Susceptibility Risk Index Mapping of Population at Tuberculosis Epidemic Risk

Abdul Rauf Abdul Rasam¹, Wan Nor Syahirah Jumali¹, Ilham Abdul Jalil¹, Lalu Muhamad Jaelani²

¹ College of Built Environment, Universiti Teknologi MARA, Selangor, Malaysia, ²Departemen Teknik Geomatika, Institut Teknologi Sepuluh Nopember, Indonesia

rauf@uitm.edu.my, lmjaelani@geodesy.its.ac.id Tel: +60355444434

Abstract

This paper presents the spatial susceptibility risk mapping of tuberculosis (TB) using a geographical information system (GIS) index model and satellite remote sensing (RS) imagery. GIS and RS-based index approach is proposed as an alternative method in identifying potentially high-risk areas in Klang, Selangor. The level of risk for the selected socio-spatial factors and the risk map was classified into five-scale from level 1, which is no risk to level 5, which indicates high risk by applying an overlay analysis and a weighted linear combination. The risk index map shows that a high concentration of TB cases is located in the district's urban and crowded areas.

Keywords: GIS-risk index model; remote sensing; susceptibility mapping; tuberculosis

eISSN 2514-7528 ©2023. The Authors. Published for AMER & cE-Bs by e-International Publishing House, Ltd., UK. This is an open-access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/). Peer-review under the responsibility of AMER (Association of Malaysian Environment-Behaviour Researchers), and cE-Bs (Centre for Environment-Behaviour StudiesCollege of Built Environment, Universiti Teknologi MARA, Malaysia.
DOI: https://doi.org/10.21834/jabs.v8i24.423
1.0 Introduction
Tuberculosis (TB) is an old disease that had affected humans for thousands of years but remained unknown until 24 March 1882 when Dr Robert Koch discovered that the bacillus, subsequently named Mycobacterium tuberculosis, is the leading cause of the disease. The disease spreads when individuals who are sick with TB release the bacteria into the air, known as pulmonary TB, and it is known as extrapulmonary TB when the effects spread to other parts of the body.

Abstract.
A relatively small proportion (5–10%) of the estimated 1.7 billion people infected with M. tuberculosis will develop TB disease during their lifetime. However, the probability of developing TB disease is much higher among people living with HIV; it is also higher among people affected by risk factors such as undernutrition, diabetes, smoking and alcohol consumption. A diagnostic test is a standard method for diagnosing TB disease, including sputum smear microscopy, developed more than 100 years ago. TB resistance to first- and second-line anti-TB drugs can be detected using rapid tests, culture methods and sequencing technologies. Rapid molecular tests were initially endorsed by WHO in 2010 as culture-based methods; the latter take up to 12 weeks to provide results but remain the reference standard. It is undeniable that this medical technique offers accurate results, but the medical geography of the disease needs to be considered because TB disease is also related to environmental factors.

Geospatial technology, such as a geographical information system (GIS), can map, analyze and manage the datasets of tuberculosis cases. Integrating the GIS and Multi-Criteria Decision Making (MCDM) process can enhance the technology’s technical capabilities. Integrating GIS capabilities with the MCDM methodology offers more fantastic performance and decision-making capacity when solving spatial final decision problems (Abdul Rasam et al., 2020; Abdul Rasam et al., 2019; Maris et al., 2008; Rajab et al., 2020). In general, input map layers are prepared, and two different MCDM methods are introduced into GIS: Weighted Linear Combination (WLC) and Analytical Hierarchy Process (AHP).

GIS tools and a spatial scan statistic were also used in India to investigate statistically significant hotspots of TB (Tiwari et al., 2006). Similarly, a local researcher 2016 (Abdul Rasam et al., 2016) used the GIS-MCDM approach with selected risk factors to define risk areas of TB in Shah Alam. TB dynamics in Malaysia are unique due to the geographical characteristics of the affected areas. Therefore the study aims to define potential high-risk areas of local tuberculosis cases in Klang, Selangor. This study comprises three main objectives, namely, i. to determine the risk factor of the local TB, ii. to calculate the level of risk for every factor influencing the local TB, and iii. to create a susceptibility risk mapping of TB in the study area.

2.0 Literature Review
The Global Tuberculosis (TB) Report by WHO (2021) stated that The TB epidemic is strongly affected by changes in society and the economy, as well as by health-related risk
factors like being undernourished, having diabetes, being infected with HIV, having a problem with alcohol, or smoking. Low socio-economic status and environmental factors are risk factors for tuberculosis in 22 countries worldwide (Harling et al., 2014; WHO, 2023). The spread of TB in Malaysia is complicated by the way the disease moves and the unique features of the areas where it is found. Since finding a risk factor framework that contributes to local TB incidence is essential for understanding local TB and helping to find missing cases on the ground (Abdul Rasam et al., 2021), knowledge-driven methods like GIS-based index models are often used to figure out how likely a disease is to spread (Pfeiffer et al., 2008).

To reach global goals for TB disease burden reduction and better access to TB prevention, diagnosis, and treatment services, we need to make progress towards universal health coverage (UHC) and take steps to address health-related risk factors and the broader social and economic causes and effects of TB (WHO, 2023).

A Geographic Information System (GIS) is a computer system that looks at and shows information based on location. The GIS modelling process can show a part of the real world. Maps are the best way to show how modelling works. Maps are complicated tools that hold much information. Some of that information is clear from the symbols that explain the map. Based on the index values, the index model for an increasing unit region figures out the index value and makes a graded chart. An index model is similar to a binary model in that both need multi-criteria calculations and rely on overlapping data processing operations. An index model is similar to a binary model in that both need to use multi-criteria calculations and rely on overlapping data processing operations.

The main thing that goes into making an index model, whether it's based on vectors or rasters, for calculating index values. A common way to figure out the index value is to use a weighted linear combination. The weighted linear combination of the analytical hierarchy process involves a three-level evaluation. The first way is to examine how important each criterion is compared to the others. The performing ratio was estimated for each pair of criteria as part of this method.

The standardization data for each parameter is the second method. A linear transformation is a common way to ensure that all the data are identical. For example, the formula can turn data in intervals or ratios into a standard scale from 0.0 to 1.0. The third method is the index value, found for each unit area, by adding the weighted criterion values and dividing that number by the total number of weights. The weighted linear combination method is one of many ways to do it. Most of the time, these different ways deal with issues like the independence of factors, the weights of criteria, the collection of data, and the standardization of data (Chang, 2018).

The techniques of the GIS index model have been applied in disease and health studies (Soni et al., 2022; Alirez, 2022; Rakibul et al., 2021; Sameer & Sandy, 2021). Soni et al. (2022) highlighted that Analytical Hierarchy Process (AHP) and GIS play an essential role in making multi-criteria decisions and identifying the corona concern zone of larger populated areas across the country in a single platform which can be further helpful for better control, planning, and management during several pandemic outbreaks. Rakibul et
al. (2021) added that a GIS could predict the confirmed case numbers and specific locations where the outbreak would happen with higher statistical precision.

In Malaysia, many researchers have also been exploring the capabilities of GIS-MCDM for disease clustering and pattern analysis (Mohammad et al., 2023; Ridzuan et al., 2021; Zaini et al., 2022; Abdul Rasam et al., 2016). Specifically, the studies suggested that the GIS approach can use to analyze the spatial distribution of disease cluster cases, including to determine the hotspot location of the disease cluster to examine the spatial distribution of the factors affecting the disease cluster. By using a GIS-based disease vulnerability mapping, it can help decision-makers to take proper actions as early as possible, mainly in highly disease-vulnerable governorates, to control the risk associated with the potential outbreak of the virus and accordingly to protect social life and to sustain economic conditions (Sameer & Sandy, 2021; Couceiro et al., 2011).

3.0 Methodology
Figure 1 shows the methodology used in this study, starting from project planning and followed by data collection, data input, data processing, result, and analysis.

3.1 Project planning
Understanding the steps involved in the project process is the first step in the methodology. Project planning includes selecting hardware and software used for data processing and analysis. Klang is selected as the study area because it is one of the most rapidly developing areas and some factors lead to the occurrence of TB disease in the area.

3.2 Data input of risk factors

Table 1: TB risk factors according to previous studies

<table>
<thead>
<tr>
<th>No Factors</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Land use, health care, population</td>
<td>Including the third dimension: a spatial analysis of TB cases in Houston Harris County. Tuberculosis (De Queiroga, 2012; Feske et al., 2011)</td>
</tr>
<tr>
<td>2. Land use/urban, built-up area, socio-economic</td>
<td>A spatial analysis of social and economic determinants of tuberculosis</td>
</tr>
<tr>
<td>3. High-risk group, population, type of housing</td>
<td>Investigation of space-time clusters and geospatial hot spots for the occurrence of tuberculosis in Beijing (Li et al., 2012)</td>
</tr>
<tr>
<td>4. Health care, population, high-risk group, socio-economic</td>
<td>Spatial patterns of pulmonary tuberculosis incidence and their relationship to socio-economic status in Victoria, Brazil (Maris et al., 2008; Macei et al., 2010; Nair, Yakun, 2014)</td>
</tr>
<tr>
<td>5. Urban area, factory, type of housing</td>
<td>Spatial distribution of tuberculosis and relationship with living conditions in an urban area of Campina Grande (De Queiroga, 2012)</td>
</tr>
</tbody>
</table>
Table 1 shows the critical risk factors of TB used in this study, as suggested by previous studies. The factors or input layers are urban, factory, population, type of house and hospital centre. These layers are ranked on a five scale and entered in the attribute table of the ArcMap. Each factor also consists of sub-criteria, as shown in Table 2. The ranks of each sub-criterion were on the map based on the historical causes of the disease.

![Flowchart of the research methodology.](image)

Table 2: TB risk factors and scale measurement

<table>
<thead>
<tr>
<th>Rank of Factors</th>
<th>Scale of Risk Point Data</th>
</tr>
</thead>
<tbody>
<tr>
<td>The rank of land use</td>
<td>Location = Urban (5), rural (1), other (0) and semi-urban (3)</td>
</tr>
<tr>
<td>The rank of built-up</td>
<td>Distance of factory = 250m (5), 500m (4), 600m (3), 750m (2), 1000m (1)</td>
</tr>
<tr>
<td>Rank of population</td>
<td>Family No &lt; 2 (1), 3-5 (2), 6-7 (3), 9-11 (4), &gt;11 (5)</td>
</tr>
<tr>
<td>The rank of house type</td>
<td>Type of House = Flat (5), Apartment/Condominium (4), Village (3), Terrace (2), Semi-D (1), Other (0)</td>
</tr>
<tr>
<td>Rank of health care centre</td>
<td>Distance of health care centres = 0.5Km (1), 1km (2), 1.5Km (3), 2km (4), &gt;2.5Km (5)</td>
</tr>
</tbody>
</table>

(Source: Abdul Rasam et al., 2016)
3.2 Data processing and analysis

The datasets of the risk factor are stored in the personal geodatabase in ArcMap to avoid duplication or exchanges during the data processing. The next step is buffering to create a zone around a map feature measured in units of distance or time. A buffer is useful for proximity analysis. For example, as shown in Figure 2, the input feature of hospital layers was selected to buffer the healthcare centres in a range of 500m. Health care centres locations (TB Facilities) in Klang include the main hospital, Hospital Besar Tengku Ampuan Rahimah, and other clinics or private hospitals. The geocoding of TB cases and overlay analysis was conducted in ArcGIS Online. The cases were divided into several tables to simplify data processing.

Figure 2: Industrial factory buffer in four different buffering zone

The risk level of land use was determined based on land classification, whether the areas are urbanized or not (Table 3). Urban areas surround Klang because Klang is one of the cities with rapid physical development. As viewed on Google Earth, many buildings and factories are located in Klang. According to previous studies, the risk for diseases increases when there are more urbanized areas.

Table 3: Five-scale land use (urbanization) risk level

<table>
<thead>
<tr>
<th>Type of Land Use</th>
<th>Level of Risk</th>
</tr>
</thead>
<tbody>
<tr>
<td>City / Urban Area</td>
<td>5</td>
</tr>
<tr>
<td>Semi-Rural Area</td>
<td>3</td>
</tr>
<tr>
<td>Rural Area</td>
<td>1</td>
</tr>
<tr>
<td>Others</td>
<td>0</td>
</tr>
</tbody>
</table>

(Source: Abdul Rasam et al., 2016)
The population of an area is another influential factor in the spread of the disease. The population is determined by the number of family members, which is also related to the type of housing or where they live. For example, for a terrace house, it can be estimated that the total number of family members is between 6 to 8. For the type of house, the class category is referred to Google Maps as a guideline for risk scaling.

Furthermore, the distance of the factory refers to the distance from the factory's location to the housing areas. Every distance was determined by using buffering tools. Weight overlay tools were utilized because it is one of the most common approaches for overlay analysis to solve multi-criteria problems such as site selection and suitability models (Figure 3). After the data input, the next stage was to export the vector to the raster, reclassify the data, weight overlay, and raster calculator. This stage covered all the risk factors that contribute to TB cases.

![Figure 3: GIS weighted overlay analysis for TB susceptibility mapping](image)

### 4.0 Results

#### 4.1 The determination of the local TB risk factors

Based on Table 1, the five factors contributing to the tuberculosis disease were selected according to the local conditions and data availability. The five factors, i.e. urbanization, population, type of housing, built-up area and health care centre, are selected for this study. Local researchers have also used similar factors and models to define the risk areas of TB in Shah Alam (Abdul Rasam et al., 2019; Abdul Rasam et al., 2016). The model is built in three main stages: frame development, data processing Abdul Rasam et al., 2016, and risk analysis and modelling. The model includes eight risk factors: urbanization, factory size, socio-economic status (SES), risk group, human mobility, house type, distance to healthcare centres and population.

#### 4.2 The Susceptibility risk mapping of the local TB cases

Figure 4 illustrates actual TB cases consisting of 700 hundred cases in Klang overlaid with the susceptibility map of TB cases on a 5 risk-scale (5 for high-risk in red and 1 for low-risk
in green). The result shows that the cases of TB are primarily in urban areas. For instance, the top three locations with many cases are Pandamaran, Taman Saujana 3 and Teluk Pulai.

These three locations are in urbanized areas, covering industrial areas and dense environments. These potential risk areas have been coloured in red, meaning that the areas are risky for TB transmission and infection. Each location contributes to the five factors influencing TB epidemics in the states.

Figure 4: Susceptibility risk mapping of TB cases in Klang using a five-scale from level 0/1, which is no/lowest risk (dark green) to level 5 for highest-risk (red).

5.0 Discussion
Based on Table 1, land use and cover are the main elements of a risky environment that relate to physical factors such as geology, climate, and physical surroundings (such as nursing homes and hospitals), biological factors such as insects that transmit the agent, and socio-economic factors such as crowding, sanitation, and the availability of health services. Physical environment and ecology are also among the significant determinants of TB incidence. Environmental factors are extrinsic factors which affect the agent and the opportunity for TB exposure (Nana Yakam et al., 2014; Tadesse et al., 2013; Wang et al., 2012; Roza et al., 2012). In Malaysia, urban and crowded areas are potential areas for the spread of TB cases. Some experts (Tadesse et al., 2013; Roza et al., 2012) also suggested that future research combines these factors with other factors, such as various socio-economic and environmental factors, for significant model findings, especially on the high incidence of TB areas.

Physical development is related to a country's physical development planning, economic growth, and development. For this study, planning on the location may be indirectly related to urban buildings such as factories and commercial buildings that are important to reduce TB occurrences in terms of the condition or pattern of settlement in urban cities and crowded environment. Even though there is no clear relationship between the factory distances with TB occurrences, qualitative findings indicate a relationship
between workplace in TB healthcare and control as found in Bangladesh (Zafar Ullah et al., 2012; Lima et al., 2019). These same researchers also added that engagement occurred because the workers had inadequate knowledge regarding its causation, transmission, and prevention (Lima et al., 2019).

The Canadian Tuberculosis Committee (CTC) in 2007 stated that housing conditions are used as socio-economic indicators of health and well-being. Housing condition is selected based on the local TB condition related to house types. Poor housing quality and overcrowding are associated with poverty, specific ethnic groups, and increased susceptibility to disease. Crowding, poor air quality within homes because of inadequate ventilation, and smoke contribute to poor respiratory health and are associated with the spread of tuberculosis (TB). Housing or residential building is also an environmental element that indicates TB risk factors.

Besides, limited public health resources in terms of funding, healthcare techniques (Goswami et al., 2012) and healthcare workers are among the main indirect factors in TB cases. Thus, there is a need for high commitment and time, and the results for attending clinical-based testing techniques (Goswami et al., 2012), occupational health skills and knowledgeable staff for handling TB systematically (Yazdani-Charati et al., 2014). Moreover, an innovative method needs to be introduced for an alternative solution. For example, when healthcare employees do not know appropriately about TB (Yazdani-Charati et al., 2014), they will be exposed to higher risks of getting TB from patients/suspected cases during treatment.

Regarding the risk map of TB cases, as shown in Figure 4, the three locations that recorded the highest number of cases are located in urbanized areas, covering industrial areas and dense environments. These potential risk areas have been coloured in red, meaning that the areas are risky for TB transmission and infection. Each location contributes to the five factors influencing TB epidemics in the states.

Based on the images obtained from Google Earth, Pandamaran, primarily an industrial area, has the highest number of TB cases. It is important to note that rapid development in that area heavily influences the possibilities of TB incidence. It is also similar to Taman Saujana 3 and Teluk Pulai, whereby the same factors would spread local TB cases. Physical environment and ecology are also among the significant determinants of TB incidence. Environmental factors are extrinsic factors which affect the agent and the opportunity for TB exposure (Nana Yakam et al., 2014; Tadesse et al., 2013; Wang et al., 2012; Roza et al., 2012; Wong et al., 2007).

On the contrary, low-risk areas such as Kampung Bukit Badak have fewer tuberculosis cases, possibly because there is a health care centre nearby 1.5km. This situation means that the closeness of the area to the health care centre makes it less risky for TB incidence. Limited public health resources in terms of funding, healthcare techniques (Goswami et al., 2012) and healthcare workers are among the main indirect factors in TB cases. Thus, there is a need for high commitment and time, and the results for attending clinical-based testing techniques (Goswami et al., 2012), an occupational health skills and knowledgeable staff for handling TB systematically (Yazdani-Charati et al., 2014).
Physical development is related to a country’s physical development planning, economic growth, and development. In urban areas, the incidence of TB is generally higher than in rural areas, as revealed by previous research (Jalil et al., 2021; Abdul Rasam et al., 2016), making urban TB control particularly challenging. Tendencies to higher TB pressure in urban than rural areas could be attributed to high population growth, cramped housing and working environments, and shifts in lifestyle correlated with urban living. For this study, the location selection may be indirectly related to urban buildings such as factories and commercial buildings that are important to reduce TB occurrences, explicitly focusing on the condition or pattern of the settlement in urban cities and crowded environment. Other possible factors that require further investigation are potential high-risk groups such as the poor foreign-born community in urban and rural areas. The effect of a high influx of immigrants comes from geographic factors such as low-income communities, overcrowded areas, and high economic sectors in urban areas, as Pelabuhan Klang reported, which could also influence the worldwide spread of TB cases.

6.0 Conclusion
This study identified potentially high-risk areas of local tuberculosis cases in Klang, Selangor, using the selected socio-environmental risk factors, GIS, and spatial analysis. The possible risk factors contributing to local tuberculosis are urbanization, population, type of houses, built-up areas, and location of health care centres. Urbanization or built-up areas influence the rapid development in Klang year by year, significantly influencing the spread of tuberculosis. The population dynamics also indirectly influence local tuberculosis occurrences because the number of people living in a house or a building will affect the infection rate among the community members. Other factors include the types of houses and the location of factories and healthcare centres in the area. Using these selected risk factors and the GIS index model in a five-scale, a relevant susceptibility risk map of TB in Klang is produced. The risk map is used to identify TB transmission hotspots in high-risk communities in Klang for transmission reduction strategies in high-burden settings. However, these selected risk factors might be combined with other possible factors, such as social behaviour patterns of individuals, for a comprehensive TB risk of infection model as being conducted by this study.

Acknowledgement
The authors gratefully acknowledge the assistance of the Ministry of Higher Education (MOHE) and Universiti Teknologi MARA Selangor through the Fundamental Research Grant Scheme (FRGS) 600-IRMI/FRGS 5/3 (093/2019). The author also wishes to thank ReNeU UiTM and ILD UiTM for facilitating the writing and publication workshop and the College of Built Environment for supporting this research.
Competing Interest and Consent
We declare no conflicts of interest to disclose regarding this manuscript. The research is registered in the National Medical Research Register, Malaysia (ID: NMR R-15-2499-24207). The authors also thank the Selangor State Health Department for providing TB datasets used in this study.

Article Contribution to Related Field of Study
This susceptibility map could be beneficial in targeting the TB transmission hotspots in the high-burden communities of the district in Klang, Selangor.

References


Feske, ML., Teeter LD, Musser JM, Graviss EA. (2011).Including the third dimension: A spatial analysis of TB cases in Houston Harris County. Tuberculosis. 2011;


