

양병윤* · 황철수**

Abstract : The proportion of overweight and obese individuals in the United States has been continuously increasing up to recently. Many studies related to obesity have concentrated on jurisdictional levels of aggregation, making it very difficult to clearly illustrate at risk regions. In other words, little research has been conducted in relation to spatial patterns considering spatial dependency and heterogeneity by spatial autocorrelation models over space. In response, this research analyzes spatial patterns between overweight/obesity and risk factors, such as high blood pressure and diabetes, over space. Specifically, the Moran's I and Geary's C will be conducted for global and local measures. What is more, the Ordinary Least Square (OLS) linear regression and Geographically Weighted Regression methods will be applied to identify spatial dependency and spatial heterogeneity. Data provided by the Behavioral Risk Factor Surveillance System (BRFSS) have Body-Mass Index (BMI) rates, containing 4 rates of under, healthy, overweight, and obesity. In addition, high blood pressure and diabetes rates in the United States will be used as independent variables. Lastly, we are confident that this research will be beneficial for a decision maker to make a prevention plan for obesity.

Key Words : overweight and obesity, spatial autocorrelation, spatial dependency and heterogeneity

요약 : 미국에서 과체중과 비만은 최근까지 지속적으로 증가하고 있다. 그러나 과체중과 비만을 일으키는 요인에 대한 연구에 비해 이들 현상이 갖는 공간적 분포, 즉 공간의존성과 이질성에 대한 연구는 미미한 실정이다. 이에 본 연구는 과체중과 비만, 그리고 위험요인(고혈압과 당뇨병) 사이의 공간적 분포를 분석하였다. Moran's I와 Geary's C 분석 결과, 비만은 공간적으로 무작위 분포를 갖고 있지만, 저체중 · 정상 · 과체중은 공간적 유사성을 지니고 있음을 확인할 수 있었다. 또한 일반최소자승법에 의한 선형 회귀분석과 잔차도분석을 통해 고혈압과 당뇨는 비만과 공간적 의존성을 갖는 것으로 파악할 수 있었다.

주요어 : 과체중, 비만, 공간적 자기상관관계, 공간적 의존성과 이질성

1. Introduction

1) Background and Problem Statement

In recent years, the rapid rise of obesity and overweight has been widely recognized as a public health crisis. Obesity means a condition that results from a chronic energy imbalance whereby intake exceeds expenditure. In other words, obesity is having a very high amount of body fat in relation to lean body mass, or a Body Mass Index (BMI) of 30 or higher, which indicates a simple index of weight-for-height that is commonly used to classify underweight, overweight and obesity in adults (World Health Organization, 2004). According to the Center for Disease Control and Prevention (CDC), there are 20.8 million people, or 7.0% of the population of the United States, who have diabetes. 16.3% of adolescents (ages 12 to 19) and 15.3% of

* Ph.D. Candidate, Department of Geography, University of Georgia(조지아대학교 지리학과 박사과정)

** Associate Professor, Department of Geography, Kyung Hee University(경희대학교 지리학과 부교수)(hcs@khu.ac.kr)

children (ages 6 to 11) were obese in 2006. In addition, the National Health and Nutrition Examination Survey (NHANES) shows that, among adult men, the prevalence of obesity was 33.3% from 2005 to 2006. Also, according to the CDC report, 28.2% of Georgians were obese by the standards of the BMI in 2007. (See table 1) In a recent report, the CDC provided geographic patterns among low-income, preschool-aged children, as well as obesity rates by race and ethnicity. According to their reports (Obesity by Race/Ethnicity 2006~2008, 2009), the non-Hispanic black population had the highest prevalence, followed by Hispanics, and non-Hispanic whites. In more detail, higher prevalence was found in the Midwest and the South. Prevalence ranged from 23.0% to 45.1%, and 5 states (Alabama, Maine, Mississippi, Ohio, and Oregon) had a prevalence of >= 40%.

As stated above, this rate of obesity and overweight raises health concerns. To illustrate, obesity can cause coronary heart disease, type 2 diabetes, cancers, hypertension (high blood pressure), dyslipidemia, stroke, sleep apnea, respiratory problems, and so forth. Obesity accounts for approximately 300,000 deaths a year in the United States, and prevalence rates have been increasing over the past decade (Maddock, 2004).

In response to these crises, questions have been raised about how we can figure the relationship between BMI rates and these risk factors over space. Although GIS-based obesity maps show fundamental statistical information among different population groups, it is hard to clearly identify vulnerable regions. Thus, a lack of spatial pattern studies for identifying vulnerable areas with respect to obesity, with limited information provided by a public organization, like the CDC, has been reported. To be specific, the Behavioral Risk Factor Surveillance System (BRFSS) established in 1984 by the CDC provides general information with statistic reports with a GIS map, and a web enabled analysis tool (WEAT). The GIS map contains BRFSS data that is mapped for both the states and metropolitan/micropolitan statistical areas (MMSAs).

On the web site, a class of BRFSS data on

Table 1. 2007 U.S. Obesity Trend (CDC, 2007)

2007 State Obesity Rates							
State	%	State	%	State	%	State	%
Alabama	30.3	Illinois	24.9	Montana	21.8	Rhode Island	21.4
Alaska	27.5	Indiana	26.8	Nebraska	26.0	South Carolina	28.4
Arizona	25.4	Iowa	26.9	Nevada	24.1	South Dakota	26.2
Arkansas	28.7	Kansas	26.9	New Hampshire	24.4	Tennessee	30.1
California	22.6	Kentucky	27.4	New Jersey	23.5	Texas	28.1
Colorado	18.7	Louisiana	29.8	New Mexico	24.0	Utah	21.8
Connecticut	21.2	Maine	24.8	New York	25.0	Vermont	21.3
Delaware	27.4	Maryland	25.4	North Carolina	28.0	Virginia	24.3
Washington DC	21.8	Massachusetts	21.3	North Dakota	26.5	Washington	25.3
Florida	23.6	Michigan	27.7	Ohio	27.5	West Virginia	29.5
Georgia	28.2	Minnesota	25.6	Oklahoma	28.1	Wisconsin	24.7
Hawaii	21.4	Mississippi	32.0	Oregon	25.5	Wyoming	23.7
Idaho	24.5	Missouri	27.5	Pennsylvania	27.1		

the GIS map can be changed by type, such as Equal-Interval, Quantiles, Standard Deviations, and Natural Breaks. It can give people differences between BRFSS data on the map. However, it cannot explain spatial relationships between one BRMS data type and another one having strong dependency. That is, the GIS maps provided to BRFSS cannot explain spatial dependency over space. Furthermore, the web enabled analysis tool provides cross tabulations from the BRFSS data for analysis and logistic analysis that can analyze the BRFSS data using logistic regression. The analysis can explain the possibilities and differentiates between different BRFSS data, but it cannot explain spatial dependency and differentiation over space. Hence, spatial analysis for identifying spatial dependency and hhe rogenedencyre required over space. That is, cis ter analysis is needed to examine the SS data for BRFSbesity spatial patterns, taking into account the overall, as well as local, variability in the data. It would be worthwhile to analyze spatial patterns of obesity in the United States.

2) Objective

Public health in the Unite States is necessarily concerned with spatial aspects of environmental pollution, the spread and control of infectious diseases, and the delivery of critical health care and social services (Gatrell and Loytonen 1998; Walsh *et al.*, 1997). Thus, geographic mappings of public health will increasingly be required to interpret spatial relationships or patterns.

The aim of this research is to recognize spatial dependency and differentiation of The United States based on Spatial Autocorrelation (SA) analysis, Ordinary Least Square (OLS) regression, and a Geographically Weighted Regression (GWR) model. In other words, this study is to quantify the spatial dependency between obesity rates and diabetes/high blood pressure on the global and local scales. Prevalence of overweight and obesity in one health region are likely to be correlated with provenience in nearby regions, indicating the presence of clusters (Bailey and Gatrell, 1995). In addition, spatial heterogeneity will be analyzed to recognize spatial relationships between dependent and independent values under non-stationarity. To do this, GWR will be applied. Accordingly, to reveal spatial dependency and heterogeneity in this research, SA, OLS, and GWR methods will be conducted with Arc-GIS 9.X version and Geoda (Anselin, 1994).

2. Study Area

1) Study Area and Data Collection

The study area is The United States. Data applied in this research come from the Behavioral Risk Factor Surveillance System (BRFSS), which is a state-based system of health surveys that collects information on health risk behaviors, preventive health practices, and health care access primarily related to chronic disease and injury. The BRFSS is a collaborative project of the Centers for Disease Control and Prevention (CDC) and U.S. states and territories. The BRFSS, administered and supported by CDC's Behavioral Surveillance Branch, is an ongoing data collection program designed to measure behavioral risk factors in the adult population (18 years of age or older) living in households. Their ways to collect data depend on a computer -assisted telephone interview (CATI) system. The questionnaire has two sections: a core section, concerning health status, diabetes, demographics and cancer survival and an optional section, concerning adult asthma history, inadequate sleep and healthy days, with "yes" or "no" questions.

In this research, as noted in the above introduction, I will use high-risk factors as variables, such as rates of high blood pressure and diabetes with obesity, to identify spatial dependency and heterogeneity over space. Specifically, rates of BMI for obesity and overweight will be used as dependent variables for regression models and spatial autocorrelation. Rates of high blood pressure and diabetes will be used as independent values.

The CDC provides GIS data, which include both polygon and point shape files. The polygon shape file contains the 50 states, and the point shape file includes the metropolitan/micropolitan statistical areas (MMSAs) sampled. In this research, polygon data are used for analysis of polygon patterns in terms of obesity and overweight. The reason why points are not considered is because each state has different numbers of sample points and locations generated by computer- each telephone interviews. These cannot be satisfpolywith fair conditions over space. In other words, the spatial distributd locf points is not fair. Although sample cities can be compared in this way for the quantitative level, analysis for spatial pattern is not suitable for this research. The polygon shape file has an Albers Conic projection tied to the United States Datum of 1983 (NAD 83).

2) Variances for Analysis

Adult obesity is associated with a variety of diseases, including heart disease, diabetes, high blood pressure and stroke, high cholesterol, certain types of cancer, arthritis, and breathing problems. Hence, as mentioned in the previous chapter, the overweight and obesity in terms of Body Mass Index (BMI) will be used as dependent values. High blood pressure and diabetes rates will be used as independent values because the two values have a strong association with obesity.

BMI is a measure of an adult's weight in relation to his or her height, and defined as an adult's weight in kilograms divided by the square of his or her height in meters. BMI provides a reliable indicator of body fatness for most people and is used to screen for weight categories that may lead to health problems. To illustrate, levels of different household income, consumption of fruit and vegetables, and physical activity can be considered as independent values because these factors can allow the dependency of obesity and overweight to change. However, in this research, these factors will not be considered.

In an interpretation of BMI, if your BMI is less than 18.5, it falls on the "underweight" range. If your BMI is 18.5 to 24.9, it falls within the "normal" or healthy weight range. If your BMI is 25.0 to 29.9, it falls within the "overweight" range. Finally, if your BMI is 30.0 or higher, it falls within the "obese" range (Macera *et al.*, 2005). In this research, the data set BMI two classes for BMI rate rangef Ist range ur Bromet .l.to 24.9, and second range is 33.1 to 40.4. Hence, the class will be estimated to analyze spatial dependency.

3. Methodology

1) Spatial Pattern, Spatial Dependency, and Spatial Heterogeneity

A spatial pattern is a static concept since a pattern only shows how geographic objects are distributed at one given time. The spatial pattern can generally be categorized as clustered, dispersed, or random. Spatial statistics provide useful tools for describing and analyzing how various geographic objects occur or change across the study area and over time. These statistics are formulated specifically to take into account the location attributes of the geographic objects (Lee, Wong, 2001).

In this research, when analyzing spatial pattern, observed patterns with a theoretical pattern will be compared for spatial dependency. To do this, the similarity or dissimilarity of any pair of neighboring polygons, or polygons within a given neighborhood, which may be defined by a certain distance, will be estimated by spatial statistics. In other words, when these similarities and dissimilarities are summarized for an entire spatial pattern, essentially the magnitude of spatial autocorrelation or spatial dependency essentially is measured (Odland, 1988).

Spatial dependency is based on Waldo Tobler's "First Law of Geography," whereby "everything is related to everything else, but closer things are more related". It can be assessed by spatial autocorrelation. To analyze spatial dependency, Moran's I can be applied. The method quantifies the extent to which similar and dissimilar geographic features are clustered (Mitchell, 2005).

In doing this conceptual methodology, polygon pattern analysis will be conducted for spatial dependency and heterogeneity analyses. To be specific, Global and local measures of spatial association will be identified. For instance, Global measures are used to test general patterns in data and identify statistically significant patterns of high (hot spots) or low (cold spots) attribute values or outliers within the study area. The local measures of spatial association (Local Moran's I or LISA (Local Area Indicators of Spatial Association)) are well known types of local measures of spatial association (Hwang, 2008; Mitchell, 2005). In addition, spatial heterogeneity, known as spatial structure, non-stationarity, will be conducted referring to differences in the mean, variance, and covariance structures, including spatial auto correlation within a spatial region (LeSage, 1999). In the following chapter, these methodologies will be introduced briefly.

2) Global Moran's I and Geary's C (Spatial Autocorrelation)

Spatial autocorrelation, known as spatial dependence, spatial interaction, or local interaction can be loosely defined as a similarity or dissimilarity measure between two values of an attribute that are near by spatially (Anseline, 1995). It can be measured by various indexes, of which probably the most well-known is Moran's I statistic (Moran 1948). The Global Moran's I function calculates a Z score value that indicates whether the clustering or dispersion could be the result of random chance or is statistically significant. To determine if the Z score is statistically significant, it is compared to the range of values for a particustatconfidence level. For example, at a significance level of 0.01, a Z score would have to exceed ±2.58 standarscoeviations to be statistically significant. Tandaalso helpdadetermine whether to reject the null hypothesnda(n.e., obesnty, overweight and neighborhood characteristics are evenly distributed across the study area) for spatial pattern analysis (Mitchell, 2005). When interpreting, 'high-high' indicates positive spatial autocorrelation, 'low-high' indicates negative spatial autocorrelation, 'low-low' indicates positive spatial autocorrelation, 'high-low' indicates negative spatial autocorrelation, and 'not significant' indicates that there is no spatial autocorrelation.

Geary's C also is a measure of spatial autocorrelation, and the method can define hot or cold spots in study area. Both Geary's C and Moran's I determine attribute similarity, albeit in different ways. While Geary's C is based on the difference between attribute values, Moran's I is based on deviations from the mean.

3) OLS and GWR (Spatial Heterogeneity)

While spatial autocorrelation is in-line with

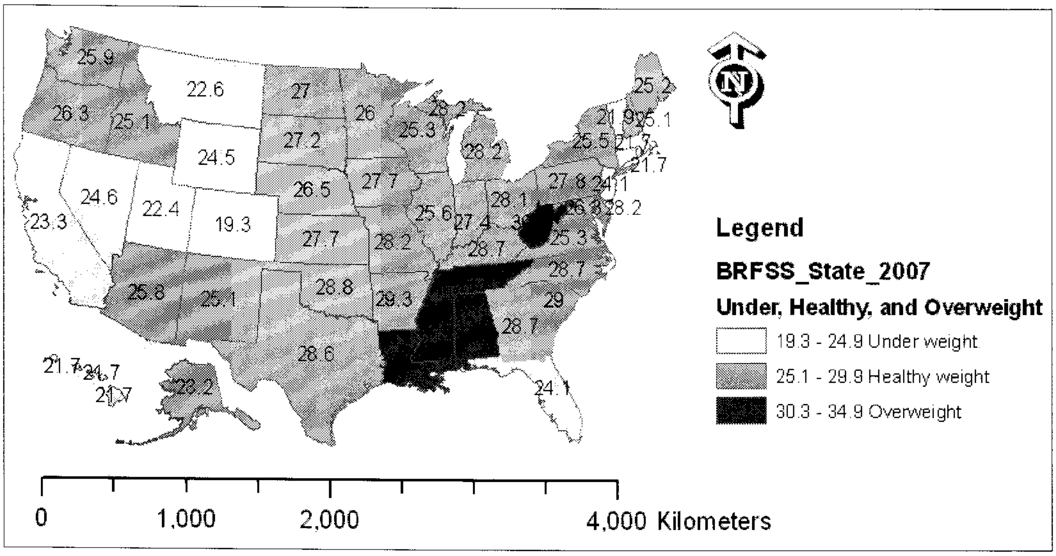


Fig 1. Under, Healthy and Overweight rate (BMI)

the First Law of Geography (Tobler 1970), spatial heterogeneity is related to spatial differentiation (Anselin, 1996). In the second crucial analysis, we will use Ordinary Least Squares (OLS) linear regression and Geographically Weighted Regression (GWR) to analyze spatial differentiation between dependent and independent variances over space.

OLS performs global Ordinary Least Squares linear regression to generate predictions to model a dependent variable in terms of its relationships to a set of explanatory variables. The regression analysis often begins with exploratory data analysis, identifying spatial clusters and spatial outliers, and diagnosing possible misspecification of spatial aspects of the statistical models, all of which can help better specify regression models (Anselin, 2002).

Geographically Weighted Regression (GWR) is a local spatial statistical technique for exploring spatial non-stationarity, defined as when the measurement of relationships among variables differs from location to location (Fotheringham *et al.*, 1998).

4) Spatial Weight Matrix

To account for spatial autocorrelation in lattice data analysis, it is necessary to establish a neighborhood structure for each location by specifying those locations on the lattice that are considered its neighbors (Anselin, 1988). Therefore, a Spatial Weight Matrix (SWM) corresponding to the neighborhood structure should be created. To do this, in this research, we will use "queen's case" contiguity weight metrics of first order. Queen's case has diagonal direction.

4. Quantifying Spatial Patterns (Spatial Autocorrelation)

1) Geographical Mapping for Overweight and Obesity Rates

As noted, BRFSS data for underweight and obesity had two columns: the first class includes a BMI of 19.3 to 24.9%, which contains a range of under, healthy, and overweight.

As shown in 〈figure 1〉, the State of Georgia's range in Class I is 28.7%, and it ranks 19th out of all states in the United States America. The

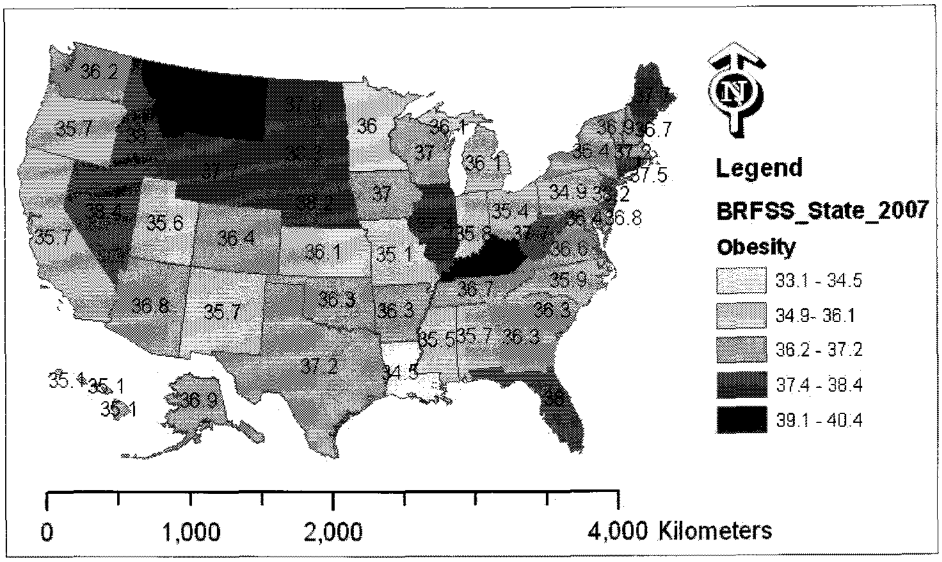


Fig 2. Obesity rates from 33 to 40.4 (BMI)

value's age used a population over 18 years. As seen above, Georgia is within the range of healthy weight in Class I of BRFSS data. In fact, according to the World Health Organization, BMI has 5 classes, as noted in chapter 2.2. Underweight is 18.5 to 24.9, healthy weight is 25.0–29.9, overweight and obesity ranges from 30 to over 40. However, unfortunately, BRFSS has only two columns (Class I and Class II). Hence, this study will not consider specific classes, like WHO BMI classes. Thus, only two classes will be considered in this research. As shown in ⟨Figure 2⟩, Class II has a range of 33.1 to 40.4, containing BMIs considered to identify overweight and obesity. The State of Georgia has a rate of 36.3%.

2) Spatial Autocorrelation (Global and Local Measures)

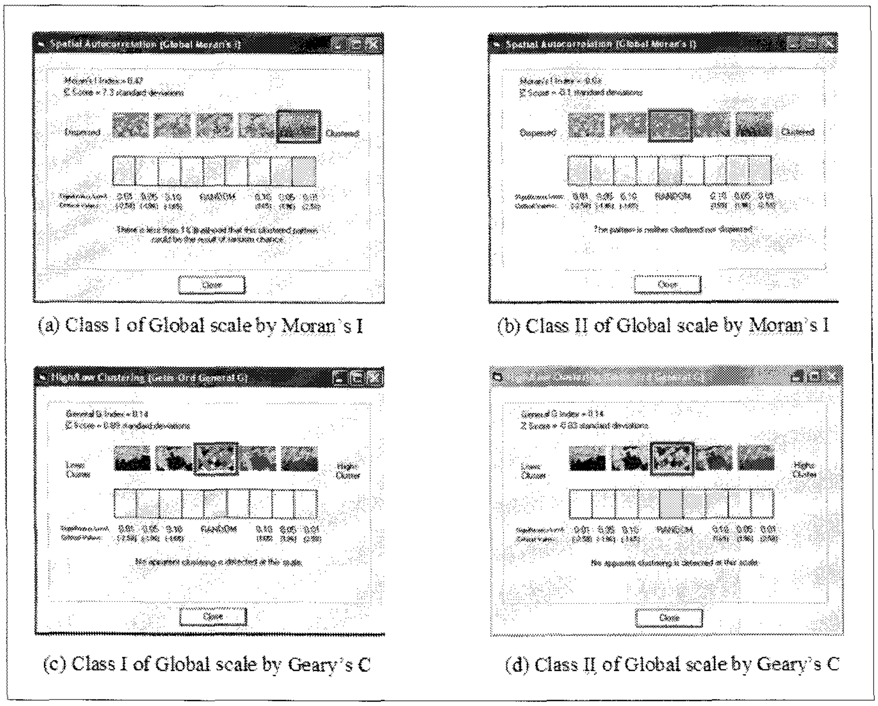
In this step, spatial dependency between under, healthy, over, and obesity weight will be conducted by spatial autocorrelation. Spatial dependency finds whether attribute values are correlated or not over space. Also, spatial autocorrelation describes how an attribute is distributed over space. To do this, Moran's I and Geary's C ratio is considered for global measure.

As shown in ⟨figure 3⟩, Class I has been computed by Moran's I and Geary's C. Class I has a range of 19.3 to 34.9. As noted in ⟨table 1⟩ below, Class I has a positive spatial autocorrelation of 0.47. This means that Class I has similarity and significance. In other words, positive spatial autocorrelation indicates clustering of high values. Class II has a negative spatial autocorrelation of −0.03, which means that distribution of the obese population shows random patterns.

In the Geary's C analysis, even if Class I and

Table 2. Spatial Autocorrelation of Class I and Class II in United States in 2007.

	Moran's I	P-value	Geary's C	P-value
Class I (Under, Healthy, and Overweight)	0.47	7.3	0.14	0.89
Class II (Obesity)	−0.03	−0.1	0.14	−0.03



(a) Class I of Global scale by Moran's I

(b) Class II of Global scale by Moran's I

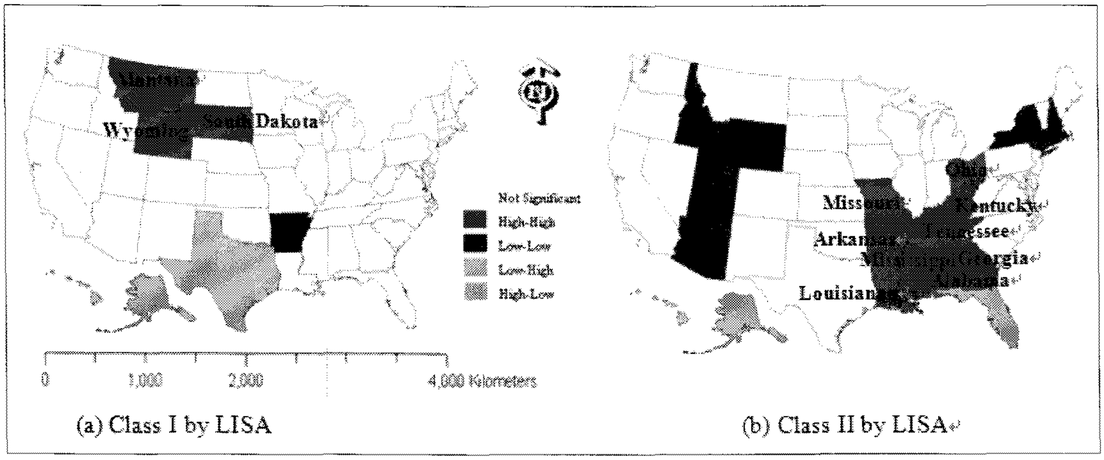
(c) Class I of Global scale by Geary's C

(d) Class II of Global scale by Geary's C

Fig 3. Measures of Spatial Autocorrelation (Geary's C and Moran's I)

Class II are represented as positive spatial autocorrelation as shown in ⟨table 2⟩, its distribution, which is based on measures of spatial autocorrelation shows random patterns.

In the next analysis, Local Indicators of Spatial Association (LISA) are conducted for local measures. While global measures of SA calculate a single statistic that summarizes the pattern for the study area, local measures of SA calculate a statistic for each feature. In analytical approaches, 'High-High' indicates clustering of high values of Class I and Class II rates, which mean positive spatial autocorrelation. 'Low-High' indicates that low values are adjacent to high



(a) Class I by LISA

(b) Class II by LISA

Fig 4. LISA Measures of Spatial Autocorrelation (Moran's I).

values of Class I and Class II rates, which means negative spatial autocorrelation. 'Low-Low' indicates clustering of low values of Class I and Class II rates, which mean positive spatial autocorrelation. 'High-Low' indicates that high values are adjacent to low values of Class I and Class II rates, which mean negative spatial autocorrelation. 'Not significant' indicates that there is no spatial autocorrelation.

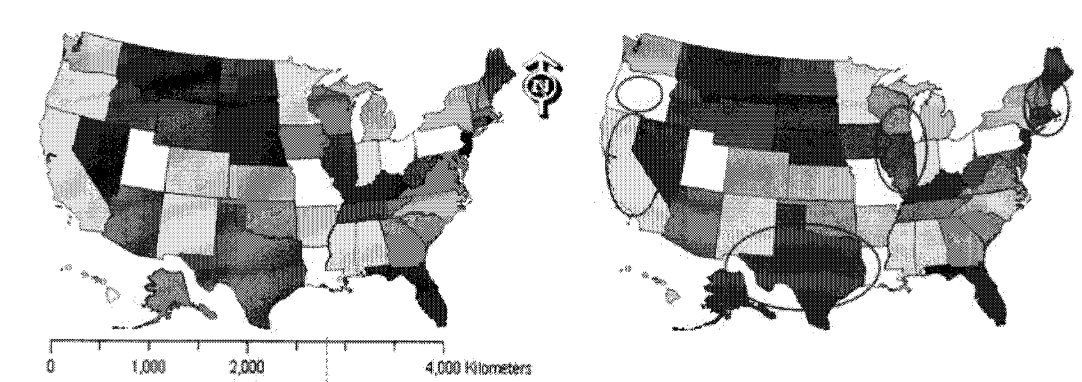
As shown in ⟨figure 4⟩, Class I shows that Montana, South Dakota, and Wyoming in the Midwestern/ the Western region of the United States has High-High, which means high significance. Class II also shows that Ohio, Kentucky, Missouri, Arkansas, Tennessee, Louisiana, Mississippi, Alabama, and Georgia in the Southeastern region have clustering with much significance. This means that all variables that are dependent and independent values are spatially and locally related to a high degree.

To analyze LISA, a spatial weight matrix based on the queen's method with first order is used for weighted values. Finally, in the Global and Local measures, these results indicate that under, healthy, overweight and obesity suggests no-randomness in the overall spatial pattern.

3) Ordinary Least Squares (OLS) Linear Regression Analysis

Diabetes is one of the most troublesome aspects because of the relationship between diabetes and obesity. High blood pressure also has a strong relationship with obesity. So, in this step, an Ordinary Least Squares (OLS) linear regression model is considered to analyze spatial pattern. To do this, Class II, indicating only obesity rates, is used as a dependent variable, and the rates of high blood pressure and diabetes are considered as independent values.

As shown in ⟨Figure 5⟩, figure (a) shows a residual map for spatial dependency between high blood pressure and diabetes with obesity by OLS. The analysis is for spatial dependency between high blood pressure and diabetes with obesity. Results for the diagnostic test show that the Akaike info criterion (AIC) is 175.262 and the R-squared value is 0.014413. That is, the dependency has a lower relationship over space, but the relationship between high blood pressure and diabetes has strong possibilities with obesity rate. The result for possibility shows that high blood pressure is 0.954 and diabetes is 0.733 in relation to obesity respectively. Additionally, in



(a) **Residual map** for spatial dependency between high blood pressure and diabetes with obesity by OLS

(b) **Spatial Lag model** for high blood pressure and diabetes with obesity

Fig 5. OLS and Spatial Lag Model

interpretation to improve R-squared and AIC results, spatial lag models are chosen because the spatial lag model has more possibility than the spatial error model. So, the left map of ⟨figure 5⟩ shows the results considering the spatial lag model. These results shown above demonstrate that there are a few differences. As can be shown in the red circles of ⟨figure 6⟩, only some areas have been changed. Also, AIC and R-squared values have slightly been improved or decreased.

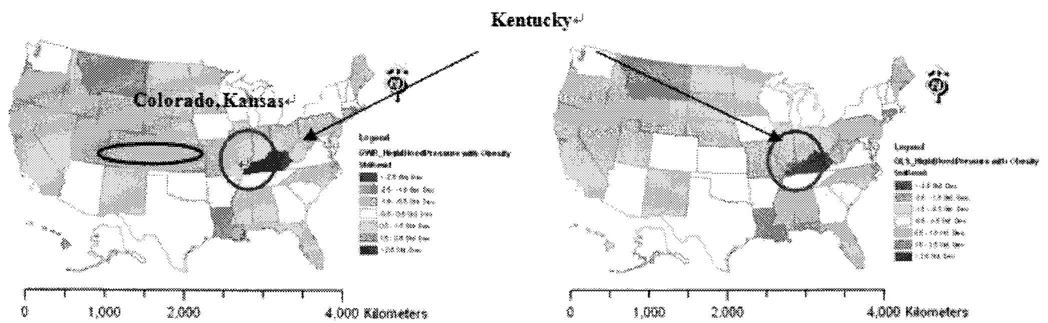
In summary, by OLS analysis, this research's assumption is that obesity would have strong spatial relationships with high blood pressure and diabetes over space in stationarity. As shown in ⟨figure 6⟩, a regression residual map shows spatial dependency over space. Therefore, this assumption has perusable results for spatial dependency between high blood pressure and diabetes with obesity.

5. Comparison for GWR and OLS

While an OLS model can recognize spatial dependency between dependent and independent variables, a Geographically Weighted Regression model can recognize spatial heterogeneity over space. In other words, GWR uses spatial non-stationarity, which means the same In mulus provotis a different response in different parts of the study region. Accordingly, non-stationarity is more realistic.

To be specific, GWR analysis uses different weights over space by changing band width in a fixed kernel. As shown in ⟨figure 6⟩, in GWR analysis, high blood pressure value has been considered as independent, and obesity rates containing Class II are used as dependent values. As noted, OLS and GWR have different characteristics. OLS is sationarity, and GWR is non-stationarity. When comparing OLS with GWR results, the residuals of OLS and GWL are -0.05 and - 0.05 with a z-score of -051, respectively. This has negative autocorrelation. In other words, although high blood pressure and obesity have high possibility, each state has randomness. Also, in comparison, both OLS and GWR have the same Moran's I values and Z-scores. As shown in a blue circle of ⟨figure 6⟩, only two states' (Colorado and Kansas) residual maps have been changed. Even if theoretically the GWR method is more realistic, there is not much difference in the results in this study.

6. Conclusion



(A) OLS analysis for high blood pressure with obesity

(B) GWR analysis for high blood pressure with obesity

Fig 6. Comparisons between GWR and OLS.

1) Summary and Discussion

For many years, overweight and obesity have been studied as important risk factors for cardiovascular and other chronic diseases with major implications for individuals and populations (National Institute of Health, 2004). Conditions of overweight/obesity increase the risk of diabetes and heart disease, contribute to greater risk for hypertension and insulin resistance as well as impact the psychosocial state and quality of life of individuals and families (Raine, 2004). Nowadays, The United States has a crisis of overweight and obesity. So, we have to find new plans to control this crisis. Secondly, even if we recognize this crisis, it is problematic that only a few studies have been performed to analyze thatial dependency and heterogeneity ostudies h. Also, BRFSS providedave the CDCdituves only general information on the webdiite. So, intellectual mapping shouldave projected for publicould. In response, the results from this study can explain shatial hattl ds that suggestuggesrandomness in the ostuall shatial hattl d. Shis crially, in global measuremn on obesity has randomness, buon undero, ealthyn the wverweight have sign criance or ggesrandomness, whrih means similnic diostudies h. In LISAes h. Also,l mappi (un ealthyn ealthy, and overweight) has a slightly higher Moran's I (0.486) value than the Moran's I (0.47) value in Class II.

In OLS analysis for spatial dependency between high blood pressure and diabetes with obesity, the results show that the possibility of high blood pressure is 0.954 and diabetes is 0.733. In addition to this result, a Lag Model has been chosen as the best way to improve R-squared or AIC, but only R-squared has slightly been increased.

In the last analysis of a comparison between GWL and OLS, this comparison is to recognize spatial dependency and spatial differentiation over space for high blood pressure with obesity. The result describes that there is no similarity in the United States for spatial relationship, but two variables have strong possibility. In addition, unfortunately, in the comparison between OLS and GWR, we couldn't find much difference or improvement.

All things considered, this research is beneficial for predicting spatial patterns related to public health. It is worthwhile work to understand geographical trends for obesity and overweight with high risk factors that can cause developing complications. In other words, spatial relationship over space can be displayed by spatial autocorrelation methods in global and local measures. Although this study has some limited conditions, such as data availability, we are confident that the research can help decision makers to propose plans for controlling and preventing obesity.

2) Limitation and Future Study

There are two issues that need to be addressed. The first one is data availability. As revealed in this research, the CDC provides two GIS shape files with a wealth of health information. The data come from the computer aid telephone survey. The data in the attribute tables have a slightly different format from that of the original survey questionnaire. So, the data should be examined and refined before running. Also, the data have a limited number of samples and those were only generated by available sites, which responded to the computer aid telephone survey. That is, the data cannot be representative for each state. In the near future, data that can be representative are required for spatial pattern analysis in each state with constant conditions.

Secondly, the Modifiable Aerial Unit Problem (MAUP) occurs when the results of statistical analysis are highly influenced by the scale as well as the shape of the aggregation. More specifically, the scale effect refers to the fact that, when the same data are aggregated at different scales, the results of statistical analysis are disparate over scale. Therefore, the smallest units of analysis available are needed.

In future studies, other independent values can be used for identifying spatial pattern by spatial autocorrelation. For instance, a physical activity, consumption of fruit and vegetables, and household income data can be considered because physical inactivity is a major risk factor for cardiovascular disease, stroke, hypertension, diabetes, and obesity. In contrast, physical activity can decrease rates of overweight and obesity. Also, consumption of fruit and vegetables can decrease rates of obesity. So, more various and specific data will be analyzed in the near future for identifying spatial dependency and heterogeneity over space.

REFERENCE

Anselin, L., 1988, *Spatial Econometrics: Methods and Models*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Anselin, L., 1994, Exploratory spatial data analysis and geographic information systems, in Longley, P., Brooks, S., Macmillan, B., McDonnell, R. (Eds.), *GeoComputation: A Primer*, Willey, New York, 77-94.

Anselin, L., 1995, Local indicators of spatial association — LISA, *Geographical Analysis*, 27, 93-115.

Anselin, L., 1996, The Moran scatterplot as an ESDA tool to assess local instability in spatial association, in M. Fischer, H. J. Scholten, & D. Unwin (Eds.), *Spatial Analytical Perspectives on GIS*, London, England: Taylor & Francis, 111-125.

Anselin, L., 2002, Under the hood: issues in the specification and interpretation of spatial regression models, *Agricultural Economics*, 27, 247-267.

Bailey, T.C. and Gatrell, A.C., 1995, *Interactive Spatial Data Analysis*, Addison Wesley Longman Limited, Malaysia.

Fotheringham, A. S., Brunsdon, C. and Charlton, M. E. ,1998, Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis, *Environment and Planning A*, 30, 1905-1927.

Gatrell, T., and Loytonen, M., 1998, *GIS and Health*, London: Taylor & Francis.

Hwang, C.S., 2008, The application of spatial statistics for land-use and land-cover change followed by substantial development, *The Geographical Journal of Korea*, 42(4), 647-658.

Lee, J. and Wong, D.W.S., 2001, *Statistical Analysis with ArcView GIS*, John Wiley & Sons, Inc.

LeSage, J. P., 1999, A spatial econometric examination of China's economic growth, *Geographic Information Sciences*, 5, 143-153.

Macera, C.A, Ham, S.A., Yore, M.M., Jones, D.A., Ainsworth, B.E., Kimsey, C.D., and Kohl III, H.W.. 2005, Prevalence of physical activity in the United States: behavioral risk factor surveillance system, preventing chronic disease, *Public Health Research, Practice, and Policy*, 2(2), 1-10.

Maddock, J., 2004, The relationship between obesity and the prevalence of fast food restaurants: state-level analysis, *American Journal of Health Promotion*, 19(2), 137-143.

Mitchell, A., 2005, *The ESRI Guide to GIS Analysis Volume 2: Spatial Measurements and Statistics*, ESRI Press, Redlands, California.

Moran, P., 1948, The interpolation of statistical maps, *Journal of the Royal Statistical Society B*, 10, 243-251.

National Institute of Health, 2004, *A Report of the NIH Obesity Research Task Force*, U.S. Department of Health and Human Services, Bethesda, MD.

Odland, J., 1988, *Spatial Autocorrelation*, Newbury Park, CA: Sage Publications.

Raine, K.D., 2004, Overweight and obesity in Canada: A population health perspective, *Canadian Population Health Initiative and Canadian Institute for Health Information*, Ottawa ON, Canada.

Tobler, W., 1970, A computer movie simulating urban growth in the Detroit region, *Economic Geography*, 46, 234-240.

Walsh, S.J., Page, P.H. and Gesler, W.M., 1997, Normative models and healthcare planning: network-based simulations within a geographic information system environment. *Health Services Research*, 32, 243-60.

World Health Organization, 2004, *The World Health Report 2004: Changing History*, WHO.

Correspondence : Chulsue Hwang, Department of Geography, Kyung Hee University, #1, Hoegi-dong, Dongdaemun-gu, Seoul, 130-701, Korea (e-mail: hcs@khu.ac.kr, phone: +82-2-961-9313)

• **교신** : 황철수, 130-701, 서울특별시 동대문구 회기동 1번지, 경희대학교 이과대학 지리학과 (이메일: hcs@khu.ac.kr, 전화: 02-961-9313)

(접수: 2010.6.28, 수정: 2010.7.26, 채택: 2010.9.3)