Fakulti Teknologi Maklumat dan Sains Kuantitatif ISSN:1823-0822 Jilid 7, Bil. 1, 2005

UNIVERSITI TEKNOLOGI MARA

JURNAL TEKNOLOGI MAKLUMAT DAN SAINS KUANTITATIF

Kandungan **Muka Surat**

Universiti Teknologi MARA

Dasar Penerbitan:
Jurnal Teknologi Maklumat dan Sains Kuantitatif diterbitkan oleh Fakulti Teknologi Maklumat dan Sains Kuantitatif, Universiti Teknologi MARA. Sumbangan penulisan adalah berkaitan dengan teori, amali, metodologi serta falsafah aspek-aspek kefenomenan dan epistemologi dalam sains matematik dan pengkomputeran. Tujuan utama jurnal ini adalah untuk mengenengahkan bahan/karya yang menunjuk dan mempersembahkan keselarasan serta keharmonian dalam teknologi maklumat dan sains kuantitatif. kandungan makalah yang dimuatkan dalam jurnal ini tidak semestinya mencerminkan pandangan dan pendirian rasmi jurnal ini.

Alamatkan semua sumbangan kepada:
Urusetia Panel Penyunting,
Jurnal Teknologi Maklumat dan Sains Kuantitatif
Fakulti Teknologi Maklumat dan Sains Kuantitatif
UiTM Shah Alam

Tel	**03-5543 5329**
Faks	**03-5543 5501**
e-mail	mshahar@tmsk.uitm.edu.my

DARI MEJA KETUA PENYUNTING

Alhamdulillah, dapat kita terbitkan Jurnal Teknologi Maklumat dan Sains Kuantitatif Jilid 7, Bil.1, 2005. Saya rasa pencinta ilmu menanti-nanti terbitan kali ini.

Seperti biasa jurnal terbitan sesuatu tahun itu, hanya dapat dihantar untuk percetakan dua atau tiga bulan berikutnya. Kadangkala, penulis yang telah menghantar balik artikel yang telah diwasitkan itu tertunggu-tunggu juga adakah artikelnya diterbitkan kali ini. Sememangnya pihak penyunting mengamalkan prinsip giliran FIFO (first in first out), tetapi kadangkala ianya tidak boleh dilakukan. Ini kerana sesuatu bidang pengkhususan itu mempunyai dua atau tiga artikel sekaligus. Jadi pihak penyunting berkemungkinan akan melewatkan salah satu daripada artikel sebidang itu kemudian. Justeru itu, giliran FIFO masih dilakukan dalam bidang yang sama.

Dalam keluaran yang lepas, saya ada mengatakan bahawa minat penulis akan terhakis apabila maklumbalas tentang penerimaan sesuatu artikel untuk diterbitkan itu lambat. Saya hanya boleh memberi nasihat kepada penulis supaya bersabar, sebab ini begantung kepada pewasit yang menilai itu sibuk atau tidak, sanggup atau tidak dan sebagainya. Percayalah, kesabaran itu akan menjadi kita penulis yang berdisiplin.

Akhir kata, saya harap semua penulis-penulis semasa dan yang akan datang tetap gigih untuk menulis supaya karya kita dapat dimanfaatkan oleh para ilmuwan yang lain dalam bidang kita iaitu Teknologi Maklumat dan Sains Kuantitatif

Terima kasih.

Ketua Penyunting.
Prof. Dr. Mohd Sahar Sawiran

Jurnal Tek. Maklumat & Sains Kuantitatif ISSN 1823-0822 Jilid 7,Bil 1, 2005

A Preliminary Study On The Collaborative Use Of Statistical Modeling In A GIS Study Of Asthmatic Morbidity

Mohammad Said Zainol, Sayed Jamaludin S. Ali, Zainal Mat Saat

Fakulti Teknologi Maklumat dan Sains Kuantitatif, Universiti Teknologi MARA,
40450 Shah Alam, Selangor
e-mail: said@tmsk.uitm.edu.my

Abstract

This paper reports an ongoing effort in the collaborative use of statistical modeling techniques and GIS in an epidemiological related research. The problems of increasing asthmatic prevalence and morbidity and their relationship to environmental pollutants in the state of Perlis are being investigated. The main hypothesis is that asthmatic prevalence and morbidity increases where and when levels of specific air pollutants increase.

The widely common conjecture that exposure to air pollutants has adverse effects upon respiratory diseases, particularly asthma, is translated into a belief that air pollutants correlate with asthmatic prevalence and morbidity, which in turn calls for the construction of models of statistical relationships between levels of air pollutants and asthmatic prevalence. Outputs of this model provide an interesting dimension to the use of GIS technology in the research into asthmatic morbidity and prevalence.

Keywords: Air pollutants, Asthma Morbidity, Geographical Information System, Statistical Models

1. Introduction

The geographic distribution of residents suffering from asthma in Malaysia varied across the country. An asthma surveillance program conducted by the Disease Control Division of the Ministry of Health identified incidence of acute exacerbation of asthmatic cases in 5 locations in the country [1]. A study conducted by the Ministry of Health, Malaysia found that the state of Perlis has the highest percentage of asthmatic cases [5].

Works involving the use of GIS in health research, linking air pollution models to GIS for the purpose of defining areas of exposure have been reported. Gatrell et al. [2] studied the use of modern point pattern methods to explore and model disease risks. Methods for detecting disease clustering were also described. The result suggested that disease clusters could not be investigated unless their sizes and boundaries coincide at least roughly with the spatial units for which the data have been encoded. There were tendency for cases to cluster or aggregate more than the population at risk.

Martin [3] reported the use of statistical methods for spatial epidemic modeling, which include use of spatial regression, tests for spatial randomness and techniques of map smoothing. These analyses portrayed patterns of disease rates on a choropleth map and showed rates with different statistical reliability in different areas. Rogerson [4] developed a spatial version of the chi-square goodness-of-fit statistics which were used to test for spatial clustering. The approach was able to filter disease cases and the people at risk for areas that can be controlled in both size and shape. The technique enable the computation of the likelihood that clusters exist at particular locations.

2. Objective

The current study is a preliminary stage of an ongoing research into investigating the incidence of asthma among the population using a combination of GIS and statistical techniques. The association between asthmatic disease and the environment will be investigated using a GIS analysis to examine disease patterns and disease rates at various levels of spatial resolution. It is complemented by a collaborative use of statistical analyses to identify causal factors of asthmatic problems.

The current preliminary stage reported in this paper is a research to test the hypothesis that non-environmental factors such as physical stress, emotional stress and physical surroundings are not significant predictors to the incidences of asthma among the population. It is an attempt to isolate these factors from those which are conjectured to contribute to asthmatic incidence, i.e., environmental pollutants.

3. Methodology

3.1 Study Area

The study area, the state of Perlis, covers an area of 810 square kilometres. A large portion of the state is low lying and well under 61 meters. The state capital is Kangar. Arau, the *Royal Town* is 10 km away. There are 22 districts (or mukim) in Perlis. The economic activity of the state is predominantly agriculture which made up about 65.3% of the land use, with a small industrial sector.

The climate of Perlis is tropical monsoon. Temperature is relatively uniform within the range of 21°C to 32°C throughout the year. Humidity is consistently high on the low lands ranging between 82% to 86% per annum. The mean rainfall is between 2,032 mm to 2,540 mm with the wettest months from May to December. During the months of January to April the weather is generally dry and hot.

There are several reasons for choosing the area. Firstly, the state of Perlis has been identified as having the highest percentage of respiratory and asthmatic cases in Malaysia. Secondly, the required map and information on asthmatic patients are available and accessible. Thirdly, it is an area of environmental interest;. the area, although small, is concentrated in its industrial activities. There is a large cement factory, a fairly vast sugar cane plantation complete with a refinery, as well as a vast padi plantation. The fairly well known post-harvest open burning activities of both the sugar cane and the padi plantations has invited numerous conjectures on their roles in the aggravation of asthmatic problems among the population. However, to date there has been no thorough study to support the conjectures.

3.2 Data

The implementation of the study involves the process of data acquisition among a sample which consists of (a) persons registered as asthmatic patients in clinics plus (b) a sample of non-patients selected randomly from among the neighbors of the sampled patients. The data acquisition process involves the planning and execution of a sample survey to capture data pertaining to factors and dimensions of asthmatic problems, profiles of respondents (category of asthmatic problems, demographic profile) as well as geographical and environmental profiles of location of residence of patients.

Data on asthmatic patients who seek treatment at the local hospital from January 2003 through March 2004 were collected. Patients' addresses were geo-coded and added to a GIS layer of census tract. Incidence rates were calculated for each census tract. Maps were created using Mapinfo and ArcView. They will later be used for identifying areas at risk of asthmatic disease. Smoothed rate maps will be produced to identify spatial patterns of asthmatic problems. The results will then be used to produce GIS maps which will be useful in asthmatic disease prevention programs in specific areas and help community groups understand the impact of air pollutants on respiratory diseases.

3.3 Variables of Interest

Variables of interest up to the current stage of study include, among others, category of asthmatic problems, demographic, geographical and environmental profiles (Refer Table 1).

3.4 The Asthmatic Database of the Perlis Health Department

An important data source for this research is the filled questionnaires which form part of the records of more than 1000 patients between ages of 1 to 80 years kept in the Perlis Health Department. Apart from allowing access to the data source, the Perlis Health Department has also assisted in

(i) The collection/compilation of data about frequency and seriousness of asthma and allergies in the population from various age categories under different living conditions.
(ii) The collection of basic epidemiological data, in order to make predictions about variations in frequency and seriousness of these illnesses in future years.
(iii) The development of a framework for future research, examining links with genetics, lifestyle, environmental factors and medical care.

Table 1: List of Variables Measured in Preliminary Survey

Group I	Group 2
1. Demographic profile of Respondents	1. Distance from suspected pollutant source
2.. Asthma incidence density	2. Pollutants (e.g., carbon monoxide, sulphur dioxide) measures
3. Standard of living profile of Respondents	3. Environmental measures (e.g., rainfall, humidity, temperature, pressure)
4. Health profiles (Related to Asthmatic Problems)	4. Geographical and environmental profiles of location of patients and controls
5. Factors related to asthmatic problems	
6. Stress factors	

4. Analysis

4.1 GIS Analysis

Visualization in GIS is important for better understanding, while statistical tests provide new understanding of associations between epidemiological and environmental phenomena. Figure 1 shows the perspective view of the terrain of the study area. The currently available health data is not yet complete, and do not cover the whole of the state of Perlis.

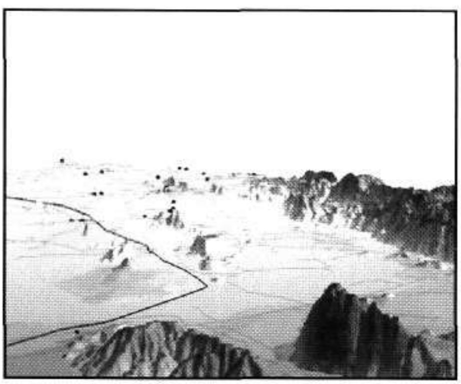

Figure 1: Perspective View of Terrain in the Study Area

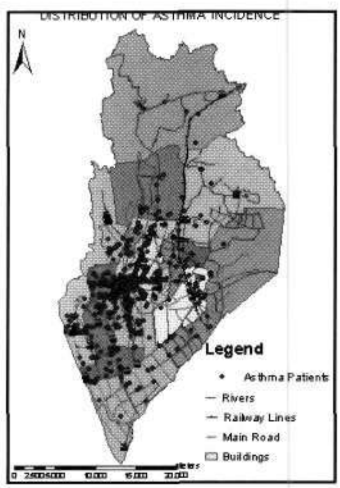

Figure 2: Distribution of Asthma Incidence in Study Area

In the next stage of this ongoing study a GIS database will be built by collecting and converting topographic maps, land use maps and other related map data into a GIS system. A GIS spatial analysis will be used to examine disease patterns and disease rates at different levels of spatial resolution. The detection of clusters will also be carried out in order to investigate the likelihood of their existence at particular locations. A logistic regression analysis will be used to estimate the probability of occurrence of asthma among residents at a particular location.

A frequency map of the distribution of asthmatic incidence within the study area has been constructed using data obtained in this preliminary stage. Figure 2 shows a spatial distribution of the asthmatic incidence in the area under study while Figure 3 shows the location of factories of the area. On closer scrutiny it becomes evident that asthmatic incidence tends to cluster in the Western and in the middle of the region. The North-East is the location of the cement factory and the sugar refinery, conjectured to be the sources of chemical emission and burning activities. Other activities such as quarrying, rice milling and other smaller industries are distributed over the region. The extreme western part of the high asthmatic incidence area is centred at the relatively high populated area of the capital town of Kangar.

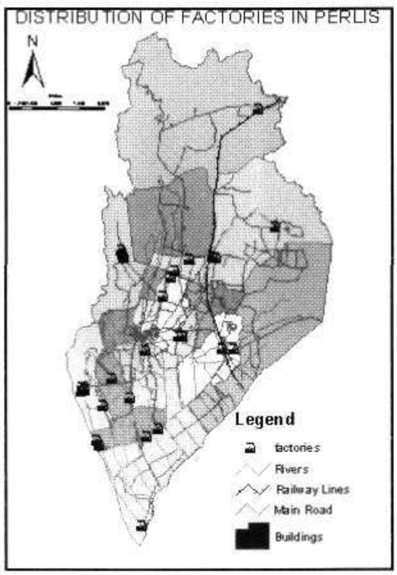

Figure 3: Location of Factories in Study Area

4.2 Spatial Analysis

The preliminary analysis of the data takes place at address-based levels since the asthma patients can be located at exact geographical locations. The first step is to find any spatial disease patterns. The results of this analysis show that the number of cases with diagnosis of asthma, asthma symptoms during the last 12 months is more than expected. In these cases a further analysis is definitely worthwhile.

The point data at address level have the following properties: Since their complete address is known, a pair of xy-coordinates is attached to every patient. There more than 100 attribute values known at each point, information about disease symptoms, the environment, lifestyle and related attributes, including stress factors. The information about disease symptoms is mainly bivariate: the individuals have the symptom or they do not have [2].

Because of confidentiality, raw point data cannot be visualized as they stand but has to be aggregated to small areas. The research region will be divided into small administrative areas called "mukim."

The epidemiologist of the Health Department has constructed two broad hypotheses as a starting point for the spatial analysis on this level. First, that there is an obvious relationship between the air pollution (conjectured to originate from the cement factory, sugar refinery/ plantation and the padi plantation), and different allergy/asthmatic symptoms. Second, that there could be a relationship between social status and different allergy/asthmatic symptoms.

The study aims to investigate if there is any clustering in the 21 symptoms. A variety of methods exist to detect clusters and clustering in a point map [2]. Wartenberg and Greenberg [6] describe a strategy to select an appropriate method of cluster detection. First, the selection of the data type: the location of an event, the distance between all pairs of events, the nearest-neighbor distance between events, or the distance to a fixed point.

After finding an eventual spatial pattern in the data, this pattern will be compared with the spatial pattern of possible causal factors. Correlation, covariance and regression methods are often used to detect relations between variables. Since disease data often have a binomial character, logistic regression will be used.

Table 2: Results of the Tests of Differences Between the Patients and the Control Group With Respect to the Prevalence of 21 Symptoms of Asthma.

Symptoms	Chi-Square	p-Value	Symptoms	Chi-Square	p-Value
1 Wheezing	361.2	0.000	12 Severe cough when awaken	96.9	0.000
2 Short of breath	383.8	0.000	13 Persistent cough	67.7	0.000
3 Whistling breath	265.9	0.000	14 Continuous cough for more than 3 months	14.7	0.000
4 Tight chest	290.2	0.000	15 Phlegm on most mornings	57.1	0.000
5 Short of breath during the day	290.4	0.000	16 Phlegm day and night	41.3	0.000
6 Short of breath after heavy work	277.9	0.000	17 Continuous phlegm for more than 3 months	100.9	0.001
7 Awaken due to short of breath	300.1	0.000	18 Heavy sneezing with watery nasal	44.5	0.000
8 Asthmatic attack	599.0	0.000	19 Sneezing with itchy /watery eyes	59.9	0.000
9 Nose allergy	32.2	0.000	20 Persistent skin rashes	14.2	0.000
10 Eczema	15.5	0.000	21 Itch under armpit, under knee, neck, eyes	12.3	0.002
11 Awaken due to severe cough	123.1	0.000			

* (Note: All p-values are less than 0.05)

4.3 Statistical Analysis

The results of the analysis on whether non-environmental factors (non-environmental factors investigated in this study include emotional and physical stress, physical surrounding, and socio-economic and habitual factors) contribute to asthmatic problems among the population yield the following results:

a. there were significant differences between the patients and the control group in the prevalence of 21 symptoms related to asthma. Hence it is concluded that respondents representing the control group has no significant asthmatic problems (Refer to Table 2).

b. Analysis on whether stress factors contribute to the asthmatic problems among the population showed no significant difference in the effect of each of the 39 emotional and physical stress factors on the patients and those in the control group. One-Way ANOVA analysis between each of the 39 stress factors and respondents' category (whether they are patients or control) yielded p-values greater than 0.05 in all cases (Refer Figure 4). This means that in all these cases the null hypothesis of no difference in the effect of stress between the patients and the control group are accepted. Hence it is concluded that these factors has no causal effect on the asthmatic problems.

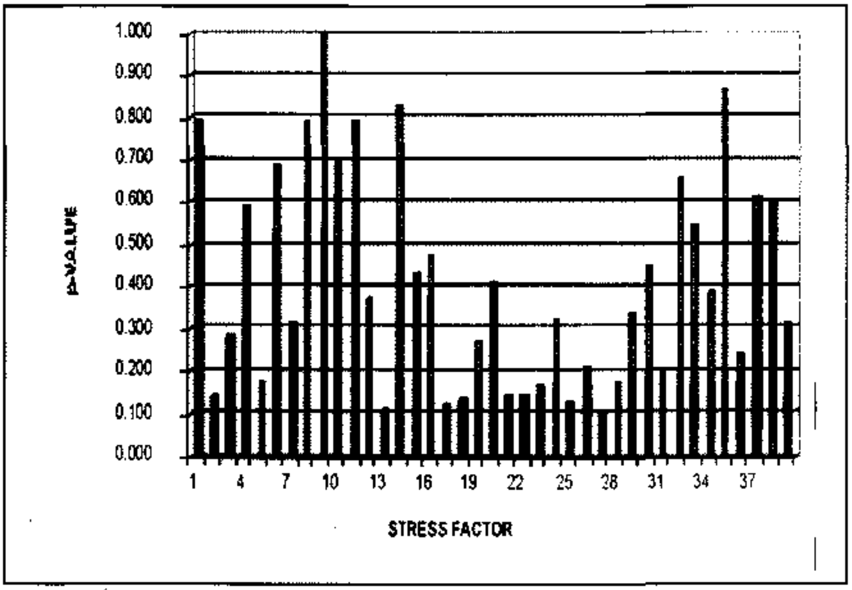

Figure 4: Plot of p-Value for each of the 39 Stress Factors
(Note: All p-values exceed 0.05)

In the next stage of the research, measurements of particulate matter (air pollutant index) will be obtained. These will then be used to obtain census tracts which in turn will be used to examine the spatial distribution of asthmatic illness and its relationship to areas of elevated particulate matter.

5. Conclusion

The use of spatial and statistical analysis in health research projects is useful, but there are still important issues to be resolved. Spatial statistical analysis, although complex, is able to produce statements about spatial patterns in epidemiological data. The current study is part of an ongoing project. Digitized basic map coverage, georefenced patients' data, visualized queries on different aggregation levels and a fundamental statistical analysis have also been made.

Statistical analysis on the contribution of stress factors to asthmatic problem showed no significant difference in their effect between the patients and the control group. It is thus concluded that these factors has no causal effect on the asthmatic problems. The investigation into the effects of environmental factors on asthmatic problems in the area will be carried out in the next stage of the study. It will involve, among others, investigations into the conjectured sources of pollution (the cement factory, the sugar refinery, the sugar cane and padi plantations), the contents of pollutants in the ambient air and their role in asthma morbidity in Perlis and will conclude with the development of a respiratory health profile of the state of Perlis.

References

1. Daud, A.R., Rozlan, I., Doraisingam & Ikhsan, M.S., 2002. *Environmental Health: Malaysian's Perspective*, Disease Control Division (NCD), Ministry of Health Malaysia Publications, **1(2)**, pp. 4-8.

2. Gatrell, A.C., Bailey, T.C., Diggle, P.J. & Rowlingson, B.S.1996. *Spatial point pattern analysis and its application in geographical epidemiology*. Transactions, Institute of British Geographers, No. 21, pp. 256-274.

3. Kulldorff, Martin. 1998. Statistical methods for spatial epidemiology: Tests for randomness. *GIS and Health GISDATA 6*, eds. A. Gatrell & M. Loytonen, Taylor & Francis, Series Editors: Ian Masser & Francois Salge.

4. Rogerson, P.A.1999. The detection of clusters using a spatial version of *the chi-square goodness-of-fit statistics*. Geographical Analysis, No. 31, pp. 130-147, 1999.

5. Rozlan, I.2002. *The Study on Asthma Admissions in Malaysia*. Disease Control Division (NCD), Ministry of Health Malaysia, 1(1), pp. 10-17.

6. Wartenberg, D. & Greenberg, M.1990. *Space-time models for the detection of clusters of diseases*. Spatial Epidemiology, ed. R.W. Thomas R.W. Pion: London, pp. 17-34, 1990.