Spatial analysis of the regional variation of hypertensive disease mortality and its socio-economic correlates in South Korea

Seong-Yong Park,1 Jin-Mi Kwak,2 Eun-Won Seo,2 Kwang-Soo Lee2

1Korea Health Promotion Foundation, Seoul; 2Department of Health Administration, College of Health Sciences, Yonsei University, Seoul, South Korea

Abstract

This paper presents a cross-sectional study based on the cause of death statistics in 2011 extracted from all 229 local governments in South Korea. The standardised hypertensive disease mortality rate (SHDMR) was defined by age- and sex-adjusted mortality by hypertensive diseases distinguished by International Classification of Disease-10 (ICD-10). Variables taken into account were the number of doctors per 100,000 persons, the proportion with higher education (including university students and high school graduates), the number of recipients of basic livelihood support per 100,000 persons, the annual national health insurance premium per capita and the proportion of persons classified as high-risk drinkers. Ordinary least square (OLS) regression and geographically weighted regression (GWR) were applied to identify the potential associations. The statistical analysis was conducted with SAS ver. 9.3, while ArcGIS ver. 10.0 was utilised for the spatial analysis. The OLS results showed that the number of basic livelihood recipients per 100,000 persons had a significant positive association with the SHDMR, and the proportion with higher education had a significant negative one. GWR coefficients varied depending on region investigated and some regional variables had various directions. GWR showed higher adjusted R^2 than that of OLS. It was found that the SHDMR was affected by socio-economic status, but as the effects observed were not consistent in all regions of the country, the development of health policies will need to consider the potential for regional variation.

Introduction

South Korea is starting to see a shift from acute diseases to chronic ones (Ministry of Health and Welfare, 2010). Hypertension, a typical chronic disease, is not accompanied by severe pain but requires careful management since it can result in complications, such as stroke, renal failure, heart failure and cardiovascular disease if neglected (Kim et al., 2005).

Studies of community healthcare projects managing hypertensive disease (and the factors affecting its level) have focused on personal health and a wide range of behavioural characteristics, such as age, family and personal history of hypertension, alcohol intake, smoking, body mass index (BMI), blood levels of cholesterol and neutral fat, salt intake, physical activity, mental stress and obesity. These studies analysed the relation between these factors and hypertension (Nicholson et al., 1983; Kam et al., 1991; Moreira et al., 1998; Wannamethee et al., 1998; He et al., 2000; Masahiko et al., 2000; Oh et al., 2000; Ishikawa et al., 2002; Player et al., 2007; Eom et al., 2008; Kim et al., 2009; Eom, 2009; Iang et al., 2009; Wen et al., 2010; Matyas et al., 2011; Centers for Disease Control and Prevention, 2012; Odell et al., 2012) with the majority of them conducted at the individual level.

Hypertension is known as the principal cause of death in the area of cardiovascular diseases (Jacqmin-Gadda et al., 2013). Researches on hypertension have recently also paid attention to socio-demographic characteristics at the community level. Schlundt et al. (2006) analysed the differences with respect to hypertension between regions by using the geographic clustering method. Lawes et al. (2008) contend that the financial burden of hypertension varies depending on the income level and age distribution in each country, while Egan et al. (2002) also suggest the need to consider given demographic characteristics and regional diversity in investigating the burden of this disease. Kim (2010) has analysed the socio-economic conditions, lifestyle, community health status, physical environment, and healthcare system, as well as the demographic characteristics of the community to identify the factors having an impact on the management of hypertension.
Most previous studies on hypertensive disease correlates have been focused at the individual level. Research on hypertensive disease needs to consider the inter-regional variation and influential factors, as well as the inter-regional differences, taking into account the demographic and socio-economic characteristics of each community when pursuing effective disease control (Macintyre and Ellaway, 2000). Our study therefore aimed to examine the relations among the variables representing demographic, socio-economic, and health and behavioural characteristics of a community and the mortality caused by hypertension. We analysed the relationship between the standardised hypertensive disease mortality rate (SHDMR) and regional factors using traditional ordinary least square (OLS) regression and geographically weighted regression (GWR).

Materials and Methods

Study area

South Korea is located in north-eastern Asia. The eastern areas of Korean Peninsula are mountainous, while the Southwest is characterised by plains. The capital, Seoul, is located in the north-western part of the country. The population is about 50 million with over 20 million residing in Seoul and surrounding provinces.

In this study, all local governments in South Korea as of 2011 were included. The different levels investigated were the Gun, which corresponds to a county-level administrative region and usually consists of rural areas; the city, which denotes large and smaller towns; and the Gu, which is an urban district (a large city can have multiple Gus). There were 251 local governments in 2011. However, if data for a local government, e.g. a Gu of a city, were not available, these data were incorporated into the City this Gu is part of. For this reason, the final count of local governments analysed amounted to 229.

Study variables

On the basis of a literature review, we selected five variables representing different regional characteristics and analysed the relationship of these variables with the mortality rate caused by hypertension of various kinds.

Medicine and healthcare

Chun et al. (2008) report that medical institutions, doctors’ ability and the trust in doctors have an impact on blood pressure control among hypertension patients. Kim (2010) notes that regions at the lower end of healthcare supply are more likely to be exposed to progressing hypertensive disease and this author empathises the importance of broadening medical resources to settle this issue.

Education and socio-demography

Many existing studies indicate that the educational factor affects hypertension (Kim et al., 2005; Campbell et al., 2006; Kim et al., 2009; Jang et al., 2009). Demaio et al. (2013) feel that individuals with lower levels of knowledge are more likely to develop hypertension, while Eom et al. (2008) find that high school graduates are less likely to be affected by hypertension.

General economy

Kim et al. (2005) report that financial status might affect hypertension prevalence, which is relatively higher in the economically unprivileged bracket in agricultural regions. Lawes et al. (2008) contend that lower-income countries have a greater burden of hypertensive disease than developed ones.

Health insurance

The annual per capita health insurance premium is known to be a variable closely associated with the income level and it has therefore been widely used as a proxy for income. Lee et al. (1989) report that the insurance premium level is positively related to healthcare use and that the lower-income brackets have a relatively greater need to settle the issue of unmet healthcare. Kim et al. (2009) find that people in lower-income brackets are significantly more likely to be affected by hypertension than those in higher-income brackets.

Health-related behavioural issues

Kam et al. (1991) include history of alcohol intake, obesity status and smoking habits as hypertension risk factors, while Moreira et al. (1998) note that alcohol intake might affect blood pressure but that the volume of alcohol intake has a greater impact on the blood pressure level than the frequency of intake.

Mortality rates caused by hypertensive disease

The dependent study variable used in this study was the mortality rate due to hypertensive disease per 100,000 persons with rates standardised in terms of age and sex. The plural, hypertensive diseases, refers to the various organ-specific complications caused by hypertension and includes primarily heart failure, renal failure, stroke and cardiovascular disease.

The 2011 statistical data about causes of death, provided by the Department of Population Trends in the Division of Social Statistics - Statistics Korea (http://www.kosis.kr), were used to examine the mortality of hypertensive diseases coded in the major classes I10-I13 according to the International Classification of Disease-10 (ICD-10). In this study, the hypertensive diseases were defined with following subcodes: benign hypertension (I100); malignant hypertension (I101); unspecified hypertension (I109); hypertensive heart diseases accompanied by (congestive) heart failure (I111); hypertensive heart diseases not accompanied by (congestive) heart failure (I119); hypertensive renal diseases accompanied by renal failure (I120); hypertensive renal diseases not accompanied by renal failure (I129); hypertensive heart and renal diseases accompanied by (congestive) heart failure (I130); hypertensive heart and renal diseases accompanied by heart failure (I131); hypertensive heart and renal diseases accompanied by (congestive) heart and renal failure (I132); unspecified hypertensive heart and renal diseases (I139).

The standardised hypertensive disease mortality rate (SHDMR) is estimated by the following equation:

\[
SHDMR = \frac{\sum_{i,j} (N_{ij} \times O_{ij})}{\sum_{j} N_{j}} \times 100,000
\]

where \( i \) is the age group (0-4, 5-9,...,95-99, ≥100; \( O_{ij} \) the age group, observation of hypertensive disease mortality of \( j \) sex; \( n_{i} \) the age group, population of \( j \) sex; and \( N_{j} \) the age group of the standard population, population of \( j \) sex.
Regional variables

Medicine and healthcare

The number of doctors was identified in the 2011 statistical data concerning healthcare use for each region, which are supplied by the National Health Insurance Service (NHIS; http://www.nhis.or.kr). As dentists are less likely to be associated with hypertension they were excluded and only medical doctors (specialists, residents, general practitioners, and interns) and doctors of traditional Korean medicine (specialists, special interns, general interns, and general practitioners) included. Conversion into the number of doctors per 100,000 persons was made to control the effects of the population size.

Education and socio-demography

The proportion of the population with higher education (including university students and high school graduates) was used as a socio-demographic variable. Of the census data from Statistics Korea (http://www.kosis.kr); here the 2010 data were analysed since the census was conducted at five-year intervals.

General economy

The economic characteristics were measured on the basis of the number of recipients of basic livelihood security support per 100,000 persons. The basic livelihood security programme aims to assist people with no economic ability to lead a life at least at a basic level of living; this can be an index representing the economic status of the region. In general, the region with a larger number of basic livelihood security recipients is likely to have a relatively lower economic level than other regions (Kim, 2010; Seok and Kang, 2013; Park and Lee, 2014). The number of basic livelihood security recipients was identified in the annual statistical report from each municipal, provincial, and metropolitan government in 2011 and was converted into the number of livelihood security recipients per 100,000 persons.

Health insurance

The insurance premium index was drawn from the 2011 statistical data concerning healthcare use for each region, which were supplied by NHIS (http://www.nhis.or.kr). An annual per capita insurance premium per region was obtained by dividing the sum of the regional National Health Insurance premiums by the number of people in each region.

Health-related behavioural issues

High-risk alcohol intake rate was used as a proxy variable for health and behavioural characteristics and was based on data from the 2011 community health research provided by the Department of Chronic Disease Management, Disease Control and Prevention Center, Ministry of Health and Welfare (https://www.chs.cdc.go.kr). The rates used in this study refer to the percentage of men drinking ≥7 glasses of strong alcohol a month or that of women drinking ≥5 glasses of alcohol a month as reported by Kim (2010).

Analytic methods

Pearson’s correlation coefficients were calculated from correlation analysis. This technique analyses the relationship between two variables and measures the strength of association between them. Most studies on hypertension employ multivariate regression analysis based on OLS for individuals. This study employed GWR, a type of spatial analysis that can reflect spatial relations and the heterogeneity of regions (Fotheringham et al., 2002; Helbich et al., 2013). In the traditional OLS equation, the dependent variable was modeled as a function of a set of independent variable and the equation as used by Brunsdon et al. (1996) is then:

\[ y_i = a_0 + \sum_{k=1}^{m} a_k x_{ik} + \epsilon_i \]

where \( y_i \) is the dependent variable; \( x_{ik} \) the independent variable; \( \epsilon_i \) the error term with zero means; and \( a_k \) determined from a sample of n observations.

GWR is a technique based on the traditional regression model and the coefficients in the GWR model are specific to each location. The GWR equation is, following Brunsdon et al. (1996):

\[ y_i = a_{ik} + \sum_{k=1}^{m} a_{ik} x_{ik} + \epsilon_i \]

where \( a_{ik} \) is the value of the \( k^{th} \) parameter at location \( i \).

Contrary to OLS that infers a global regression coefficient for all the analysis targets (Fotheringham et al., 1996, 2002), the GWR spatial statistic analysis method suggested by Brunsdon et al. (1996) can infer local regression coefficients for each region. In other words, GWR can easily determine association between neighbouring regions in terms of variables and analyse the differences with respect to the influence of regional characteristics for each region. It can also combine spatial data and attribute data that contain non-spatial characteristics. It was therefore expected that spatial homogeneity and heterogeneity between regional units, which has been less researched in the field of public health, could be explored in this way.

To confirm that GWR had greater explanatory power for a model than OLS, the Akaike Information Criterion (AIC) for each of the two statistical approaches OLS and GWR was compared. AIC shows goodness-of-fit for a model (Hurvich et al., 1998; Charlton and Fotheringham, 2009) and a difference in AIC between OLS and GWR would tell which statistical approach would be preferable.

The degree of similarity of a variable with reference to its spatial location based on location and value can be evaluated by spatial autocorrelation appraisal (Huo et al., 2012). Global Moran’s I analysis was used to assess the global spatial autocorrelation of the model. It assesses whether the patterns were clustered, dispersed, or random where a value close to +1 means a potent spatial autocorrelation and indicates thus that the values of attributes have strong similarities (Anselin, 1995). The variables in this study were statistically processed using SAS ver. 9.3 (SAS Institute, Cary, NC, USA), while the spatial analysis tool ArcGIS ver. 10.0 (ESRI, Redlands, CA, USA) was used for mapping and spatial analysis.

Results

Table 1 represents descriptive statistics showing the general characteristics of the regional factors. It should be noted that the average of SHDMR was 14.29, ranging from 1.75 to 52.68. The standard deviation (SD) of annual per capita insurance premiums was larger than the mean. For the reliable comparison of hypertensive disease mortality between regions, hypertensive disease mortality rate for age- and sex-
groups in each region was calculated and applied to the corresponding age- and sex-groups of the entire Korean population. All values for each age and sex group were summed to produce the adjusted rate.

Figure 1 shows the distribution of SHDMR in each region. Yeongyang Gun, North Gyeongsang Province (52.68 persons) had the highest mortality rate due to hypertensive diseases followed by Euisongun, South Gyeongsang Province (49.70) and Cheongsong Gun, North Gyeongsang Province (47.75). Seocho Gu, Seoul (1.76) had the lowest mortality rate followed by Hanam City, Gyeonggi Province (2.51) and Jinju City, South Gyeongsang Province (2.55).

Table 2 represents the results of the correlation analysis among the regional characteristics. Since no variable had a correlation coefficient exceeding 0.8, no serious problem with multicollinearity should be expected (Chan, 2003). Of the variables with significant correlation coefficient, there was a statistically significant negative correlation between the number of basic livelihood security recipients per 100,000 persons and the proportion of people with higher education (coefficient=-0.696); the same was true between the number of basic livelihood security recipients per 100,000 persons and annual per capita insurance premiums (coefficient=-0.211). There was a statistically significant positive correlation between the number of doctors per 100,000 persons and the proportion of people with higher education (coefficient=0.252); between the number of doctors per 100,000 persons and annual per capita insurance premiums (coefficient=0.350); and between the proportion of people with higher education and annual per capita insurance premiums (coefficient=0.276).

Table 3 presents the results from OLS regression and GWR. The OLS analysis found significance in the proportion of people with higher education and in the number of basic livelihood security recipients per 100,000 persons at the 95% significance level. The proportion of people with higher education was negatively related to the mortality rate for hypertension. The number of basic livelihood security recipients per 100,000 persons was positively associated with mortality. Since no variable had a variance inflation factor (VIF) exceeding 10, the problem of multicollinearity was not expected.

The explanatory power of GWR (48.5%) was slightly greater than that offered by OLS regression (46.6%). AIC was estimated to be 1559 by OLS regression and 1554 by GWR. Since the difference with respect to goodness-of-fit between the two models exceeded 4, it was concluded that GWR is the best approach (Lee and Choi, 2013). In addition, GWR
had lower levels of Sigma (♂) and residual squares than OLS. Koenker statistics (Neki and Osirike, 2013) had significant P-values, which implies that there was no fixed type of relationship between the dependent variables and the regional characteristics and that they varied by region.

Figure 2 shows the distribution of regression coefficient and local R² for each region as estimated by GWR. Figure 2A to 2E shows regional variation in regression coefficients. The correlation among the numbers of doctors per 100,000 persons, annual per capita insurance premiums, and the mortality rates by hypertensive diseases by region varied, i.e., it was sometimes positive and sometimes negative. Figure 2F shows that the coefficient of determination (local R²) ranges from 0.3650 to 0.5439 by region, and that the level of explanatory power of a GWR model varies by region. Usually, local R² decreased toward the Southeast and increased toward the northern regions.

Discussion

In this study, the relations between the regional characteristics and the SHDMR in 100,000 persons were analysed in 229 basic local governments of different levels as of 2011. The variations in gender and age among the regions were controlled to reflect the differences in the population structure in each region, whereas OLS regression and spatial analysis were used for the analysis.

Previous studies on the relationship between regional characteristics and disease occurrences have employed multivariate analysis based on OLS regression. However, this approach has a limitation in reflecting the variation on the effect of regional characteristics by location; that is, the OLS analysis assumes identical relational characteristics regardless of the characteristics of the regions because it estimates one regression coefficient for all the regions together. GWR analysis, on the other hand, presumes an analysis model by using a geographically weighted function based on the assumption that a region has more similar characteristics to neighbouring regions than to distant ones. On the basis of these results, regression coefficient and local R² were estimated for each region (Brunsdon et al., 1996); that is, a GWR model would be expected to be useful for the reflection of geographical characteristics in the field of healthcare that explores social phenomena that vary with regard to type and number of diseases by region. GWR can thus be used to make a comparison between dependent and explanatory variables and in that way find out and classify regions with differences. In this study, GWR was successfully employed to confirm the variation in the size and direction of the relations between the regional characteristics and the mortality rate from hypertensive diseases for each region. Compared with OLS regression, GWR had a smaller residual variate and a greater explanatory power in general (Griffith, 2008). Lastly, the results of the analysis can be presented in the form of maps to allow policy-makers to understand them with ease.

The variable representing the economic level of the region, i.e. the number of basic livelihood security recipients in the population of 100,000 persons was positively related with death caused by hypertensive diseases. It was expected that the lower the economic level of the region, the higher the likelihood of having problems caused by hypertensive diseases. This result was consistent with the finding that a lower-income region can be expected to have a heavier burden due to hypertensive diseases (Lawes et al., 2008) and that the social and economic burden of hypertension lower the quality of life (Eom, 2009).

Since hypertensive diseases do not generally cause pain or awareness of the condition, it is very likely to be neglected and therefore result in premature death.

The proportion of higher education levels was negatively related with SHDMR. In previous research, the percentage of proportion was often used as an index to reflect the socio-demographic characteristics. The group with a lower proportion of proportion was more likely to have little knowledge about hypertension or to believe that the condition of hypertension cannot be cured by medication. They also tended to use folk remedies that can give short-term relief rather than be the proper treatment for hypertension (Kim et al., 2003).

The results of the GWR analysis make important suggestions for policy-makers in the field of healthcare aimed at reducing the mortality of hypertensive diseases. However, implementing a uniform policy with no regional characteristic taken into account cannot be less effective in avoiding death caused by hypertensive diseases. The result that the relations between death caused by this condition vary by the regional unit suggests the need to differentiate policy compositions and pro-

Table 3. Results of ordinary least square and geographically weighted regression.

<table>
<thead>
<tr>
<th>Variable</th>
<th>OLS Standardised regression coefficient</th>
<th>VIF</th>
<th>Mean (SD)</th>
<th>Minimum value</th>
<th>Maximum value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Doctors  (n)°</td>
<td>0.0012 *</td>
<td>1.40</td>
<td>0.0008 (0.0006)</td>
<td>-0.0008</td>
<td>0.0019</td>
</tr>
<tr>
<td>Educated people (%)</td>
<td>-0.3714*</td>
<td>2.40</td>
<td>-29.06 (5.37)</td>
<td>-40.16</td>
<td>-19.01</td>
</tr>
<tr>
<td>Basic livelihood recipients  (n)°</td>
<td>0.3823*</td>
<td>2.33</td>
<td>0.0017 (0.0002)</td>
<td>0.0011</td>
<td>0.0023</td>
</tr>
<tr>
<td>Annual NHI premium per capita</td>
<td>0.0081</td>
<td>1.23</td>
<td>0.00002 (0.00005)</td>
<td>-0.00007</td>
<td>0.00009</td>
</tr>
<tr>
<td>High-risk drinkers (%)</td>
<td>0.0497 *</td>
<td>1.01</td>
<td>0.1147 (0.0339)</td>
<td>0.0293</td>
<td>0.1773</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>46.63</td>
<td>-</td>
<td>48.49</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>1559</td>
<td>-</td>
<td>1554</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>σ</td>
<td>7.18</td>
<td>-</td>
<td>7.05</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Residual squares</td>
<td>11,508</td>
<td>-</td>
<td>10,876</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Moran’s I</td>
<td>0.03*</td>
<td>-</td>
<td>0.02</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>Koenker statistics</td>
<td>35.0*</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td></td>
</tr>
</tbody>
</table>

OLS, ordinary least square; GWR, geographically weighted regression; VIF, variance inflation factor; SD, standard deviation; NHI, National Health Insurance; AIC, Akaike Information Criterion. °Per 100,000 persons.

*P<0.05.
Figure 2. Coefficients and local $R^2$ by region by geographically weighted regression. A) Number of doctors per 100,000 persons; B) proportion of educated people; C) number of basic livelihood recipients per 100,000 persons; D) annual National Health Insurance premium per capita; E) proportion of high-risk drinkers; F) local $R^2$. 
grammes to meet the regional situations. The results of the analysis presented here imply the need to give priority to regions with a higher proportion of the group with a lower living standard in implementing a health policy related to hypertension. The number of doctors available, annual per capita insurance premiums and high-risk alcohol intake rate were not statistically significant in the OLS analysis but were either positively or negatively related by region in the GWR analysis. It is necessary to make an in-depth analysis of the reasons for this result.

In contrast to existing studies, this investigation employed spatial analysis in examining the regional characteristics related to death caused by hypertensive diseases. It is significant in that it made an approach to the problems at the regional level, while the previous research was mostly conducted at the individual level. It is therefore expected that this study would contribute to the development of measures against the problems associated with hypertension at the regional level.

This study has some limitations, however. First, we used only a limited number of variables that represent regional characteristics. A number of known factors that affect hypertension, including the availability of medical institutions, the regional deprivation index, smoking, obesity and stress, were not used due to the possibility of multicollinearity or the possibility of statistical errors in the process of developing a GWR model. It is necessary to add and analyse variables that can overcome these limitations and still accurately reflect the regional characteristics. Second, analysis of administrative districts smaller than City, Gun and Gu, i.e. Eup, Myeon and Dong, would be expected to give the most detailed results. In reality, however, the variables that represent the regional characteristics are collected in the units of City, Gun, Gu; therefore, it is impossible to make a better analysis using these units as they are more detailed. However, it is expected that more in-depth regional research could be conducted if data in the unit of smaller regions were available. Third, it is necessary to avoid the ecological fallacy. This study explored the relations with the regional characteristics in the analysis unit of the population of basic local governments, instead of individuals, and has therefore a limitation in interpreting personal characteristics. Researchers are expected to contribute to making more effective policies in coping with problems caused by hypertensive diseases if they were to secure data at the individual level and compare the results with those at the group level.

Conclusions

This study has shown that the proportion of educated people and the number of basic livelihood security recipients per 100,000 persons are associated with a mortality rate due to hypertensive disease at a statistically significant level, and that there is a regional variation of these associations. The findings presented here emphasise the need to take regional differences into account if the burden of hypertensive diseases should be effectively reduced.

It would be desirable to examine the regional situation with respect to hypertension thoroughly before developing and implementing policies and management programmes so that they reflect the regional characteristics. As a way to solve this problem, the regional characteristics should be analysed in depth and the findings used to develop approaches that can be adjusted to fit each region by considering the local conditions.

Further research into the inclusion of variables that can reflect the regional characteristics in more detail would complement the regional characteristics used in this study and lead to a better understanding of how to approach the situation. It is expected that such information would support the development of policies that suit each region according to their needs, which would result in a more effective management of the hypertensive diseases in South Korea.

References


