Investigation of asthmatic disorders using GIS technology

S. Jamaludin S.A, M. S. Zainol & Z. M. Saat
Universiti Teknologi MARA (UiTM), Shah Alam, Selangor, Malaysia

Abstract

With environmental health moving into a critical stage in Malaysia, environmental management becomes more relevant and challenging. There is increasing evidence that exposure to air pollutants has adverse effects on public health, in particular on cardiac and respiratory disease. The increasing number of asthmatic and other respiratory cases has affected the health and quality of life of the population. With the belief that air pollution contributes to the rise in asthmatic and respiratory cases, there is a need to explore the relationship between air quality and respiratory disease.

As yet, there is no known report of the use of a Geographical Information System (GIS) in epidemiological research in Malaysia. It is our belief that making spatial statistical techniques available for epidemiologists in GIS applications can assist epidemiological research. An attempt by the Ministry of Health, Malaysia to monitor the incidence of acute exacerbation of asthmatic cases in the year 2002 did not seem to give the required results. It did not provide an accurate picture of the effect of air pollution on public health.

This paper describes an ongoing attempt at combining the use of GIS and statistical analysis in an epidemiological related research. The problems of increasing incidence of asthmatic problem and its relationship to environmental pollutants are being investigated. The task is also to investigate whether respiratory disease incidence increases, particularly when levels of specific air pollutants are elevated.

Keywords: air pollutants, asthma, Geographical Information System, spatial statistical analysis.
1 Introduction

Studies conducted by the Ministry of Health, Malaysia found escalating incidence of asthma within the country and identified the state of Perlis as having the highest percentage of respiratory and asthmatic cases [5]. On the other hand, 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]. However the program did not produce the required results due to insufficient data and the health data areas being not geographically represented.

Some 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. Disease incidence were then linked to pollution. Gatrell et al. [2] studied the use of modern point pattern methods in exploring and modelling disease risk. Second-order methods were also included for detecting disease clustering. 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] applied statistical methods for spatial epidemic modelling, which include use of spatial regression, test for spatial randomness and map smoothing techniques. 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 in a test for spatial clustering. This 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 indicated the likelihood that clusters exist at particular locations.

2 Objective

The current study is a preliminary stage of an ongoing research into identifying factors which contribute to the increasing incidence of respiratory problems among the state population using a combination of GIS and statistical techniques. The association between respiratory 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 statistical analyses in order to identify causal factors of respiratory problems. The current preliminary stage reported in this paper is a research into finding facts to test the hypothesis that non-environmental factors such as physical stress, emotional stress and physical surroundings are not significant causes to the incidences of asthma among the population. It is an attempt to try to isolate these factors from the set which are thought to be causal to the problems, i.e., environmental factors.
In the current study data on asthma 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 respiratory disease. Smoothed rate maps will be produced to identify spatial patterns of respiratory problems. The results will then be used to produce GIS maps for the Health Department of the state of Perlis which will be useful in respiratory disease prevention programs in specific areas and help community groups understand the impact of air pollution on respiratory disease.

3 Geographical area of study

The study area, the state of Perlis is the smallest and northern most state in Peninsular Malaysia which 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 state is relatively poor with a fairly high rate of unemployment and inadequately developed infrastructure and water supply. The economic activity of the state is predominantly agriculture which made up about 65.3% of the land use, with an insignificant industrial sector.

There are several reasons for choosing the area. Firstly, the state of Perlis has been identified to have 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 since the area, although small, is concentrated in its industrial activities. 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 speculations on their roles in the aggravation of respiratory problems among the population. However, to date there have been no study to support the speculation.

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.

4 Data acquisition methodology

The implementation of the study involves the process of data acquisition among a sample which consists of (a) a sample of persons registered as 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 problem, profiles of respondents (category of respiratory problems, demographic profile) as well as geographical and environmental profiles of areas. Data acquisition also includes (i) the dissipation of information about the impending survey to the respondents and (ii) getting the permission and cooperation of the guardian of other data sources (in this case the Perlis Health Department).

4.1 Measured variables

Variables measured for the study up to the current stage include, among others, category of respiratory problems, demographic, geographical and environmental profiles. The variables measured in the survey for the preliminary study are as follows:

Table 1: Variables in the survey.

<table>
<thead>
<tr>
<th>Variable Group</th>
<th>Section in Questionnaire</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Demographic Profile of Respondents</td>
<td>A</td>
</tr>
<tr>
<td>2. Geographical and Environmental Profiles of Areas of Residence</td>
<td>B</td>
</tr>
<tr>
<td>3. Standard of Living Profile of Respondents</td>
<td>C</td>
</tr>
<tr>
<td>4. Health Profiles (Related to Respiratory Problems)</td>
<td>D1 – D3</td>
</tr>
<tr>
<td>5. Causal/Trigger Factors of Respiratory Problems</td>
<td>D4 – D6</td>
</tr>
<tr>
<td>6. Stress Factors</td>
<td>D7</td>
</tr>
</tbody>
</table>

4.2 Survey sample size

The size of the sample for the survey was based on the chosen level of precision (B) of the estimate of the proportion (p) of respondents falling into a particular category or having a certain perception of an issue investigated by the survey, and the actual proportion of respondents agreeing to an issue or having a certain perception of an issue raised in the survey. The size of the sample n is calculated using the following expression

\[ n = \frac{p(1-p)z_{\frac{B}{2}}^2}{B^2} \]

For example if the precision level B is chosen to be 0.03 and the value of p obtained from the pilot survey is 0.6 (i.e., 60%), then

\[ n = 0.6(1-0.6)\left(\frac{1.96}{0.03}\right)^2 = 1024 \]

Thus if the proportion obtained in the survey is 0.60, one is 95% confident that the true proportion in the overall population is between 0.57 (0.60-0.03) and 0.63 (0.60+0.03), i.e., between 57% and 63%. In this preliminary stage a total of 599 respondents were selected.

4.3 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. The Perlis Health Department is involved in the study in the following manner:

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

5 GIS analysis

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

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 their existence at particular locations. A logistic regression analysis will be applied to estimate the probability of occurrence of respiratory problems 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. Figures 1 shows a spatial distribution of the asthmatic incidence in the area under study while Figure 2 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.

5.1 Spatial analysis at the address level

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-neighbour 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.

6 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.

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. 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 respiratory problems.

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 respiratory illness and its relationship to areas of elevated particulate matter.

7 Conclusion

The use of spatial and statistical analysis in health research projects is increasing, but there are still important issues to be resolved. Spatial statistical analysis is also found to be quite complex, but produces significant statements about spatial patterns in epidemiological data. In the current study is part of an ongoing project, we have digitised basic map coverages, georefeneced patients’ data, visualized queries on different aggregation levels and made a fundamental statistical analysis.

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 respiratory problems. The investigation into whether environmental factors have significant effects on asthmatic problems in the area will be carried out in the
next stage of the study. It will involve, among others, investigations and 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.

References


