A stochastic model for relative risk estimation of leptospirosis in Malaysia

Sufi Hafawati Ideris¹, Muhammad Rozi Malim² and Norshahida Shaadan³
¹,²,³ Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Shah Alam, Selangor

E-mail: rozi@tmsk.uitm.edu.my

Abstract. The relative risk of a disease is the observed probability that a member of an exposed group will develop the disease relative to the expected probability that a member of a susceptible group will develop the same disease. The estimation of relative risk is important for disease mapping; it is a method used to illustrate the geographical distribution of a disease occurrence for identifying areas that need more attention. Better estimates of risk would subsequently produce more accurate maps of disease risk. The study on relative risk estimation of leptospirosis in Malaysia is very scarce. Most of the related studies involved only the demographic of the disease. Furthermore, most of the mathematical modelling and statistical analyses used for disease transmission models have been deterministic; do not consider the potential of random effects. Thus, the objective of this study is to propose a discrete-time discrete-space stochastic model for relative risk estimation of leptospirosis in Malaysia based on a SIR-SI transmission model. The proposed model was demonstrated using Malaysia leptospirosis dataset (2012-2016) to estimate and analyse the expected relative risks of leptospirosis for all states. The results showed that the averages of estimated relative risks are between 0.340 and 2.898. Kelantan and Terengganu are the two most vulnerable states of leptospirosis for every epidemiology year from 2012 to 2016.

1. Introduction
Leptospirosis is a worldwide infectious disease that involves human and animal populations. It is more common in tropical and sub-tropical regions with high rainfall. Since 2010, leptospirosis is considered as a notifiable disease in Malaysia [1]. Leptospirosis can cause a wide range of symptoms including high fever, headache, and muscle aches. It is often difficult to diagnose clinically as it can appear to be very similar to many other tropical diseases. Current prevention and control strategies such as vector surveillance and control for leptospirosis are only implemented after the cases have been reported. On that account, disease risk maps are crucial to identify high and low risk areas to prevent the spread of the disease. Accurate and reliable risk maps rely on the relative risk estimation model.

The mapping of disease incidence has long been a part of disease distribution study in human population and public health [2]. Environmental pollution, climate change, rapid population and urbanization, human migration and poverty are the contributing factors of disease outbreaks. Disease mapping can be defined as the visual representation of summary measures of disease counts [3]. The aims include simple description, hypothesis generation, allocation of health care resources, assessment of inequalities, and estimation of background variability in underlying disease risk. This mapping method could be a fundamental tool to show the areas affected by the epidemic.
This study focuses on stochastic modelling of the relative risk estimation of leptospirosis in Malaysia. In Malaysia, the leptospirosis infection is frequently related to occupational activities [1, 4-6]. Most studies were basically retrospective with case studies such as biological infection, factors of infection, outbreaks and cases reported, and treatment and management of leptospirosis. In disease mapping studies, some common approaches are used to estimate relative risk; Standardized Mortality Ratio (SMR) [7] and Poisson-gamma [8]. However, these traditional methods offer simple risk structures with simple algorithms that do not include spatial correlations between neighbouring areas, and no disease transmission was analysed. Moreover, it is difficult to intervene with other parameters (such as random effects) that contribute to the risk. Also, SMR is not reliable for small counties with small populations. To overcome these drawbacks, this study introduced a Bayesian approach that includes prior information on the variability rates of disease and the likelihood of infection.

Based on the idea of Bayesian approach, a stochastic model for relative risk estimation of leptospirosis, based on SIR-SI transmission model (Susceptible-Infected-Recover for human; Susceptible-Infected for rat), was proposed. Firstly, the deterministic compartmental SIR-SI model for leptospirosis was discussed and represented mathematically by a system of differential equations. Then this model was adapted to develop the corresponding discrete-time discrete-space stochastic SIR-SI model. Finally, the proposed model was demonstrated using Malaysia leptospirosis dataset.

2. Leptospirosis disease

Leptospirosis is an infectious disease caused by pathogenic spirochete bacteria. It can be transmitted directly or indirectly from animals to humans [4]. The bacteria are spread through the urine of infected animals, which can get into water and soil, and survive for weeks to months. Mammals like rats, mice, rodents, dogs, cattle, sheep, pigs and horses are sources of infection. These infected animals, though recovered from leptospirosis, may become carriers and subsequently excrete the bacteria into the environment. The bacteria require a host for survival and reproduce. However, animals that carry the bacteria as the hosts will not be harmed. In fact, human will be the dead-end host for these bacteria as this disease normally spreads from animal to human, while human to human case is very sparse.

2.1. Transmission of leptospirosis infection

The leptospirosis infection can be transmitted in three modes; direct contact with urine or other body fluids of infected animals, indirect contact with soil, food and water contaminated with the urine of infected animals, and inhalation of droplets of infected urine. People who working outdoors or with animals are at high risk to contract with the disease. Susceptible humans who have been infected by leptospirosis would suffer various clinical symptoms ranging from asymptomatic to severe multiple organs syndrome that lead to death without appropriate treatment. The incubation period is about 5 to 14 days and between 2 to 30 days of infection depending on patient’s antibody [9]. During the first day of infection, patient usually does not show any symptom. However at the final stage, this systematic infection would harm the main organs including central nervous system.

Clinically, the infection can be divided into two types; anicteric leptospirosis (90% of cases) and icteric leptospirosis (10%). A patient with anicteric leptospirosis usually undergoes two phases of symptoms; acute and immune. In the acute phase, the patient normally has flu-like symptoms that would last about 3 to 8 days. The symptoms would gradually improve and patient seems to have recovered (immune phase). During this phase, the patient would fall sick again and worse due to the malfunction of multiple organs until 30 days or more. A patient with icteric leptospirosis would show symptoms that are mostly similar to anicteric, but it can develop into a fatal condition.

2.2. Leptospirosis Infection in Malaysia

According to Costa et al. [10], it was estimated that 1.03 million cases and 58,900 deaths due to leptospirosis each year. Despite the significant health and economic burden to human and livestock production, leptospirosis is still a neglected and under reported disease [1]. In Malaysia, leptospirosis is endemic and found in both human and animal populations, as well as the environment [6].
Leptospirosis is not a new disease in Malaysia. The first human leptospirosis case had been diagnosed by William Fletcher in 1925. Since 1986, no investigations on human leptospirosis were made, only a study on retrospective of human leptospirosis was carried out until 1998 by Institute for Medical Research (IMR). After years, human leptospirosis cases in Malaysia became more and more significant [6, 11]. The number of reported cases had increased from 263 in 2004 to 7806 in 2014.

3. Disease mapping and transmission model

3.1. Disease mapping and spatial data analysis

Disease mapping is used by epidemiologists to understand the geographical distribution of a disease. The aim is to identify and visualize the spatial estimation of relative risk across a region. It is also used to classify all states into different components by using colour shading to demonstrate high and low levels of risk. In disease mapping, there are two types of spatial data; case event and tract count [12].

The case event data basically correspond to disease occurrences which represent the detail of locations at a particular point in time. When all small spatial scales become the interest, this type of data would be crucial. However, the data are difficult to obtain due to certain factors that are hard to acquire with accurate details and perhaps because of the confidentiality issues of data. Conversely, count data (the number of cases in a given period) is the most common format of data used in the study of spatial epidemiology. Count data arise from an aggregation of case events into small areas and fixed periods [12]. This type of data often presented in spatial aggregated form.

3.2. Common approaches for relative risk estimation in disease mapping

In spatial analysis of disease mapping, there are common approaches that used to estimate the relative risk; SMR, Poisson-gamma model, log-normal model, and mixed model.

3.2.1. Standardized morbidity ratio (SMR). SMR is the ratio of observed count to expected count within tracts; i.e. an estimate of relative risk within each tract. It is the most common approach used to estimate the relative risks in disease mapping. Other approaches that have the same concept as SMR are Standardized Incidence Ratio (SIR) and Standardized Hospitalization Ratio (SHR). SMR basically compares the observed incidence with the expected incidence, which has been used for the analysis of counts within tracts [13].

In a disease mapping, the study area is divided into \( M \) mutually exclusive regions \((i = 1, 2, \ldots, M)\). Each region has its own observed number of cases \((O_i)\) and expected number of cases \((E_i)\). The relative risk \( \theta_i \) can be estimated as \( r_i = \frac{O_i}{E_i} \) [14]. The expected number of cases \( E_i \) is calculated as \( E_i = N_i \sum O_i / \sum N_i \), where \( N_i \) is the population size of region \( i \). The standardization is done based on the total population at risk assuming everybody is equally at risk. Consequently, the relative risk can be estimated using \( r_i = \theta_i = \frac{O_i \sum N_i}{N_i \sum O_i} \); i.e., the probability of a person within region \( i \) contracts the disease divided by the probability of a person in the population contracts the disease. However, SMR has several drawbacks that make interpretation difficult. It is a reliable measure of relative risk for large regions, but may be unreliable for small areas. SMR would be zero if no cases were observed (indicates that there is no risk in the infected region though everybody has the chance to be infected). To overcome this, many researchers used an alternative method such as Poisson-gamma model.

3.2.2. Poisson-Gamma model. Poisson-gamma model is one of the earliest Bayesian methods used in disease mapping analysis to estimate relative risk especially when the data are typically scarce. By using Bayes’ rule, the combination of observed data and some prior knowledge would produce a posterior distribution. This distribution describes the probability distribution and derives inference about unknown parameter.

In a Bayes’ method, the relative risk estimation is based on a posterior distribution that defined as the product of likelihood function and prior distribution. The likelihood model may follow normal, binomial or Poisson function. In disease mapping, the geographical distribution follows tract-count data
as the disease distribution and Bayesian as the method of analysis. Therefore, the probability distribution is often assumed to follow a Poisson where for each count tract there will be a different expectation. Analytically, in Bayes’ method, the likelihood is determined based on the observed data which is proportional to the joint probability distribution function of the sample data [15]. A prior distribution gives assumptions and prior knowledge for a parameter before the current data are examined. A prior distribution is usually assumed as a single gamma distribution with fix hyper-parameters. In Bayesian model, the parameter is examined by the mean of the posterior distribution.

Within Bayesian framework, a posterior distribution naturally considered the parameter \( \theta \) as a random variable. However, it is often the case that the posterior distribution can be very difficult to sample from a complex model. A posterior distribution can be sampled by a simulation technique called Markov Chain Monte Carlo (MCMC). It is a computer driven sampling method that allows one to characterize a distribution without knowing all the mathematical properties by randomly sampling values out of the distribution [16]. In Bayesian modelling, a prior is usually chosen for each parameter that needs to be inferred to generate the samples of posterior distribution and estimate the posterior quantities of interest. By using MCMC, the modelling of data will be more realistic when the chain is long. However, a major shortcoming of gamma prior is no possibility of allowing spatial correlation between risks in nearby areas and the covariate adjustment is also difficult.

4. Leptospirosis transmission model

In a disease infection modelling, the transmission could be constructed using SI, SIS, SIR, SEIR, or SEIRS models, depending on the disease. These types of models are categorized under direct transmission (‘S’ – susceptible, ‘I’ – infected, ‘E’ – exposed and ‘R’ – recovered).

Basically, there are two types of leptospirosis transmission; direct and indirect. Both transmissions can be represented by two types of models; deterministic and stochastic. Both models are useful in the study of infectious diseases at the population scale. Compared to a deterministic model, a stochastic model depends on among individual chance variation in risks of exposure, disease, and other factors. It is used when chance fluctuations or known heterogeneities are important as in small or isolated populations. A stochastic model for leptospirosis transmission can be developed from a deterministic model for both discrete-time and continuous-time problems. A stochastic model has a significant advantage; it allows follow-up of each individual in a population on a chance basis.

There are three basic steps to build a deterministic model. Firstly, it is important to have a complete picture of the biology of a disease. Second, the data collection is based on the demographic, epidemiologic, and biologic characteristics of the infection and the population. Finally, select a suitable model. The epidemic process is deterministic where the behaviour of people in a study region is determined by the history as well as the rules which describe the deterministic model [17]. The population will be divided into compartments (subgroups) with assumptions about the nature and time rate of transfer from one compartment to another. This compartmental model can be analysed using either difference equations or differential equations. Difference equations in a deterministic model describe the transitions between different compartments using discrete-time steps, while differential equations consider the transitions for continuous rather than discrete-time intervals. After a deterministic formula has been derived either by difference or differential equations based on a deterministic or compartmental model, the next step is to calculate the stochastic mean of population.

5. Common models for leptospirosis transmission

Various compartmental models were used for leptospirosis disease mapping. Unlike other diseases, leptospirosis is barely distinguished due to the non-specific presentation of symptoms, and this caused confusing with malaria, influenza, dengue and other similar diseases. For leptospirosis, the incubation period is usually 10 days but can range from 2 to 30 days. After all the symptoms have healed, the patient will not experience any symptoms, referred as asymptomatic phase (first phase). In the second phase (if occurs), it is more severe; the patient may have kidney or liver failure or meningitis. If the infection is not treated, it may cause jaundice, haemorrhage, and eventually die due to cardiac failure.
Basically, based on the disease transmission, the leptospirosis transmission starts from susceptible stage [18]. As a susceptible person exposed to leptospira pathogens, it will move to subclinical disease. After the symptoms are onset, the person will be on clinical disease stage. The final stage is recovery or death. The compartmental model for leptospirosis is separated between human and animal [19]. For human, there are three compartments; susceptible, infected and immune; for animal, there are two compartments, susceptible and infected. Figure 1 shows the leptospirosis dynamic transmission.

The study by Triampo et al. [19] was basically based on a simple deterministic model that solved numerically. The relative risk was measured by the basic reproduction number where secondary cases were raised from a primary case in a population. Similarly, a study by Pongsumpun [20] showed that the mathematical model for leptospirosis transmission consists of two compartmental models; human and rat populations. Human population was categorized into juvenile and adult humans. The compartment is allotted into susceptible, infectious and recovered. For rat population, the compartment is divided into susceptible and infectious. Unlike human, rat population has no recovery from infection; the compartment stops until infectious.

Based on the compartmental models, each compartment is deterministic and categorized into susceptible, infectious and recovered according to the biological flow of a disease. Equations are then derived for each compartment using numerical analysis. Result from the analysis would help to answer whether the epidemic occurred or getting fade in a community. However, a deterministic model can be enhanced to be more realistic and random that mimics the real situation by extending the equations to represent the corresponding stochastic model.

6. SIR model for disease transmission

The SIR infectious model begins with the development of compartmental infectious dynamic model. Figure 2 shows the compartmental model for disease transmission; the population is divided into compartments with time rate of transfer from one compartment to another. The terminology ‘SIR’ describes individuals from susceptible class $S$ to infectious class $I$, and eventually to recovered class $R$.

Susceptible is classified for individuals who will be the subject of the disease but are not infected yet. Individuals classified as infectious are the ones who infected with the disease and at the same time transmitting the disease to others. As the infected individuals have recovered or perished due to the disease, they will be classified as recovered. Based on the model in Figure 2, a deterministic model can be derived as follows. To make it relevance to this study, the superscript ($h$) represents the human population so that the compartmental models for human and animal populations can be distinguished.
Another approach of compartmental and deterministic SIR model is based on Lawson [13] (Figure 3). For $i = 1, 2, \ldots, M$ regions and $j = 1, 2, \ldots, T$ time periods, the compartment $S_{i,j}^{(h)}$ stands for susceptible population in region $i$ at the beginning of period $j$, $I_{i,j}^{(h)}$ is the infected cases in region $i$ for period $j$, $R_{i,j}^{(h)}$ represents the recovered cases in region $i$ during period $j$, and $\delta$ represents the hazard for an infectious person being recovered. The arrow from susceptible to infected represents the rate of new infections.

\begin{equation}
S_{i,j}^{(h)}(t) = \mu - \lambda S_{i,j}^{(h)} I_{i,j}^{(h)} - \mu S_{i,j}^{(h)}; \quad I_{i,j}^{(h)}(t) = \lambda S_{i,j}^{(h)} I_{i,j}^{(h)} - \mu I_{i,j}^{(h)}; \quad R_{i,j}^{(h)}(t) = \gamma I_{i,j}^{(h)} - \mu R_{i,j}^{(h)}. \quad (1)
\end{equation}

7. Proposed stochastic SIR-SI model for leptospirosis and relative risk estimation

Basically, the compartments are the sequence in an endemic model with common phases such as susceptible, infected and recovered. One compartment moving to the next is known as progression. In this section, the compartmental model for leptospirosis transmission is discussed, followed by the development of deterministic SIR model for human population and SI model for rat population. Next, the proposed discrete-time discrete-space stochastic SIR-SI model for leptospirosis is described.

7.1. Compartmental and deterministic SIR-SI models for leptospirosis

At any given time, there are a number of susceptible, infected and recovered humans from a human population and a number of susceptible and infected rats from a rat population. Lawson [13] used the term susceptible-infective-removed for the SIR model of direct transmission; susceptible represents the ‘at-risk’ population, infective represents disease morbidity, and removed represents recovery. In addition for rat transmission, vector-borne infection is included in this study; at any given time, there are a number of susceptible and infectious rats. An infected rat will not recover by itself and the leptospira pathogens remains in the animal body until it dies. Therefore, the dynamic transmission for animal starts with susceptible rat population and ends with infected cohort of disease transmission. Figure 4 shows the SIR-SI disease transmission for leptospirosis for human and animal interaction.
Let ‘S’ be the susceptible, ‘I’ is the infected and ‘R’ is recovered for \( i = 1, 2, \ldots, M \) study regions and \( j = 1, 2, \ldots, T \) time periods. The compartment \( S_{i,j}^h \) represents the total number of susceptible humans at time \( j \), \( I_{i,j}^h \) refers to the total number of humans infected with leptospirosis at time \( j \), and \( R_{i,j}^h \) is the total number of recovered leptospirosis cases at time \( j \). \( S_{i,j}^r \) and \( I_{i,j}^r \) respectively represent the total number of susceptible and infected rats at time \( j \). Furthermore, \( I^h \) represents the rate of humans who leave the group because of natural death, and \( \mu^r \) represents as the rate of rats that leave the group because of natural death. These rates, for both human and rat, are assumed to be equal for both death and birth. \( N_i^h \) and \( N_i^r \) respectively denote the total populations for human and rat, and the recovery rate for human is \( \gamma^h \).

The compartmental model of leptospirosis for discrete-time intervals (as illustrated in Figure 4) is constructed mathematically in the system of difference equations. In this study, the deterministic model comprises of two related groups, human and rat populations. The deterministic models for human and rat populations, based on Figure 4, are respectively can be written as follows

\[
\begin{align*}
S^h &= \mu^h N^h - (\mu^h + \beta^h S^h) I^h - \delta^h S^h; \\
I^h &= \beta^h (S^h - \gamma^h I^h) R^h; \\
R^h &= \gamma^h I^h + \delta^h S^h - \mu^h R^h. \\
\end{align*}
\]

(3)

\[
\begin{align*}
S^r &= \mu^r N^r - (\mu^r + \beta^r S^r) S^r; \\
I^r &= \beta^r S^r - \delta^r I^r; \\
\end{align*}
\]

(4)

\( \mu^h \) and \( \mu^r \) are assumed to be constant, where \( N_i^h = N_i^{h-1} + I_{i,j}^h + R_{i,j}^h \) (human population) and \( N_i^r = S_{i,j}^r + I_{i,j}^r \) (rat). Another assumption is the rate of death \( (\mu^h) \) is constant for all susceptible, infected and recovered population subgroups. Also, the rate of death is equal to the rate of birth.

Infection occurs between infected rats and susceptible humans, whereby infected humans will not infect other susceptible humans. Infected rats are assumed do not have recovery phase from leptospirosis. It will carry the pathogens in its body and eventually perish due to the disease. Infected humans are assumed to have recovery or immune phase. However, the transmission from recovery to susceptible is not considered in this study since the objectives are accentuated on the disease transmission from susceptible humans to infected humans. The transmission is assumed to hold until recovery phase. These discrete-time discrete-space deterministic formulas will be used as the link to stochastic population mean to construct the discrete-time discrete-space SIR-SI stochastic model.

### 7.2. Discrete-time discrete-space stochastic SIR-SI model for leptospirosis

The stochastic model is constructed using the deterministic model formulation to provide the approximation of stochastic mean. This study includes new terms to represent the number of newly infected leptospirosis for human and rat populations. The time interval is Poisson stationary and approximates the stochastic mean. This study includes new terms to represent the number of newly infected leptospirosis for human and rat populations. The time interval is Poisson stationary and approximates the stochastic mean. This study includes new terms to represent the number of newly infected leptospirosis for human and rat populations. The time interval is Poisson stationary and approximates the stochastic mean.

For human leptospirosis, \( I_{i,j}^h = \) total number of newly infected humans,

\[ S_{i,j}^h = \mu^h N_{i,j}^h + S_{i,j-1}^h - \mu^h S_{i,j-1}^h - I_{i,j-1}^h; \quad I_{i,j}^h \sim \text{Poisson}(\lambda_{i,j}^h); \quad \lambda_{i,j}^h = [\exp(\beta^h + \varepsilon^h)]\left(\frac{\rho^h}{N_{i,j}^h} I_{i,j-1}^h S_{i,j-1}^h\right); \]

(5)

The discrete-time discrete-space stochastic model for rat population is assumed as non-stochastic because of the unavailability and inadequate data for rat population. The equations are as given as follows; \( I_{i,j}^r = \) total number of newly infected rats with leptospirosis,

\[ S_{i,j}^r = \mu^r N_{i,j}^r + S_{i,j-1}^r - \mu^r S_{i,j-1}^r - I_{i,j}^r; \quad I_{i,j}^r = \left(\frac{\rho^r}{N_{i,j}^r}\right) I_{i,j-1}^r S_{i,j-1}^r; \quad \text{and} \quad I_{i,j}^r = (1 - \mu^r) I_{i,j-1}^r + I_{i,j}^r. \]

(6)
The mean of $\lambda^{(b)}_{ij}$ will follow a Poisson distribution with $c^{(b)}_{ij}$, $\beta_0$ represents the constant terms to describe the overall rates of the process, and $c^{(b)}_i$ represents the random effects to absorb the residual spatial variation of populations. The intrinsic conditional autoregressive (CAR) priors are applied to fit the model for random effects $c^{(b)}_i$. The CAR-normal distribution is used as prior to allow for spatial dependence between the random effects $c^{(b)}_i$ in nearby areas. The CAR prior model can be stated as

$$[c_i | c_j, i \neq j, \sigma^2] \sim N(\bar{c}_i, \sigma^2),$$

where mean $\bar{c}_i = \frac{1}{\sum_j \kappa_{ij}} \sum_j c_i \kappa_{ij}$ and variance $\sigma^2 = \frac{\kappa}{\sum_j \kappa_{ij}}$; $\kappa_{ij}$ is equal to 1 if $i$ and $j$ are adjacent, and 0 if they are not adjacent.

7.3. Relative risk estimation based on stochastic SIR-SI model

An alternative method of relative risk estimation based on the disease transmission model for leptospirosis is introduced. In Bayesian analysis, posterior distribution is defined as the product of likelihood functions and prior distributions. Likelihood is obtained from the observed data while prior distribution is based upon prior assumptions where both are functioning to produce posterior distribution. The posterior expected means of the number of leptospirosis infections are as follows;

$$\lambda^{(b)}_{ij} = \exp(\beta_0 + c_i) \left( \frac{\lambda^{(b)}_{ij}}{\kappa_{ij}} \right) I_{i,j} \sum_{j=1}^{n} \lambda^{(b)}_{ij}, \text{ and } \bar{\lambda}^{(b)}_{ij} = \frac{1}{n} \sum_{k=1}^{n} \lambda^{(b)}_{ijk}.$$

Next, the relative risk of leptospirosis can be determined using $\lambda^{(b)}_{ij} = c^{(b)}_{ij} / e^{(b)}_{ij}$. To approximate the posterior expected relative risk, the equation could be written as $\lambda^{(b)}_{ij} = \frac{1}{n} \sum_{k=1}^{n} \left( \lambda^{(b)}_{ijk} / e^{(b)}_{ij} \right)$.

The posterior expected relative risk for leptospirosis is equal to the expected mean number of infections divided by the expected mean number of new infections. The value of relative risk is defined based on Samat and Percy [14]. A zero value of relative risk means that people within a region have no infection risk. However, the value should not be zero since everyone has the chance to catch with the disease. If the value of relative risk is close to 1, then there is no significant difference between people in a region contracts the disease and people in the general population contracts the disease. Conversely, if the value is greater than 1, the people within a study region are more likely to contract the disease compared to the general population. Conversely, a value below 1 indicates that the people in a study region are less likely to contract the disease compared to the general population.

8. Analysis and results

In this section, the proposed model was demonstrated using Malaysia leptospirosis dataset. The dataset was obtained from Disease Control Department, Ministry of Health (MOH) Malaysia. The data are in the form of number of cases for all states in Malaysia from 2012 to 2016. The discrete-time discrete-space stochastic SIR-SI model for leptospirosis was applied on the dataset using a Bayesian software called WinBUGS. The expected relative risks of leptospirosis for the period 2012-2016 for all states in Malaysia are summarised in Table 1 and illustrated in Figure 5.

Based on Table 1, the averages of the estimated relative risks are between 0.340 and 2.898. Majority of the states (10 out of 15) are more likely to contract the disease with average expected relative risks greater than one. Figure 5 clearly shows that Kelantan and Terengganu are at the top two positions for every epidemiology year (2012-2016) with average expected relative risks of 2.898 and 2.198, respectively. The northern states of Peninsular Malaysia (Perlis, Kedah and Pulau Pinang) are less likely to contract the disease compared to the general population with average expected relative risks less than 0.5. The three states in East Malaysia (Sabah, Sarawak, and Labuan) are also at risk of leptospirosis with average expected relative risks of 1.306, 1.019, and 0.987, respectively.
Table 1. Expected Relative Risks of Leptospirosis for all States for the Period 2012-2016.

| States/ Year            | 2012    | 2013    | 2014    | 2015    | 2016    | Average |
|-------------------------|---------|---------|---------|---------|---------|---------|
| Perlis                  | 0.658   | 0.504   | 0.273   | 0.243   | 0.369   | 0.409   |
| Kedah                   | 2.158   | 1.691   | 0.916   | 0.823   | 1.260   | 1.370   |
| Pulau Pinang            | 0.663   | 0.520   | 0.281   | 0.251   | 0.375   | 0.418   |
| Perak                   | 1.621   | 1.286   | 0.686   | 0.620   | 0.932   | 1.029   |
| Selangor                | 1.868   | 1.462   | 0.794   | 0.719   | 1.086   | 1.186   |
| K. Lumpur/ Putrajaya    | 1.610   | 1.257   | 0.685   | 0.619   | 0.923   | 1.019   |
| Negeri Sembilan         | 1.851   | 1.449   | 0.816   | 0.710   | 1.076   | 1.180   |
| Melaka                  | 2.392   | 1.864   | 1.059   | 0.905   | 1.418   | 1.528   |
| Johor                   | 0.542   | 0.424   | 0.227   | 0.203   | 0.303   | 0.340   |
| Pahang                  | 1.573   | 1.231   | 0.665   | 0.601   | 0.896   | 0.993   |
| Terengganu              | 3.472   | 2.707   | 1.469   | 1.324   | 2.019   | 2.198   |
| Kelantan                | 4.540   | 3.561   | 1.924   | 1.738   | 2.725   | 2.898   |
| Sabah                   | 2.061   | 1.617   | 0.875   | 0.789   | 1.186   | 1.306   |
| W.P. Labuan             | 1.244   | 0.974   | 0.527   | 0.472   | 1.720   | 0.987   |
| Sarawak                 | 1.641   | 1.227   | 0.676   | 0.617   | 0.934   | 1.019   |

Figure 5. Expected relative risks of leptospirosis for all states in Malaysia (2012-2016).

9. Discussion and conclusion

This paper has proposed a discrete-time discrete-space stochastic SIR-SI model for relative risk estimation of leptospirosis. This model considers the disease transmission from vector (rat population) to human population. The model was constructed using discrete-time discrete-space deterministic formulas. The proposed stochastic model was demonstrated using Malaysia leptospirosis dataset for the period 2012-2016 via WinBUGS software.

The results revealed that, with two third of the states are more likely to contract the disease, Malaysia as a whole is always exposed to the risk of leptospirosis infection. With the average expected relative risks of greater than two, Kelantan and Terengganu are the two most vulnerable states with leptospirosis compared to the general population. This leads to a general conclusion that leptospirosis cases are more likely to occur within high rainfall and flooding areas such as Kelantan and Terengganu. Since Terengganu is located next to Kelantan, it shows that the risks are also affected by neighbouring areas.

The proposed stochastic model considers the leptospirosis transmission in human population, and at the same time enables the spatial correlation between neighbouring areas. This could solve the problems...
associated with SMR when no observed count data are available in certain regions. Moreover, this model could overcome the drawback of a Poisson-Gamma model where there is no possibility for allowing covariate adjustment and spatial correlation between neighbouring areas.

Acknowledgements
The authors would like to thank and acknowledge the Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA (UiTM), Shah Alam for providing a special fund to attend and share this paper at ICoAIMS2019.

References
[1] Tan W L, Soelar S A, Mohd M A, Hussin N, Cheah W K, Verasahib K and Goh P P 2016 Southeast Asian J. Trop. Med. Publ. Health 47 434.
[2] Koch T 2017 Cartographies of Disease: Maps, Mapping, and Medicine New Expanded Edition (2nd Edition) (Rhedlands: ESRI Press).
[3] Richardson S, Thomson A, Best N and Elliott P 2004 Environ. Health Perspect. 112 1016.
[4] Thayaparan S, Robertson I D, Fairuz A, Suut L and Abdullah M T 2013 Malays. J. Pathol. 35 123.
[5] Wahab Z A 2015 Proc. Local Authority Confer. Environ. Health 2015 Labuan.
[6] Garba B, Bahaman A R, Khairani-Bejo S, Zakaria Z and Mutalib A R 2017 EcoHealth 14 389.
[7] Lawson A B, Browne W J and Rodeiro C L 2003 Disease mapping with WinBUGS and MLwiN (Chichester: John Wiley & Sons).
[8] Kismiantini 2009 Proc. First Int. Seminar on Sci. Tech. 2009 Uni. Islam Indonesia Jogjakarta.
[9] Lim J K, Murugaiyah V A, Ramli A S, Abdul Rahman H, Mohamed N S F, Shamsudin N N and Tan J C 2011 WebmedCentral Infectious Diseases 2011 2 WMC002764.
[10] Costa F, Hagan J E, Calcagno J, Kane M, Torgerson P, Martinez-Silveira M S, Stein C, Abela-Ridder B and Ko A I 2015 PLoS Neglected Tropical Diseases 9 e003898.
[11] Bahaman A R 2017 Veterinary Research Communications 29 26.
[12] Lawson A B, Banerjee S, Haining R P and Ugarte M D 2016 Handbook of Spatial Epidemiology (Boca Raton: Taylor & Francis).
[13] Lawson A B 2006 Statistical Methods in Spatial Epidemiology (Chichester, England: Wiley).
[14] Samat N A and Percy D F 2012 J. App. Stat. 39 2029.
[15] Lawson, A B and Williams F L 2003 An Introductory Guide to Disease Mapping (Chichester: John Wiley).
[16] Ravenzwaaij D V, Cassey P and Brown S D 2016 Psychonomic Bull. Rev. 25 143.
[17] Brauer F 2012 Deterministic Compartmental Disease Transmission Models Springer.
[18] Trotter H and Philippe P 2000 Internet J. Infect. Dis. 1(2).
[19] Triampo W, Baowan D, Tang I M, Nuttavut N, Wong-Ekkabut J and Dounghawee G 2008 Int. J. Med. Health Sci. 2 22.
[20] Pongsumpun P 2012 Int. J. Math. Comp. Phy. Elec. Comp. Eng. 6 1639.