Risk Assessment and Factors Associated with Lung Cancer using GIS in Mae Ka Subdistrict, Muang District, Phayao Province, Thailand

Jeefoo, P.,[1*] Preedapirom, W.[2] and Rojanavasu, P.[3]

[1]Research Unit of Spatial Innovation Development (RUSID), Geographic Information Science, School of Information and Communication Technology, University of Phayao, 19 Moo 2, Mae Ka, Muang, Phayao 56000 Thailand, E-mail: phaisarn.je@up.ac.th,* p.jeefoo@gmail.com

[2]Division of Physiology, School of Medical Sciences, University of Phayao, 19 Moo 2, Mae Ka, Muang, Phayao 56000 Thailand, E-mail: watcharaporn.pr@up.ac.th, watcharaporn.pree@gmail.com

[3]Computer Engineering, School of Information and Communication Technology, University of Phayao, 19 Moo 2, Mae Ka, Muang, Phayao 56000 Thailand, E-mail: pornthep.ro@up.ac.th

DOI: https://doi.org/10.52939/ijg.v17i5.2003

Abstract

Cancer is a major health problem in the developing countries. Variations of its incidence rate among geographical areas are due to various contributing factors. This study was performed to assess the spatial patterns of lung cancer incidence in the Mae Ka subdistrict, Muang district, Phayao province, based on lung cancer registry data and to determine geographical clusters. In this cross-sectional study, the cases of lung cancer were recorded from 2015 to 2020. Crude incidence rate was estimated based on age groups and sex in the province of the Mae Ka subdistrict. It uses spatial autocorrelation analysis (SAA) techniques to provide insight into the patterns, in terms of their geographical distributions and hotspot identification. Spatial autocorrelation analysis was performed in measuring the geographic patterns and clusters using GIS. In addition, local indicators of spatial association (LISA) and kernel density (KD) estimation were used to detect lung cancer hotspots using data at village level. Factors associated with the incidence of lung cancer was analyzed for behavior risk factors. Analysis of the spatial distribution of lung cancer shows significant differences from year to year and between different areas. The hotspot maps showed spatial trend patterns of lung cancer diffusion. Villages in the northern part revealed higher incidence. Furthermore, the spatial patterns during the years 2015, 2017 and 2019 were found to represent spatially clustered patterns, both at global and local scales. However, a clear spatial autocorrelation is observed, which can be of grate interest and importance to researchers for future epidemiological studies, and to policymakers for applying preventive measures.

1. Introduction

Lung cancer is the most frequent solid tumor and the leading cause of cancer-related deaths in both developing and developed countries, which is a major public health problem worldwide (Siegel et al., 2012 and Bray et al., 2013). Early state (I/II) non-small-cell lung cancer (NSCLC) and small-cell lung cancer (SCLC) are considered to have a better 5-year survival rates (45% and 31%, respectively) than that at advanced stage (III/IV) (Luchtenborg et al., 2014). Thus, to identify certain population who may have susceptibility to lung cancer and diagnose lung cancer in early stage are crucial to improve treatment outcome. Since the genetic characteristics has been proved to contribute to lung cancer development, many molecular epidemiological studies have been conducted to evaluate the relationship between lung cancers and the genetic variety, such as single nucleotide polymorphisms (SNP) in genes which may be involved in lung cancer development (Xu et al., 2016).

Cancer is one of the most significant health problems all over the world and in Thailand. Environmental factors, notably nutrition and smoking, are known to be involved in more than half of cancer cases. Furthermore, cancer is a disease whose frequency increases depending on various factors. The most important issue in controlling cancer disease in the accurate cancer registry in a country. Unless accurate statistical data are attained, it is impossible to know which cancer has what significance and to make strategic plans, above all in relation to realistic human researchers

(Tarik, 2013). Lung cancer is one of the most common types of cancer and the most important cause of cancer death worldwide (Southern Europe, Central and Eastern North America, and Southeast Asia), which imposes a major burden of burden (Torre et al., 2015). About 17% and 9% of common cancers are in man and woman, respectively, and 19% of cancer deaths are related to lung cancer (Hanley et al., 2014). Lung cancer is the deadliest cancer due its low chances of survival, despite a recovery in the survival of some types of cancer in recent years, the 5-year survival of lung cancer is relatively low and is mainly due to the late diagnosis and the advanced stages of the disease (Abazari, 2015).

Tobacco use through cigarette smoking has been established as a principal risk factor for lung cancer. However, 10% to 20% of all lung cancer cases occur in never smokers, who are defined as people who smoked fewer than 100 cigarettes in their lifetime (Zhu, 2020). Environmental risk factors including exposure to air pollution, radon, asbestos, uranium, and diesel exhaust can also lead to lung cancer (Eckel et al., 2016, Benbrahim-Tallaa et al., 2012 and Garshick et al., 2006). In fact, the International Agency for Research on Cancer (IARC) of the World Health Organization (WHO) has classified air pollution as a Group 1 carcinogen with clear evidence of being carcinogenic to humans, accounting for more than 230,000 lung cancer deaths per year worldwide (IARC, 2020 and Lelieveld et al., 2015). These environmental exposures are likely to contribute to lung cancer development in never smokers and can increase lung cancer risk in smokers (Cruz et al., 2011).

Geographical information systems (GIS) have been extensively used to research public health issues in recent years. GIS are potentially powerful resources for community health for many reasons including their ability to integrate data from disparate sources to produce new information, and their inherent visualization (mapping) functions, which can promote creative problem solving and sound decisions with lasting, position impacts on people's live (Buckeridge, 2002 and Maged and Kamel, 2004). This new approach to the epidemiology has come into its own during the last few years as it has become increasingly clear that, from the geological point of view (Xiao-Nong, 2009).

Spatial autocorrelation is an assessment of the correlation of a variable with reference to its spatial location and it deals with the attributes and the locations of the spatial feature (Nakhapakorn and Jirakajohnkool, 2006 and Xiao-Ni et al., 2011). There are two popular indices for measuring spatial autocorrelation applicable to a point distribution: Geary's C Ratio and Moran's I Index. Both indices measure spatial autocorrelation for interval and ratio attribute data (David and Wong, 2005 and Olalekan, 2009). The local indicators of spatial association (LISA) statistics can also be used to identify influential locations in spatial association analysis (Ching-Lan, 2011). The goal of this article is to identify spatial patterns of lung cancer based on a hypothesis, which also revealed previously unsuspected patterns leading to the formulation of additional theories. The spatial analyses were used to investigate spatial patterns of lung cancer. In addition, the LISA was used to indicate the level of spatial autocorrelation that enabled the location of hotspot zonations of lung cancer in Mae Ka subdistrict, Muang district, Phayao province from 2015 to 2020.

2. Materials and Methods

2.1 Study Area: Mae Ka Subdistrict, Muang District, Phayao Province, Thailand

Mae Ka subdistrict, a subdistrict in the northern part of Thailand (Figure 1), had the first highest lung cancer morbidity rate in Thailand in 2020, therefore Mae Ka was selected as the study area because of the high number of cases. Mae Ka subdistrict comprises 18 villages. The Mae Ka subdistrict covers an area of 131 square kilometers with geographical location between 2092000 N to 2116000 N and 584000 E to 612000 E. The subdistrict has a population of about 13,120 people (Department of Province Administrator, 2020). It is mostly covered with forested mountain, with an approximate elevation of 332 meters about mean sea level. Methodology is summarized in the flowchart shown in Figure 2.

2.2 Data Preparation

All new lung cancer patients living in Mae Ka subdistrict based on data at addresses listed in cancer registries, a total of 30 cases were collected from the cancer registries population-level database from the cancer registries Lampang Cancer Hospital (LCH) to analyze the spatial distribution of lung cancer patients. The study of spatial patterns of lung cancer covers the 18 villages for the years 2015 through 2020. Data represented only the patients and were filled in the official form by the LCH. The form provided data for each patient's address, age, gender, the dates of the symptoms and the dates of the death. For village data, locations of each village of Mae Ka subdistrict were collected from the Department of Provincial Administrative (DPA), Thailand. Village point locations were confirmed for accuracy by overlaying on high resolution WorldView-3 satellite images with 31 centimeters resolution. However, risk factors of lung cancer questionnaire were used village's sample population. Sample size, this research was determined the sample size using a case-to-control ratio of 1:2 (Kelsey et al., 1996).

2.3 Data Analysis

2.3.1 Lung cancer affected villages

Data from all the lung cancer cases were geocoded using village location from the address of the patient e.g. 56011005,…,n (56 = province code, 01 = district code, 10 = subdistrict code, and 05 = village code). Initial assessment for geographical accuracy at the village level revealed sufficient information to study the spatial pattern of the disease and allowed us to use the patient address as the location of the infection (Jeefoo et al., 2011 and Jeefoo, 2012). Mapping incidence is the first step in spatial analysis of a disease, but mapping, as always with any ratio, need to be made carefully. Villages with a small number of inhabitants are more variable than villages with high numbers of inhabitants, and ratios may also reflect this difference in statistical variability (Ashley et al., 2011 and Jeefoo, 2012). While a small population density occurs generally in large areas, mapping reinforces this difference and may give a false view of observed reality. To overcome this problem, an empirical Bayes smoothing (EBS) method based on the idea of pooling information across villages was developed (Jeefoo, 2016). Essentially, rates were smoothed and thus stabilized by borrowing strength from other spatial units (Anselin, 2005). The lung cancer incidence rate ($IR_{Lung\text{-}Cancer}$) per year were adjusted by EBS function and converted to the lung cancer morbidity rate by multiplying by 1000 (Jeefoo et al., 2011).

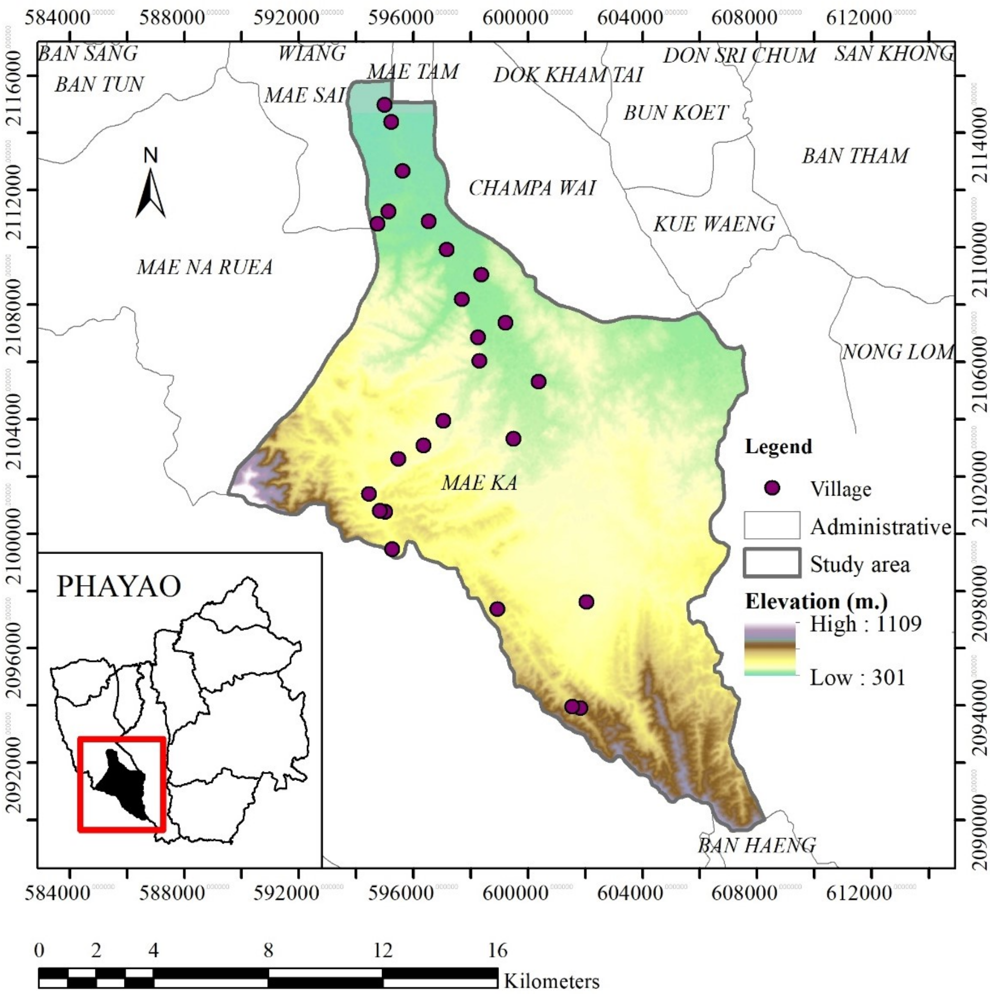


Figure 1: Study area: Mae Ka subdistrict, Muang district, Phayao province, Thailand

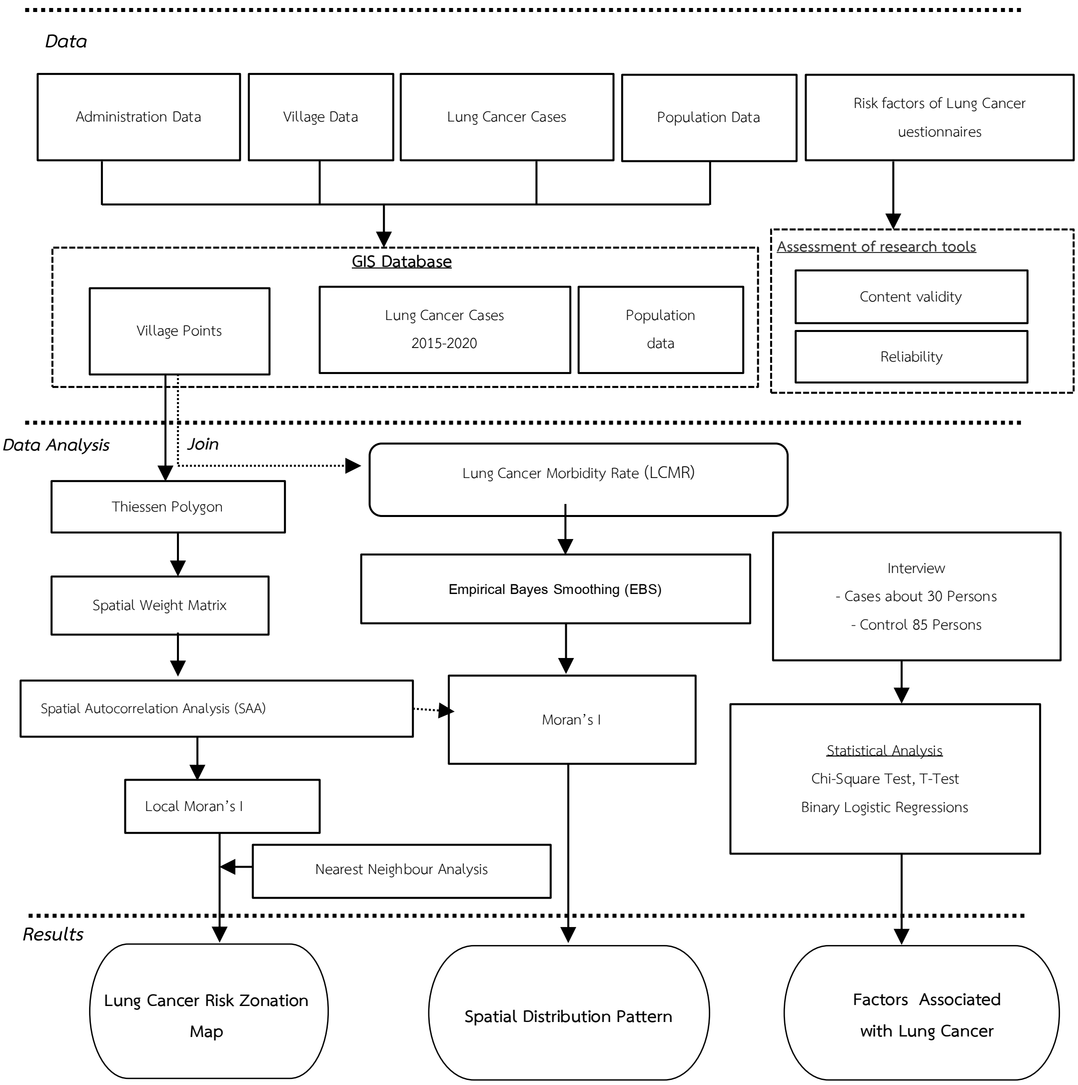


Figure 2: Flowchart of methodology

Table 1: Numeric Scales of Moran's I ratio

Spatial Patterns	Moran's I
Clustered pattern in which adjacent or nearby points show similar characteristics	$1 > E(I)$
Random pattern in which points do not show particular patterns of similarity	$1 \sim = E(I)$
Dispersed pattern in which adjacent or nearby points show different characteristics	$1 < E(I)$

2.3.2 Spatial pattern analysis

Spatial autocorrelation analysis (SAA) was applied to detect spatial patterns of lung cancer in Mae Ka subdistrict (Kristen, 2011). The village locations and the annualized lung cancer morbidity rate at each of these villages were used in the analyses. SAA was used to measure and test how villages were clustered/dispersed in space with respect to their lung cancer morbidity rate.

To evaluate autocorrelation in lung cancer spatial distribution, Moran's I was used by setting the significance level is 0.01 and the indices were evaluated by simulation (999 permutation tests) (Anselin et al., 2006). All these global indices measure the spatial patterns of lung cancer. The interpretation of the indices values is presented in Table 1.

2.3.3 Hotspot detection

Hotspot is defined as a condition indicating some form of clustering in a spatial distribution (Osei and Duker, 2008). Hotspot detection can be useful, even if the global pattern is not clustered. Moreover, clusters of cases that occur randomly can also have an influence on the spread of an infectious disease. This section describes the methods for detecting hotspots of lung cancer by considering both the location of the villages and lung cancer morbidity rate. Previous spatial analyses evaluated only global patterns. Local indicators of spatial association (LISA) can be used to determine locations of clusters or hotspots. The LISA method was carried out in order to find the lung cancer case hotspot patterns (clustered/dispersed/random) at the local level. In this study, local Moran's I value was used to examine the local level of spatial autocorrelation in order to identify villages where values of the lung cancer morbidity rate were both extreme and geographically homogeneous (Jepsen et al., 2009) (the variability of the local indices is evaluated by simulation and the spatial pattern of the villages is not influencing the result). This led to identification of so-called lung cancer hotspots, where the value of the index was extremely pronounced across localities, as well as those of spatial outliers. Firstly, the standardized values of lung cancer morbidity rate were calculated using the spatial weight matrix that defined a local neighborhood around each geographic unit and 999 permutation tests by setting the significance level as 0.01. Secondly, a Moran scatterplot was produced with a spatial lag of lung cancer morbidity rate on the vertical axis and a standardized lung cancer morbidity rate on the horizontal axis.

Once a significance level was set values could also be plotted on a map to display the specific locations of hotspots: locations with high values with similar neighbours (high-high) and potential outliers (Getis and Ord, 1992). The last, to compare hotspot locations with lung cancer disease spatial distribution, kernel density (KD) interpolation was used to create a continuous surface representing the density of lung cancer morbidity rate across the study area (Osei and Duker, 2008).

2.3.4 Factors associated with the incidence of lung cancer

The study of factors related to lung cancer incidence in Mae Ka subdistrict. It is an analytical research by a retrospective study case-control study, with a sample ratio of 1:2, consisted of case-control study 30 patients with lung cancer registered at Lampang Cancer Hospital, and controls, which consisted of 85 peoples without lung cancer, totally 115 peoples. The cases in which both groups had similar characteristics, namely sex, age (5 years) and living in the same village. Conducted field trips to collect data with interview forms between May and June 2021.

2.3.5 Software

Various software's, namely SPSS, GeoDa, and ArcGIS (www.esri.com), were used in this study. SPSS was calculated of lung cancer patients and factors associated, GeoDa was used for hotspot detection and spatial pattern and ArcGIS was used for creating the spatial analysis and mapping.

3. Results

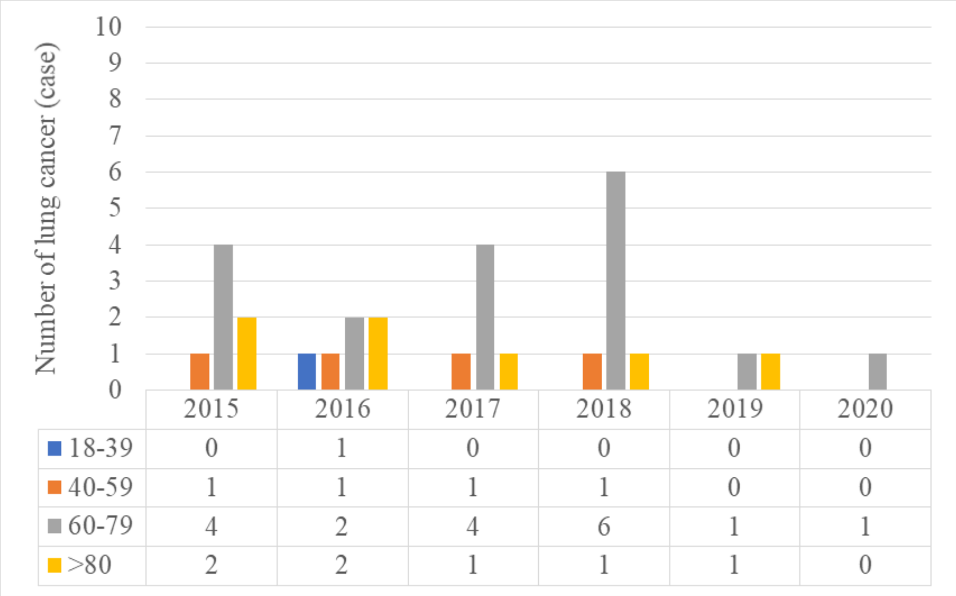
3.1 General analysis

General information of lung cancer patients in Mae Ka subdistrict contains distribution information of patients based on personal information consist of cancer cell types, year of diagnosis, and patient housing. In 2018, the total number was 8 cases, which is the highest recorded incidence for the current decade. After 2018, a slow decrease was seen in the number of lung cancer until 2020, when another increase occurred. The lowest occurrence was in 2020 (1 case). As shown in Table 2, in total 30 cases were reported, including 20 males and 10 females. During the highest lung cancer incidence in year 2018, 6 male and 2 female patients were suspected cases. There were slightly more male patients (66.67%) than female patients.

Table 2: Number of lung cancer classified by gender over the years 2015-2020

Gender	Year						Percentage (%)
	2015	2016	2017	2018	2019	2020	
Male	5	3	4	6	2	0	66.67
Female	2	3	2	2	0	1	33.33
Total	7	6	6	8	2	1	100.00

ISSN 2673-0014 (Online) / © Geoinformatics International



	2015	2016	2017	2018	2019	2020
18-39	0	1	0	0	0	0
40-59	1	1	1	1	0	0
60-79	4	2	4	6	1	1
>80	2	2	1	1	1	0

Figure 3: Shows the age group of lung cancer patients

Table 3: Global indices of spatial autocorrelation

Year	**Moran's I**	**Pattern**
2015	0.2513^{g}	Clustered
2016	-0.1061^{g}	Dispersed
2017	0.1386^{g}	Clustered
2018	-0.134^{g}	Dispersed
2019	0.3292^{g}	Clustered
2020	-0.0624^{g}	Dispersed

The epidemiological data collected from 2015-2020 was classified into several demographic groups such as gender, age groups. Lung cancer distribution based on age of the patients was also determined. The age distribution of lung cancer cases observed was different from the general population age distribution in the Mae Ka subdistrict (Figure 3). The highest incidence was in the 60-79 years age group with a percentage of 60% (18 cases), while incidence in the older than 80 years age group was 23.37% (7 cases). In the population younger than 59 years, the incidence was only 16.33% (5 cases).

3.2 Spatial Pattern Analysis

Table 3 gives the global spatial autocorrelation analysis for annualized morbidity rate of villages in Mae Ka subdistrict from 2015 through 2020 showed that the Moran's I (-0.1341 to 0.3292) values were significant (significance < 0.01) for each year, implying that distribution of the affected villages with lung cancer was somewhat spatially autocorrelated (low clustered) though the overall tendencies were not so strong. The global spatial autocorrelation analysis with Moran's I indices showed that the spatial distribution of Lung Cancer Morbidity Rate (LCMR) was clustered for 2015, 2017 and 2019 years (Table 3). The highest of Moran's I value was confirmed 0.3292 and in 2019. Global Moran's I indices measured the autocorrelation in the incidence villages with low lung cancer morbidity rate or on cases, which were therefore less sensitive to clustering.

3.3 Hotspot Detection

The map in Figure 4 shows the locations with significant local Moran statistics and classifies those locations by type of association (LISA cluster map). The outputs from LISA represent the spatial autocorrelation of lung cancer at the village level. The study only focused on the univariate spatial distribution and the location of any significant clusters or spatial outliers in the LCMR data.

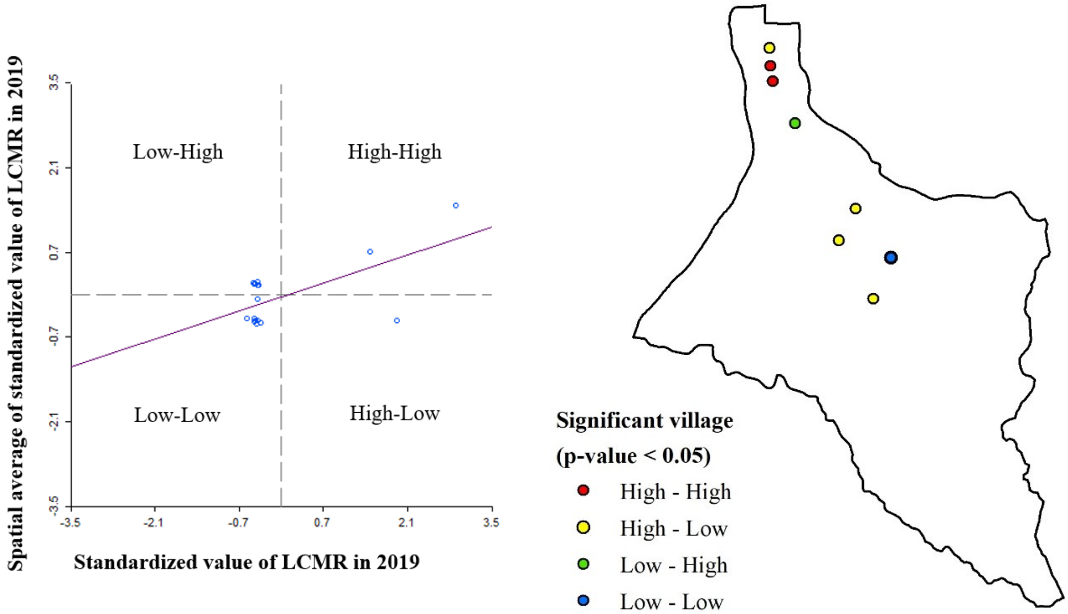


Figure 4: Moran scatter plot matrix and LISA cluster (hotspot) maps of LCMR for p-value < 0.05 for the year 2019

On the right hand panel of Figure 4, the sample LISA cluster map of LCMR in the years of 2019 is shown, depicting the locations of significant local Moran's I statistics, classified by type of spatial association as the red and yellow locations indicating spatial clusters (high surrounded by high, and low surrounded by low), the green and blue indicating spatial outliers (low surrounded by high, and high surrounded by low). There were some outstanding spatial clusters of LCMR coving specific areas in each year. The clustered villages with high LCMR (hotspot: red colored point) was found to cover the conurbations in the north of Mae Ka subdistrict which is Ban Mae Tam Bun Yong. The standardized values of LCMR in each village are displayed in spatial scatter plot to contrast observed value with their spatial average (spatial averaged adjacent values), and to detect outliers by obtaining the significant level as 0.01.

The lung cancer hotspots (high – high values) were illustrated by interpolating the values over the space by using Kernel Density (KD), as shown in Figure 4. These maps show clear spatial patterns of lung cancer that were concentrated in north (Ban Mae Tam Bun Yong village) of the study area during 2015, 2017, and 2020 while in 2016, 2018 and 2019 they were mostly spread in the middle (Ban Bua, Ban Mae Ka Luang, and Ban Mae Ka Hua Thung villages) of Mae Ka subdistrict. The highest density of clustering of hotspots occurred within the local areas of Ban Mae Tam Bun Yong village for the year 2015, 2017, and 2020, Ban Bua village in 2016, Ban Mae Ka Luang in 2018 and Ban Mae Ka Hua Thung in 2019 (Figure 5).

3.4 Factors Associated with the Incidence of Lung Cancer

The sample population interviewed in 18 villages, Mae Ka subdistrict, totaling 85 cases, divided into 27 males, representing 31.76%, and 58 females, representing 68.24% (1:0.5). The most common age group is over 60 years old, accounting for 58.82%, followed by 40-59 years old, accounting for 41.18%. In addition, 29 people are engaged in agriculture, accounting for 34.12%, followed by trading accounted for 21.18% and unemployment accounted for 18.82% of congenital disease, found that 29.41% had congenital disease and a history of cancer was 5.88%. Moreover, family history of cancer was found at 28.24% respectively (Table 4).

Studying behavioral risk factors is essential to definitively explain the causes of lung cancer. This research study found that 42 people did not drink alcohol, accounting for 49.41%, followed by drinking once a month or less, 15 people accounted for 17.65%, did not drink alcohol, 12 people accounted for 14.12%. There were 12 people who drank 2-3 and 2-4 times a week, representing 14.12%. Finally, 4 peoples were drinking alcohol four times a week, representing 4.71%. Therefore, the alcohol consumption factor of the sample population was in the level was low compared to the non-alcoholic group.

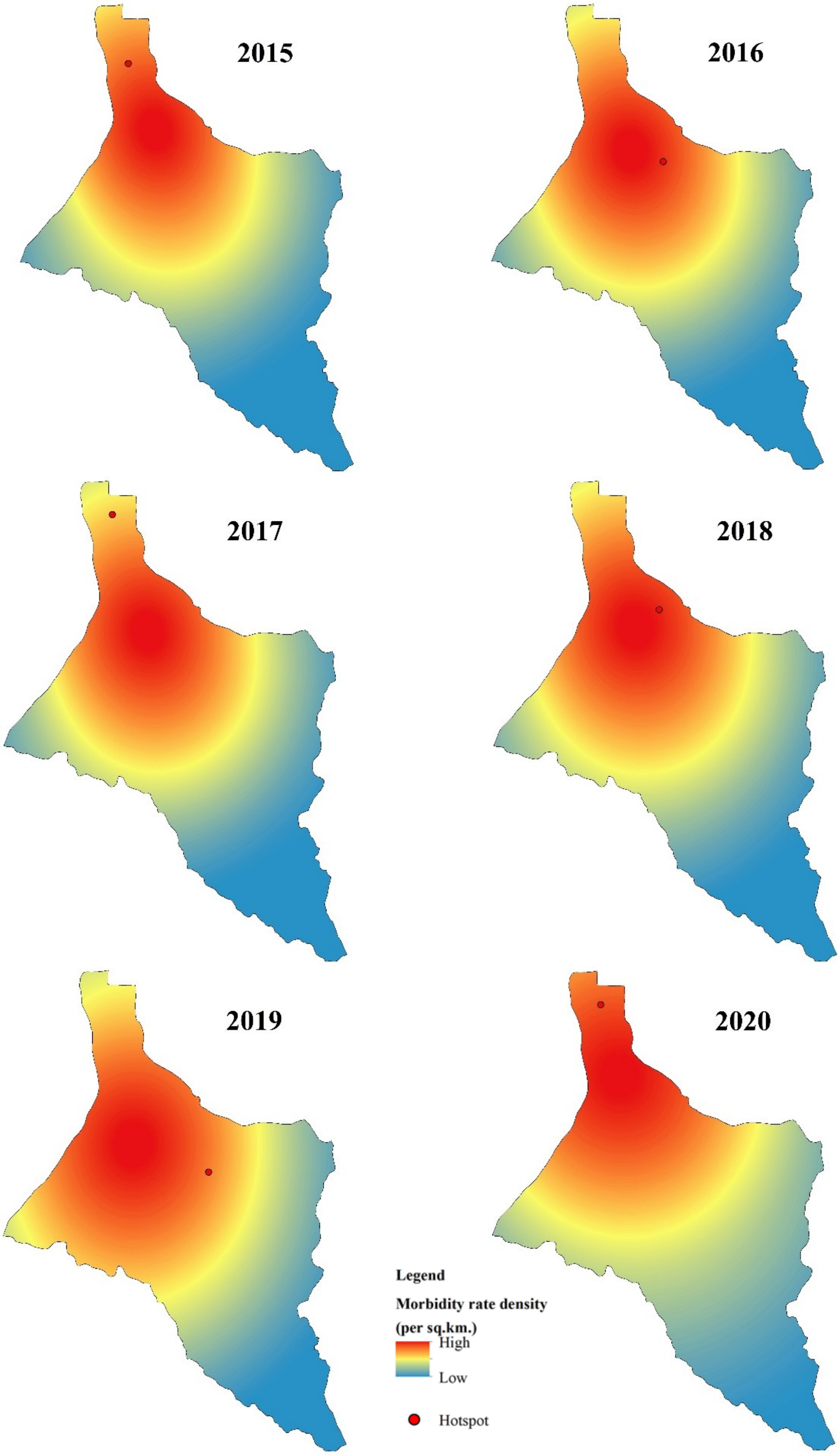


Figure 5: Hotspots of Lung Cancer during 2015 to 2020

Table 4: Personal factors

Age	Description	Total (People)	Percentage (%)
	< 39 year	0	0.00
	40 - 59 year	35	41.18
	> 60 year	50	58.82
	Total	85	100.00
Gender			
	Male	27	31.76
	Female	58	68.24
	Total	85	100.00
Occupation			
	Unemployed	16	18.82
	Agricultural	29	34.12
	Trade	18	21.18
	Personal business	2	2.35
	General employee	13	15.29
	Government service	0	0.00
	Other	7	8.24
	Total	85	100.00
Congenital disease			
	No underlying disease	60	70.59
	Have a congenital disease	25	29.41
History of lung / respiratory disease			
	No	80	94.12
	Yes	5	5.88
Family history of cancer			
	No	61	71.76
	Yes	24	28.24

Smoking is one of the causes of lung cancer. Which in cigarettes contains many carcinogens. From this study, it was found that 55 people were non-smokers, representing 64.71%, followed by 18 people, representing 21.18%, and had smoked, but quit, there were 12 people, accounting for 12%. It was also found that in the family that smoked, there were 28 people, representing 32.94%. For respiratory protective equipment, it was found that they were worn every time while working, a total of 57 people, representing 67.06%, had never worn a number of 23 people, representing 27.06%, and were worn occasionally while working, totaling 5 people, representing 5.88% respectively. For burning leaves, it was found that the sample population was engaged in burning leaves 14 people representing 16.47% and not burning leaf scraps a number of 71 people representing 83.53%. In addition, the highest cooking fuel was charcoal, 56 people, representing a hundred 65.88% per person, followed by wood with 15 people or 17.65%. Smoke from smoking in the house was found to never 58 people accounted for 68.24% and had 27 people accounted for 31.76%. Smoke from work found that a number of 59 people accounted for 69.41% volatile in the area. Working, found that there were 17 people, representing 20%. Finally, the smoke from the workplace found that there were 21 people, accounting for 24.71% respectively (Table 5).

Table 5: Behavioral risk factors

Factors	Description	People	Percentage (%)
Drinking alcohol			
	Never drink	42	49.41
	Used to drink	12	14.12
	Once a month or less	15	17.65
	2-4 times/week	6	7.06
	2-3 times/week	6	7.06
	4 or more times/week	4	4.71
Smoking			
	Never smoked	55	64.71
	Use to smoke	12	14.12
	Currently still smoking	18	21.18
Smoke from family members			
	No smoke	57	67.06
	Smoke	28	32.94
Respiratory protective equipment			
	Never wear respiratory protection	23	27.06
	Occasionally wear respiratory protection while working	5	5.88
	Always wear respiratory protection when working	57	67.06
Burning leaves			
	No burning	71	83.53
	Burned	14	16.47
Cooking fuel			
	Wood	15	17.65
	Charcoal	56	65.88
	Cooking gas	9	10.59
	Electric stove	5	5.88
	Biomass gas	0	0.00
Smoke from smoking inside the house			
	There is no smoke from smoking inside the house	58	68.24
	There is smoke from smoking inside the house	27	31.76
Smoke from vehicles			
	Never get smoke from the vehicle	26	30.59
	Used to get smoke from vehicles	59	69.41
Smoke from agricultural burning			
	Never	71	83.53
	Ever	14	16.47
Agricultural chemicals			
	Never	68	80.00

	Ever	17	20.00
Dust at work			
	Never	64	75.29
	Ever	21	24.71

4. Discussion

Estimates of lung cancer incidence rate ($IR_{Lung\text{-}Cancer}$) accounted for the variability in population distribution. Bayesian autoregressive smoothing models are often uses in spatial epidemiology (Kathryn, 2012). The Bayesian smoothing technique addresses the issue of heterogeneity in the population at risk, and it is therefore recommended for use in explorative mapping of disease/incidence rates. The study showed that spatial distribution patterns of lung cancer were significantly clustered, and identified the lung cancer hotspots in Mae Ka subdistrict, Phayao province, Northern Thailand. Spatial autocorrelation coefficients indices whether the values of a variable influence each other and measure the strength of their association (Adrian et al., 2003). Kernel density estimation illustrated variation in the grouping of lung cancer locations across the study area, and strongly confirmed the visible pattern on the point location map. These maps show clear spatial patterns of lung cancer that were concentrated in north (Ban Mae Tam Bun Yong village) of the study area during 2015, 2017, and 2020 while in 2016, 2018 and 2019 they were mostly spread in the middle (Ban Bua, Ban Mae Ka Luang, and Ban Mae Ka Hua Thung villages) of Mae Ka subdistrict. The highest density of clustering of hotspots occurred within the local areas of Ban Mae Tam Bun Yong village for the year 2015, 2017, and 2020, Ban Bua village in 2016, Ban Mae Ka Luang in 2018 and Ban Mae Ka Hua Thung in 2019.

The sample population interviewed in 18 villages, Mae Ka sub-district, totaling 85 cases, divided into 27 males, representing 31.76%, and 58 females, representing 68.24%. The most common age group is over 60 years old, accounting for 58.82%. In addition, 29 people are engaged in agriculture accounting for 34.12%, followed by trading accounted for 21.18% and unemployment accounted for 18.82% of congenital disease, found that 29.41% had congenital disease and a history of cancer was 5.88%. Moreover, family history of cancer was found at 28.24%. Studying behavioral risk factors is essential to definitively explain the causes of lung cancer. This research study found that 42 people did not drink alcohol (49.41%). Drinking once a month or less with 17.65%, did not drink alcohol with 14.12%. There were 14.12%e who drank 2-3 and 2-4 times a week. Finally, 4.71% were drinking alcohol four times a week. Smoking is one of the causes of lung cancer. From this study, it was found that 64.71% were non-smokers followed 21.18% had smoked. It was also found that in the family that smoked, there were 32.94%. For respiratory protective equipment, it was found that they were worn every time while working representing 67.06% had never worn a number of 27.06% and were worn occasionally while working representing 5.88% respectively. For burning leaves, it was found that the sample population was engaged in burning leaves 16.47% and not burning leaf scraps 83.53%.

In addition, the highest cooking fuel was charcoal 65.88%, followed by wood with 17.65%. Smoke from smoking in the house was found to never, 58 people accounted for 68.24% and had 27 people accounted for 31.76%. Smoke from work found that a number of 59 people accounted for 69.41% volatile in the area. Working, found that there were 17 people, representing 20%. Finally, the smoke from the workplace found that there were 21 people, accounting for 24.71% respectively. However, there were some limitations in the study, epidemiological data were unknown for some villages, which were either new or restructured and therefore did not figure in older data from other source.

5. Conclusion

This study also demonstrated that the use of spatial autocorrelation, spatial statistics, cluster detection methods, GIS, and factors associated with the incidence of lung cancer can aid health planners in appropriately assessing and identifying spatial disparities in risk in populations so as to better guide evidence-based health planning decisions. The results showed that proposed methods and tools can be beneficial for public health officers to visualize and understand the distribution and trends of diffusion patterns of diseases and to prepare waning and awareness to the masses. This paper explored the spatial patterns of lung cancer from 2015 to 2020 in Mae Ka subdistrict and show that the spatial distribution is clustered and dispersed, which areas of lung cancer epidemic were densely clustered, and highlight the spatial trends of the hotspots in the study area. Analysis of the spatial distribution of lung cancer shows significant differences from year to year and between different areas. The hotspot maps showed spatial trend patterns of lung cancer diffusion. Villages in the northern part revealed higher incidence. Furthermore, the spatial patterns during the years 2015, 2017 and 2019 were found to represent spatially clustered patterns, both at global and local scales. GIS can also be used as an effective tool to manage and monitor lung cancer and related routine activities. The spatial modeling capacities offered by GIS can help to understand the spatial variation in the incidence of disease, and its covariation with environmental factors with health care. An understanding of epidemiological principles and methods is required to structure studies and interpret results for proper socioeconomic development at various levels of the society.

Acknowledgements

This research study was funded by the University of Phayao capital contract no. FF64-RIM009 for fiscal year 2021. We thank the Lampang Cancer Hospital (LCH), Ministry of Public Health (Thailand), for providing the lung cancer data.

References

Abazari, M., Gholamnejad, M., Roshanaei G., Abazari, R., Roosta, Y. and Mahjub, H., 2015, Estimation of Survival Rates in Patients with Lung Cancer in West Azerbaijan, the Northwest of Iran. *Asian Pacific Journal of Cancer Prevention*, Vol. 16(9), 3923-3926.

Adrian, E., Jose, M. I. and Torres, M. E., 2003, Spatial Analysis of Genetic Diversity as a Tool for Plant Conservation. *Biological Conservation*, 351-365.

Anselin, L., 2005, *Spatial Statistical Modeling in a GIS Environment. In GIS, Spatial Analysis, and Modeling*; Environemntal Systems Research Institute (ESRI) Press: Redlands, CA, USA, 93-111.

Anselin, L., Syabri, I. and Kho, Y., 2006, GeoDa - An Introduction to Spatial Data Analysis. *Geographical Analysis*, Vol. 38(1), 5-22.

Ashley, C. E., Carnes, E. C., Phillips, G. K. Padilla, D., Durfee, P. N., Brown, P. A., Hanna,T. N., Liu, J., Phillips, B., Carter, M. B., Carroll, N. J., Jiang, X., Dunphy, D. R., Willman, C. L., Petsev, D. N., Evans, D. G., Parikh, A. N., Chackerian, B., Wharton, W., Peabody, D. S. and Brinker, C. J., 2011, The Targeted Delivery of Multicomponent Cargos to Cancer Cells by Nanoporous Particle-Supported Lipid Bilayers. *Nature Materials*, Vol. 10, 389-397.

Benbrahim-Tallaa, L., Baan, R. A., Grosse, Y., Lauby-Secretan, B., El Ghissassi, F., Bouvard, V., Guha, N., Loomis, D. and Straif, K., 2012, Carcinogenicity of Diesel-Engine and Gasoline-Engine Exhausts and Some Nitroarenes. *Lancet Oncol*, Vol. 13, 663-664.

Bray, F., Ren. J. S., Masuyer, E. and Ferlay, J., 2013, Global Estimates of Cancer Prevalence for 27 Sites in the Adult Population in 2008. *Int. J. Cancer*, Vol. 132, 1133-1145.

Buckeridge, D. L., 2002, Making Health Data Maps: A Case Study of a Community/University Research Collaboration. *Social Science & Medicine,* Vol. 55(7), 1189-206.

Ching-Lan, C., 2011, Using Spatial Analysis to Demonstrate the Heterogeneity of the Cardiovascular Drug-Prescribing Pattern in Taiwan. *BMC Public Health*, Vol. 11(1), 1-9.

Cruz, C. S. D., Tanoue, L. T. and Matthay, R. A., 2011, Lung Cancer: Epidemiology, Etiology, and Prevention. *Clin. Chest Med.*, Vol. 32, 605-644.

David, W. S. and Wong, J. L., 2005, *Statistical Analysis of Geographic Information with ArcView GIS and ArcGIS*. John Wiley & Sons, Inc. 260-264.

Department of Province Administrator, 2020, DOPA – Thailand [https://stat.bora.dopa.go.th/-new_stat/webPage/statByYear.php].

Eckel, S. P., Cockburn, M., Shu, Y. H., Deng, H., Lurmann, F. W., Liu, L. and Gilliland, F. D., 2016, Air Pollution Affects Lung Cancer Survival. *Thorax*, Vol. 71, 891-898.

Garshick, E., Laden, F., Hart, J. E., Smith, T. J. and Rosner, B., 2006, Smoking Imputation and Lung Cancer in Railroad Workers Exposed to Diesel Exhaust. *Am. J. Ind. Med.*, Vol. 49, 709-718.

Getis, A. and Ord, J. K., 1992, The Analysis of Spatial Association by use of Distance Statistics. *Geographical Analysis*, Vol. 24, 189-207.

Henley, S. J., Richards, T. B., Underwood, J. M., Eheman, C. R., Plescia, M., McAfee, TAJMM, 2014, Lung Cancer Incidence Trends among Men and Women-United States, 2005-2009, Vol. 63(1), 1-5.

IARC, International Agency for Research on Cancer. Available online: https://www.iarc.-who.int/ (accessed on 28 August 2021).

Jeefoo, P., Tripathi, N. K. and Souris, M., 2011, Spatio-Temporal Diffusion and Hotspot Detection of Dengue in Chachoengsao Province, Thailand. *International Journal of Environmental Research and Public Health*, Vol. 8, 51-74.

Jeefoo, P., 2012, Space-Time Analysis Tools of Dengue Epidemics in Chachoengsao Province, Thailand. *International Journal of Geoinformatics*, Vol. 8(3), 9-13.

Jeefoo, P., 2016, Analyzing Spatial Clustering and Hotspots Detection of HIV/AIDS Prevalence using GIS Technology. *International Journal of Geoinformatics*, Vol. 12(1), 65-73.

Jepsen, M. R., Simonsen, J. and Ethelberg, J. S., 2009, Spatio-temporal Cluster Analysis of the Incidence of *Campylobacter* Cases and Patients with General Diarrhea in a Danish County, 1995–2004. *International Journal of Health Geographics*, Vol. 8, 1-12.

Kathryn, T. M., 2012, Application of Bayesian Spatial Smoothing Models to Assess Agricultural Self-Sufficiency. *International Journal of Geographic Information Science*, Vol. 26(7), 1213-1229.

Kelsey, J. L., Whittemore, A. S., Evans, A. S. and Thompson, W. D., 1996, *Methods in Observational Epidemiology*. Oxford University Press.

Kristen, H., 2011, Adjusting for Sampling Variability in Sparse Data: Geostatistical Approaches to Disease Mapping. *International Journal of Health Geographics*, Vol. 10, 1-17.

Lelieveld, J., Evans, J. S., Fnais, M., Giannadaki, D. and Pozzer, A., 2015, The Contribution of Outdoor Air Pollution Sources to Premature Mortality on a Global Scale. *Nature*, Vol. 525, 367-371.

Luchtenborg, M., Riaz, S. P., Lim, E., Page, R., Baldwin, D. R. and Jakobsen, E., 2014, Survival of Patients With Small Cell Lung Cancer Undergoing Lung Resection in England, 1998-2009. *Thorax*, Vol. 69, 269-273.

Maged, N. and Kamel, B., 2004, Towards Evidence-Based, GIS-driven National Spatial Health Information Infrastructure and Surveillance Services in the United Kingdom. *International Journal of Health Geographics*, Vol. 3, 1-50.

Nakhapakorn, K. and Jirakajohnkool, S., 2006, Temporal and Spatial Autocorrelation Statistics of Dengue Fever. *Dengue Bulletin*, Vol. 30, 177-183.

Olalekan, A. U., 2009, A Trend Analysis and Sub-Regional Distribution in Number of People Living with HIV and Dying with TB in Africa, 1991 to 2006. *International Journal of Health Geographics*, Vol. 8, 1-8.

Osei, F. B. and Duker, A. A., 2008, Spatial and Demographic Patterns of Cholera in Ashanti Region-Ghana. *International Journal of Health Geographics,* Vol. 7, 1-10.

Siegel, R., Naishadham, D. and Jemal, A., 2012, Cancer Statistics, 2012. *CA Cancer J. Clin.,* Vol. 62, 10-29.

Tarik, T., 2013, Investigating Environmental Health Studies using Geographical Information Systems: Cancer Case Study. *Pol. J. Environ. Stud.*, Vol. 22(5), 1505-1517.

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J. and Jemal, A., 2015, Global Cancer Statistics, 2012. *CA: A Cancer Journal For Clinicians*, Vol. 65(2), 87-108.

Xiao-Nong, Z., 2009, Spatial Epidemiology in Zoonotic Parasitic Disease: Insights Gained at the 1st International Symposium on Geospatial Helth in Lijiang, China, 2007. *Parasites & Vectors*, Vol. 2, 10.

Xiao-Ni, H., 2011, Spatial Pattern Analysis of Heavy Metals in Beijing Agricultural Soils Based on Spatial Autocorrelation Statistics. *International Journal of Environmental Research and Public Health*, Vol. 8, 2074-2089.

Xu, X. L., Zheng, J., Mao, W. M. and Ling, Z. Q., 2016, RRM1 *151A>T, RRM1-756T>C, and RRM1-585T>Gis associated with increased susceptibility of lung cancer in Chinese patients. *Cancer Medicine*, Vol. 5(8), 2084-2090.

Zhu, Y., Mckeon, T. P., Tam, V., Vachani, A., Penning, T. M. and Hwang, W. T., 2020, Geographic Differences in Lung Cancer Incidence: A Study of a Major Metropolitan Area within Southeastern Pennsylvania. *Int. J. Environ. Res. Public Health*, Vol. 17, 9498, 1-14.