MULTIPLE SCLEROSIS DEATH RATES and County-level Contextual Characteristics: A New England Case Study

RJ Malavasi[1]
Sr. Risk Analyst at The Hartford Insurance Group

Yunliang Meng[2]
Central Connecticut State University

ABSTRACT

There is growing interest in the role that environmental, demographic, and socio-economic factors play in the prevalence of the disease Multiple Sclerosis (MS) within the United States. Existing empirical evidence examining the spatial association between the environment and MS death rates, however, remains ambiguous. The objective of this research is to examine the relationship between MS death rates and contextual characteristics at a county level in the New England region. The analysis shows that predictors such as available sunlight, race, and access to healthcare are associated with MS death rates in the region. Most importantly, the association between MS death rates and all major explanatory variables in our analysis significantly varied over space, highlighting the need for local and context-specific MS prevention and intervention programs.

Key words: Multiple Sclerosis, New England, Geographically Weighted Regression

Introduction

Multiple Sclerosis (MS) is the most common inflammatory disorder of the central nervous system and a leading cause of disability in young adults (Milo and Kahana 2010). Most researchers suggest the disease to be autoimmune, but others categorize the disease as an infectious-viral disease or a neurodegenerative disease (Milo and Kahana 2010). Common symptoms of MS patients include paresthesias or numbness, motor weakness, monocular visual disturbances, incoordination, diplopia, dizziness and vertigo. To maintain functioning and alleviate symptoms, MS patients must undergo lifelong treatment. However, the disease can still lead to disability and in many cases, early death. Mortality in patients with MS is significantly increased compared with the general population. The most common cause of death for MS patients is multiple sclerosis itself, or complications related to the disease. For example, Koch-

Henriksen et al.'s (1998) research findings suggest that among the 8,142 MS Danish patients documented within the period 1951-1993, 2,803 (or 34.4 percent) died from multiple sclerosis and/or its complications.

Today, MS affects approximately 2.0 to 2.5 million people around the globe and with prevalence rates of less than five cases per 100,000 within Latin America and Asia and up to 200 cases per 100,000 people in Europe and Northern America (Milo and Kahana 2010). While a wide range of symptoms and types of MS are known to the medical realm, the exact causes of MS are still unknown. The disease is so vastly complex that the cause is not correlated directly to one instance or exposure, but seems to have intertwined relations to many different risk factors with varying intensities. There are many factors that may be the causes of the disease, but several demographic (i.e. age, race, and sex), socio-economic (i.e. access to healthcare) and environmental factors (i.e. available sunlight and latitude) have become essential aspects of what may cause or relate to the disease (Minden et al. 2007; Conradi et al. 2011; Kakalacheva and Lunemann 2011; Bove and Chitnis 2014).

Public health research has shown that MS is not spatially evenly distributed, but instead tends to be geographically patterned. For example, results from Mayer's (1981) study suggests that high latitude areas are particularly notable for clusters of high MS prevalence due to the lack of exposure to sunlight. Local demographic, socio-economic, and environmental correlates of MS prevalence also differ across geographic space. Given this evidence, empirical research is needed to investigate the spatial difference in MS death rates and identify contextual characteristics that underlie existing spatial differences. Therefore, the objective of this research is to examine the relationship between MS death rates and contextual characteristics at county level in the New England region.

This article is organized as follows. Section Two provides a review of risk factors associated with the MS disease. Following this, Section Three presents the methodology outlining a description of the study area, variables used in this study and the local analysis method – Geographically Weighted Regression (GWR). In Section Four, the results of a local analysis using GWR are discussed. The final section presents concluding remarks.

Risk Factors Associated with MS Disease

Available Sunlight

Sunlight is an important protective factor. Different research suggests that receiving a good amount of sunlight throughout life, which leads to preferable vitamin D levels, can prevent or defer MS (Freedman et al. 2000; Kakalacheva and Lunemann 2011; Bäärnhielm et al. 2012). Low exposure to sunlight can increase the risk of MS (Van der Mei et al. 2003; O'Gorman et al. 2012). A study conducted by Mayer (1981) takes into consideration the incidence rates (number of cases per 10,000 people) of Multiple Sclerosis by country while also taking into consideration the approximate latitude of each country. The countries typically with higher

MS rates are those that are higher latitudinally or ones located further away from the equator. It should be noted that available sunlight hours per day and latitude have an inverse linear correlation (Kakalacheva and Lunemann 2011). Therefore, the amount of available sunlight holds answers or hints towards the reasons why MS disease is more likely to occur in higher latitudes.

Age

MS typically presents itself in young adults between twenty and fifty years of age, with a peak occurrence at thirty years of age according to the National Multiple Sclerosis Society, although MS can also occur in children as well as in seniors (Compston et al. 2006). Some researchers find that young age may be a protective factor. For example, Conradi et al. (2011) discover that living in an urban area while of the age of zero to six may be a protective factor against MS. Age is also a factor that differentiates the development of the disease. Harding et al. (2013) noticed that children diagnosed with MS took longer to develop secondary-progressive disease (thirty-two years) than adults diagnosed with MS (eighteen years), and they also took longer to reach disability milestones.

Gender

Gender is a risk factor that has its roots in the development of the MS disease. The National MS Society reported in the 1940s that the female to male ratio was around 1:1 and was on the rise in females. Recent empirical studies show that the increased prevalence in females coincides with the conclusion from the National MS Society. Research done by Milo and Kahana (2010) indicates that the ratio of MS prevalence is 2:1 in females to males. Bove and Chitnis (2014) discovered that the ratio is approaching and even exceeding 3:1 in some regions, mostly located in western societies.

Race

Race is another important factor. Although MS is present in the three principal racial groups in the world (White, Asian-Mongol, and Black), it tends to be unequally distributed. Empirical research has pointed towards people of White European descent having the highest risk of Multiple Sclerosis (Kingwell et al. 2013), but there are still other races and ethnicities that suffer from MS as well. For example, recent research suggests that minority populations in the United States, such as African Americans and Hispanic Americans, have a higher incidence of MS compared with their ancestral countries of origin (Amezcua et al. 2015). Another study suggests that the prevalence of the disease, especially the classical, Western form, is on the rise in northeast Asia (Kira 2006).

Access to Healthcare

Access to healthcare is an essential contributing factor to MS risk. Limited access to quality health care can have far reaching effects on the physical, social, psychological, and economic well-being of people with chronic disabling conditions such as MS (Neri and Kroll 2003). As a key for access to healthcare, health insurance allows people to have access to routine and unforeseen medical care and protects them against the high cost of medical bills. People without health insurance are less likely to visit doctors, have usual sources of care, receive preventive services, and undergo recommended tests or fill prescriptions. They also have high out-of-pocket medical expenditures which may stop them from receiving future treatments. Therefore, the MS risk associated with people without health insurance is higher, since they are less likely to be diagnosed as MS patients in time and more likely to miss recommended tests or fill prescriptions and even undertake delayed treatment (Minden et al. 2007; Wang et al. 2016).

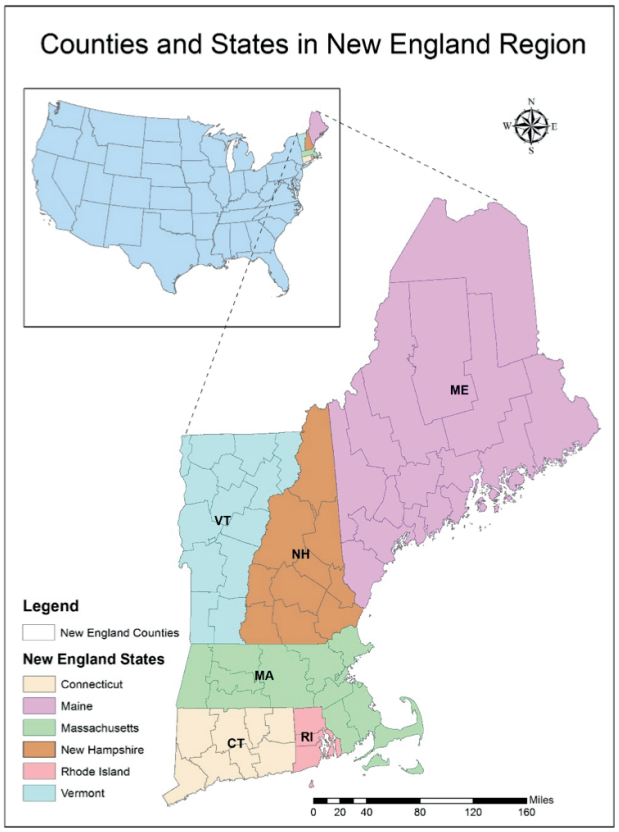

Figure 1. Study Area.

Methodology

Study Area

The study area (see Figure 1) consists of the region of New England within the United States of America. New England is comprised of six states that are Maine, Massachusetts, New Hampshire, Vermont, Rhode Island, and Connecticut which consist of a total of sixty-five counties. With a population of over fourteen million, the region has a fairly large concentration of people that have a wide and varied distribution within its boundaries. The New England region is an important study area for MS research, since its unique geographic location – a middle latitude region (between 40°59'N and 47°28'N), and an increased prevalence of MS has been documented in the region recently (Murray 2005).

Data

The MS death data used in this study was acquired through the Center for Disease Control and Prevention (CDC) WONDER database (https://wonder.cdc.gov/). However, we encountered the problem of suppressed data, because the CDC does not report all sub-national data if the MS death counts fall below a pre-determined cut-off value determined by CDC. In this study, ten counties or 14.9 percent of the total sixty-seven counties in New England have no reported MS death data because the CDC intentionally deleted the data before disseminating them. The MS death data consisted of 2,687 reported MS death incidents in New England Region during 1999 and 2014 and the MS death rate in a county was calculated by using the MS death number divided by total population in the county and then multiplying by 100,000. The descriptive statistics for the dependent variable – MS death rates is shown in Table 1 below.

	Min	Max	Average	Interquartile Range	Standard Deviation
MS death rates: MS death per 100,000 people	0.8	2.6	1.65	0.6	0.43

Table 1. Descriptive Statistics for the Dependent Variables – MS Death Rates.

In this study, the data used to model the death rates of MS are demographic and socio-economic variables together with an environmental factor – available sunlight. Specifically, the demographic variables such as overall population, age groups, gender, and race were taken from the 2010 Census provided by U.S. Census Bureau through the American FactFinder website (https://factfinder.census.gov/). The age variable was determined using the number of people aged between twenty and fifty divided by the total population in each county and then multiplying the result by 100, since this age group has a higher MS risk. The gender variable was measured using the total female population divided by the total population in each county from the census and then multiplying the result by 100, because the female population has a higher risk for MS. The race variable was quantified using the number of non-Hispanic White population divided by the total population in each county and then multiplying the result by 100, since non-Hispanic White people are at greater risk for the disease. The socio-economic variable – access to healthcare was determined by using the uninsured population data reported through the recent five-year estimates (2009-2013) from the US Census Bureau's American Community Survey (ACS). The uninsured population data was used because it is a data set collected over sixty months to ensure that it is the most reliable data and largest sample size for the region of New England. The access to healthcare variable was measured using the number of uninsured population divided by the total population in each county and then multiplying the result by 100. When it comes to the available sunlight variable, it was quantified by the yearly average sunlight based on records from January 1979 to December 2011. This data was acquired through the Center for Disease Control and Prevention (CDC) WONDER database, but was gathered by the North America Land Data Assimilation System (NLDAS) (Daily Sunlight

insolation for years 1979-2011 on CDC). All variables were collected and measured at the county level and ArcMap 10.3.1 (ESRI 2015) was used to join the data variable with the New England counties shapefile which was downloaded from US Census Bureau's website (https://www.census.gov/en.html). The descriptive statistics for each explanatory variable are shown in Table 2 as below.

Variables	Min	Max	Average	Interquartile Range	Standard Deviation
Age: % of people aged between 20 and 50	19.2	26.1	22.9	1.7	1.4
Gender: % of female population	45.2	55.2	49.5	3.1	2.2
Race: % of non-Hispanic White population	54.8	99.5	89.6	8.6	7.7
Access to healthcare: % of uninsured population	2.5	15.2	8.4	5.0	3.2
Available sunlight: daily sunlight insolation in KJ/m^2	13,133.0	15,731.8	14,306.0	916.3	679.4

Table 2. Descriptive Statistics for the Explanatory Variables.

GWR Model Building

Many early contextual studies on disease patterns are criticized by their use of a conventional global regression modeling technique, such as Ordinary Least Squares or OLS regression (Moore and Carpenter 1999), because the technique violates some important assumptions (e.g., independence of observations and spatial stationarity of the relationship between independent and dependent variables) when spatial data are used in the studies. GWR (Brunsdon et al. 1996; Fotheringham et al. 2002) relaxes these assumptions and enables the analysis of spatially clustered data. GWR is often considered as an extension of OLS regression, since it allows local instead of global parameters to be estimated, hence making it possible to model spatial variations within the data (Fotheringham et al. 2002). Unlike OLS regression, which produces a single global model across space, GWR simultaneously conducts multiple regressions for different data units so that there is one regression model per spatial data unit (e.g. a county) (Hipp and Chalise 2015). In a GWR model, observations near a particular data unit will have more influence in the estimation than observations distance away (Hipp and Chalise 2015). Given the weaknesses of OLS regression and strength of GWR, this research uses GWR for analyzing the spatial non-stationarity relationship between MS death rates and county contextual characteristics in the New England region.

The first step is to examine the dependent variable and explore its spatial heterogeneity. If the MS death rates are not spatially clustered, there is no need to build a spatially explicit model. The Moran's *I* Index (Anselin 1995) provided by ArcMap 10.3.1 (ESRI 2015) was used to identify the clustering of MS death rates across counties in the New England Region. Moran's *I* ranges from −1.0, perfectly dispersed (e.g., a checkerboard pattern), to a +1.0, perfectly clustered. In this study, a Moran's *I* score (0.202) and *p* value (0.0049) were generated, indicating

MS death rates are spatially clustered and the result is statistically significant. A Local Moran's *I* Cluster Analysis of MS death rates was conducted, and the results are shown in Figure 2 below. The map demonstrates five different types of spatial clustering: (1) high-high, for counties with high MS death rates that are in close proximity to counties with high MS death rates; (2) low-low, for counties with low MS death rates that are in close proximity to counties with low rates; (3) high-low (known as spatial outliers), counties with high MS death rates, but are proximate to counties with low rates; (4) low-high (also known as spatial outliers), for counties with low MS death rates, yet are in close in proximity to counties with high rates; (5) not significant, for counties where there is no significant spatial clustering. As illustrated in Figure 2, a low-low spatial cluster was found in most counties located in Connecticut and those within Greater Boston region; while a high-high spatial cluster is overlapped with counties located in the Southwestern Maine and Northern New Hampshire and Vermont.

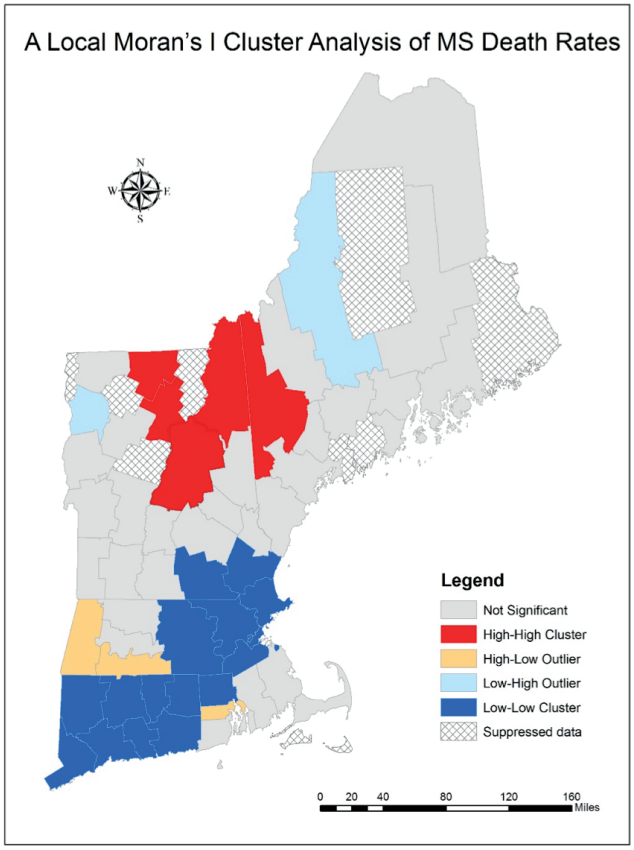

Figure 2. A Local Moran's *I* Cluster Analysis of MS Death Rates.

OLS multivariate model (Aiken and West 1991) in SPSS 22 was then used to conduct initial data exploration and model specification. Two factors motivated the decision to first specify the OLS model: 1) to identify independent variables significantly correlated with the dependent variable (MS death rates) before specifying the GWR model; and 2) the GWR software used for spatial analysis does not provide a variance inflation factor (VIF) to assess multicollinearity. If the standard regression equation in the investigation of MS death rates is given by:

$$Y_i = \beta_0 + \sum_k \beta_k x_{ki} + \varepsilon_i$$

where Y_i is the MS death rate at county i, β_0 is a constant term (i.e., the intercept), and β_k measures the relationship between the independent variable x_k and Y for the set of i counties, and ε_i is the error associated with county i. It should be noted that $i \in C = \{1,2,...,n\}$ which is the index set of locations of n observations (i.e. all counties in the New England Region).

It should be noted that the above equation "results in one parameter estimate for each variable included" (Cahill and Mulligan 2007). The summary of the OLS analysis result is presented in Table 3 as below. In the OLS regression, included were only variables significantly correlated with the dependent variable – MS death rates. The OLS model is significant ($F = 3.984$, $p < 0.05$). The adjusted R^2 value is 0.21 which means that the model explained 21.0 percent of the variance in county-level MS death rates. The VIF for all variables was less than 4.0, a commonly used cutoff point, suggesting no multicollinearity was detected among the independent variables (Table 3).

Variables	β	SE	p value	VIF
Intercept	2.555	1.886	> 0.05	-----
Available sunlight: yearly average sunlight	-1.7×10^{-4}	0	< 0.05	1.386
Race: % of White Population	1.642	0.813	< 0.05	1.234
Access to healthcare: % of Uninsured Population	0.618	2.086	< 0.05	1.161

Table 3. Results from Ordinary Least Square Model of County-Level MS Death Rates.

As shown in Table 3, there is a positive and significant relationship between MS death rates and the percentage of White population as well as between MS death rates and the percentage of uninsured population. In other words, the higher the percentage of White people or people without insurance coverage in a county, the higher MS death rate in that county. As negative and significant relationship is detected between MS death rates and available sunlight. In other words, the lower the available sunlight in a county, the higher the MS death rate in the county. Age and gender are insignificantly related to MS death rates in this study. The residuals of the OLS model were spatially auto-correlated (Moran's $I = 0.12$; $p < .05$). In other words, the OLS model overestimated MS death rates for some counties, while it underestimated the outcomes for some other counties.

Then, the same set of variables was then used to specify a GWR model using the GWR4 software http://geodacenter.asu.edu/gwr). GWR is a modeling technique used to explore spatial non-stationarity (Brunsdon et al. 1996). The "main characteristic of GWR is that it allows regression coefficients to vary across space, and so the values of the parameters can vary between locations" (Mateu 2010, 453). In other words, instead of estimating a single parameter

for each variable, GWR estimates local parameters. By estimating a parameter for each data location (i.e. county) in the New England Region, the GWR equation would only alter the OLS equation as follows:

$$Y_i = \beta_{0i} + \sum_k \beta_{ki} x_{ki} + \varepsilon_i$$

where Y_i is the MS death rate at county i, β_{0i} is the constant term at county i, x_{ki} is the explanatory variable (i.e. available sunlight, race, or access to healthcare) at county i, β_{ki} is the value of the parameter for the corresponding explanatory variable at county i, and ε_i is the error term at county i. It should be noted that $i \in C = \{1,2,...,n\}$ which is the index set of locations of n observations (i.e. all counties in the New England Region).

GWR becomes useful when "a single global model cannot explain the relationship between some sets of variables" (Brunsdon et al. 1996, 281). In the GWR model, a continuous surface of parameter values is estimated under the assumption that locations nearer to i will have more influence on the estimation of the parameter $\hat{\beta}_i$ for that location. Consequently, GWR allows researchers to explore "spatial non-stationarity by calibrating a multiple regression model which allows different relationships to exist at different geographical locations" (Leung et al. 2000, 9). The GWR model was used to explore the macro-level spatial non-stationarity of the statistical relationship among MS death rates and the predictors including available sunlight, race and access to healthcare.

While conducting GWR, the adaptive kernel was used, which was produced using the bi-square weighting function. The adaptive kernel uses varying spatial areas, but a fixed number of observations for each estimation. It is the most appropriate technique when the distribution of observations varies across space. In this case, observations (counties) are much smaller and closer together in the South and Southeast than they are in the North. Finally, a process that minimizes the Akaike Information Criteria (AIC) was used to determine the best kernel bandwidth. The parameter estimates and t values produced by the software were exported and mapped using ArcMap 10.3.1 (ESRI 2015).

Results and Discussion

A Local Moran's *I* cluster analysis (Anselin 1995) was conducted for the residuals of the GWR as a diagnostic for the collinearity in GWR residuals. There is no violations of residual independence. The GWR model generated β coefficients for each county (See Table 4 and Figure 3 as below). The direction of the relationships among the dependent variable and the predictors was consistent as expected in all counties included in the study and the MS death rates at the individual county-level were significantly clustered (Moran's *I* value: 0.202 and $p < 0.05$). This clustering and the relationships suggest that local socio-economic contexts and environment attributes are associated with MS death rates and that the amplitude of such contexts and environments varies across the New England region.

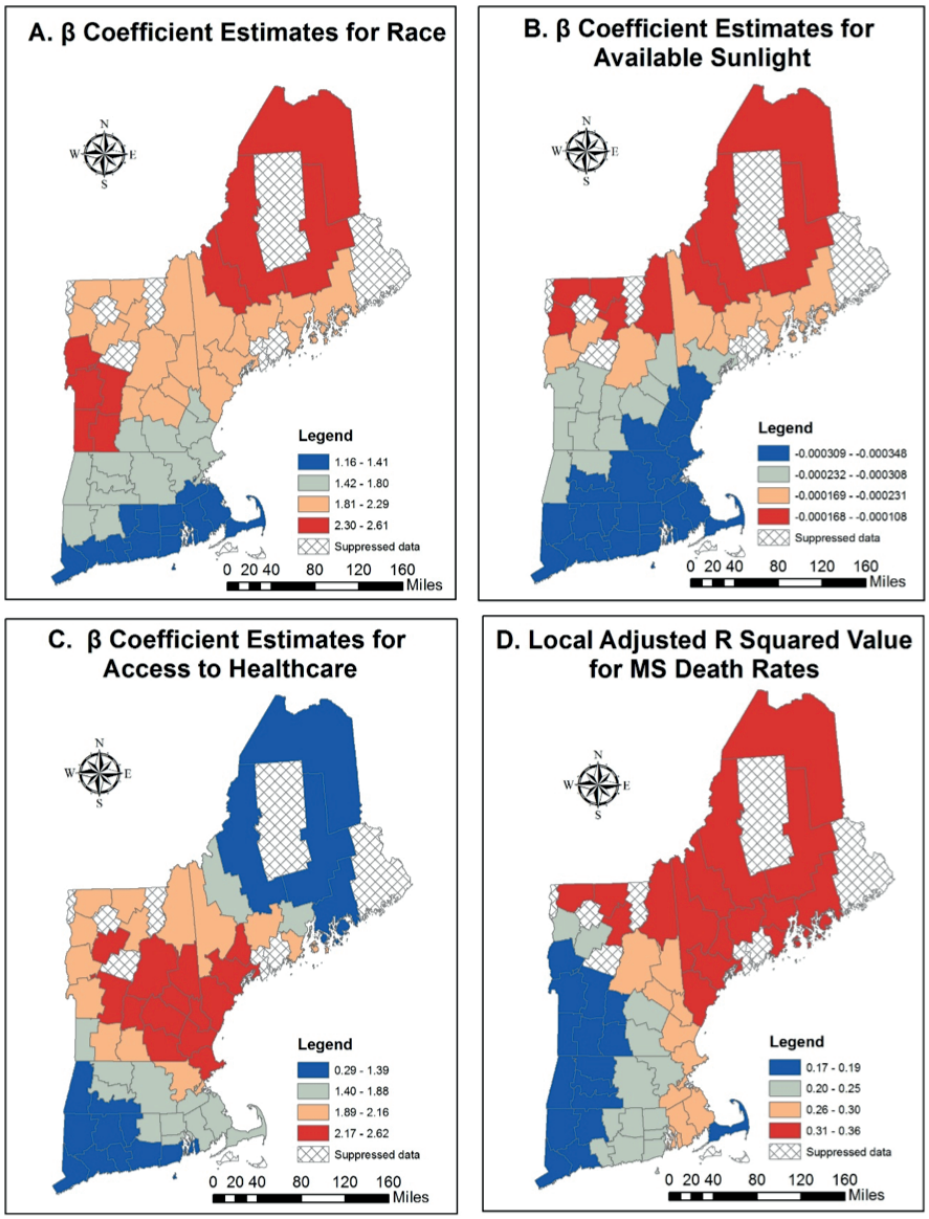

Figure 3. Spatial variations in β coefficient estimates in New England counties for explanatory variables: race (map A), available sunlight (map B), access to healthcare (map C), and local R squared value (maps D) from the GWR model .

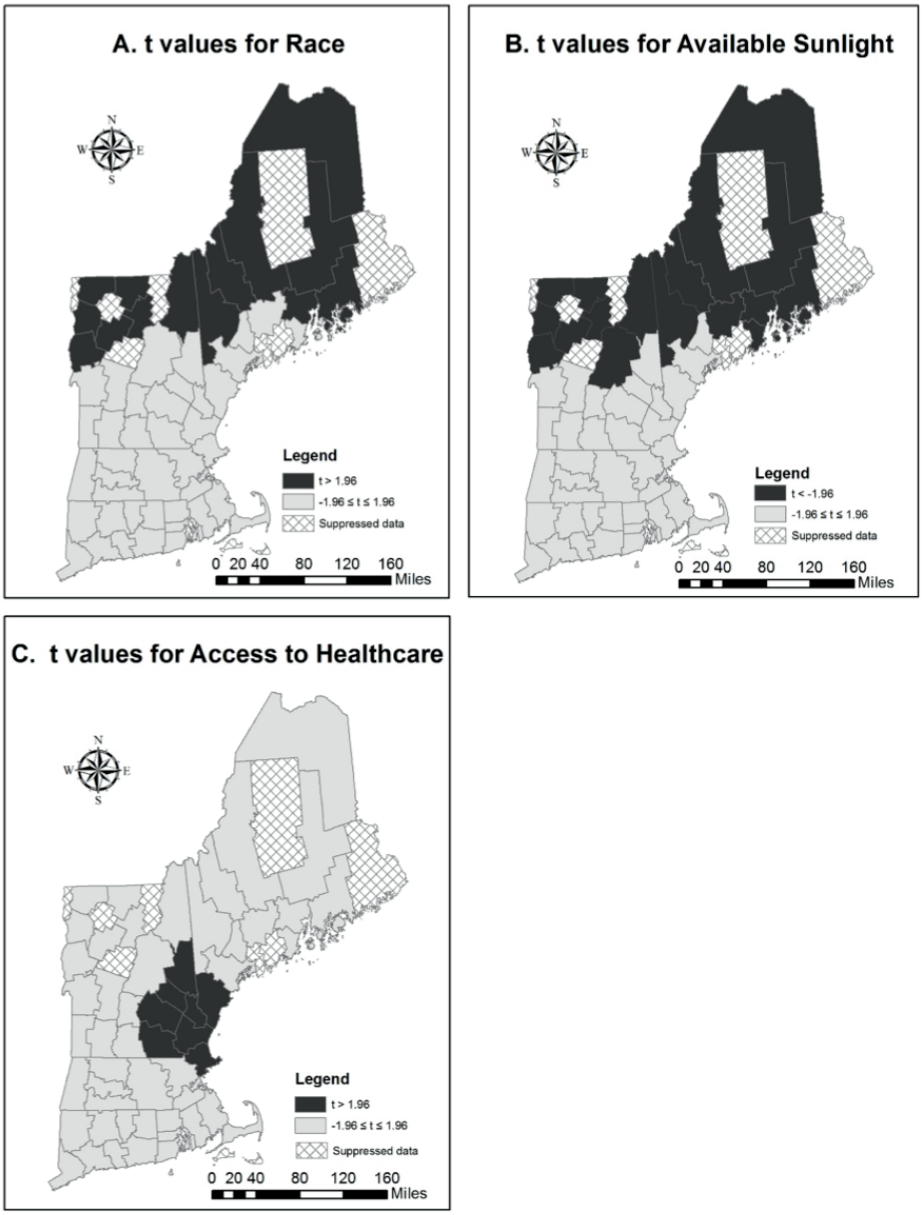

Figure 4. Spatial variations in t values in New England counties for explanatory variables: race (map A), available sunlight (map B), access to healthcare (map C). Note: 1.96 is the cut-off value for t-test. When $|t| > 1.96$, the β coefficient estimate for a variable is significant at a significance level of 0.05.

Variables	β coefficients		Percentage of Counties by 95% of t Statistics		
	Min	Max	$t^* < -1.96$	$-1.96 \leq t^* \leq 1.96$	$t^* > 1.96$
Intercept	-0.74	0.96	-----	-----	-----
Available sunlight: yearly average sunlight	-3.48×10^{-4}	-1.08×10^{-4}	24.5	75.5	0.0
Race: % of White Population	1.16	2.61	0.0	80.7	19.3
Access to healthcare: % of Uninsured Population	2.1	2.6	0.0	86.0	14.0

Table 4. Results from GWR Model of New England County-Level MS Death Rates.

* 1.96 is the cut-off value for t-test. When $|t| > 1.96$, the β coefficient estimate for a variable is significant at a significance level of 0.05.

As shown in Table 4, available sunlight, defined by yearly average sunlight, negatively associated with county-level MS death rates. This finding supports previous research which suggests the farther away from the equator a person lives and the less sun exposure they could have the higher the risk of developing MS which can lead to more deaths caused by the disease. As demonstrated in Table 4, race, defined by the percentage of White population, is positively associated with county-level MS death rates. This finding is consistent with previous research findings which suggest that MS, so far, has been found to be most prevalent among Caucasians, particularly those of European ancestry. As illustrated in Table 4, access to healthcare, defined by the percentage of uninsured population, is positively associated with county-level MS death rates.

The direction of the relationships between the dependent variable and the explanatory risk factors in Table 4 provides insights regarding MS disease intervention in the region. The negative relationship between MS death rates and yearly average sunlight confirms the importance of sun exposure. Recent research suggests the risk of having a preliminary symptom of MS decreased by 30.0 percent for every 1,000 kilojoules of exposure to ultraviolet light (Lucas et al. 2011). Given the fact that yearly available sunlight is a constant in a place, people living farther North in the New England region should have more sunlight exposure by spending more daylight hours outside.

The positive relationship between MS death rates and the percentage of White population should not be mistakenly explained as African Americans, Asians or Latinos are risk free. In fact, for example, given the false impression that African Americans are less likely to develop MS disease, research shows they are less likely to receive care from a neurologist specializing in MS or to go to an MS clinic (Khan et al. 2015). Also, research shows that, for African Americans with multiple sclerosis, the MS disease progresses much faster than it does for White peers (Khan et al. 2015). Therefore, further research is needed to examine the relationship between MS correlates and race-specific MS prevalence rates if race-specific MS data is available.

The positive relationship between MS death rates and the percentage of uninsured population is particularly important because the United States does not have a uniform health system, has no universal health care coverage, and only recently enacted legislation mandating healthcare coverage for almost everyone. In 2014, there were 283.2 million people in the United States. 89.6 percent of the U.S. population had some types of health insurance while the rest of the population didn't (Smith and Medalia 2015). Given the fact that the Affordable Care Act (or ACA) may be repealed in the near future, a significant increase of MS prevalence or death in the New England region may happen in the long run.

As shown in Table 4 and Figure 3, the change in magnitude of the coefficients suggests spatial non-stationarity of the relationship between MS death rates and the predictors. The variation in parameter estimates from GWR suggests thUe necessity to apply this spatial statistical tool to future MS studies that would be restricted by using global OLS models, since GWR provides insights on how a particular explanatory variable influences MS death rates across the study area. As demonstrated in Figure 3A and 4A, the race variable had the highest impact in counties located in northern Maine and southern Vermont. As shown in Figure 3B and 4B, available sunlight had the greatest effect in counties located in northern New Hampshire, Maine and Vermont. Counties in these areas have a relatively less available sunlight with disproportionately high rates of MS death. As illustrated in Figure 3C and 4C, in counties located in southern New Hampshire together with Essex County, Massachusetts, and York County Maine, access to healthcare had a greater association with MS death rates than any other variable. The GWR results are potentially useful in targeting priority areas for MS disease prevention and intervention, and for informing local health planning and policy development, since they suggest that preventive policies should be informed by an understanding of MS death's contextual factors. In particular, these factors should be examined locally, and different policies aimed at preventing and reducing MS death should be applied in different counties of the New England region. For example, different prevention policies could be recommended for the counties located in northern New Hampshire, Maine and Vermont as well as the counties in southern New Hampshire. A policy designed to increase access to healthcare may be sufficient to reduce MS prevalence or death in southern New Hampshire. However, the same approach is unlikely to be effective in the counties located in northern New Hampshire, Maine and Vermont. In those areas, regeneration initiatives and MS prevention/intervention programs aimed at increasing people's outdoor time and early and frequent screening of MS disease should be considered as tools for preventing and reducing MS death.

The importance of using spatial statistical tools such as GWR in future MS studies can also be confirmed by the adjusted R^2 value (see Figure 3D). The adjusted R^2 for the GWR model ranged from 0.17 to 0.36, with an average of 0.27, while the adjusted R^2 in the OLS model was 0.21. Explicitly, the OLS R2 of 0.21 masks a wide distribution of local associations between the independent variables and MS death rates. In other words, without GWR, it would be unable to estimate the variance of local associations. In a county such as Penobscot, Maine, the GWR model explained 36.0 percent of the variance in MS death rates. However, in counties such as Litchfield, Connecticut and Berkshire, Massachusetts, the model did not explain much of the

variance (from 17.0 to 20.0 percent), a spatial variation that would have been neglected with the OLS model alone. In addition, the local adjusted R^2 value can be used as goodness-of-fit measure for the GWR model. As shown in Figure 3D, the local adjusted R^2 values increased throughout the New England region from southwest to northeast, indicating the GWR model better explains the variance in MS death rates in counties in the northeast of the New England region.

This study is not without limitations. The first group of limitations is associated with geographic boundary effect, such as Modifiable Areal Unit Problem (MAUP) and edge effect. It should be noted that the statistical relationships drawn from areal data must be carefully interpreted. Robinson (1950) long ago suggested the data scale/boundary problem and clearly explained that inferring individual level relationships from macro-level correlations is inappropriate. In this study, county boundaries were used as the units of analysis and the relationships between MS death rates and contextual characteristics at the county level thus cannot be interpreted as and/or applied to individual level relationships. In addition, GWR model is limited by the edge effect, whereby counties located on the edges of New England do not have the 360° influence of counties in the region's interior.

The second group of limitations is related with data availability. In this study, encountered are missing data and time alignment problem. For example, ten counties in New England have no reported MS death data because CDC intentionally suppressed the data. MS death counts were collected by the CDC during 1999 and 2014, but the benchmark – residential population at each county was collected by the Census Bureau in the year of 2010. In addition, the local R^2 values accounted for 17.0 to 36.0 percent of county-level MS death rates, which means that other risk factors associated with MS disease need to be added into the GWR model. For example, a defected gene on chromosome six is an important risk factor. People whose close relatives have MS are more likely to develop the disease (Donati and Jacobson 2002). However, no organization or government agency collects and releases personal MS genes data or data regarding people whose close relatives have MS at the county level or a finer scale.

Conclusion

This study analyzed the spatial distribution and correlations of MS death rates in the New England region. To be specific, this research incorporates demographic, socio-economic and environmental correlates with MS death rates. The relationship between the disease and predictors, such as available sunlight, race, and access to healthcare, are not new (Freedman et al. 2000; Neri and Kroll 2003; Kakalacheva and Lunemann 2011; Kingwell et al. 2013), but little research was done to investigate the spatial heterogeneity in the relationship. This study fills the gap by illustrating that there is a significant association between MS death rates and the explanatory variables and that this relationship has a spatial but nonstationary association which highlights the need for local and context-specific MS prevention and intervention programs. In other words, using GWR, public health researchers and practitioners can gain an understanding of health-related issues and respond to the notion that "all health is local" (Gebreab and Diez-Roux 2012). The results of this study can also be used by local public health departments to

tailor messages and materials for their target audiences. For MS research, this study presented an initial and exploratory step in this direction, but much more in-depth work remains before public health researchers and practitioners understand why these spatial variations exist and why predictors, such as available sunlight, race, and access to healthcare, have very low explanatory effect in some counties, but explain up to 36.0 percent of MS death rates in other counties. Further explored should be MS deaths caused by the migration from low risk to high risk counties in the New England region (as shown in Figure 2) if longitudinal MS data are available. Such research can help average people to make better migration decisions (i.e. avoiding high risk areas) if MS deaths are found to be correlated to risk factors such as time of migration, race, gender, and age.

RJ Malavasi[1] is currently a Sr. Risk Analyst at The Hartford Insurance Group as part of the GIS team. He is trained as a spatial analyst and his studies focus on Geographic Information Systems (GIS), spatial statistics, and a variety of computer science technologies. He received his undergraduate and post graduate degrees in Geography from Central Connecticut State University. Email: rj.malavasi@yahoo.com.

Dr. Yunliang Meng[2] is Associate Professor at the Central Connecticut State University, Geography Department. His research interests lie in the field of GIS, Urban Geography, Quantitative Analysis, Social and/or Environmental Justice. His publications cover a wide range of topics including participatory planning, racial profiling/crime pattern analysis, and disease pattern modeling. Email: mengy@ccsu.edu.

References

Aiken, L.S., and West, S.G. (1991). *Multiple regression: Testing and interpreting interactions*, Newbury Park: Sage.

Amezcua, L., Conti D.V., Liu L., Ledezma K., and Langer-Goulda A.M. (2015) Place of birth, age of immigration, and disability in Hispanics with multiple sclerosis, *Multiple Sclerosis and Related Disorders*, 4(1), 25-30.

Anselin, L. (1995). Local indicators of spatial association – LISA, *Geographical Analysis*, 27(2), 93-115.

Bäärnhielm, M., Hedström, A. K., Kockum, I., Sundqvist, E., Gustafsson, S. A., Hilbert, J., Olsson, T., and Alfredsson, L. (2012) Sunlight is associated with a decreased multiple sclerosis risk: no interaction with human leukocyte antigen-DRB1*15, *European Journal of Neurology*, 19(7), 955-962.

Bove, R., and Chitnis, T. (2014) The role of gender and sex hormones in determining the onset and outcome of multiple sclerosis, *Multiple Sclerosis Journal*, 20(5), 520-526.

Brunsdon, C., Fotheringham, A.S., and Charlton, M.E. (1996) Geographically weighted regression: A method for exploring spatial nonstationarity, *Geographical Analysis*, 28(4), 281-298.

Cahill, M., and Mulligan, G. (2007) Using geographically weighted regression to explore local crime patterns, *Social Science Computer Review*, 25(2), 174-193.

Smith, J.C., and Medalia, C. (2015) Health Insurance Coverage in the United States: 2014, *REPORT NUMBER P60-253*, Washington, D.C.: U.S. Census Bureau.

Compston, A., McDonald, I., Noseworthy, J., Lassmann, H., Miller, D., Smith, K., Wekerle, H., and Confavreux, C. (2006) *McAlpine's Multiple Sclerosis*, London: Churchill Livingstone.

Conradi, S., Malzahn, U., Schröter, F., Paul, F., Quill, S., Spruth, E., Harms, L., Bergh, F.T., Ditzenbach, A., Georgi T., Heuschmann, P., and Rosche, B. (2011) Environmental factors in early childhood are associated with multiple sclerosis: A case-control study, *BMC Neurology*, 11(1), 123.

Donati, D., and Jacobson, S. (2002) Chapter 6: Viruses and Multiple Sclerosis. In K.A. Brogden and J.M. Guthmiller (Eds.), *Polymicrobial Diseases*, Washington D.C.: Press, 99-125.

Environmental System Research Institute (ESRI) (2015), ArcGIS Desktop: Release 10.3.1, Redlands, CA.

Fotheringham, A.S., Brunsdon, C., and Charlton, M., (2002) *Geographically weighted regression: The analysis of spatially varying relationships*, Chichester: Wiley.

Freedman, D.M., Dosemeci M., and Alavanja M.C.R. (2000) Mortality from multiple sclerosis and exposure to residential and occupational solar radiation: A case-control study based on death certificates, *Occupational and Environmental Medicine*, 57(6), 418-421.

Gebreab S.Y. and Diez-Roux A.V. (2012) Exploring racial disparities in CHD mortality between blacks and whites across the United States: a geographically weighted regression approach, *Health Place*, 18(5):1006-1014.

Harding, K.E., Liang K., Cossburn M.D., Ingram G., Hirst C.L., Pickersgill T.P., Te Water Naude J., Wardle M., Ben-Shlomo Y., and Robertson N. P. (2013) Long-term outcome of paediatric-onset multiple sclerosis: a population-based study, *Journal of Neurology, Neurosurgery, and Psychiatry*, 84(2), 141-147.

Hipp, J.A., and Chalise, N. (2015) Spatial analysis and correlates of county-level diabetes prevalence, 2009-2010, *Preventing Chronic Disease*, 12:140404. DOI: http://dx.doi.org/10.5888/pcd12.140404.

Kakalacheva, K., and Lünemann, J. D. (2011) Environmental Triggers of Multiple Sclerosis, *FEBS Letters*, 585(23), 3724-3729.

Kingwell, E., Marriott, J.J., Jetté, N., Pringsheim, T., Makhani, N., Morrow, S.A., Fisk, J.D., Evans, C., Béland, S.G., Kulaga, S., Dykeman, J., Wolfson, C., Koch, M.W., Marrie, R.A. (2013) Incidence and prevalence of multiple sclerosis in Europe: a systematic review, *BMC Neurology*, 13, 128.

Kira, J. (2006) Epidemiology of multiple sclerosis in Japanese: with special reference to opticopsinal multiple sclerosis, *Clinical Neurology*, 46(11), 859-862.

Koch-Henriksen, N., Brønnum-Hansen, H., and Stenager, E. (1998) Underlying cause of death in Danish patients with multiple sclerosis: results from the Danish Multiple Sclerosis Registry, *Journal of Neurology, Neurosurgery, and Psychiatry*, 65(1), 56-59.

Lucas, R.M., Ponsonby, A.L., Dear, K., Valery, P.C., Pender, M.P., Taylor, B.V., Kilpatrick, T.J., Dwyer, T., Coulthard, A., Chapman, C., Van Der Mei, I., Williams, D., McMichael, A.J., (2011) Sun exposure and vitamin D are independent risk factors for CNS demyelination, *Neurology*, 76(6), 540-548.

Mateu, J. (2010) Comments On: A Gernal Science-Based Framework for Dynamical Spatio-Temporal Model, *Test*, 19, 452-455.

Mayer, J. (1981) Geographical clues about multiple sclerosis, *Annals of the Association of American Geographers*, 71(1), 28-39.
Milo, R., and Kahana, E. (2010), K. Multiple sclerosis: geoepidemiology, genetics and the environment, *Autoimmunity Reviews* 9(5), A387-A394.
Minden, S.L., Frankel, D., Hadden, L., Hoaglin, D.C. (2007) Access to health care for people with multiple sclerosis, *Multiple Sclerosis Journal*, 13(4), 547-558.
Moore, D.A., and Carpenter, T.E. (1999) Spatial analytical methods and geographic information systems: use in health research and epidemiology, *Epidemiologic* Reviews, 21(2), 143-161.
Murray, T.J. (2005) *Multiple Sclerosis: The history of a disease*, New York: Demos Medical Publishing.
Neri, M.T., and Kroll, T. (2003) Understanding the consequences of access barriers to health care: experiences of adults with disabilities, *Disability and Rehabilitation*, 25(2), 85-96.
O'Gorman, C., Lucas, R., Taylor, B. (2012) Environmental risk factors for multiple sclerosis: a review with a focus on molecular mechanisms, *International Journal of Molecular Sciences*, 13(9), 11718-11752.
Omar, K., Williams, M.J., Amezcua, L., Javed, A., Larsen, K.E., and Smrtka, J.M. (2015) Multiple sclerosis in US minority populations, *Neurology Clinical Practice*, 5(2), 132-142.
Robinson, W.S. (1950) Ecological correlations and the behavior of individuals, *American Sociological Review*, 15(3), 351-357.
Van der Mei, I.A., Ponsonby, A.L., Dwyer, T., Blizzard, L., Simmons, R., Taylor, B.V., Butzkueven, H., and Kilpatrick, T. (2003) Past exposure to sun, skin phenotype, and risk of multiple sclerosis: Case-control study, *BMJ*, 327, 1-6.
Wang, G., Marrie, R.A., Salter, A.R., Fox, R., Cofield, R.S., Tyry, T., and Cutter, G.R. (2016) Health insurance affects the use of disease-modifying therapy in multiple sclerosis, *Neurology*, 87(4), 365-374.