보건학 석사 학위논문

Spatiotemporal Analysis of National Scrub Typhus Occurrence: 2001-2010

2001-2010:

대한민국 쯔쯔가무시증 발생의 시공간 분석

2012 년 12 월

서울대학교 대학원

보건학과 보건학전공

김 지 혜

Spatiotemporal Analysis of National Scrub Typhus Occurrence: 2001-2010

지도 교수 조 성 일

이 논문을 보건학 석사 학위논문으로 제출함

2012년 10월

서울대학교 보건대학원
보건학과 보건학 전공
김 지 혜

김지혜의 보건학 석사 학위논문을 인준함

2012년 12월

위 원 장	정 해 원	(인)
부위원장	김 호	(인)
위 원	조 성 일	(인)

Abstract

Jihye Kim
Department of Public Health
The Graduate School of Public Health
Seoul National University

Orientia tsutsugamushi is a causative agent of scrub typhus, a zoonotic disease, distinguished by symptoms of sudden high fever and macular rash. The number of incidences had exploded in Korea over the past decade. This study was performed to investigate the spatial and temporal factors associated with scrub typhus occurrences in Korea. The study population of the study is all the scrub typhus cases reported within Korea from 2001 to 2010.

Among the scrub typhus cases, more cases were reported among females, older age groups, and agricultural and fishery occupational group. The study not only observed associations with scrub typhus incidence counts, but also with crude incidence rates and age-gender standardized incidence rates.

Moran's I and LISA were observed for each year to understand spatial autocorrelations. The proportions of dry-field use, paddy-field use, and orchard use were considered as areal characteristics, and were analyzed through spearman rank correlation test and negative binomial regression analysis. There were significant clustering of scrub typhus occurrences with positive linear trend and although all types of land use had significant results, highest correlation coefficient was observed between scrub typhus occurrences and the proportions of paddy field use in PHC level regions. The results from negative binomial regression analysis also support with significant level of results explaining the effect size of paddy field proportions.

Keywords : scrub typhus, spatiotemporal analysis
Student Number : 2011-22098

Table of Contents

List of Tables

List of Figures

Introduction

I-A. Scrub typhus in Korea

Scrub typhus is an infectious disease that occurs in a wide scope, across eastern Asia and the western Pacific region, called Tsutsugamushi Triangle[1,2]. Korea is located in the heart of this endemic window. It is known that major proportion of scrub typhus patients are engaged with farming and forestry industry. This suggests people in agricultural and forestry industry to have higher risk of being infected, and this argument has been supported by studies claiming 93.9% of occupational infectious diseases reported among Koreans in agricultural and fishery industry is scrub typhus [3,4].

The influence of scrub typhus had been investigated in a great extent from numbers of countries affected by the disease, but people are still prone to the infection since there is no effective vaccine developed up to this date due to diverse antigenic variations of the causative agent[5]. External preventive measures, such as limiting outdoor activities and wearing long sleeves and pants when outdoor activities are absolutely necessary, have been suggested along with early detection and treatment as secondary prevention[4,5].

Scrub typhus has been declared as class III nationally notifiable communicable disease in Korea. The diseases in class III are characterized as infectious diseases that have potentials for intermittent epidemic, therefore surveillance system to overlook occurrences, and development of preventive measures to minimize its effect are required. When diagnosed, physicians are obligated by law to report the cases of class III infections immediately to designated public health centers and the reports are accumulated at Korea Center for Disease Control and Prevention (KCDC) through hierarchical report system[6].

After having a scrub typhus patient reported from Japan in 1923, there was an endemic occurrence during World War II in Southeast Asia region affecting about 40,000 people. In Korea, a report of scrub typhus patient was found in 1927 near southern coast, and in 1951, 6 UN military officers were reported as scrub typhus patients. Cases of scrub typhus were under the radar until they re-appeared in 1986 and the number of patients had undergone explosive increase towards the end of 1990's[7].

I-B. Ecology of scrub typhus

Scrub typhus is an infectious disease, which occurs in specific regions, including India, Australia, Astrakhan (central Asia), and many countries in Asia, but most occurrences are reported from Southeast Asia. Main clinical symptom is high fever, and eschar, cough, poor appetite, lymphadenopathy, headache, rash, hepatomegaly, vomiting, drowsiness, abdominal pain, Splenomegaly, and neck stiffness have also been reported[8]. The incubation period of scrub typhus is known to be about 5-20 days, averaging 10-12 days after the first bite from rickettsia[9]. Although complete cure of scrub typhus is still uncertain, scrub typhus is considered to be curable with antibiotics[10]. However scrub typhus is characterized by vasculitis and perivasculitis, which may involve various organs, such as lung, liver, heart, spleen, and central nervous system[1]. In addition, encephalitis, interstitial pneumonia, myocarditis and pericarditis, acute renal failure, acute hepatic failure, and acute hearing loss are subjected to occur among scrub typhus patients and many studies have listed acute respiratory distress syndrome (ARDS) as a serious complication[11]. Moreover, there had been numbers of experiments that have shown persistence of viable *O. tsutsugamushi* observed for as long as 610 days after inoculation[10].

I-C. Associated factors with scrub typhus

The causative agent of scrub typhus is *Orientia tsutsugamushi*, which is a gram negative obligatory intracellular parasite, using chiggers as hosts. Cells of *Orientia tsutsugamushi* are short rods with diameters of 0.5-0.8μm and lengths of 1.2-3.0μm. The organisms are found only in the eastern hemisphere[12]. About 1 million new cases of scrub typhus are identified each year and 1 billion people are estimated to be at risk[13].

Besides ecological factors, studies have shown that many of the scrub typhus patients are reported from skilled agricultural, forestry and fishery occupational group, and this is also apparent in studies from Taiwan and

Japan[14,15]. Furthermore, a case-control study had revealed that patients involved with activities like gathering chestnuts, harvesting fruits, and harvesting in dry fields had increased odds ratio of becoming infected with scrub typhus[16].

In addition to individual factors, spatial correlations of scrub typhus occurrence were suggested after observing clustered occurrences in the eastern section of Taiwan[15].

I-D. Spatiotemporal relationship

The importance of spatial modeling has been gaining increased recognition within the field of public health due to its wide use understanding population at risk and risk factors through disease distribution simulations, leading to development of proper and effective interventions[17,18].

Factors that could influence differences among regions include environmental factors, host exposure, and pathogen viability. There had been studies within Korea encompassing environmental factors and pathogen viability. For associations between scrub typhus occurrence and environmental factors, one of the studies indicated that incidences of scrub typhus has significant correlations with summer average temperature and relative humidity[14]. For explanations to associations between scrub typhus occurrence and pathogen viability, a study showed that the population of *L. pallidum* and *L. scutellare* was high during months with high disease occurrence in Korea[19]. For explanations on host exposure, a study from Laos revealed that there was a spatial correlation with scrub typhus occurrences, having more people infected by scrub typhus in the periphery of the countries, whereas murine typhus occurrence was more prevalent in the central region[20]. Likewise, spatial analysis is not a novel approach in understanding characteristics of scrub typhus occurrences. Nonetheless, when compared with immense studies on ecological pathways of scrub typhus infections, epidemiological investigations on pathways of host exposure are still limited, especially in explaining spatial clustering of scrub typhus occurrences. A descriptive analysis of scrub typhus occurrence in Korea suggested that the areas with high scrub typhus incidences are likely to have high land-use of paddy-field and dry-field[7]. However statistical analyses on spatial characteristics that are associated with scrub typhus incidences are yet to be investigated.

I-E. Objectives

First goal of the study is to examine spatial and temporal trend of scrub typhus occurrences in Korea through statistical analyses.

Second goal of the study is to examine the associations between types of land use and scrub typhus occurrence.

With both goals achieved, the purpose of the paper is to understand spatial correlation of scrub typhus occurrence and spatial characteristics that influence the correlation, and further aid in developing effective preventive measures to protect the population, especially those who are at higher risk due to occupational characteristics.

Methods

II-A. Study population

All registered residents of Korea during the observation period of 2001 to 2010 were considered as potential study population. Among the population at large, people who had been reported to KCDC as scrub typhus patients during the period were considered as study population.

II-B. Data collection

This study is a population-based retrospective observation study, involving data collected through 10 years of observation period. The three sets of data utilized for the analysis were scrub typhus incidence database from KCDC, registered residence database from Korea Statistical Information Service (KOSIS), and area database from KOSIS and local autonomous government sectors. All data sets were organized with spatial information on 253 regional levels, each of which has public health centers (PHC), and this regional level is to be abbreviated as PHC level, hereafter.

Temporal occurrence data

Infectious disease database of KCDC and the database of national registered residents were utilized for disease occurrence analysis. Infectious disease database is established through national surveillance system, monitored by KCDC. The database contains accumulated information on cases of infectious diseases reported within Korea.

For each case, reported date was utilized to determine date of disease occurrence, and besides the reported date, gender, age, occupation, probable origin of disease, reported public health center, and reported region were taken into account to understand the characteristics of the study population.

National resident registry

National resident registry collected during the ten years of observation period was utilized to incorporate population factor to the analysis through calculating crude and age-gender standardized incidence rates. Each yearly dataset had numbers of registered residents on 253 PHC levels, distinguished by gender and 5-year interval age groups. The specific format of resident registry dataset was obtained in order to calculate age-gender standardized incidence rates. Crude incidence rate was calculated for each year on PHC level by dividing the total yearly occurrence of scrub typhus by the total numbers of registered residents of that year on corresponding PHC level. In addition, age-gender standardized incidence rate was calculated for each PHC level (methods for age-gender standardization are explained in Appendix B). Both rates were calculated per 100,000 persons.

Spatial data

Statistics Korea provides national areal data through KOSIS. There were two levels of regional divisions recorded for each data. The smallest regional division was categorized by dividing Korea into 253 regions, each of which has own PHC, where information on scrub typhus incidences are gathered and transferred to the province level, then to KCDC. The next level of regional division was categorized by dividing Korea into 16 regions, and this level is to be described as province level, hereafter.

In addition to areal information of patients, four sub-types of area data were utilized for spatial analysis. Four sub-types of area information contain PHC level measurements on whole land area, dry-field area, paddy-field area, and orchard area across the observation period, 2001-2010.

According to legislation from Ministry of Land, Transport, and Maritime Affairs, dry-field is defined as land that does not necessarily use water continuously to cultivate vegetation including grain, horticultural crop (excluding those that bear fruits), medicinal crop, seeding, ornamental trees, and edible bamboo sprouts; paddy-field is defined as land that continuously uses water to cultivate vegetation including rice, and lotus; and orchard is defined as land used to cultivate large quantities of fruits including apple, pears, chestnut, walnut tangerine and the land used to keep the fruits.[21]

Four sub-types of area data were first collected through KOSIS but if information on specific regions or time frame were unavailable, each local autonomous government sector was contacted to obtain information required for the analysis. After obtaining the data, the measurement units, i.e.km^2 and

$m^{2,}$ were unified, and percentage used for each land type was calculated by having whole land areas as denominators and each dry-field area, paddy-field area, and orchard field area of that region as numerators.

II-C. Statistical theory

In this study, explanatory variables and disease occurrences were explored through descriptive statistical methods, and relationships among the variables had been explored through analytical statistical methods.

Yearly, monthly, and regional distributions of Scrub Typhus were observed with reported incidence counts, and dependent variable was further investigated with calculating crude incidence rates per 100,000 person-year and age-gender standardized incidence rates per 100,000 person-year.

Age, gender, occupation of individual patients, and probable origin of infection were considered as basic demographic variables.

Primary outcome

Primary outcomes of the study were measures of spatial autocorrelations. For spatial analyses, geographic information system (GIS) technique was utilized for disease occurrence simulation across the nation, and in order to understand spatial autocorrelation, Local Indicators of Spatial Association (LISA) and Moran's I were observed for each year.

LISA is a statistical index, which indicates significant spatial clustering around geographical observations[22].

Moran's I is a measure of spatial autocorrelation, which is related to Pearson correlation coefficient[23]. Pearson correlation coefficient, denoted as *p,* is a measure of average association between x_i and y_i.

$$\rho = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{[\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2]^{1/2}}$$

[$\bar{x}$ and $\bar{y}$ represent the sample means]

Extending from the theories of Pearson correlation coefficient equation, Moran's I formula incorporate spatial characteristics through using weights, indicating neighboring function. The simplest form of weights application is setting w_{ij} = 1 for neighboring observations and 0 for others.

$$I = \frac{n}{S_0} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

[w_{ij} represent weight between observation i and j;

S_0 represent the sum of all w_{ij}]

Moran's I had been calculated for all three dependent variables, incidence counts, crude incidence rates, and age-gender standardized incidence rates. Linear regression was performed with Moran's I values calculated for each year to analyze the trend of Moran's I values throughout the observation period.

Secondary outcome

Spatiotemporal associations of three dependent variables, incidence counts, crude incidence rate, and age-gender standardized incidence rate, were further examined by observing correlation with three potential associated factors. These potential associated factors include proportions of dry-field land use, proportions of paddy-field land use, and proportions of orchard land use.

For each type of land use, proportions of land use was calculated by dividing areal values of each field type by whole areal values for each PHC level. In order to compare associations, Spearman rank correlation test was performed not only for three potential associated factors but also for combined proportions of dry-field and paddy-field, orchard and paddy-field, dry-field and orchard, and dry-field, paddy-field, and orchard. In addition, negative binomial regression coefficients on crude incidence rates and standardized incidence rates were obtained to observe extent of effects of proportions of land use on scrub typhus occurrence.

Utilities

SAS 9.3 was used for statistical analysis and Open GeoDa and Quantum GIS were used for geographical simulation and spatial analysis.

Results

III-A. Variables of Interest

During the ten years of observation period, a total of 46,671 cases of scrub typhus were reported nationwide [Table 1]. Most of the cases were estimated to have been infected within Korea. Among scrub typhus cases, 63.7% were reported from females, occupying about twice as many as that of the males.

The age variable had normal distribution with a mean value of 59.5 and standard deviation of 15.4. Cases are reported from all seven age groups, however, majority of cases were reported from patients who are 60 years old or above; in fact, 90.4% of the cases were reported from people who are 40 years old and over. This age distribution was apparent in both males and females.

A large proportion of patients had occupations in the industry of agriculture and fishery and it was consistent for both genders. The next most frequent occupation among females was housewives, followed by others and unemployed. For males, it was in the order of others, unemployed, and office works, which includes office jobs, service workers, sales associates, specialists, and legislators or government officials.

Table 1. Basic demographics of scrub typhus patients reported in Korea between 2001 and 2010

			N (%)
All			46,671
Probable origin of the infection			
Within Korea			46,665 (99.9)
Other countries			6 (0.1)
Gender			
male			16,941 (36.3)
female			29,730 (63.7)
Age	**All**	**Male**	**Female**
mean ± standard deviation	59.5 ± 15.4	57.4 ± 16.6	60.8 ± 14.5
0-19	1,031 (2.2)	556 (3.3)	475 (1.6)
20-29	1,046 (2.2)	577 (3.4)	469 (1.6)
30-39	2,041 (5.1)	1,170 (6.9)	1,231 (4.1)
40-49	6,128 (13.1)	2,468 (14.6)	4,660 (12.3)
50-59	9,836 (21.1)	3,376 (19.9)	6,460 (21.7)
60-69	13,240 (28.4)	4,578 (27.0)	8,662 (29.1)
70 +	12,989 (56.2)	4,216 (24.9)	8,773 (29.5)
Occupation	**All**	**Male**	**Female**
agriculture and fishery workers	16,736 (35.9)	6,547 (38.7)	10,189 (34.3)
housewives	7,653 (16.4)	39 (0.23)	7,614 (25.6)
others	11,056 (23.7)	4,936 (29.14)	6,120 (20.6)
unemployed	6,840 (14.7)	2,564 (15.13)	4,276 (14.4)
office workers *	2,122 (4.5)	1,336 (7.89)	786 (2.6)
students	864 (1.9)	487 (2.87)	377 (1.3)
indoor manual workers**	656 (1.4)	558 (3.3)	98 (0.3)
outdoor workers***	744 (1.6)	474 (2.8)	270 (0.9)

* office workers include office jobs, service workers, sales associates, specialists, and legislators or government officials
** indoor manual workers include assembly line workers, technicians, and machine operators
*** outdoor workers include construction workers and military force

III-B. Temporal analyses

Incidence counts

Monthly trend of each year reflects consistent peaks appearing in regular intervals. Although the heights vary, constant appearance of peaks during September to December each year suggests strong seasonal trend [Figure 1-1].

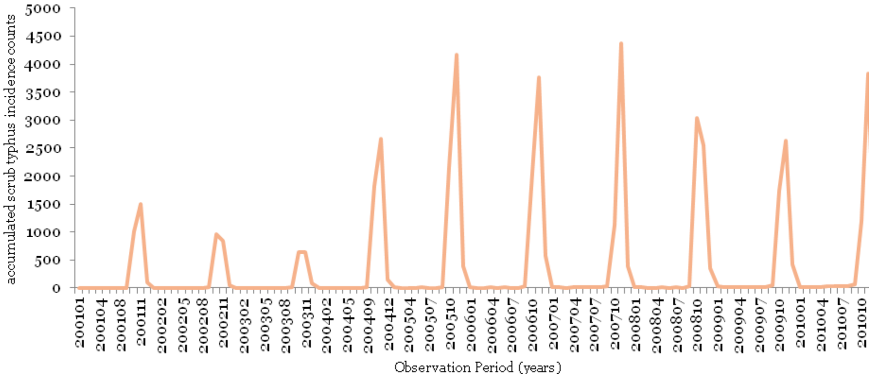


Figure 1-1. Monthly distribution of scrub typhus occurrence

Scrub typhus is reported throughout a year in Korea. When all the occurrences of the observation period are accumulated by months, most occurrences are reported between September and December [Figure 1-2]. The trend shows that the occurrence begins to rise in September, exploding in October and November, and after reaching the highest peak in November, the occurrence drastically drops in December.

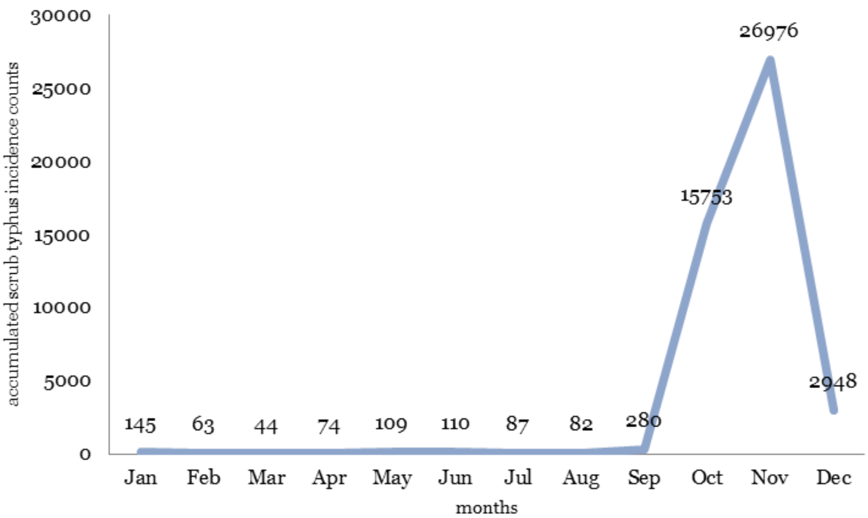


Figure 1-2. Monthly trend of accumulated scrub typhus occurrence

At the beginning of the observation period, annual number of reports started with 2,637 cases and gradually decreased until 2003, dropping down to 1,415 cases reported per year. However, the occurrence sharply inclines close to 5,000 cases in 2004, peaking in 2005, then reaching plateau from year 2006. Scrub typhus occurrence in 2005 was about 4.8 times the occurrence in 2003, having 5,000 case differences between the years. Although acute increase did not appear after 2005, data from 2010 is still nearly 4 times the data from 2003 [Figure 1-3].

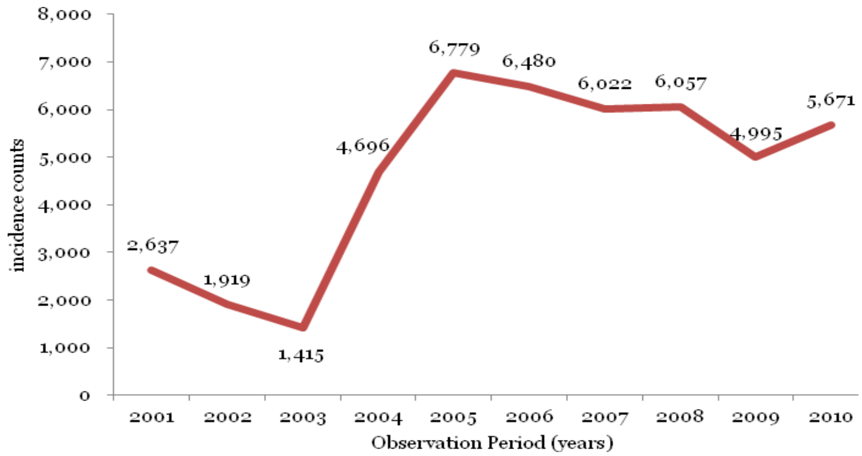


Figure 1-3. Yearly trend of scrub typhus incidence counts

Crude and age-gender standardized incidence rates

Similar trend was observed with yearly trend of crude incidence rates [Figure 1-4]. Although age-gender standardized incidence rates displayed similar trend, the values differed slightly when compared with the trends with crude incidence rates and incidence counts [Figure 1-5].

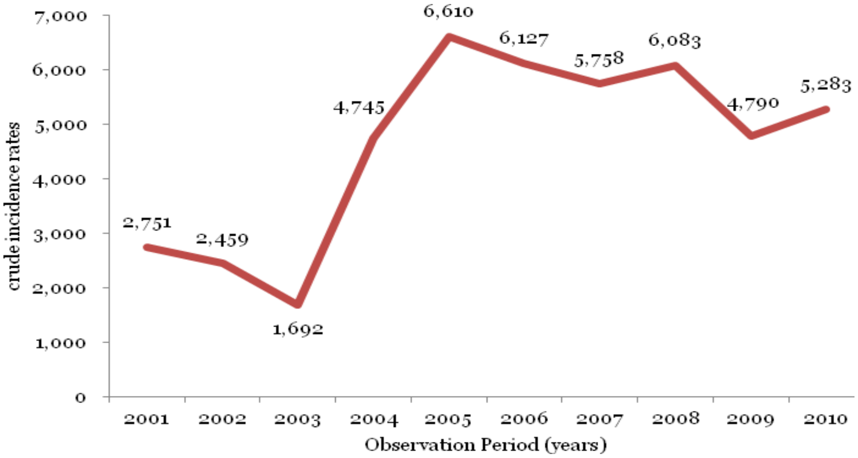


Figure 1-4. Yearly trend of scrub typhus crude incidence rates

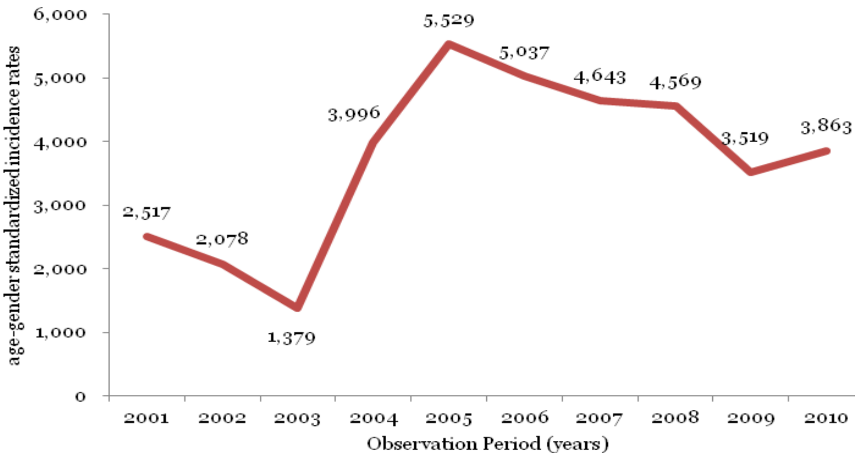


Figure 1-5. Yearly trend of age-gender standardized scrub typhus incidence rates

Regional distribution was examined by dividing the accumulated numbers of incidence reports into sixteen provincial level regional divisions [Figure 1-6]. Visible differences were observed between regions. Jeju had the lowest reports, followed by Gangwon and Incheon while Jeonbuk had the highest reports, followed by Chungnam and Gyeongnam. Three of the highest incidence regions are geographically located nearby, suggesting spatial association with high scrub typhus incidence.

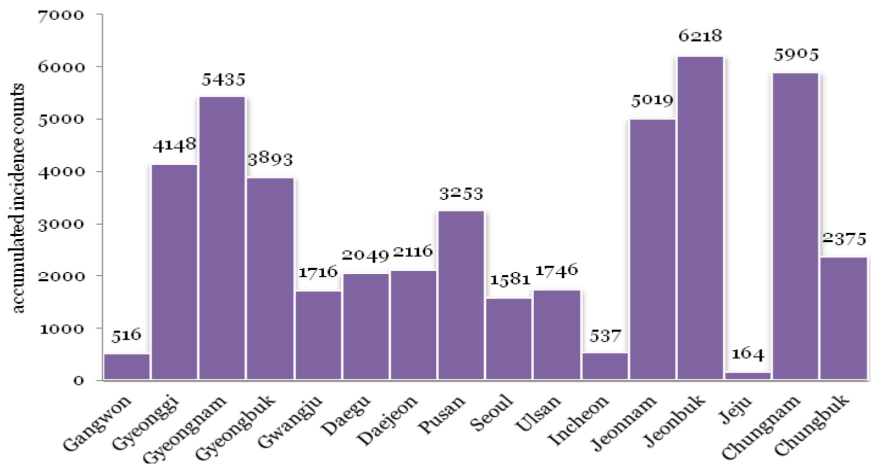


Figure 1-6.Accumulated regional distributions of scrub typhus occurrence

III-C. Spatial analyses

GIS simulation

First step of spatial analyses was to observe GIS simulation on disease occurrence distribution across the nation. GIS simulation was obtained for all three dependent variables, incidence counts, crude incidence rates, and age-gender standardized incidence rates.

For GIS simulation with incidence count data, the map exhibited lighter tone for years between 2001 and 2003 since the numbers of incidence was very low compared to the years to come [Figure 2-1]. As number of incidence increases, disease occurrence became more apparent in the lower left half of the nation with specific dark regions. However, beginning from 2006, colors of some darkest regions began to fade away.

For GIS simulation with crude incidence rates, the color begins to appear more readily beginning in 2004, when the incidence counts sharply arose [Figure 2-2], and colors became more apparent in the lower central regions of the nation. Also, some of the darkest regions shown with incidence count data were not the darkest regions for crude incidence count data.

For GIS simulation with age-gender standardized incidence rates, the colors were even more apparent between 2001 to 2003 than the maps of crude incidence rates and incidence count data during the time frame [Figure 2-3]. Beginning from 2004, lower left side of the nation had more colors than the upper right side of the nations.

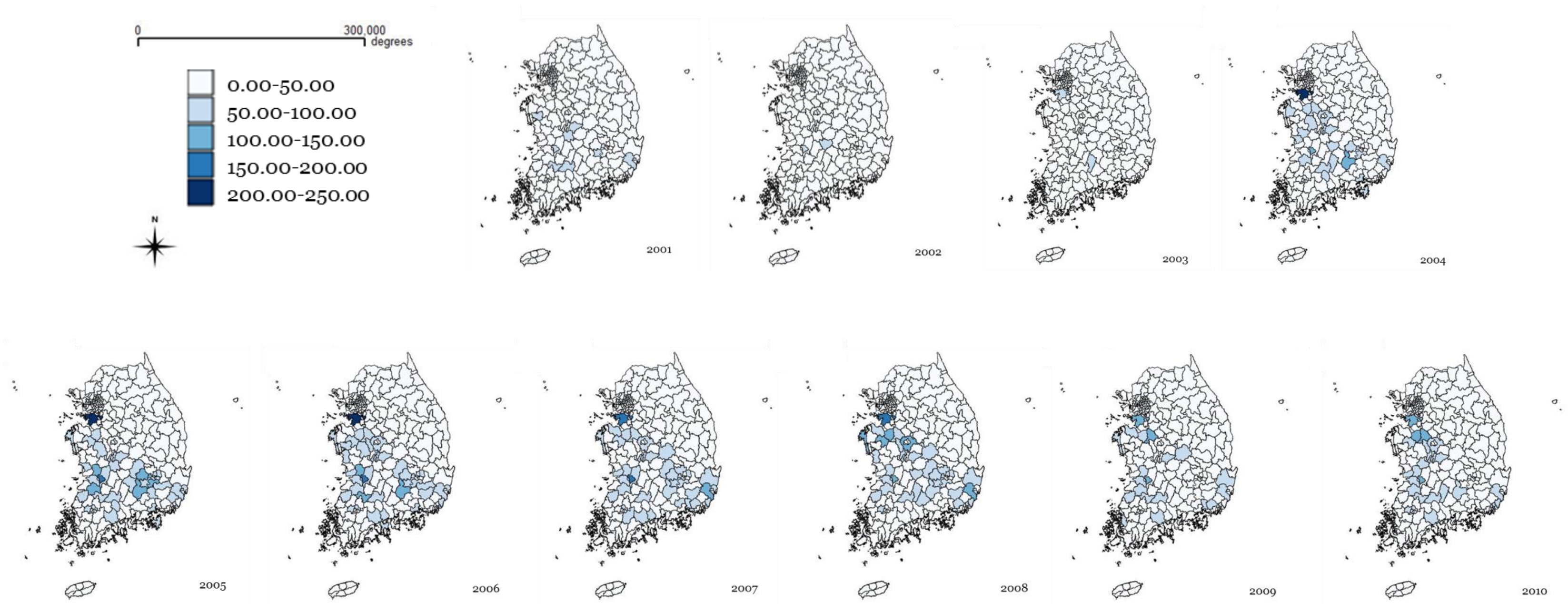


Figure 2-1. Scrub typhus incidence counts GIS simulation

Figure 2-2. Scrub typhus crude incidence rates GIS simulation

Figure 2-3. Scrub typhus age-gender standardized incidence rates GIS simulation

LISA cluster map and significance map

In order to observe significant clustering pattern of disease occurrence, LISA cluster maps and significant maps were observed for three dependent variables. LISA significance map exhibits areas with high incidence neighboring high incidence, areas with low incidence neighboring low incidence, areas with high incidence neighboring low incidence, and areas with not significant results.

For LISA cluster maps of incidence counts, the nation exhibits clear distinction between low-low and high-high clusters of scrub typhus incidences [Figure 3-1]. Although spatial separation is already apparent in earlier years, there are ‘not significant’ areas mixed in Gangwon region. However, as years pass by, upper right region, including Gangwon, part of Kyunggi and Seoul, is all colored in blue, indicating low-low clustering of scrub typhus occurrence. On the other hand, diagonal linear band in the middle region of the nation is colored in red, indicating high-high scrub typhus occurrence clustering in the region. LISA significance maps were consulted to understand significant level of LISA cluster maps and areas with high-high and low-low scrub typhus incidences exhibited results on significant level [Figure 3-2].

LISA cluster maps of crude incidence rate data display less numbers of regions with low-low and high-high clustering when compared with incidence counts, especially during the earlier years in the observation period [Figure 3-3]. Although cluster pattern of crude incidence rate data is similar with that of incidence count data, distinct clustering is more apparent among high-high regions than low-low regions for crude incidence rate data. For LISA significance maps, regions that showed high-high and low-low clusters for cluster map had significant level [Figure 3-4].

For LISA cluster map with age-gender standardized incidence rates, although spatial clustering of many regions were indicated as ‘not significant’ during 2001-2003, low-low cluster and high-high cluster became more apparent after the explosive increase in scrub typhus increases [Figure 3-5]. LISA significance maps had similar trend as those mapped with incidence count data and crude incidence rate data [Figure 3-6].

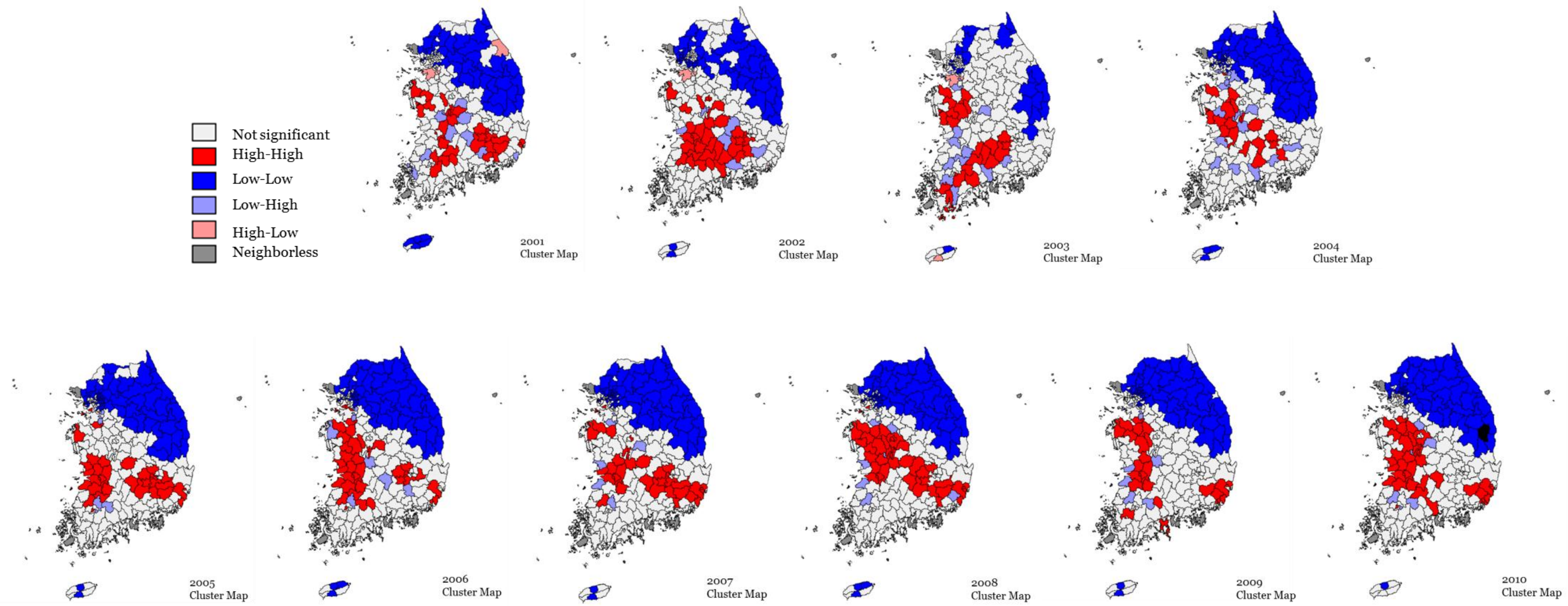


Figure 3-1. Scrub typhus incidence counts LISA cluster map

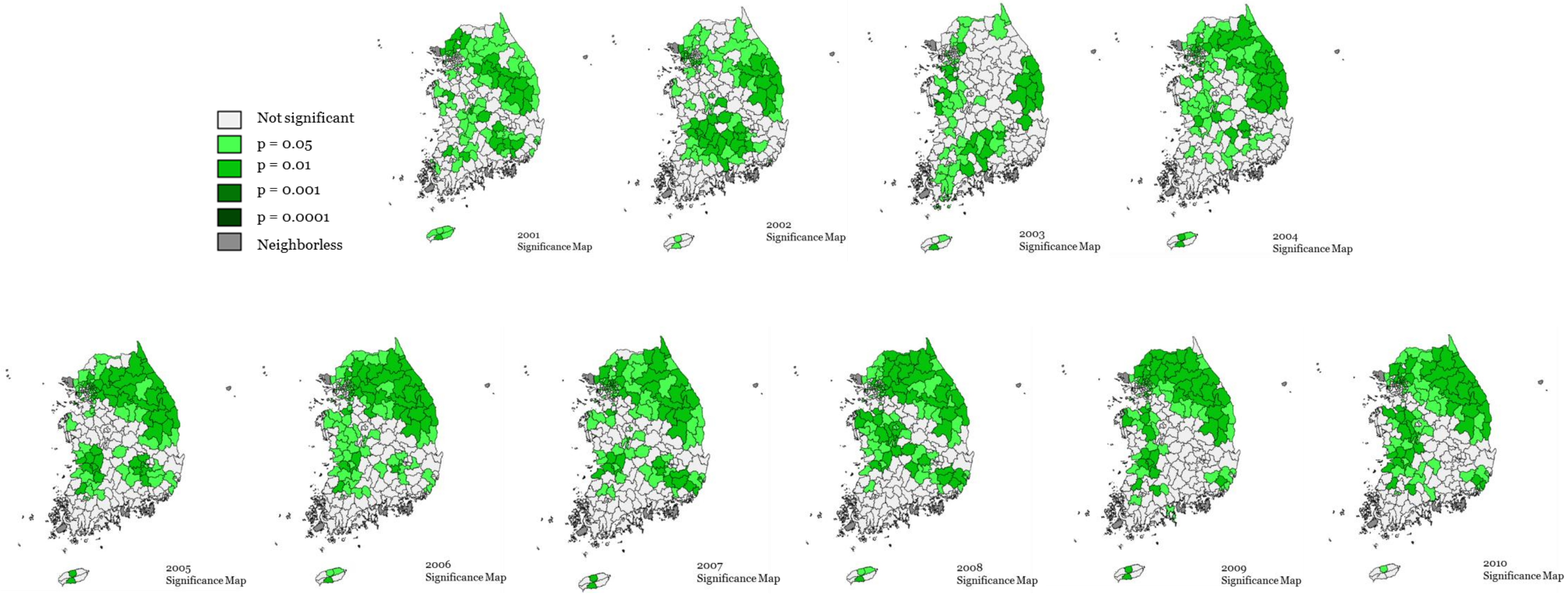


Figure 3-2. Scrub typhus incidence counts LISA significance map

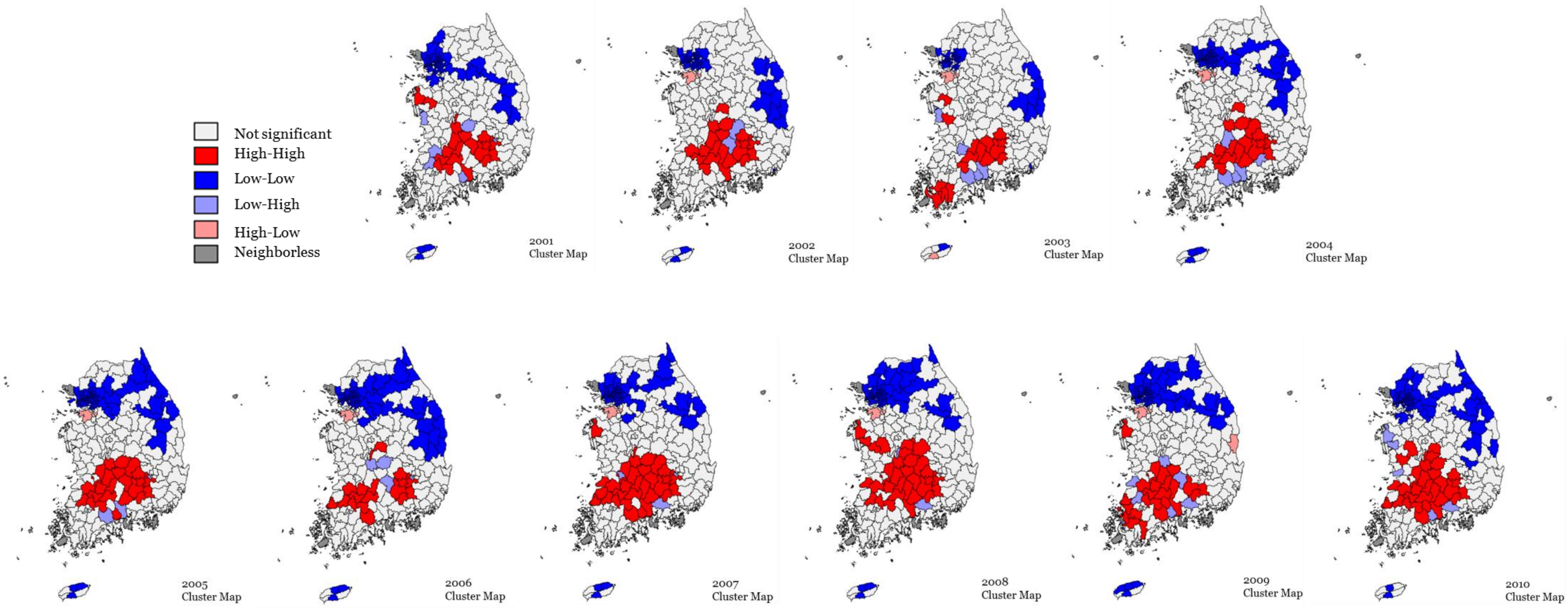


Figure 3-3. Scrub typhus crude incidence rates LISA cluster map

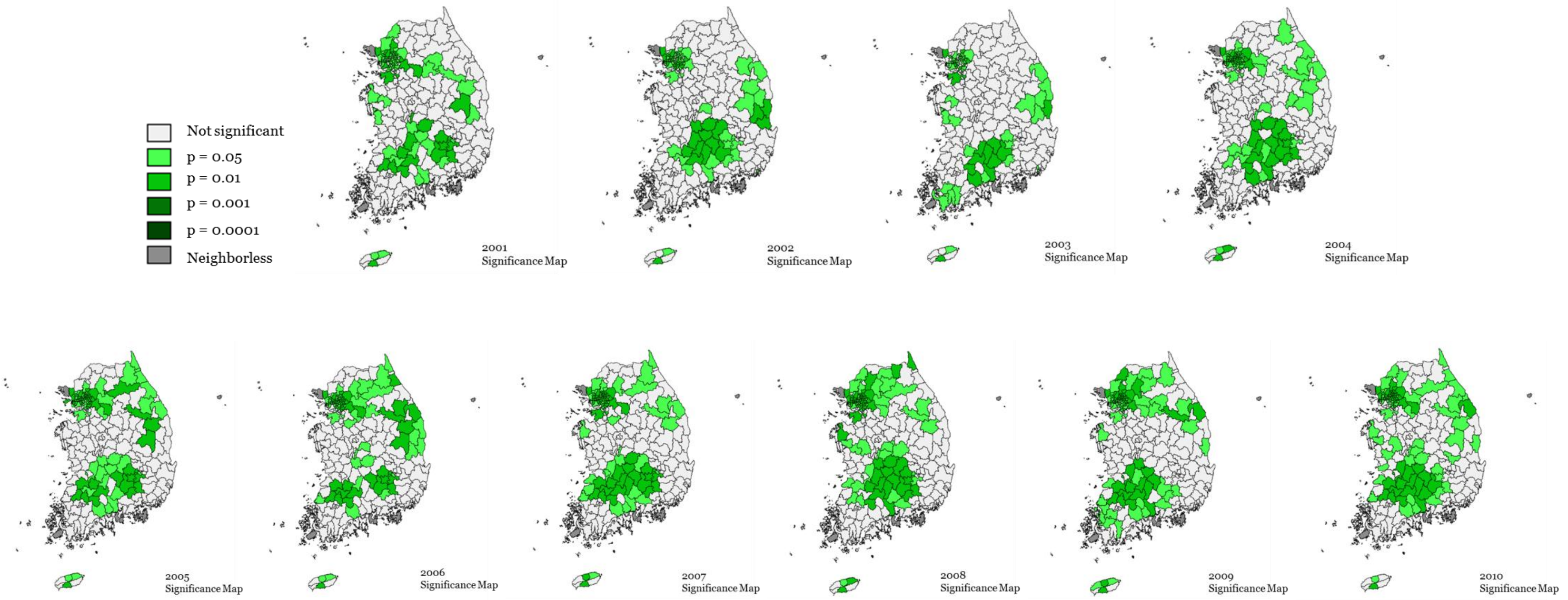


Figure 3-4. Scrub typhus crude incidence rates LISA significance map

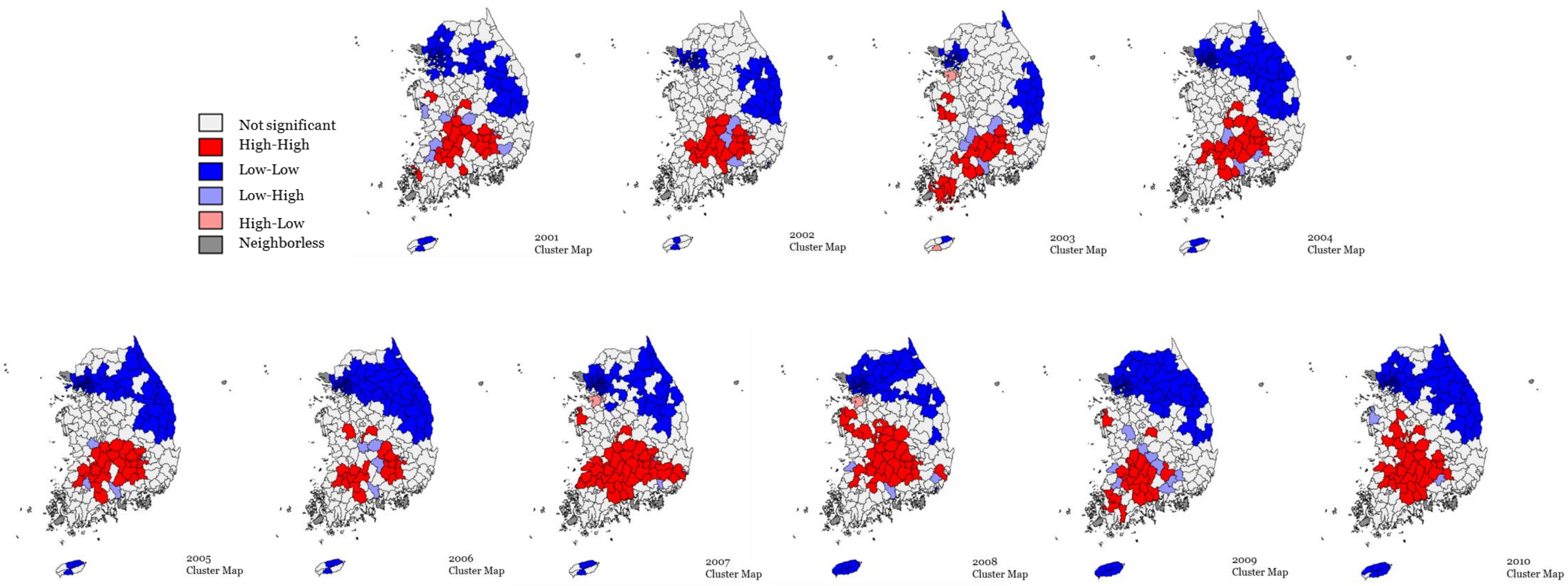


Figure 3-5. Scrub typhus age-gender standardized incidence rates LISA cluster map

Figure 3-6. Scrub typhus age-gender standardized incidence rates LISA significance map

Moran's I

Moran's I was calculated for each year and for all three dependent variables to determine spatial autocorrelation [Appendix A].

Moran's I calculated with incidence count data fluctuates throughout the first half of the observation period, however towards the second half, the trend line becomes relatively smooth [Figure 4-1].

The Moran's I values calculated with incidence rates appear more consistent compared to that of incidence count data, and exhibited more visible positive linear trend [Figure 4-2].

Moran's I calculated with age-gender standardized incidence rate present similar trend with the trend drawn by crude incidence rates, with slightly different values of Moran's I [Figure 4-3].

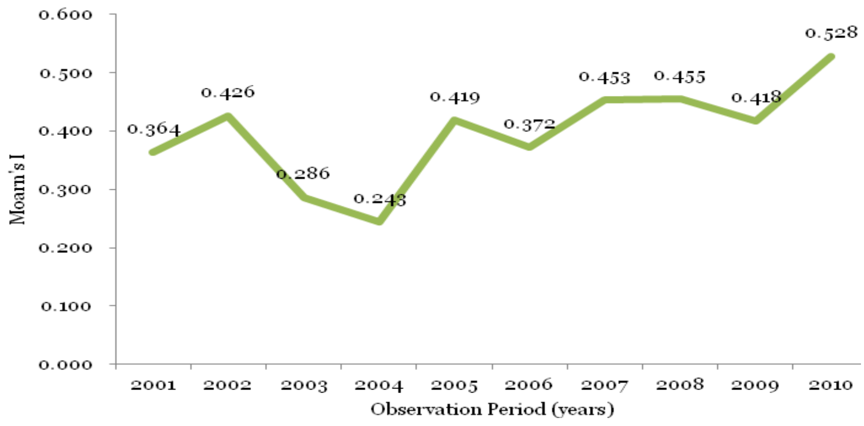


Figure 4-1. Yearly trend of Moran's I calculated using scrub typhus incidence counts

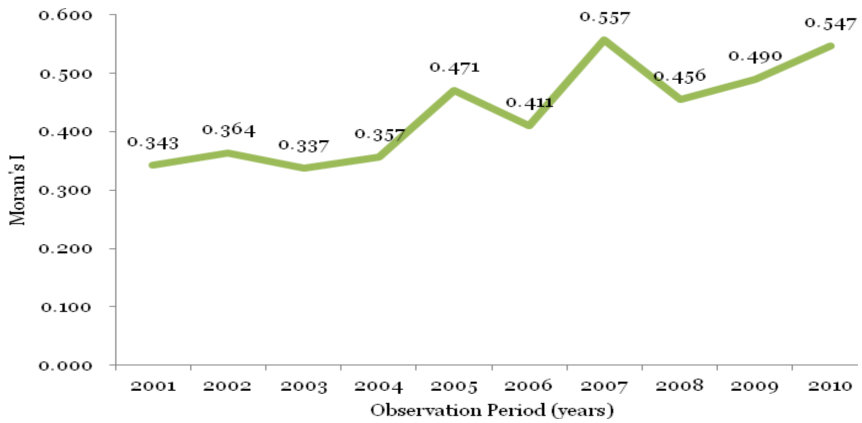


Figure 4-2. Yearly trend of Moran's I calculated using scrub typhus crude incidence rates

Figure 4-3. Yearly trend of Moran's I calculated using age-gender standardized scrub typhus incidence rates

In order to examine the validity of the trends, first, normality test for the Moran's I distribution for each type of dependent variable was performed. Once the normality of observations was evident, logistic regression was performed with all tree datasets. All three data had normal distribution and although coefficients were very small, all three had significant positive linear trend of Moran's I throughout the observation period [Table 2].

Table 2. Comparison of linear regression results between three types of dependent variables						
	Incidence Count		**Crude Incidence Rate**		**Standardized IR**	
	score	***p*-value**	**score**	***p*-value**	**score**	***p*-value**
coefficient	0.018	0.045	0.020	0.002	0.020	0.004
R^2	0.412		0.720		0.660	
adjusted R^2	0.339		0.685		0.617	

III-D. Land Use and its Effect

After obtaining the national areal information, proportions of land use for each type (dry-field, paddy-field, and orchard) were calculated by PHC level. Spearman rank correlation coefficients were obtained, and compared between seven types of land uses and three types of dependent variables [Table 3].

Seven types of proportional land use include, dry-field, paddy-field, orchard, dry-field and paddy-field combined, orchard and paddy-field combined, dry-field and orchard combined, and dry-field, paddy-field, and orchard combined.

When incident counts were considered as occurrence data, all types of land use had significant correlation results. Among the types of land use, paddy-field had the largest correlation coefficient, followed by land proportions of orchard and paddy-field combined, dry-field and paddy-field combined, and dry-field, paddy-field, and orchard combined. All types of land use also had significant correlation results when crude incidence rates were considered as occurrence data. Among the types of land use, proportion of paddy-field use had the largest correlation coefficient followed by orchard and paddy combined, dry-field and paddy-field combined, dry-field, paddy-field, and orchard combined.

When age-gender standardized incidence rates were considered as occurrence data, all types of land use had significant correlation results as well. Among all types of land use proportions, paddy-field had the highest correlation coefficient, followed by orchard and paddy-field combined, dry-field and paddy-field combined, and dry-field, paddy-field, and orchard combined.

Table 3. Spearman rank correlation coefficients by types of occurrence and land use

Occurrence	Land Use	Correlation coefficient	*p*-value
Incidence Counts	Dry-field	0.105	<0.001
	Paddy-field	0.401	<0.001
	Orchard	0.224	<0.001
	Dry-field + Paddy-field	0.325	<0.001
	Orchard + Paddy-field	0.398	<0.001
	Dry-field + Orchard	0.115	<0.001
	Dry-field + Paddy-field + Orchard	0.320	<0.001
Crude Incidence Rate	Dry-field	0.222	<0.001
	Paddy-field	0.469	<0.001
	Orchard	0.283	<0.001
	Dry-field + Paddy-field	0.403	<0.001
	Orchard + Paddy-field	0.464	<0.001
	Dry-field + Orchard	0.225	<0.001
	Dry-field + Paddy-field + Orchard	0.397	<0.001
Age-gender Standardized Incidence Rate	Dry-field	0.244	<0.001
	Paddy-field	0.509	<0.001
	Orchard	0.305	<0.001
	Dry-field + Paddy-field	0.437	<0.001
	Orchard + Paddy-field	0.497	<0.001
	Dry-field + Orchard	0.244	<0.001
	Dry-field + Paddy-field + Orchard	0.427	<0.001

In order to understand the relationship between dependent variables and the type of land use, regression test was performed. Since all three dependent variables had variance that is greater than the mean, they do not fit the assumptions for Poisson regression model. Therefore, negative binomial regression model was considered for the analysis. The results from negative binomial regression analysis display significant effect size of paddy-field for all three types of dependent variables. Proportions of field use had significant level of negative effect size towards scrub typhus occurrence.

Table 4. Comparison of negative binomial regression results between three types of dependent variables

	Estimate	Standard error	95% CI	Chi-square	*p*-value
Count					
mean, variance					18.452, 610.21
Negative binomial regression model					
Paddy field use rate	0.070	0.004	0.063, 0.078	323.86	<0.001
Field use rate	-0.054	0.007	-0.068, -0.039	54	<0.001
Orchard use rate	0.083	0.027	0.030, 0.137	9.4	0.002
Crude incidence rate					
mean, variance					18.506, 1047.03
Negative binomial regression model					
Paddy field use rate	0.082	0.005	0.071, 0.092	236.73	<0.001
Field use rate	-0.006	0.010	-0.025, 0.014	0.31	0.576
Orchard use rate	-0.049	0.028	-0.104, 0.007	2.99	0.084
Age-gender standardized incidence rate					
Mean, variance					14.115, 494.351
Negative binomial regression model					
Paddy field use rate	0.077	0.005	0.068, 0.087	253.48	<0.001
Field use rate	-0.017	0.009	-0.035, 0.000	3.66	0.056
Orchard use rate	-0.034	0.028	-0.088, 0.020	1.52	0.218

IV. Discussion

Conclusion

The data gathered through national surveillance system of KCDC shows strong seasonal trend of scrub typhus occurrences in Korea and high increase in the number of cases reported per year. Descriptive analysis shows that during the ten years of observation period, females, older people and people in agriculture and fishery industry were more likely to be affected by scrub typhus. Considering the ecology of the causative agent, *Orientsia tsutsugamushi*, people with higher frequency of outdoor activities are more prone to be exposed to the vectors.

During the ten years of observation period, scrub typhus incidences exhibited strong seasonal trend, most occurrences appearing from September to December. Although the time frame in which reports of scrub typhus appears did not change, the number of incidences appearing each year changed drastically in the past decade, especially between the years 2004 and 2006. In addition, when all the incidences were analyzed by province regional level, scrub typhus incidence varied drastically by regions.

Moran’s I and LISA observations support the argument that there are spatial autocorrelations on scrub typhus occurrences. Furthermore, linear regression coefficients support the argument that Moran’s I values exhibit positive linear trend during the observation period. The study further investigated on proportions of land use, which was hypothesized to influence spatial autocorrelations and found out that there are strong correlations between scrub typhus incidences and proportions of paddy-field use in PHC level regions.

Negative binomial regression coefficients display significant effect size for scrub typhus occurrences and proportions of paddy field use, indicating each unit of increase in field size, there is an increase in scrub typhus incidence counts, crude incidence rates, and age-gender standardized incidence rates.

Design of the study

In the database, there were five dates, infected, diagnosed, and reported dates (on physician, PHC, and provincial levels), available to be used for incidence analysis. Although infected date was first expected to be the most plausible date data to be utilized to determine the date of incidence, since scrub typhus does not have prominent symptoms immediately after infection, infected dates are usually unknown and inaccurate. Reported infected dates are based on estimation from the diagnosed dates. Following the infected date, diagnosed date was considered next as a measure to determine the date of infection, but there were missing values for diagnosed dates. Since scrub typhus is required by law to be reported immediately, reported dates were also considered as logical date data to be utilized. Among three levels of physician, PHC, and province, physician level was utilized for disease occurrences for it is the first level in the reports system and there was no missing value for the variable among scrub typhus cases. There were three types of areal information available for scrub typhus incidences in the database and these are patient address information, reported public health center, and reported province. Patient address information is recorded by the patients, and reported health center and province is recorded by which public health center reports the occurrence. There are 253 public health centers and 16 province divisions in Korea. Although patient address was the smallest measurement of spatial data and it could serve as point estimates of disease occurrence, there was no method to verify validity of patient address. Therefore it was not considered as an accurate areal variable to be utilized, and instead, public health center level was determined to be the most precise and specific address to locate scrub typhus occurrence.

Infectious diseases are often analyzed using the rates as measurement for the outcome variable since the person and the time involved is important considering the communicable characteristics of infectious diseases. Ecological evidence shows that scrub typhus does not transmit between individuals, but transmitted due to external factors, such as mice and rickettsia. Therefore person and time might not be as important as other infectious diseases that have R_o larger than 0. Nonetheless, after analyzing patient characteristics, three possible dependent variables were decided to be incorporated for thorough investigation. The three variables include, incidence count data (number of scrub typhus occurrence reported), crude incidence rates per 100,000 person-year, and age-gender standardized incidence rates per 100,000 person-year.

Interpretation of results

This study investigates characteristics of scrub typhus infection propagation, and having almost all cases presumed to be infected within Korea, it reflects validity of choosing study population to investigate spatial correlation of disease distribution within Korea. There were six cases presumed to be infected from outside of Korea, however, since the variable itself is based on an assumption, six cases were also considered as study population.

The result shows that people in older ages are more likely to be reported with scrub typhus. Although it cannot be concluded from this study, this could be explained by physiological characteristics, such as weak immune system, of older people and also by changes in social structures in Korea. Most of the younger people living in the rural areas are moving out to urbanized areas, acquiring non-agricultural occupations. Therefore mean ages of people involved in agricultural and fishery industry are rising and with the effect of weakened immune system collaborating, people in older ages are automatically at high risk of being infected by scrub typhus.

As it was evident through GIS simulations, there were dark spots becoming lighter as years go by. Since numbers of scrub typhus incidence did not drop drastically, having less scrub typhus incidence is insufficient reasoning to explain decrease in scrub typhus incidences in specific once-concentrated areas. Since many people in agricultural and fishery industry were reported with infections with scrub typhus beginning in 2004, KCDC has performed preventive measures in certain areas with high reports. This could be the reason for having lighter spots showing on areas that had high concentrations of scrub typhus infection during the rise period, from 2004 to 2006, of scrub typhus.

Moran’s I values calculated with three types of occurrence data exhibited high spatial autocorrelation of scrub typhus occurrences. Although this does not explain the high increase in scrub typhus during the rise period, it supports the argument that there is a spatial autocorrelation of scrub typhus incidences. This study attempted to investigate spatial characteristics that are hypothetically influential to scrub typhus incidence, to contribute in development of more plausible and effective preventive measures.

Since there was no report on explosive increase in scrub typhus cases in other countries, Korea is unique in experiencing high rise of scrub typhus in short period of time. Therefore understanding the society, types of vectors, environmental, and host factors that are unique in Korea would aid in understanding the unusual phenomena.

Limitations of the study

The limitations on this study include, limited precision and specificity on location of infection and patient characteristics. The national data are obligated to protect privacies of citizens, therefore, many patient information were unavailable for the analysis.

Another limitation for the study is difficulty in obtaining data that could explain possible pathways of disease transmission. For example, rodents act as vehicles in delivering *O. tsutsugamushi* to humans, however, national level data describing numbers of rodents in specific region or movements of rodents were unavailable, whereas data on bigger animals like cows and chickens were available.

Future Studies

There had been different types of scrub typhus appearing in different regions in Asia. For example, although cases of scrub typhus appearing in Korea are winter types, summer types of scrub typhus appear in China[2]. In addition, scrub typhus occurs throughout a year in Japan, reporting higher rates than Korea in all seasons besides autumn, and more scrub typhus infection was found among males than females[24]. Likewise, disease and patient characteristics differ by countries. Investigation on differences between hosts, environments, and vectors of winter type and summer type of scrub typhus may potentially bring new insight in understanding scrub typhus infections. Furthermore, since results of this study support that scrub typhus incidences are associated with paddy-field ratio, understanding methods of cultivation in paddy fields may aid in understanding causal pathway of the association.

Reference

1 Jeong, Y. J., Kim, Suk, Wook, Yeh D., Lee, Jun W., Kim, Kun-il, and Lee, Sun H. Scrub Typhus: Clinical Pathologic, and Imaging Findings *RadioGraphics* **27**, 161-172 (2007).

2 Sharma Parag, K., Rakesh, Kaore, Shilpa N., Yadav, Vijay K., Sharma, Raj. Geographical Distribution, Effect of Season & Life Cycle of Scrub Typhus. *Journal of Medical Education and Research* **12**, 63-64 (2010).

3 Ogawa, M., et al. . Scrub Typhus in Japan: Epidemiology and Clinical Features of Cases Reported in 1998. *American Journal fo Medical Hygiene* **67**, 162-165 (2002).

4 Lee, K., et al. Comparison of the Awareness and Knowledge of Scrub Typhus between Case and Control Groups. *Journal of Agricultural Medicine and Community Health* **37**, 1~11 (2012).

5 Sharma, R. Scrub Typhus: Prevention and Control. *Journal of Medical Education and Research* **12**, 91 (2010).

6 Infectious Diseases Surveillance Yearbook, 2011 (Korea Center for Disease Control and Prevention, 2011).

7 Kong, W.-s., et al. . Time-Spatial Distribution of Scrub Typhus and its Environmental Ecology. *The Korean Geographical Society* **42**, 863-878 (2007).

8 Huang, C.-T., Chi, Hsin, Lee, Hung-Chang, Chiu, Nan-Chang, and Huang, Fu-Yuan. Scrub Typhus in Children in a Teaching Hospital in Eastern Taiwan, 2000-2005. *Southeast Asian Journal of Tropical Medicine and Public Health* **40** (2009).

9 Frequently Asked Questions: Scrub Typhus.

10 Chung, M.-H., Lee, Jin-Soo, Baek, Ji-hyeon, Kim, Mijeong, and Kang, Jae-Seung. Persistence of *Orientia tsutsugamushi* in Humans. *Journal of Korean Medical Science* **27**, 231-235 (2012).

11 Lee, C.-S., Hwang, Jeong-Hwan, Lee, Heung-Bum, Kwon, Keun-Sang. Risk Factors Leading to Fatal Outcome in Scrub Typhus Patients. *Americal Journal of Tropical Medicine Hygiene* **81**, 484-488 (2009).

12 Tamura, A., Ohashi, Norio, Urakami, Hiroshi, Miyamura, Sadao. Classification of *Rickettsia tsutsugamushi* in a New Genus, *Orientia* gen. nov., as *Orientia tsutsugamushi* comb. nov. *International Journal of Systematic Bacteriology* **45**, 589-591 (1995).

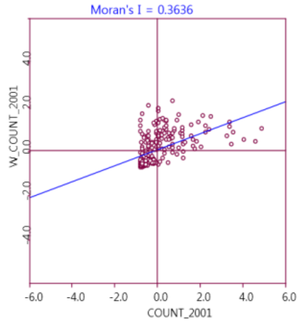
13 Base, K. B., Youn, Won Hwa, Lee, Youn Jae, Jung, Soo Jin, Hong, Kwan Hee. Massive Small Bowel Bleeding Caused by Scrub Typhus in Korea. *World Journal of Gastrointestinal Surgery* **2**, 47-50 (2010).

14 Kim, S.-H., Jang, Jae-Yeon. Correlations between climate change-related infectious diseases and meteorological fators in Korea. *Journal of Preventive Medicine and Public Health* **43**, 436-444, doi:10.3961 (2010).

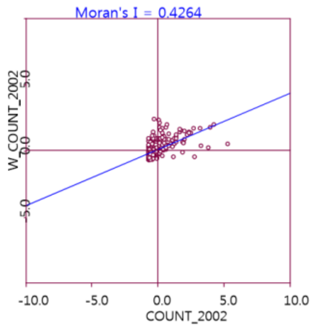
15 Chi-Chien Kuo, J.-L. H., Chia-Ying Ko, Pei-Fen Lee, Hsi-Chieh Wang. Spatial analysis of scrub typhus infection and its association with environmental and socioeconomic factors in Taiwan *Acta Tropica* **120**, 52-58 (2011).
16 Kweon, S.-S., Choi, Jin-Su, Lim, Hyun-Sul, Kim, Jang-Rak, Kim, Keon-Yeop, Ryu, So--Yeon, Lee, So-Dam, Im, Heung-Ku, and Kwon, Jun-Wook. A community-based case-control study of behavioral factors associated with scrub typhus during the autumn epidemic season in South Korea. *The American Society of Tropical Medicine and Hygiene* **80**, 442-446 (2009).
17 Tatem, A. J., Adamo, Susana, Bharti, Nita, Burgert, Clara R., Castro Marcia, Dorelien, Audrey, Fink, Gunther, Linard, Catherine, Mendelsohn, John, Montana Livia, Montgomery, Mark R., Nelson, Andrew, Noor, Abdisalan M., Pindolia, Deepam, Yetman, Greg, Balk, Deborah. Mapping populations at risk: improving spatial demographic data for infectious disease modeling and metric derivation. *Population Health Metrics* **10**, doi:10.1186/1478-7954-10-8 (2012).
18 Long, J. A., Robertson, Colin, Nathoo, Farouk S., Nelson, Trisalyn A. A Bayesian space-time model for idscrete spread processes on a lattice. *Spatial and Spatio-temporal Epidemiology* **3**, 151-162 (2012).
19 Lee, I. Y., Kim, Heung Chul, Lee, Young-Sun, Seo, Jang Hoon, Lim, Jae Won, Yong, Tae Soon, Klein, Terry A., Lee, Won Ja. Geographical Distribution and Relative Abundance of Vectors of Scrub Typhus in the Republic of Korea. *Korean Journal of Parasitology* **47**, 381-386 (2009).
20 Vallée, J., Thaojaikong, Thaksinaporn., Moore, Catrin E., Phetsouvanh, Rattanaphone., Richards, Allen L., Souris, Marc., Fournet, Florence., Salem, Gérald., Gonzalez, Jean-Paul J., Newton, Paul N. . Contrasting Spatial Distribution and Risk Factors for Past Infection with Scrub Typhus and Murine Typhus in Vientiane City, Lao PDR. *PLOS neglected tropical diseases* **4** (2010).
21 Legislation on Survey, Hydrographic Investigation, and Cadastral Mapping. (2012).
22 Anselin, L. Local indicators of spatial association - LISA. *Geographical Analysis* **27**, 93-115 (1995).
23 Paradis, E. Moran's Autocorrelation Coefficient in Comparative Methods. *R Foundation for Statistical Computing* (2011).
24 Bang, H.-A., Lee, Myeong-Jin, Lee, Won-Chang. Comparative research on epidemiological aspects of Tsutsugamushi disease (scrub typhus) between Korea and Japan. *Japanese Journal of Infectious Diseases* **61**, 148-150 (2008).

Appendix A.

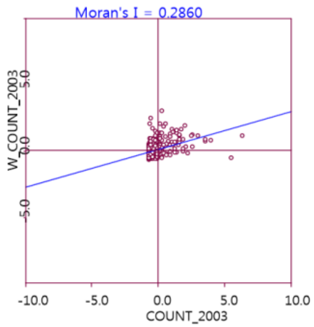
i. Moran's I maps – incidence counts



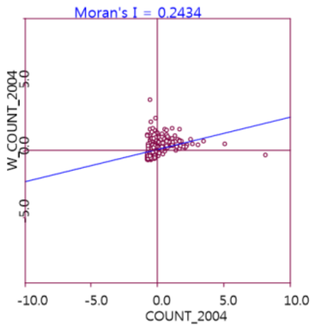
Incidence counts 2001



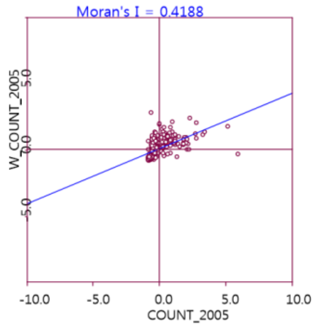
Incidence counts 2002



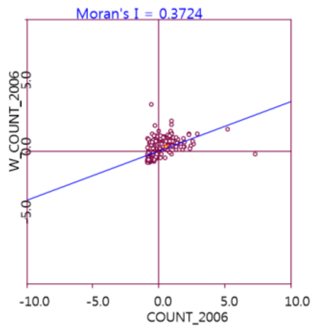
Incidence counts 2003



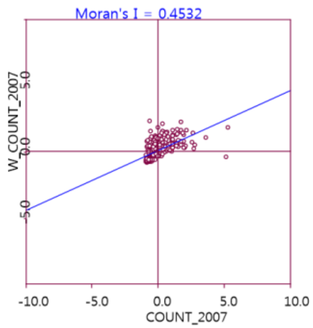
Incidence counts 2004



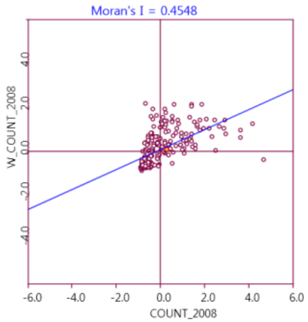
Incidence counts 2005



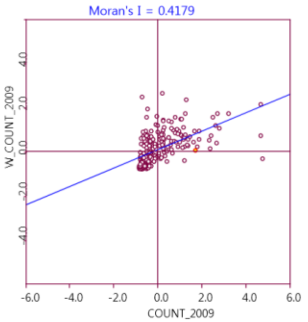
Incidence counts 2006



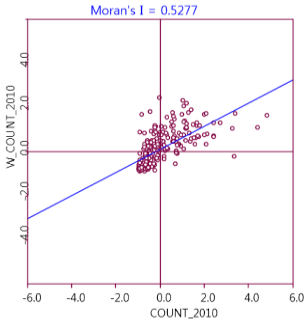
Incidence counts 2007



Incidence counts 2008

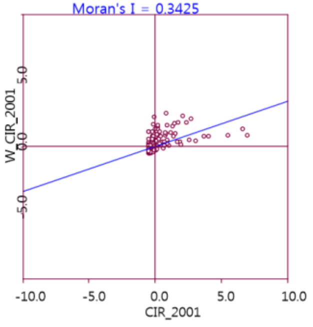


Incidence counts 2009

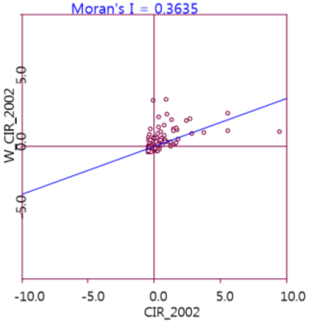


Incidence counts 2010

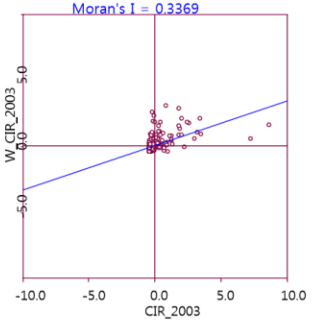
ii. Moran's I maps – crude incidence rates



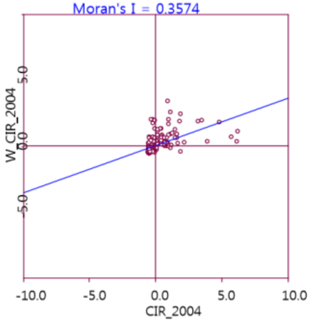
Crude incidence rates 2001



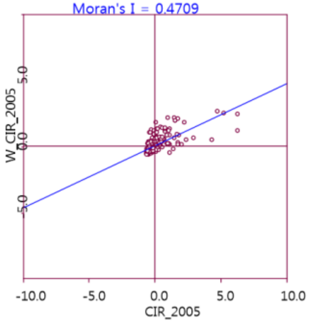
Crude incidence rates 2002



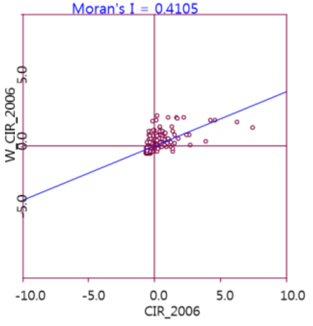
Crude incidence rates 2003



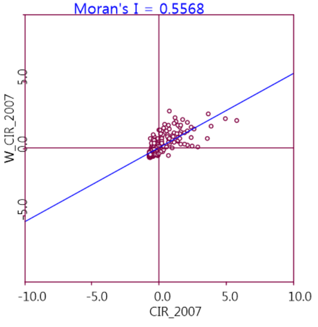
Crude incidence rates 2004



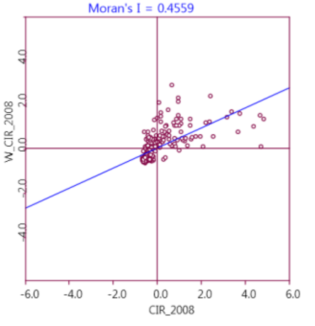
Crude incidence rates 2005



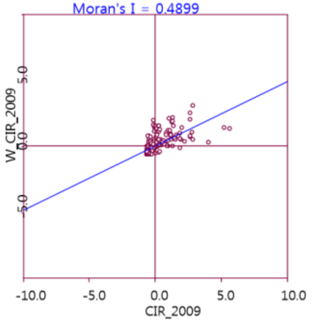
Crude incidence rates 2006



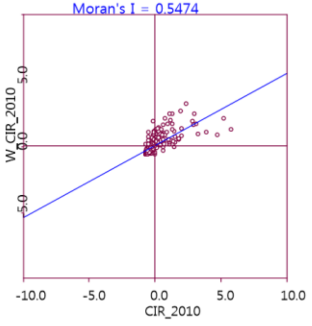
Crude incidence rates 2007



Crude incidence rates 2008

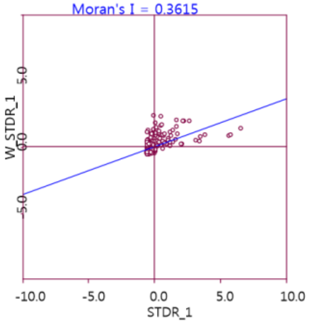


Crude incidence rates 2009

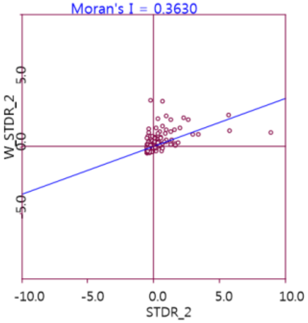


Crude incidence rates 2010

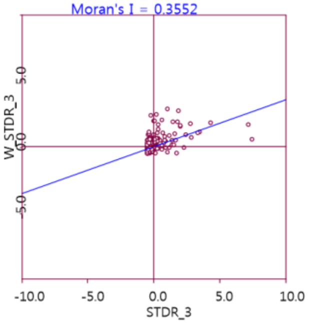
iii. Moran's I maps – age-gender standardized incidence rates



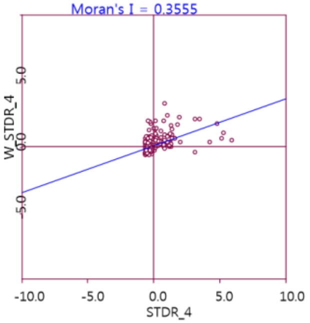
Standardized incidence rates 2001



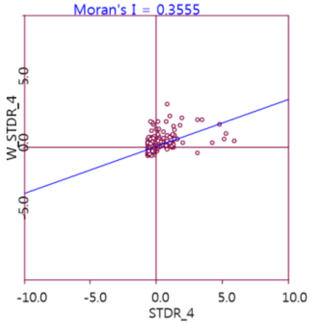
Standardized incidence rates 2002



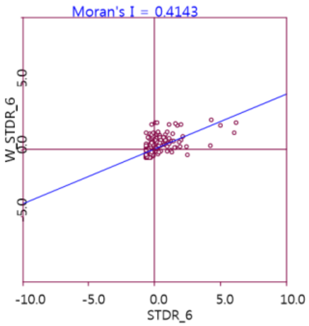
Standardized incidence rates 2003



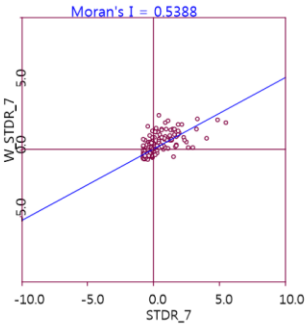
Standardized incidence rates 2004



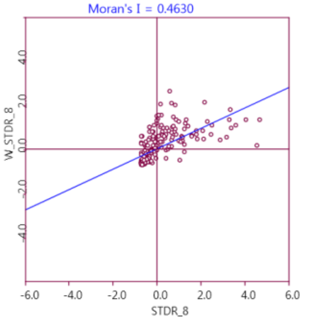
Standardized incidence rates 2005



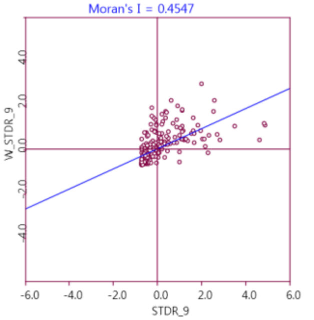
Standardized incidence rates 2006



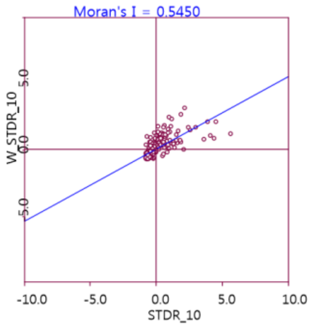
Standardized incidence rates 2007



Standardized incidence rates 2008



Standardized incidence rates 2009



Standardized incidence rates 2010

Appendix B.

Methods: age-gender standardization

Step 1. Download registered residence data from KOSIS.

The data must have variables indicating, numbers of registered residence by PHC level, age, and gender.

Example 1. Registered residence data structure

rpt_pl	rpt_hos	age_gp	rra_all_2001	rra_male_2001	rra_female_2001
전국	전국	Total	48,021,543	24,102,463	23,919,080
전국	전국	00-04	3,104,377	1,621,395	1,482,982
전국	전국	05-09	3,560,746	1,892,474	1,668,272
전국	전국	10-14	3,224,996	1,700,556	1,524,440
전국	전국	15-19	3,512,042	1,819,832	1,692,210
전국	전국	20-24	4,013,575	2,056,776	1,956,799
전국	전국	25-29	4,227,260	2,159,430	2,067,830
전국	전국	30-34	4,623,704	2,348,376	2,275,328
전국	전국	35-39	4,245,121	2,195,143	2,049,978
전국	전국	40-44	4,361,219	2,222,918	2,138,301
전국	전국	45-49	3,283,315	1,663,544	1,619,771
전국	전국	50-54	2,413,404	1,214,278	1,199,126
전국	전국	55-59	2,040,350	994,032	1,046,318
전국	전국	60-64	1,877,962	875,998	1,001,964
전국	전국	65-69	1,436,991	617,792	819,199
전국	전국	70-74	949,201	354,058	595,143
전국	전국	75-79	612,526	213,384	399,142
전국	전국	80-84	332,945	104,940	228,005
전국	전국	85-89	146,212	36,494	109,718
전국	전국	90-94	45,471	9,769	35,702
전국	전국	95+	10,124	1,274	8,850
서울	서울특별시	Total	10,263,336	5,141,741	5,121,595
서울	서울특별시	00-04	590,235	305,913	284,322
서울	서울특별시	05-09	652,524	345,287	307,237
서울	서울특별시	10-14	629,798	331,921	297,877

Step 2. Assign age groups(the same age group interval used for scrub typhus incidence) to registered residents data.

Step 3. Delete repeated measures.

Within ten years of observation period, there were provinces that had to go through division due to population expansion (examples 2, 3).

For example, although resident count was measured on province level up to year 2005 for Yongin province, beginning in 2005, due to population expansion, Yongin was measured from three PHC level divisions, Giheung, Suji, Cheoin.

However, the registered population registry has a line of Yongin population count with three PHC levels as well. If the data is used without transformation, the population in Yongin will be double-counted. Therefore repeated counts must be deleted in order to calculate precise incidence rates per 100,000 persons.

Example 2. Repeated residents information

rpt_pl	rpt_hos	age_gp	rra_all_2005	rra_male_2005	rra_female_2005
경기	용인시	total	693,660	347,191	346,469
경기	처인구	total	196,830	100,873	95,957
경기	기흥구	total	224,770	112,663	112,107
경기	수지구	total	272,060	133,655	138,405

Example 3. Yearly district division (areas with additional divisions during observational period)										
	’01	’02	’03	’04	’05	’06	’07	’08	’09	’10
Suwon										
Province level	■	■								
PHC level			■	■	■	■	■	■	■	■
Seongnam										
Province level										
PHC level	■	■	■	■	■	■	■	■	■	■
Anyang										
Province level										
PHC level	■	■	■	■	■	■	■	■	■	■
Bucheon										
Province level										
PHC level	■	■	■	■	■	■	■	■	■	■
Ansan										
Province level	■									
PHC level		■	■	■	■	■	■	■	■	■
Goyang										
Province level										
PHC level	■	■	■	■	■	■	■	■	■	■
Ilsan										
Province level	■	■	■	■						
PHC level					■	■	■	■	■	■
Yongin										
Province level	■	■	■	■						
PHC level					■	■	■	■	■	■
Cheongju										
Province level										
PHC level	■	■	■	■	■	■	■	■	■	■
Pohang										
Province level										
PHC level	■	■	■	■	■	■	■	■	■	■

* Filled areas indicate the data utilized for the analysis.

Step 4. Age group re-arrangement.

Population data has twenty age groups, whereas scrub typhus data has seven age groups. Therefore when age groups were assigned in step 2, there would have been multiple lines of data with age group designated as group 1. In order to combine population data and the scrub typhus data, age groups must match. Thus, all the population counts data lines designated age group as 1 must summed into 1 line (example 4).

Example 4. Age group re-arrangement structure

rpt_pl	agegp	age_gp	rra_male_2001	new agegp	male_count
전국	total	Total	24,102,463		
전국	1	00-04	1,621,395		
전국	1	05-09	1,892,474		
전국	1	10-14	1,700,556		
전국	1	15-19	1,819,832	1	7,034,257
전국	2	20-24	2,056,776		
전국	2	25-29	2,159,430	2	4,216,206
전국	3	30-34	2,348,376		
전국	3	35-39	2,195,143	3	4,543,519
전국	4	40-44	2,222,918		
전국	4	45-49	1,663,544	4	3,886,462
전국	5	50-54	1,214,278		
전국	5	55-59	994,032	5	2,208,310
전국	6	60-64	875,998		
전국	6	65-69	617,792	6	1,493,790
전국	7	70-74	354,058		
전국	7	75-79	213,384	7	719,919
전국	7	80-84	104,940		
전국	7	85-89	36,494		
전국	7	90-94	9,769		
전국	7	95+	1,274		

Step 5. Scrub Typhus Incidence Count Data Preparation.

Prepare scrub typhus count data in the same format (number of scrub typhus incidences by PHC level, age, and gender)

Example 5. Scrub typhus incidence count data preparation

rpt_hos_no	agegp	m1	f1	m2	f2	m3	f3
1	1	0	0	1	0	0	3
1	2	1	0	0	0	1	0
1	3	1	1	0	0	0	0
1	4	0	0	0	0	0	0
1	5	0	1	0	0	0	0
1	6	0	3	0	0	0	0
1	7	0	1	0	1	0	0
2	2	0	0	0	0	0	0
2	3	0	0	0	0	0	0
2	4	1	1	1	0	1	0
2	5	0	2	0	0	0	0
2	6	0	0	0	0	0	1
2	7	0	1	1	1	0	0
3	1	0	0	0	0	0	0
3	3	0	0	0	1	0	0
3	4	0	0	0	0	0	1
3	5	0	0	1	0	0	1
3	6	0	0	0	0	1	1
3	7	0	0	0	0	0	0

Step 6. Combine population data with scrub typhus data.

It is important to merge scrub typhus data to population data because there could be certain age groups with zero scrub typhus counts. If population data were merged to scrub typhus data, the lines with zero scrub typhus counts will be deleted, resulting in incorrect sum of registered resident counts.

Step 7. Calculate incidence rates and sum for male and female (example 6).

Example 6. Calculating incidence rates and sum for male and female

rpt hos no	ageg p	m	f1	rra male	rra female	ra_m	ra_f	male sum	female sum
1	1	0	0	71035	62531	0	0	670289	587153
1	2	1	0	55877	59606	1.79E-05	0	488780	558134
1	3	1	1	44372	45514	2.25E-05	2.2E-05	467056	495396
1	4	0	0	43809	53980	0	0	451688	527011
1	5	0	1	31247	29675	0	3.37E-05	338028	356842
1	6	0	3	12900	12400	0	0.000242	172513	163698
1	7	0	1	5353	12341	0	8.1E-05	71423	137157

Step 8. Calculate total population counts by age groups.

Step 9. Multiply each age group of rates (scrub typhus count divided by number of registered residents) with total number of population counts for specific age group.

Step 10. Sum the answers obtained through step 8 on PHC level (example 7).

All the answers from each age group should be summed to have 1 value for each PHC level.

Example 7. Sum of answers obtained through step 8.

rpt hos no	r_m_sum 01	r_m_sum 02	r_f_sum 01	r_f_sum 02
1	1500.151696	689.55113104	7710.9356182	1371.0374097
2	885.6356269	2954.1467034	4856.5105088	1616.7745455
3	0	1337.6967301	0	1169.8611614
4	4235.0213783	794.72978076	1605.1813569	3061.5187415

Step 11. Divide answers obtained through step 10 by either total number of male population, total number of female population, total female and male population. Then, multiply 100,000 to each observation (example 8).

Example 8. Age-standardized incidence rate

rpt hosno	stir_m1	stir_f1	stir_a1
1	0.7024334141	3.6029586833	2.1542284441
2	0.4146914333	2.2691729355	1.3429119377
3	0	0	0
4	1.9830131402	0.7500259883	1.3658681568

국문초록

대한민국 쯔쯔가무시증 발생에 대한 시공간 분석

쯔쯔가무시증은 *Orientsia tsutsugamushi* 라는 병인으로 인하여 발생하는 감염병으로 고열과 같은 감기 증상으로 우리나라에는 가을철 빈번하게 발생하는 감염병 중 하나이다. 특히 2003년부터 2005년까지 대한민국에서 쯔쯔가무시증의 신고건수는 폭발적이게 증가하였다. 본 연구는 지난 2001년부터 2010년까지 대한민국에 신고된 쯔쯔가무시증 감염 자료에 대한 분석으로 쯔쯔가무시증 발생과 시·공간 분석 및 공간적 관련요인들의 연관성을 분석하고자 하였다.

10년간 신고된 쯔쯔가무시증 환자의 특성으로는 여성이 많았고 농업 및 어업에 종사하는 사람들에게서, 그리고 나이가 많은 사람들에게서 많이 발생한 것으로 나타났다. 따라서 쯔쯔가무시증의 신고 건수, 신고율, 그리고 연령별 성별 표준화된 신고율에 대해 연관성을 관찰하였다.

시·공간 분석을 위해 연간 쯔쯔가무시증의 발생 추이와 Moran's I와 LISA를 관찰하였다. 또한, 논, 밭, 그리고 과수원 토지사용율을 구하여 Spearman rank correlation test와 negative binomial model을 이용하여 특정 토지사용율과 쯔쯔가무시증 발생과의 연관성을 분석하였다. 결과로는 쯔쯔가무시증은 지역적 연관성이 있는 것으로 나타났으며 Moran's I는 양의 선형추이를 나타내고 있다. 또한, 논, 밭, 그리고 과수원 토지사용율을 이용한 분석은 모두 유의하였지만 특히 논의 면적과의 연관성이 가장 높았으며 논의 면적이 쯔쯔가무시증의 발생에 유의한 영향을 주는 것으로 나타났다.

주요어 : 쯔쯔가무시증, 시공간분석

학 번 : 2011-22098

감사의 글

사랑하는 사람 혹은 본인의 연약한 육신으로 인해 가슴이 철렁하거나 먹먹해보았던 사람들에게 저의 작은 관심과 노력이 위로가 되기를 바래봅니다.

많은 기대를 가지고 시작했던 대학원 생활이 그 기대보다 더욱 유익할 수 있도록 지식의 경계를 끊임없이 자극해 주신 존경하는 조성일 교수님, 누구를 위해 공부하는가를 직접 보고 경험 하게 해주신 신기철 교수님, 제 연구의 부족한 점을 보완해주신 노영선 선생님, 논문심사위원장으로 지도해주신 정해원 교수님과 김호 교수님, 논문작업에 도움주신 이선주 선생님, 유효순 선생님과 김도형 교수님, 그리고 우리나라와 세계시민의 건강을 위해 고군분투 하시며 지도해주신 교수님들께 지식뿐만 아니라 인성까지 가르쳐 주신 것에 대해 진심으로 감사드립니다.

서로의 어려움을 이해하고 고독한 길을 묵묵히 함께 가준 만성병 역학방 교실원들을 비롯하여 무거운 눈꺼풀을 열정으로 들어올리며 고민하는 대학원 동기들과 선후배님들께 감사드립니다.

바쁘다는 핑계로 친구의 역할에 소흘해도 너그러이 이해해준 사랑하는 경진이와 아란이, my fair ladies that are 'not a fan', mr. & ms huskies, chicagoans, sli shoobas, 서초동 동창들과 그 외 특정 카테고리로 분류되지는 않지만 인생의 여정을 함께하는 친구들에게 내 삶을 풍요롭게 해주어 고맙다는 말을 전합니다.

나의 영적 위로와 소망이자 구원의 기쁨을 자극해주는 친구들교회와 그 식구들, 제게 소중한 간증입니다.

지금까지 인류의 발전을 위해 힘써주신 수많은 분들과, 나에게 capability를 준 국가, 오랜 실험 시간과 잦은 시험 등 혹독한 훈련을 통해 health에 대한 이해와 흥미를 심어준 University of Washington (WA, USA), 신뢰로 기회를 제공해준 Abbott Laboratories (IL, USA), 학문에 대한 흥미를 일깨워준 서울대학교를 비롯하여 지난 2년 간 실습의 장을 제공해준 질병관리본부, 국민건강보험공단, 고려대학교 의학통계학교실, 서울특별시 보라매병원, 한국 노바티스, 그리고 다수의 1차 의료기관 관계자 여러분들께 감사드립니다.

마지막으로 든든한 지원군, 사랑하는 가족에게 감사드립니다. 모든 경험을 가능하게 해주시고 삶의 모든 영역에 완전한 멘토이신 아빠, 사소한 필요까지 헤아려 채워주시는 아름다운 엄마, 꿈을 가지고 멋지게 살아가는 듬직한 지효를 비롯하여 의료 분야에 대한 조언을 해주신 외삼촌과 이모부, 동하네, 규하네 가족, 할머니, 그리고 정이 많은 친가 가족들께 깊은 감사드립니다.

눈물로 씨앗을 뿌리는 자는 기쁨으로 단을 거두리라는 말씀을 잡고 시작된 여정입니다. 언제나 신실하신 주님을 나의 영혼이 찬양하나이다. I am a sinner that falls each day, yet Jesus still loves me, and granted me amazing grace that allowed me to believe in the power of the cross. Give thanks to Jesus Christ my Lord and Savior for He is good; His love endures forever.

여러 사람에 대한 감사로만 가득 담긴 지금이 가능하기까지 받은 은혜가 너무나 큽니다. 앞으로 더욱 사랑하고, 격려하고, 섬기며 살아가겠습니다.

나에게 한 번 주어지는 이 순간,
김 지 혜