

고혈압, 당뇨, 뇌졸중 유병률에 대한 지역적 공간 자기상관 분석: 한국의 사례에 대하여

주성하¹, 노주환², 김창수², 허 준¹

¹연세대학교 토목환경공학과, ²연세대학교 의과대학 예방의학교실

Local Spatial Autocorrelation Analysis of 3 Disease Prevalence: A Case Study of Korea

Sungha Ju¹, Juhwan Noh², Changsoo Kim², Joon Heo¹

¹Department of Civil and Environmental Engineering, Yonsei University, Seoul; ²Department of Preventive Medicine, Yonsei University College of Medicine, Seoul, Korea

Objectives: This study aims to derive correlation between disease prevalence and geographical adjacency, by using global and local autocorrelation.

Methods: In order to derive the correlation, data provided by community health survey was utilized. The data contains disease prevalence rate for hypertension, diabete mellitus, stroke in 2012, covering the whole South Korea. Global autocorrelation analysis was implemented to derive the spatial characteristics of each disease prevalence rate, and local autocorrelation analysis was implemented to derive local spatial patterns of each disease prevalence rate. All the results are visualized into disease prevalence map. **Results:** All three diseases had significant spatial autocorrelation, and unique local clustering patterns were derived when local autocorrelation analysis was conducted. Spatial outliers, where disease prevalence rate was significantly different, were found and analyzed accordingly. **Conclusions:** The result of the study brought new insight towards spatial patterns of disease prevalence rate. The patterns of each diseases were unique, and spatial adjacency factor was found to be a grave influential factor in terms of disease prevalence rate. Also outlier regions, where disease prevalence rate is critically higher or lower and adjacent regions, were used for further analysis to figure out the reasons for disease prevalence. This study allows understanding of spatial characteristics of disease prevalence rate, thus enabling the spatial factors to be considered in terms of disease causation analysis, which can aid in decision making and resolving unbalanced medical service of community.

Key words: Hypertension, Diabetes, Stroke, Disease prevalence, Local spatial autocorrelation

서론

연구의 필요성

유병률은 지역사회의 건강 수준을 나타내는 척도 중 하나로, 지역 간 유병률의 변이는 사회행정, 공공의료 등의 분야에서 국민 건강관리의 기초 자료로 활용된다. 이에 따라, 질병관리본부에서는 지역사회건

강조사를 통해 국민의 건강행태, 건강검진 및 예방접종, 이환, 의료이용 등의 자료를 수집하여 일반 국민들에게 2008년부터 각 연도별로 공개해 왔다[1]. 여러 사회적 특성이 유병률을 포함한 사람들의 건강 수준에 관계가 있다는 연구가 보고되고 있으며[2,3] 사람들이 어디에 사는지에 대한 지역적 요소 또한 건강 수준과 관계가 있다는 연구가 보고되었다[4].

Corresponding author: Joon Heo

50 Yonsei-ro, Seodaemun-gu, Seoul, Korea

Tel: +82-2-2123-7479, E-mail: jheo@yonsei.ac.kr

Received: October 13, 2017 Revised: November 24, 2017 Accepted: November 26, 2017

No potential conflict of interest relevant to this article was reported.

How to cite this article:

Ju S, Noh J, Kim C, Heo J. Local spatial autocorrelation analysis of 3 disease prevalence: a case study of Korea. J Health Info Stat 2017;42(4):301-308. Doi: <https://doi.org/10.21032/jhis.2017.42.4.301>

© It is identical to the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permit sunrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

© 2017 Journal of Health Informatics and Statistics

질병의 발병 정보를 지도 위에 표현하고 분석하는 방법은 공공보건과 관련된 분야의 질병의 지역적 변이를 분석하는 데 있어 기본적인 방법으로[5], 각종 질병에 대한 변이 요인을 공간적 분석 방법론을 이용하여 분석한 연구가 활발히 진행되고 있다. Rosenberg et al. [6]은 전역적 공간 자기상관(global spatial autocorrelation) 분석 기법 및 지역적 공간 자기상관(local spatial autocorrelation) 분석 기법을 통하여 서부 유럽 40종의 암으로 인한 사망의 공간적 연관성을 분석하였으며, 그 결과 서부 유럽의 암 사망률이 강하게 공간적으로 연관되어 있음을 확인하였다. Mohebbi et al. [7]은 이란 카스피 지역 식도암의 지리적 변화를 모델링하기 위해 베이지안 기법을 통한 선형 모델을 도출하는 연구를 진행하였으며, Al-Anmadi and Al-Zahrani [8]은 공간 분석 방법인 지리적 가중회귀분석(geographically weighted regression)을 통하여 사우디아라비아의 암 발병률의 공간적 연관성을 도출하였고, 여성의 유방암과 남성의 전립선 암 등이 유의하게 공간적으로 상관이 존재하는 것을 확인하였다.

이러한 질병 발병의 공간 상관 분석은 암뿐만 아니라 고혈압, 당뇨, 뇌졸중 등의 심뇌혈관 질병에 대해서도 진행되어 왔다. Hipp and Chalise [9]의 연구에서는 미국의 당뇨병 발병률과 지리적 환경변수 사이의 연관성을 분석하기 위하여 지리적 가중회귀분석을 사용하였으며, 그 결과로 빈곤 수준, 유색 인종 비율, 신체 활동의 정도, 걷기 또는 자전거 타기 등의 환경변수가 당뇨병과 지역적으로 연관되어 있는 인자로 도출되었다. Green et al. [10]은 캐나다 위니펙의 당뇨 발병의 지역적 분포와 관련된 환경 인자들을 도출하기 위하여, 공간 분석 방법인 공간 통계 값(Space Scan Statistics)을 사용하였다. 또한 Adzandeh et al. [11]은 나이지리아 오순 주의 뇌졸중 유병률의 공간 패턴을 확인하기 위하여 GIS (Geographical Information System) 기법을 사용하였으며, Soljak et al. [12]은 영국의 심혈관 질환(Coronary Heart Disease), 고혈압, 뇌졸중 유병률의 지역적 변이를 분석하기 위하여 Local Moran's I 과 지리적 가중회귀분석을 수행하였다.

국내에서도 여러 질병을 공간 분석 방법을 통해 변이요인을 도출하는 연구가 진행되고 있다. Seok and Kang [13]은 변이 계수를 이용하여 고혈압 유병률의 지역 간 변이가 존재함을 확인하였으며, 지리적 가중회귀분석을 통하여 고혈압 유병률과 연관되어 있는 인자로 고위험 음주율, 중증도 이상 신체활동 실천율, 비만율, 스트레스 인지율이 도출되었다. Kim et al. [14]의 연구에서는 지리적 가중회귀분석을 포함한 여러 회귀분석과 의사결정나무를 이용하여 다년간의 고혈압 유병률의 지역적 변이를 파악하고, 지역 간 차이를 규명하고자 하였다. Oh et al. [15]은 전역 공간자기상관 분석의 척도인 Global Moran's I을 통하여 고혈압, 당뇨병 등 24종의 질병 유병률의 공간 패턴을 분석하였으며, 총 12종의 질병이 공간적으로 유의하게 연관되어 있음을 확인하였다.

연구목적

그동안의 선행연구에서는 공간 분석 수단으로 공간 자기상관도가 대표적으로 사용되어 왔으며, 지역의 변이를 파악하기 위한 지역적 공간 자기상관이나, 지리적 회귀분석 방법이 많이 사용되었다. 이들 연구의 결과는 질병들의 공간적 패턴이 지역적인 변이요인과 연관되어 있음을 나타낸다. 이는 특정 지역과 그 인접지역에서 특정 질병의 유병률이 공간적 요소로 인하여 유사한 공간패턴을 보일 수 있으며, 반대로 특정 지역과 인접한 지역이라도 서로 상이한 유병률 공간 패턴을 보일 수 있음을 말한다. 특히 지역적 공간 자기상관 분석은 질병의 공간적 연관성에 대한 탐색과 가시적인 질병의 공간 군집 패턴을 확인할 수 있다는 장점을 가지고 있다. 이를 통하여 서로 연관성이 존재하는 질병들이 비슷한 공간 분포를 가지는지, 아니면 서로 다른 고유한 공간 분포를 가지는지 가시적으로 확인할 수 있다. 따라서 본 연구는 전역적 공간 자기상관과 지역적 공간 자기상관을 모두 활용하였으며, 특히 그동안의 지역적 공간 자기상관 분석을 사용한 선행연구에서 주목하지 않았던, 주변 지역과는 상이한 유병률 경향을 보이는 이상 지역을 파악하고 이들 이상 지역이 가지는 특성을 도출하고자 하였다.

이에 따라 본 연구에서는 2012년 지역사회건강조사에서 발표한 전국 시군구별 유병률 자료 중 대표적인 성인 질환이자 질병 간 연관성이 존재하는 고혈압, 당뇨병, 뇌졸중의 유병률 데이터로부터 각 질병이 가지는 독특한 공간적 패턴을 도출하기 위하여 전역적 공간 자기상관 분석 방법을 통하여 각 질병의 대한 공간적 상관성을 확인하였으며, 지역적 공간 자기상관 분석 방법을 통하여 지역적으로 유의한 질병 클러스터(cluster) 지도를 도출하였다. 도출한 클러스터 지도를 통하여 각 질병이 가지는 공간 패턴을 확인하였으며, 주변 지역과는 다른 유병률 경향을 보이는 이상 지역을 도출하고 이들 이상 지역이 가지는 주변 지역과는 다른 특성을 살펴보았다. 본 연구의 결과는 주변 지역과 다른 발병 패턴을 보이는 지역에 대한 원인 분석에 대한 기초 자료로 활용될 수 있으며, 나아가 질병의 공간적 발병 원인을 규명하기 위한 자료로 활용될 수 있을 것이다.

연구 방법

연구 자료

유병률은 단위 인구 당 질병을 가진 사람의 비율로서, 지역 간 유병률의 변이 정도는 국가의 국민 건강관리의 기초 자료로 활용된다. 본 연구에서 사용한 유병률 데이터는 2012년 지역사회건강조사 자료를 활용하였다. 지역사회건강조사 자료는 2008년부터 보건복지부와 질병관리본부에서 매년 전국적으로 조사하여 발표하며, 지역 주민의 건강행태, 의료 이용, 사회 경제적 활동 등 총 18개 영역, 127개의 지표에 대

Table 1. List of disease prevalence and basic statistics

Disease	Mean	SD	Min	Max
Hypertension	176.04	59.19	85.01	583.36
Diabete mellitus	66.83	22.75	29.85	194.94
Stroke	84.30	24.84	25.62	216.69

SD, standard deviation.

하여 조사를 수행, 이를 표준화하여 제공한다[1]. 2012년 지역사회건강조사 유병률 자료는 2010년 총조사인구의 시군구별 인구 비율을 적용하여 2012년 시군구별 인구를 계산하고, 이를 바탕으로 표준화된 유병률을 제시하고 있다.

본 자료의 유병률 값은 각 시군구별 인구 1,000명당 유병률을 나타내며, 공간 상관 분석을 위하여 전국 230개 시군구 단위 지도 위에 표현하였다. 본 연구에서 활용한 질병 유병률 목록 및 각 질병별 유병률 수치의 대략적 분포를 파악하기 위한 기초 통계량은 Table 1과 같다.

연구 도구

전역적 공간 자기상관

공간 자기상관은 주어진 변수와 주어진 변수의 위치 사이의 상관관계를 규명하는 것이다. 다시 말해서, 공간 자기상관은 변수와 위치 사이에 특정한 패턴이 존재하는지 혹은 특정 변수가 그 위치에서 높게 나타나고 주변의 값 또한 높은 값을 나타낼 때, 이러한 위치와 변수 간의 유사성을 의미한다[16]. 예를 들어 양의 공간 자기상관도는 한 지역의 높은 값이 이웃 지역의 높은 값과 관련이 있을 때 발생한다. 이러한 공간 자기상관을 측정하는 대표적인 방법은 Global Moran's I로, 이 방법은 변수와 공간적으로 주변 지역 변수들의 가중 평균 간의 외적 연산으로 공간 상관도를 계산하며 다음과 같다[17].

$$I_{global} = \frac{n \sum_{i=1}^n \sum_{j=1}^n W_{i,j} (x_i - \bar{x})(x_j - \bar{x})}{(\sum_{i=1}^n \sum_{j=1}^n W_{i,j}) \sum_{i=1}^n (x_i - \bar{x})^2} \dots \text{(Equation 1)}$$

n: 변수 x의 개수, \bar{x} : 변수 x의 평균, j: 특정 지역 i의 인접 지역, $W_{i,j}$ 공간 가중행렬

$W_{i,j}$ 는 공간 가중행렬로서, 대부분의 공간적 연관성에 대해 다루는 분석들은 특정 지역과 그 인접 지역 사이의 관계를 표현하기 위하여 공간 가중행렬을 사용한다[16]. 본 연구에서는 인접도 기반의 공간 가중행렬을 사용하였는데, 인접도 기반의 공간 가중행렬은 주어진 변수가 속한 폴리곤과 인접한 주변의 모든 인접 폴리곤을 고려할 수 있기 때문이다[18]. 전역적 공간 자기상관도를 계산하기 위해 사용한 인접도

기반 공간 가중행렬은 Queen contiguity matrix이며, 이는 본 지역과 주변 지역의 폴리곤의 경계를 공유하거나, 모서리를 공유하는 경우 인접 지역으로 판단한다.

Global Moran's I 값은 -1에서 1 사이의 값을 가지며 1에 가까울수록 공간 분포가 균집을 이룬다는 의미이며, -1에 가까울수록 공간 분포가 분산되어 있다는 것을 의미한다. 0에 근접한 경우, 특정 공간 패턴이 나타나지 않은 무작위 공간 분포를 말한다. Moran's I은 통계적 유의성 검정을 통하여 통계적으로 유의한 값인지를 판단한다. 이는 계산된 I 값이 무작위로 형성된 분포에 의한 것이라는 귀무가설을 기각함으로써 통계적 유의성이 검정되는데, 이때 검정의 기준으로 표준화 점수(z-score)가 사용된다[19]. 예를 들어, 신뢰수준 95%에서 통계적으로 Moran's I이 유의한 값인가를 판단하기 위해서는 표준화 점수가 -1.96보다 작거나 1.96보다 커야 한다. 본 연구에서는 신뢰수준을 95%로 설정하여 전역 공간 자기상관도를 계산하였다.

지역적 공간 자기상관

앞에서 서술하였듯이, 전역 공간 자기상관은 공간상에 분포하는 데이터 전체의 균집화 정도를 나타낸다. 즉, 고혈압 유병률이 한국 전역에서 공간적인 패턴이 나타나는지 그렇지 않은지를 나타내게 된다. 그러나 전역 공간 자기상관으로는, 고혈압 유병률이 어느 지역이 균집을 이루는지, 어느 지역이 주변지역의 경향과는 다른 패턴을 보이는 이상 지역인지, 혹은 수도권 지역에 나타나는 지역적 균집이 어떤 종류를 가지는지에 대해서는 알 수 없다. 따라서 본 연구에서는 유병률의 공간 균집이 어느 지역에서 나타나는지 파악하고, 질병의 종류에 따라 지역별로 서로 상이한 패턴과 유병률의 공간적 이상 지역을 발견하기 위하여 지역적 공간 자기 상관 분석을 사용하였다.

Local Indicator of Spatial Association (LISA)는 Moran's I의 지역적 버전으로, 지역적으로 유의한 공간 균집을 제공하며 공간상에서 주어진 변수의 지역적 균집의 존재 유무를 나타내 변수와 관련된 현상을 분석하는 데 있어 지역적 공간 자기상관을 고려할 수 있게 해준다[19]. LISA는 i지역의 변수 x에 대하여 다음과 같이 표현된다[20].

$$I_{local}^i = \frac{x_i - \bar{x}}{S_x} \sum_{j=1, i \neq j}^n \frac{W_{i,j} (x_j - \bar{x})}{S_x} \dots \text{(Equation 2)}$$

n: 변수 x의 개수, \bar{x} : 변수 x의 평균, S_x : 변수 x의 분산, j: 특정 지역 i의 인접 지역, $W_{i,j}$ 공간 가중행렬

LISA 분석을 위한 공간 가중행렬은 전역 공간자기상관과 마찬가지로 인접도 기반 공간 가중행렬인 Queen contiguity matrix를 사용하였다. 그러나 전국 시군구 폴리곤 중 울릉도와 같이 주변에 인접한 행정

구역이 없는 시군구가 8개 존재하므로, 본 연구에서는 인접 행정구역이 존재하지 않는 8개 시군구는 지역적 공간 자기상관 분석에서 제외하였다. LISA의 결과는 무작위 반복 접근법을 기반으로 하여, 유의수준에 따라 통계적으로 유의한 군집을 추출한다. 본 분석에서는 몬테카를로 시뮬레이션(Monte-Carlo simulation) 반복 숫자를 9999로, 유의수준은 $p < 0.05$ (95% 신뢰수준)로 설정하였으며, LISA 알고리즘을 제시한 Anselin이 개발한 공간 분석 소프트웨어인 GeoDa v1.6.7 [21]를 통해 LISA 분석을 수행하였다. LISA를 통해 추출된 공간 군집은 지역적 상관관계를 나타내는 high-high (HH), low-low (LL)와 지역적으로 주변 지역과 유사하지 않음을 나타내는 low-high (LH), high-low (HL)로 분류되며, 다음과 같은 성격을 가지고 있다.

HH 군집: 본 지역의 변수가 높은 값을 가지면서 주변의 값 또한 높은 경향을 보임.

LL 군집: 본 지역의 변수가 낮은 값을 가지면서 주변의 값 또한 낮은 경향을 보임.

LH 군집: 본 지역의 변수가 낮은 값을 가지는데 반하여 주변의 값은 높은 경향을 보임.

HL 군집: 본 지역의 변수가 높은 값을 가지는데 반하여 주변의 값

은 낮은 경향을 보임.

HH 군집의 경우 핫스팟(hot-spot)으로, LL 군집의 경우는 콜드스팟(cold-spot)으로 볼 수 있으며, 두 경우는 주어진 변수가 공간적으로 양의 연관성을 가진다는 것을 말한다. LH 군집과 HL 군집은 이상 지역(outlier), 또는 고립지역으로 볼 수 있으며, 이 경우에는 공간적으로 음의 연관성을 가진다고 볼 수 있다. 따라서 LISA 방법론은 위치를 기반으로 한 데이터의 특이 지역을 파악하고 공간 분포의 이질성 정도를 평가하는데 알맞은 도구이며, 본 연구에서는 이를 활용하여 유병률 데이터의 지역적 군집 분포 패턴 및 이상 지역을 파악하였다.

연구 설계

본 연구에서는 고혈압, 당뇨병, 뇌졸중 유병률의 공간적 상관성을 확인하고자 전역 공간 자기상관을 파악하는 대표적인 방법인 Global Moran's I을 계산하여 각 질병이 공간적으로 연관되어 있는지 확인하였다. 더 나아가, 각 질병의 지역적인 유병률 군집 지역과 군집의 특징을 파악하기 위하여 지역적 공간 자기상관 분석을 수행하였다. 그 결과로 3종의 질병에 대한 LISA 군집 지도를 도출하였으며, 질병별로 높거나 낮은 유병률이 모여 있는 군집 지역을 확인하고 주변 지역과는 다

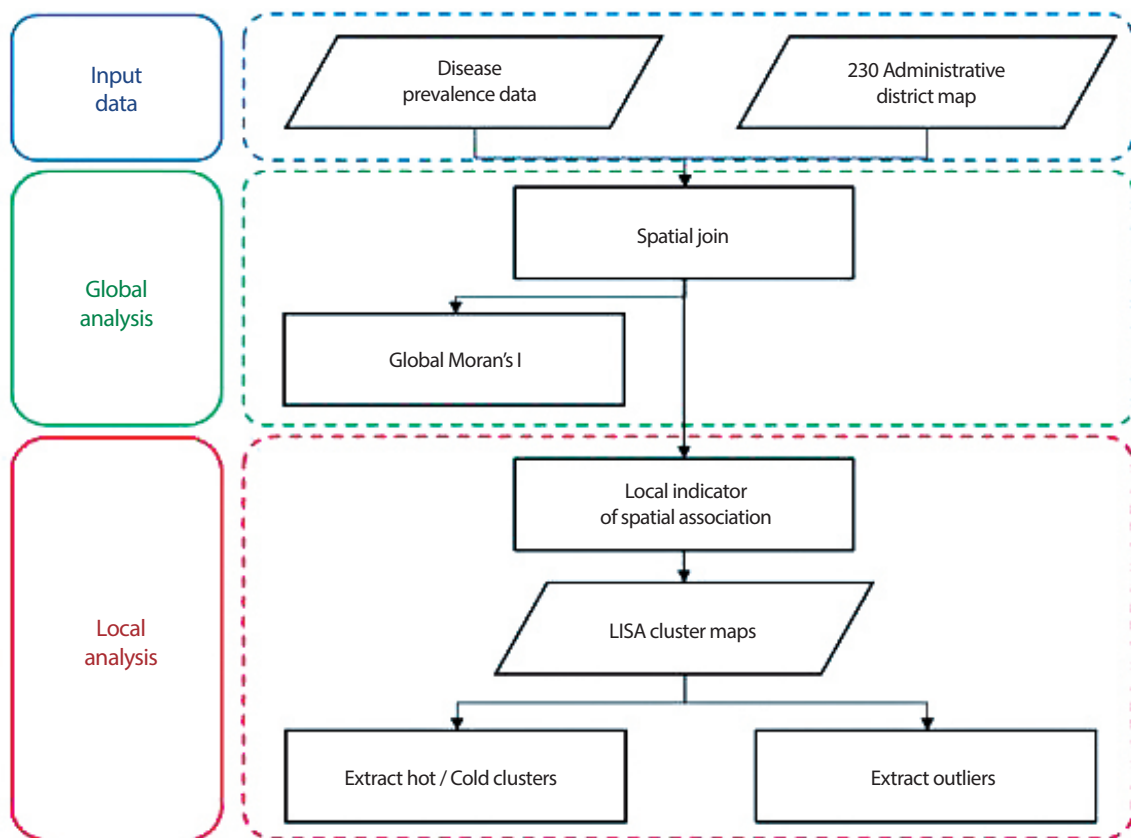


Figure 1. Flowchart.

른 경향을 보이는 이상 지역을 확인하였다. Figure 1은 본 연구의 흐름도로, 각 단계를 통하여 유병률의 공간적 상관성을 확인하고, 유병률 공간 군집 지도를 도출하는 과정을 표현하였다.

연구 결과

전역적 공간 자기상관 분석 결과

유병률과 공간 간의 상관성을 확인하기 위하여, 전국 230개의 시군구의 고혈압, 당뇨병, 뇌졸중 유병률에 대하여 Global Moran's I을 계산하였다(Table 2). 그 결과 세 질병 모두 표준화 점수가 1.96 이상으로, 95% 신뢰수준에서 통계적으로 유의한 결과값으로 도출되었다. 이는 세 질병들이 공간적으로 군집 패턴을 가지고 있으며, 각 질병 유병률과 공간 간의 상관성이 존재한다는 것을 말한다. 또한 세 질병 모두 Global Moran's I이 양수이므로, 공간적인 유병률의 분포가 군집 형태를 이룬다는 것을 확인하였다.

전역적 공간 자기상관 분석의 결과로 각 질병이 공간적으로 유의한 상관성이 존재하며 공간 군집 패턴을 가지고 있다는 것을 알 수 있었으나, 세 질병 유병률이 지역적 요소와 연관성이 있다는 사실만 확인할

수 있으며, 어떠한 성격을 군집을 이루고 있는지 시각적으로 확인할 수는 없었다. 또한 어느 지역에 공간 군집이 위치하고 있는지 확인할 수 없으며, 공간 군집 패턴과는 다른 이상 지역 또한 확인할 수 없었다. 이에 따라, 지역적 공간 자기상관 분석을 진행하여 세 질병 유병률의 실제 공간 군집의 분포 및 공간 군집의 성격과 이상 지역을 확인하였다.

지역적 공간 자기상관 분석 결과

공간적 군집 형태를 이루고 있는 것으로 파악된 세 질병에 대하여, LISA 방법론을 이용해 지역적 공간 자기상관 분석을 수행하였다. LISA 방법론의 결과로 각 질병마다 고유한 군집 지도를 도출하였으며, Figure 2는 세 질병에 대한 LISA 군집 지도이다.

빨간색 지역(HH 군집)들은 해당 유병률이 그 지역에 높게 나타나며, 주변의 지역 역시 같이 해당 유병률이 높은 지역을 의미한다. 파란색 지역(LL 군집)들은 그 지역의 유병률이 낮으며, 주변의 유병률 또한 낮다는 것을 말한다. 하늘색 지역(LH 군집)들은 그 지역 유병률은 낮으나, 주변 지역의 유병률이 높은 이상 지역을 말한다. 마찬가지로 분홍색 지역(HL 군집)들은 그 지역 유병률은 높으나, 주변 지역의 유병률이 낮은 이상 지역을 말한다. 지도에서 유병률의 지역적 공간 자기상관이 통계적으로 유의하지 않은 지역들은 흰색으로 표현하였으며, 인접지역이 존재하지 않는 일부 섬 지역들은 회색으로 표현하였다.

본 LISA 군집 지도로부터, LH 군집과 HL 군집, 즉 주변 지역과는 다른 유병률 패턴을 보이는 이상 지역들을 추출하였다. 이들 이상 지역은 주변 지역의 유병률과는 통계적으로 유의하게 반대되는 경향성을 보

Table 2. Results of global Moran's I statistic

Disease	Global Moran's I	z-score
Hypertension	0.338	8.209
Diabete mellitus	0.315	7.587
Stroke	0.386	9.267

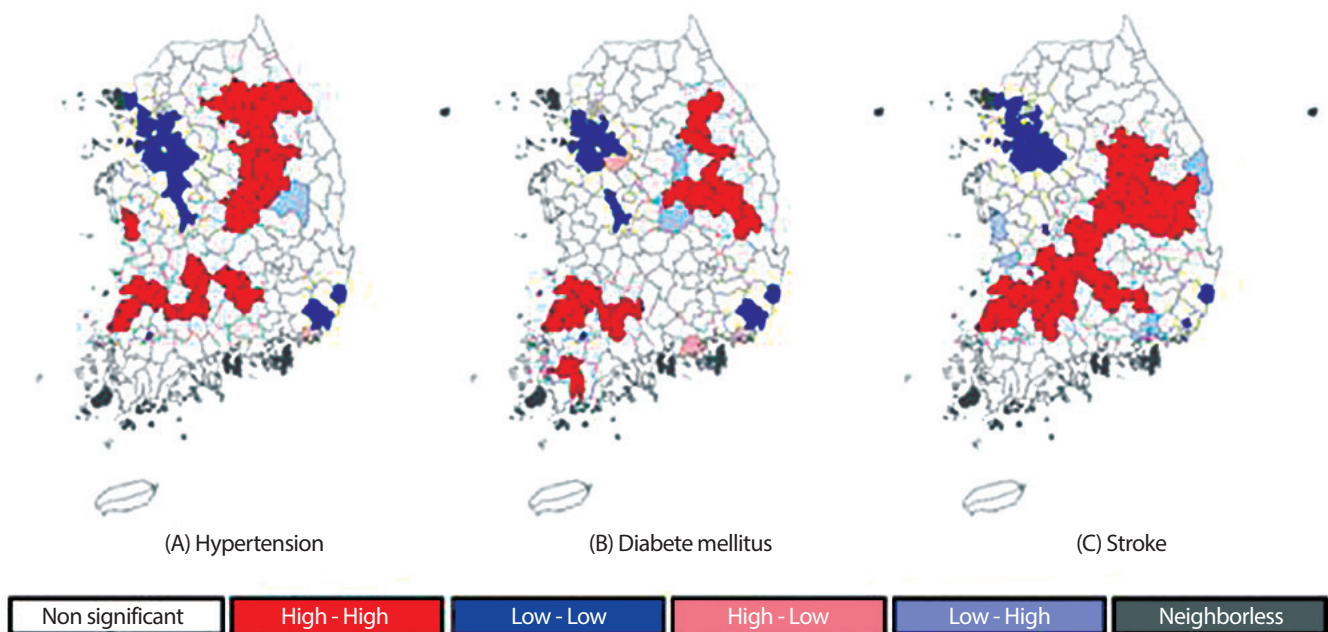


Figure 2. LISA cluster maps of three disease prevalence.

Table 3. Outlying regions for three disease prevalence

Disease	High-low outlier	Low-high outlier
Hypertension	Busan Gangseo-gu, Daegu Jung-gu	Gyeongsangbuk-do Andong-si
Diabete mellitus	Seoul, Yongsan-gu, Gyeonggi-do Anseong-si, Daegu Jung-gu, Gyeongsangnam-do Goseong-gun, Busan Gangseo-gu	Chungcheongbuk-do Jecheon-si, Gyeongsangbuk-do Sangju-si
Stroke	Seoul Gangbuk-gu, Seoul Geumcheon-gu	Chungcheongnam-do, Boryong-si, Jeonlabuk-do Gunsan-si, Gyeongsangbuk-do Uljin-gun, Gyeong-sangnam-do Chang-won-si

이는 지역들로, Table 3에 명시하였다. 각 질병 유병률은 1개 이상의 HL 이상 지역과 LH 이상 지역을 가지고 있었으며, 이상 지역 중 대구광역시 중구는 고혈압과 당뇨에게서 같은 HL 이상 지역으로 나타났다.

고찰

Figure 2의 (A) 고혈압은 대도시인 수도권과 부산에 LL 군집이 형성된 것을 확인할 수 있다. LL 군집은 고혈압 유병률이 낮은 지역들의 군집을 뜻하므로, 대도시 지역이 고혈압의 유병률이 낮다는 것을 말한다. 반대로, 내륙 지방에는 HH 군집이 형성되었는데, 이는 상대적으로 농촌 지역에서는 고혈압의 유병률이 지역적으로 유의하게 높다는 것을 말한다. 이러한 고혈압 군집 분포는 지역의 도시화 정도로 설명할 수 있다. Kim et al. [14]은 우리나라의 고혈압의 유병률의 지역 간 변이 요인을 분석하였는데, 인구밀도가 낮을수록 고혈압 유병률이 높고 경제활동비율이 높을수록 고혈압 유병률이 낮은 경향을 확인하였다. 이는 도시화 정도라는 지역적인 특징이 고혈압 군집의 분포와 연관되어 있음을 나타내며, 본 연구에서의 군집 분포와 일치하는 결과이다.

당뇨의 경우에도 앞의 고혈압과 비슷한 군집의 패턴을 가지고 있는 것으로 확인되었다. 대도시가 위치하는 지역들은 LL 군집들이 관찰되었고, 도시화 정도가 낮은 지역은 HH 군집들이 위치하고 있는 것을 확인할 수 있다. Hwang and Shon [22]은 우리나라의 당뇨 유병률과 사회·경제적 인자 사이의 연관성을 분석하였는데, 낮은 수입과 낮은 교육 수준이 높은 당뇨 유병률에 연관된 인자로 보고되었다. 낮은 수입과 낮은 교육 수준은 지역의 도시화와 연관성이 높은 인자들로, 본 연구의 당뇨 유병률 군집 패턴의 결과와 상응하는 연구결과이다. 마찬가지로 뇌졸중의 경우에도 지역의 도시화와 연관된 인자인 지역의 빈곤도와 교육 수준이 뇌졸중 사망률의 차이를 만드는 것으로 나타났으며

[23], 이러한 세 질병의 유병률 군집 분포의 특징은 지역적인 요소가 반영된 결과로 볼 수 있다. 질병 특성과 공간 군집 사이에는 연관성이 존재하며, 지역적인 특징에 따라 서로 다른 군집의 분포가 형성되고, 비슷한 지역적 특징을 가진 지역들은 유사한 공간 군집이 형성되었다고 판단된다.

한편, 이상 지역은 공간적 패턴이 뚜렷한 질병임에도 불구하고 주변 지역의 유병률 경향성과 반대되는 지역으로, 세 질병 유병률은 모두 LH, HL 이상 지역이 하나 이상 존재하였다. 지역적 특징으로 인하여 유병률 군집이 유사하게 분포하는 것과는 다르게, 이상 지역은 주변 지역과 차별되는 특정한 요소가 반영되어 나타나는 결과이다. 따라서 각 질병의 HL 이상 지역들과 LH 이상 지역들은 본 지역들만이 가지고 있는 특정한 요소가 있으며, 이러한 요소가 발현되어 주변 지역과는 다른 패턴을 보이는 것이라 설명할 수 있다. 본 분석을 통해 도출된 이상 지역들을 분석한다면 특정 질병의 발병을 증가 또는 억제시키는 요소를 도출할 수 있을 것이라 판단된다.

본 연구에서 확인된 이상 지역들 중, 부산광역시 강서구와 대구광역시 중구가 동일하게 고혈압과 당뇨병의 경우에 대하여 HL 이상 지역으로 나타났다. 앞서 언급한 것처럼 이상 지역은 주변 지역과 차별되는 특정한 요소가 반영되어 나타나는 결과이므로, 부산광역시 강서구와 대구광역시 중구는 두 질병 모두에 대하여 주변 지역과는 차별되는 특정 요소가 작용하여 이상 지역으로 나타났다고 해석할 수 있다. 이는 두 지역이 이상 지역으로 나타내게 되는 공통된 원인이 존재할 수 있다는 것을 말하며, 두 지역의 공통된 요소들과 주변 지역들과는 차별되는 요소들을 분석한다면 고혈압과 당뇨병의 공통된 발병 원인을 도출할 수 있을 것이라 판단된다.

전체적인 이상 지역의 분포가 산발적인 것과는 대조적으로 뇌졸중의 경우 LH 이상 지역인 충청남도 보령시, 전라북도 군산시, 경상북도 울진군, 경상남도 창원시 지역 모두 바다와 접해 있는 해안 지역으로 나타났다. 대부분의 해안 지역이 특정 패턴을 가지는 것은 아니지만, 뇌졸중의 LH 이상 지역이 모두 해안 지역이라는 특징은 뇌졸중의 발병과 연관된 요소 중에 바다와 관련된 요소가 존재할 수 있다는 것을 뜻한다. 해당 이상 지역들의 공통된 특징을 분석함과 더불어 해안 지방이라는 지역적 요소를 고려한다면, 뇌졸중의 발병을 억제시키는 요소를 규명할 수 있을 것이라 판단된다.

한편, 세 질병의 유병률에 대한 선행연구의 결과에 의하면, 세 질병의 유병률에 연관되어 있는 인자로는 성·연령·비만도 등 개인 인자 [13,14,24]와 사회경제적 위치, 삶의 질 등의 사회·경제적 인자 [9,13,22-25] 등으로 나타났다. 그러나 본 연구는 질병의 유병률의 지역적 공간 분포를 확인하기 위한 공간 분석을 수행하였으나, 위 세 질병의 발병에 중요한 인자로 알려진 성, 연령 등의 인자는 표준화된 유병률 데이터를

사용하였으므로 분석에 반영되지 않았으며, 지역의 사회-경제적 통계 요소를 함께 분석하지 않은 한계점을 지닌다. 또한 공간 분석 시 지형적 특징인 지역 접근성, 기후, 고도 등을 활용하지 않은 인접도 기반의 단순 분석으로, 보다 심도 깊은 지역적 특징이 반영되기 위해서는 이러한 공간적 요소, 통계적 요소와 더불어 시군구 기반의 행정구역 경계보다 더 세분화된 경계를 기반으로 한 분석이 수행되어야 할 것이다.

결론

본 연구에서는 고혈압, 당뇨병, 뇌졸중 질병 유병률의 전역적 공간 상관성을 확인하고, 더 나아가 지역적 공간 자기상관 분석을 통하여 각 질병의 고유한 공간 패턴을 확인하고자 하였다.

그 결과로 세 질병 모두에게서 공간적 상관성이 존재한다는 것을 확인하였으며, 양의 공간 상관성으로 모두 군집 형태를 이루고 있음을 확인하였다. 지역적 공간 자기상관 분석을 통해 도출된 군집 지도를 통해 세 질병 모두에게서 도시화가 많이 진행된 지역인 수도권과 부산에서 주변보다 유의하게 질병의 유병률이 낮은 군집이 확인되었으며, 내륙 지방에 주변보다 유의하게 질병의 유병률이 높은 군집이 확인되었다. 또한 주변 지역과는 다른 이상 지역들을 확인하였는데, 고혈압은 부산광역시 강서구, 대구광역시 중구, 경상북도 안동시가 이상 지역으로, 당뇨병은 서울특별시 용산구, 경기도 안성시, 대구광역시 중구, 경상남도 고성군, 부산광역시 강서구, 충청북도 제천시, 경상북도 상주시가 이상 지역으로 나타났다. 뇌졸중의 경우 서울특별시 강북구와 금천구가 주변의 낮은 유병률에도 불구하고 높은 뇌졸중 유병률을 보이는 지역으로 나타났으며, 주변의 높은 유병률에도 불구하고 낮은 뇌졸중 유병률을 보이는 모든 이상 지역은 해안과 접해있는 해안 지역으로 나타났다.

이를 통하여, 지역적 특징이 비슷한 지역은 비슷한 패턴의 공간 군집이 형성된다는 것을 확인하였으며, 서로 연관성이 존재하는 질병들은 공간 군집의 분포가 유사하게 형성된다는 것을 확인하였다. 또한, 각 질병의 유병률이 주변 지역과는 다른 이상 지역을 도출하여 각 질병의 발병 원인을 분석하기 위한 후보지역들을 제시하고, 질병의 발병 원인을 분석하는 데 있어 이러한 지역적 요소를 고려하여야 한다는 것을 확인하였다.

REFERENCES

1. Korea Centers for Disease Control and Prevention. Community health survey. Available at <https://chs.cdc.go.kr/chs/index.do> [accessed on 13 October 2017].

2. Macintyre S, Ellaway A. Ecological approaches: rediscovering the role of the physical and social environment. *Social epidemiol* 2000;9(5): 332-348.

3. Takano T, Nakamura K. An analysis of health levels and various indicators of urban environments for Healthy Cities projects. *J Epidemiol Community Health* 2001;55(4):263-270.

4. Scholten HJ, de Lepper MJ. The benefits of the application of geographical information systems in public and environmental health. *World Health Stat Q* 1991;44(3):160-170.

5. Kandwal R, Garg PK, Garg RD. Health GIS and HIV/AIDS studies: perspective and retrospective. *J Biomed Inform* 2009;42(4):748-755.

6. Rosenberg MS, Sokal RR, Oden NL, DiGiovanni D. Spatial autocorrelation of cancer in Western Europe. *Eur J Epidemiol* 1991;15(1):15-22.

7. Mohebbi M, Wolfe R, Forbes A. Disease mapping and regression with count data in the presence of overdispersion and spatial autocorrelation: a Bayesian model averaging approach. *Int J Environ Res Public Health* 2014;11(1):883-902.

8. Al-Ahmadi K, Al-Zahrani A. Spatial autocorrelation of cancer incidence in Saudi Arabia. *Int J Environ Res Public Health* 2013;10(12): 7207-7228.

9. Hipp JA, Chalise N. Spatial analysis and correlates of county-level diabetes prevalence, 2009-2010. *Prev Chronic Dis* 2015;12(8):1-9.

10. Green C, Hoppa RD, Young TK, Blanchard JF. Geographic analysis of diabetes prevalence in an urban area. *Soc Sci Med* 2003;57(3):551-560.

11. Adzandeh AE, Awoppe J, Oviyasu OI. A preliminary spatial analysis of diagnosed stroke disease in Osun state, Nigeria. *Pan Afr Med J* 2016; 25:63.

12. Soljak M, Samarasinghe E, Indulkar T, Walford H, Majeed A. Variations in cardiovascular disease under-diagnosis in England: national cross-sectional spatial analysis. *BMC Cardiovasc Disord* 2011;11(1): 12.

13. Seok H, Kang S. A study on the regional variation factor of hypertension prevalence. *Health Social Welfare Review* 2013;33(3):201-236 (Korean).

14. Kim Y, Cho D, Hong S, Kim E, Kang S. Analysis on geographical variations of the prevalence of hypertension using multi-year data. *Korean Geog Soc* 2014;49(6):935-948 (Korean).

15. Oh W, Nguyen CH, Kim S, Sohn J, Heo J. Spatial autocorrelation of

- disease prevalence in South Korea using 2012 Community Health Survey Data. *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography* 2016;34(3):253-262 (Korean).
16. Anselin L, Cohen J, Cook D, Gorr W, Tita G. Spatial analyses of crime. *Criminal Justice* 2000;4(2):213-262.
17. Le Gallo J, Ertur C. Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe, 1980-1995. *Pap Reg Sci* 2003;82(2):175-201.
18. Seabrook NR. The Obama effect: patterns of geographic clustering in the 2004 and 2008 presidential elections. *Forum* 2009;7(2):6.
19. Erdogan S. Explorative spatial analysis of traffic accident statistics and road mortality among the provinces of Turkey. *J Safety Res* 2009;40(5):341-351.
20. Anselin L. Local indicators of spatial association-LISA. *Geog Anal* 1995;27(2):93-115.
21. Anselin L, Syabri I, Kho Y. GeoDa: an introduction to spatial data analysis. *Geog Anal* 2006;38(1):5-22.
22. Hwang J, Shon C. Relationship between socioeconomic status and type 2 diabetes: results from Korea National Health and Nutrition Examination Survey (KNHANES) 2010-2012. *BMJ Open* 2014;4(8):e005710.
23. Balamurugan A, Delongchamp R, Bates JH, Mehta JL. The neighborhood where you live is a risk factor for stroke. *Circ Cardiovasc Qual Outcomes* 2013;6(6):668-673.
24. Marshall JA, Hamman RF, Baxter J, Mayer EJ, Fulton DL, Orleans M, et al. Ethnic differences in risk factors associated with the prevalence of non-insulin-dependent diabetes mellitus: The San Luis Valley Diabetes Study. *Am J Epidemiol* 1993;137(7):706-718.
25. Dinca-Panaitescu S, Dinca-Panaitescu M, Bryant T, Daiski I, Pilkington B, Raphael D. Diabetes prevalence and income: results of the Canadian Community Health Survey. *Health Policy* 2011;99(2):116-123.