JOURNAL OF INFORMATION AND COMMUNICATION TECHNOLOGY
https://e-journal.uum.edu.my/index.php/jict

How to cite this article:
Mohd Diah, I., & Aziz, N. (2022). The comparison between standardized mortality ratio, poisson-gamma and stochastic sic model for pneumonia disease mapping in Malaysia. *Journal of Information and Communication Technology, 21*(4), 549-570. https://doi.org/10.32890/jict2022.21.4.4

The Comparison between Standardised Mortality Ratio, Poisson-Gamma and Stochastic Sic Model for Pneumonia Disease Mapping in Malaysia

[*1]Ijlal Mohd Diah & [2]Nazrina Aziz
[1&2]Department of Mathematics and Statistics,
School of Quantitative Sciences,
Universiti Utara Malaysia, Malaysia
[2]Institute of Strategic Industrial Decision Modelling (ISIDM),
Universiti Utara Malaysia, Malaysia

**ijlal_ti_mohd@ahsgs.uum.edu.my,*
nazrina@uum.edu.my
*Corresponding author

Received: 30/1/2022 Revised: 5/7/2022 Accepted: 6/7/2022 Published:20/9/2022

ABSTRACT

Pneumonia is one of the primary causes of death from infectious diseases. Traditionally, its spread has been tracked based on the total number of cases reported, with no concern for geographical distribution. Disease mapping is among the ways public health and the government can monitor diseases as a preventative strategy. Clear pictures of the risk areas can be seen using this method. Relative risk estimation is a significant part of disease mapping that needs to be

considered when studying disease occurrence. This paper aimed to estimate the relative risk values for pneumonia based on three models and compare the results. The approaches used in this study were Standardised Morbidity Ratio (SMR), Poisson-gamma, and discrete time-space stochastic Susceptible-Infected-Carriers (SIC) models, which were applied in estimating the relative risk values. Results showed that Kuala Lumpur was classified as a very low-risk area for pneumonia incidence when using the SMR and Poisson-gamma models. In contrast, Selangor was identified as a very low-risk area when using the discrete time-space stochastic SIC model. Putrajaya was categorised as a very high-risk area in the results of all three types of methods. In conclusion, this stochastic SIC model demonstrated better performance than the conventional models.

Keywords: Disease mapping, Poisson-gamma, pneumonia, SIC model, SMR.

INTRODUCTION

Pneumonia is an infection that inflames the air sacs with pus or fluid in one or both lungs. Different types of microorganisms, including fungi, viruses, and bacteria, can cause pneumonia (Otieno et al., 2012). The most prevalent causes in any person who contracts pneumonia are Streptococcus pneumoniae, a pneumonia bacterium, and viruses. Among these viruses is Coronavirus disease 2019 (COVID-19), the latest virus that is currently a worldwide pandemic that can cause pneumonia and make it serious. Viruses are the most well-known cause of pneumonia in children under five years old. Although viral pneumonia is typically just moderate, in some cases, it may become the most dreadful (Mayo Clinic, 2020). According to Normandin (2021), germs that can cause pneumonia are contagious, meaning that they can easily be transferred to another person. Both viral and bacterial pneumonia are infectious and the disease can spread from one person to another through the air. If an infected individual sneezes or coughs and releases airborne droplets that contain germs in the air, another individual could inhale the airborne droplets from the infected person. It may also be transmitted to others by interaction with bacterial or virus-contaminated objects or surfaces.

The seriousness of pneumonia can range from moderate to fatal. The highest risk groups are babies and young children, particularly those

under the age of five, those 65 years of age and over, and those with compromised immune systems, which may become more dangerous if they are infected with pneumonia (Mayo Clinic, 2020). Despite being a preventable disease, pneumonia causes more deaths in children than any other disease (UNICEF, 2019). According to Seramo et al. (2022), almost one million children die from pneumonia globally, with children under five years old accounting for around 15 percent of all deaths yearly. This percentage equates to approximately 2,500 children dying each day, or 100 children dying every hour. In Malaysia, pneumonia is one of the top three diseases that caused mortality in 2019 (Department of Statistics Malaysia, 2019). Traditionally, its spread has been monitored only based on the total number of cases recorded regardless of geographical distribution. According to Samat and Percy (2012), disease mapping is one of the significant methods in public health research as it can be effective in monitoring by taking into account the geographical distribution and can also be used as a preventative measure for diseases. Disease mapping may be used to provide a visual representation of the burden of pneumonia in certain geographic regions. A good risk map relies on best-fitted mathematical models used to estimate the relative risk (Diah et al., 2016). Among the critical parts of the research on geographical distributions of disease incidence is the estimation of relative risks (Samat & Percy, 2012).

In this study, a stochastic Susceptible-Infected-Carriers (SIC) model is used as an alternative method for estimating the relative risk of pneumonia based on discrete time and space. This method is designed to solve the limitations of the traditional methodology, i.e., Standardised Morbidity Ratio (SMR) and the Poisson-gamma model, an early Bayesian approach to estimating the relative risk in disease mapping. According to Meza (2003) and Awang and Samat (2017), SMR may not be worth using directly in research since it is incapable of identifying small areas and does not consider the high diversity of different regions and spatial patterns of regions under study. Besides, when there are no observed count data or cases, SMR will be zero since the mean and variance of the model are highly influenced by E_i, as it depends on the ratio estimator (Awang & Samat, 2017). Meanwhile, the Poisson-gamma model does not support covariate adjustment and is incapable of coping with the spatial correlation between risks in neighbouring locations. Therefore, the alternative approach, i.e., the stochastic SIC model, includes the extension of the fundamental Poisson-gamma model by borrowing strength values

from geographically referenced neighbouring values. This paper first discusses the existing methods: SMR and Poisson-gamma models. Then, the deterministic SIC model for pneumonia is mathematically represented by differential equations system. The corresponding discrete time-space stochastic SIC model is then developed using this deterministic SIC model. Finally, the existing methods and the proposed model are demonstrated using the Malaysia pneumonia dataset. The comparison between these three methods is shown in the results section.

RELATED STUDIES

To solve the shortcomings of the SMR and Poisson-gamma models as mentioned previously, researchers developed a transmission model or mathematical modelling (Samat, 2012). However, the deterministic model is used in most mathematical models, particularly for pneumonia. Even though many researchers prefer this type of model, it does not consider any possibility of random effects. A random effect is an additional variance component that can be estimated on a map and associated with a specified probabilistic structure. In other words, the random effect is treated as a nuisance factor. Therefore, this study proposes a stochastic pneumonia disease transmission model that considers the parameters significant for pneumonia disease. Furthermore, the random effect of spatial prior distribution is used, which allows dependency between adjacent areas.

From previous studies, researchers came out with mathematical modelling to understand more about the transmission of pneumonia disease. According to Mcbryde (2006) and Assab et al. (2017), there are two reasons why mathematical models are helpful in controlling diseases. Firstly, it can be used to quantitatively forecast the development of an epidemic, for instance, the peak of its total size, peak time, and the repercussion of infection control interventions. This development includes non-linear interaction that occurs when multiple interventions are undertaken. Secondly, to refrain from the assumption of serial independence and to handle the interval censoring as well as the unknown number of infectious diseases, a mathematical model is used as it can provide information on trial design and statistical analysis structure. Besides, mathematical modelling can also show the progress of infectious diseases. The SIC model is the fundamental mathematical modelling for pneumonia.

Doura et al. (2000) presented two models known as the Susceptible–Infected–Carriers (SIC) model and the Susceptible–Infected susceptible–Carriers–Infected carriers (SISCIc) model, which described the dynamics of the population using streptococcus. From this basic SIC model, other researchers built other models such as the Susceptible–Carriers–Infected–Recovered (SCIR) by Otieno et al. (2012); the Susceptible–asymptomatic Infectives–symptomatic Infectives–treated Infective (SIcIiT) model by Ndelwa et al. (2015); the Susceptible–Exposed–Infectious–Recovered (SEIR) model by Kassa and Murthy (2016); and the Susceptible–Vaccinated–Carrier–Infected–Recovered (SVCIR) model by Tilahun et al. (2017). Recently, Soliman and Bueno (2018) used the basic Susceptible–Infected–Recovered (SIR) model to analyse and predict the transmission rate of pneumonia in the Philippines. Mbabazi et al. (2019) formulated a Susceptible–Vaccinated–Exposed–Carrier–Infected (SVECI) model, which was a time delay model for pneumonia. From the literature, there were only two studies on pneumonia disease that used the stochastic model (Smith et al., 1993; Melegaro et al., 2002). Therefore, in this study, the stochastic pneumonia disease transmission model is proposed, where random effects that are significant in pneumonia disease are considered in the model. The model considered in this study is the SIC model.

MATERIALS AND METHOD

The values of relative risk were calculated using WinBUGS software in this analysis. This software is a programme designed to use Markov chain Monte Carlo (MCMC) computations to implement Bayesian inference on statistical problems (Lawson et al., 2003). The results of this research are described in table and graph forms. A pneumonia risk map was employed to display the low and high-risk areas that depend on measured relative risk values. ArcGIS software was used to generate the map.

SMR Method

The SMR approach essentially divides the observed incidence by the expected occurrence (Lawson, 2006). The SMR approach was commonly applied for counts within tracts analysis, and it was calculated as follows in Equation 1:

$$\hat{\theta}_i = O_i / E_i, \tag{1}$$

In this context, O_i refers to the number of observed incidence cases or disease-related fatalities in the location, whereas E_i is the number of expected cases. Assume that the mapping of the disease study area is divided into M mutually exclusive states (i = 1, 2,…, M). Each state has a different number of observed cases, O_i, as well as the number of expected cases, E_i. In the SMR model, according to Samat (2012) and Awang and Samat (2017), the values of O_i and E_i are obtained from the existing data to estimate the relative risk, $\hat{\theta}_i$, for area i. The values can be defined as in Equation 2:

$$r_i = \hat{\theta}_i = \frac{O_i}{E_i} \tag{2}$$

The number of observed cases can be obtained from various sources, including health indicators from the Ministry of Health Malaysia's website. Based on Samat (2012), the value of expected cases, E, for the research location, I, is determined by using the formula as in Equation 3:

$$E_i = N_i \frac{\sum O_i}{\sum N_j} \tag{3}$$

where N_i is the population of location, I, and the summations (Σ) denote j = 1, 2, ..., M.

Poisson-gamma Model

Other researchers suggested the Poisson-gamma model to address the shortcomings of the Standardised Morbidity Ratio (SMR) approach (Lawson et al., 2003; Samat & Imam, 2013). The Poisson distribution was applied in this study since it is the essential model for count data. Here, the areas of study are represented by i = 1, 2,…, M, while j = 1, 2,…, T represents the times. Assume that the number of new infections, y_{ij}, follows the distribution of Poisson during a time, with mean and variance, $e_{ij}\theta_{ij}$. The number of expected new infections is stated as e_{ij}, and the relative risk is stated as θ_{ij}; thus, it can be written as in Equation 4:

$$y_{ij}|e_{ij}, \theta_{ij} \sim Poisson\ (e_{ij}\theta_{ij}) \tag{4}$$

The parameter of relative risk has a gamma prior distribution with parameters α and β, as shown in Equation 5:

$$\theta_{ij} \sim Gamma(\alpha, \beta). \tag{5}$$

One of the analysis outcomes is the expected posterior relative risk based on this Poisson-gamma model. Further discussion based on this method can be found in Diah and Aziz (2021).

Deterministic Discrete Time-Space SIC Model for Pneumonia

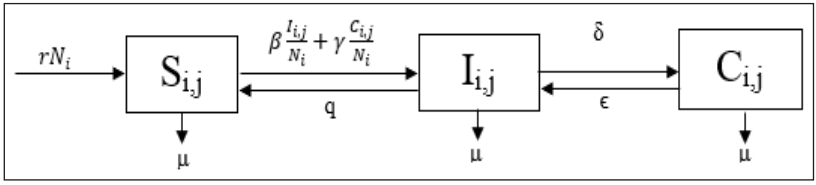
The SIC model is one of the basic models used in the studies of pneumonia disease transmission. This model is adapted from Doura et al. (2000) and classifies the population into three compartments, where $i = 1, 2,\ldots, M$ research locations and $j = 1, 2,\ldots, T$ period intervals in this study:

a) Number of susceptible individuals to the disease (susceptible) in total for location, i, at period, $j - S_{i,j}$,
b) Number of contagious (Infected or Infectious) individuals in total for location, i, at period, $j - I_{i,j}$,
c) Number of carrier (already infected and no external inoculation is needed) individuals in total for location, i, at period, $j - C_{i,j}$.

Here, $S_{i,j}$, $I_{i,j}$, and $C_{i,j}$, reflect the number of individuals in each compartment, while the total host population is $N_i = S_{i,j} + I_{i,j} + C_{i,j}$. The flow of the transmission process is shown in Figure 1.

Figure 1

SIC Model Flow Diagram



From Figure 1, μ represents the annual birth and death rates (assumed to be equal), β represents the rate at which an infected person can infect a susceptible person, and γ represents the rate at which a carrier can infect a susceptible person. The recovery rate where infectious individuals recover and move into the carrier class is represented by δ, while q denotes the recovery rate where infectious individuals recover and move into the susceptible class. ϵ signifies the rate at which carrier individuals become infected (move into the infectious class),

and N_i represents the total number of populations for the research region *i*. The following differential equations can be used, as shown in Equations 6–8, to express this model mathematically.

$$S_{i,j} = \mu N_i + qI_{i,j-1} + (1-(\beta \frac{I_{i,j-1}}{N_i} + \gamma \frac{C_{i,j-1}}{N_i}) - \mu)S_{i,j-1} \quad (6)$$

$$I_{i,j} = (\beta \frac{I_{i,j-1}}{N_i} + \gamma \frac{C_{i,j-1}}{N_i})S_{i,j-1} + (1 - q - \mu - \delta)I_{i,j-1} + \epsilon C_{i,j-1} \quad (7)$$

$$C_{i,j} = \delta I_{i,j-1} + (1 - \epsilon - \mu)C_{i,j-1} \quad (8)$$

This formulation is used to improve the model by considering the stochastic element, which is discussed in the next section.

Stochastic Discrete Time-Space SIC Model for Pneumonia

The deterministic model gives a decent estimate of the stochastic method for major epidemics for a larger sample size (Molzon, 2009; Samat & Percy, 2012). In this study, a deterministic model formulation was used to provide an estimation for the stochastic means. The current study expanded the deterministic SIC model by using a spatial prior, where the correlations between adjacent regions were considered, as discussed later in this section. A new term and notation, $\bar{I}_{i,j}$, were included in this study to represent the numbers of new infections in the research area, *i*, and study period (*j* -1, *j*). This study focused on determining the number of new infectious persons ($\bar{I}_{i,j}$), which was the model's only stochastic component, while the other terms were non-stochastic elements.

Equations 6–8 of the deterministic SIC model were used to develop the stochastic model for the transmission of pneumonia, as shown in Equations 9–14, where $i = 1, 2, \ldots, M$ research locations and $j = 1, 2, \ldots, T$ times.

$$\bar{I}_{i,j} = (\beta \frac{I_{i,j-1}}{N_i} + \gamma \frac{C_{i,j-1}}{N_i})S_{i,j-1} \quad (9)$$

$$\bar{I}_{i,j} \sim Poisson(\lambda_{i,j}) \quad (10)$$

$$\lambda_{i,j} = \exp(\beta_0 + b_i)(\beta \frac{I_{i,j-1}}{N_i} + \gamma \frac{C_{i,j-1}}{N_i})S_{i,j-1} \quad (11)$$

$$\log(\lambda_{i,j}) = \beta_0 + b_i + \log(\frac{\beta I_{i,j-1}}{N_i} + \frac{\gamma C_{i,j-1}}{N_i}) + \log(S_{i,j-1})$$

$$S_{i,j} = \mu N_i + qI_{i,j-1} + (1-\mu)S_{i,j-1} - \bar{I}_{i,j} \quad (12)$$

$$I_{i,j} = \bar{I}_{i,j} + (1-q-\mu-\delta)I_{i,j-1} + \epsilon C_{i,j-1} \quad (13)$$

$$C_{i,j} = \delta I_{i,j-1} + (1-\epsilon-\mu)C_{i,j-1} \quad (14)$$

Poisson distribution was used to model the number of new infectious persons since it is the essential model for count data. The deterministic part of Equation 7 was matched with a positive multiplicative factor using mean, $\lambda_{i,j}$, to show the spatial correlation. The number of new infectious persons was considered to follow the Poisson distribution in which the number of expected newly infectious individuals comprised some aspects of disease transmission. Consequently, the number of new infections was meant to occur in the same geographical area and a linear prediction term, which might have random effects or covariates. The transmission of new infectious cases was calculated by $\lambda_{i,j}$. For the starting point of the count data analysis, Poisson distribution was used (Lawson, 2006; Samat, 2012; Awang, 2017). The number of new infectious cases implied the number of cases.

Based on Figure 2, a linear predictor term was the key aspect of the transmission equation as it considered any covariates or a random effect and a simple direct dependence of the current number of infectious cases on the same previous spatial unit, such as in Equation 15:

$$(\beta \frac{I_{i,j-1}}{N_i} + \gamma \frac{C_{i,j-1}}{N_i})S_{i,j-1} \quad (15)$$

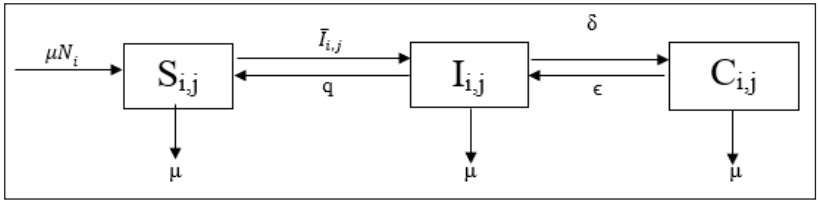
The number of new infectious persons is given as in Equation 16:

$$\bar{I}_{i,j} = (\beta \frac{I_{i,j-1}}{N_i} + \gamma \frac{C_{i,j-1}}{N_i})S_{i,j-1} \quad (16)$$

The equations for $S_{i,j}$, $I_{i,j}$, and $C_{i,j}$ were fixed or non-stochastic. Since these counts were conditional on other variables, the Poisson distribution could not be tested separately. In Equation 11 above, β_o is a constant term representing the overall process rate, while b_i refers to the random effect that absorbs the remaining spatial variance and allows dependency on adjacent areas. Figure 2 demonstrates the flow of the stochastic SIC model.

Figure 2

Stochastic SIC Model Diagram



For the random effect in this research, the conditional autoregressive prior (CAR) was applied as a prior distribution family. Besag et al. (1991) introduced this CAR model in which the probability densities values are based on adjacent areas at any given location. In estimating the relative risk for pneumonia, this study used the proposed discrete time-space stochastic SIC model.

Estimation of Relative Risk

This study used WinBUGS software to run this model, which is a programme intended to conduct Markov chain Monte Carlo (MCMC) computations in an extensive range of Bayesian statistical problem inferences (Lawson et al., 2003). This non-linear equation can be solved using computer programming as it can perform numerical analysis.

The first step in this analysis was finding the posterior distribution of the mean number of new infectious individuals. The relative risk's posterior mean would then be found from this input. In this research, the relative risk was estimated using a formula recommended by Samat and Percy (2012). Generally, using an unbiased sample mean, for $i = 1, 2, \ldots, M$ research locations and $j = 1, 2, \ldots, T$ times, the mean number of posterior expected infectiousness can be estimated as in Equation 17:

$$\tilde{\lambda}_{i,j} = \frac{1}{n}\sum_{k=1}^{n} \lambda_{ijk} \, . \tag{17}$$

where λ_{ijk} for $k = 1, 2, \ldots, n$ is formed by the posterior distribution for the expected mean number of infectious, $\tilde{\lambda}_{ij}$.Equation 18 shows the relative risk parameter, θ_{ij} is described as:

$$\theta_{ij} = \frac{\lambda_{ij}}{e_{ij}} \, . \tag{18}$$

where e_{ij} is a new expected number of infectious cases population-based in the study locations. Equation 19 shows that the relative risk of posterior expected was equal to the posterior expected mean number of new infectious, $\tilde{\lambda}_{ij}$, divided by the mean number of infectious for the population in all research locations, $\tilde{e}_{ij}$.

$$\tilde{\theta}_{ij} = \frac{1}{n}\sum_{k=1}^{n} \theta_{ijk} = \frac{1}{n}\sum_{k=1}^{n} \frac{\lambda_{ijk}}{e_{ij}} = \frac{\tilde{\lambda}_{ij}}{\tilde{e}_{ij}} \quad (19)$$

The present study defined relative risk as the conditional probability that individuals in the location get the disease divided by the conditional probability that individuals in the whole population get the disease. It is stipulated that, in this case, there has been no infection so far. This formula was applied to estimate the relative risk for mapping pneumonia using the discrete time-space stochastic SIC model.

Dataset

For this study's analysis, the dataset was provided by the Ministry of Health Malaysia and the Department of Statistics Malaysia. These three methods were applied to pneumonia data from 2010 until 2019 in the number of cases for all Malaysian states, including the three federal territories of Kuala Lumpur, Putrajaya, and Labuan. The values for β, γ, and δ were chosen to be 0.5445, 0.05, and 1, respectively, based on the study by Kassa and Murthy (2016). The annual rates of μ, q, and ϵ were individually 0.01334, 0.002747, and 0.3059, which were converted from the daily rates obtained from Samat (2012). The results were presented in the form of graphs, tables, and pneumonia risk maps, which showed the high-low risk areas of pneumonia incidents.

RESULTS

Figures 3 to 5 display the estimated results of the relative risk values for 16 states (including three federal territories) in Malaysia based on three methods used in this study. From Figure 3, there were no values of posterior expected relative risk for the epidemiology year 2010. This outcome is due to the condition referred to as moving average, in which outcomes were generated by averaging a set of previous data points. Based on Figures 3 and 4, from 2010 until 2019, a majority of states had relative risk values greater than one, implying that individuals in these states tended to be infected with pneumonia more

than individuals throughout the whole population. These states were Perlis, Kedah, Perak, Putrajaya, Negeri Sembilan, Melaka, Johor, Pahang, Terengganu, and Kelantan. Figure 5 shows the same results for states with relative risk values of more than one, but there was one additional territory, which was Labuan.

Relative risk is defined in this study as the conditional probability of someone inside a certain location becoming infected with the disease divided by the conditional probability of someone in the entire population becoming infected. For most epidemiology years based on the graphs in Figures 3 and 4, the relative risk values for Sabah, Sarawak, and Labuan were close to one. In contrast, from Figure 5, only Sarawak and Sabah had relative risk values close to one. Accordingly, there was no significant difference in the chance of persons in these states and the general population contracting pneumonia. Contrariwise, for Pulau Pinang, Kuala Lumpur, and Selangor (in Figures 3 to 5), in most epidemiology years, the relative risk values of these three states were less than one, indicating that the people within these states were less probable to get pneumonia compared to the whole population in Malaysia. Figure 3 revealed a zero value for relative risk when utilising the SMR approach since no observed count data were found in Putrajaya in 2012. However, there were values when using the Poison-gamma and stochastic SIC models. Consequently, the Poisson-gamma and stochastic SIC models could overcome the drawback of the SMR model.

Figure 3

Time Series Plot for Relative Risk Estimation using SMR for Malaysian States

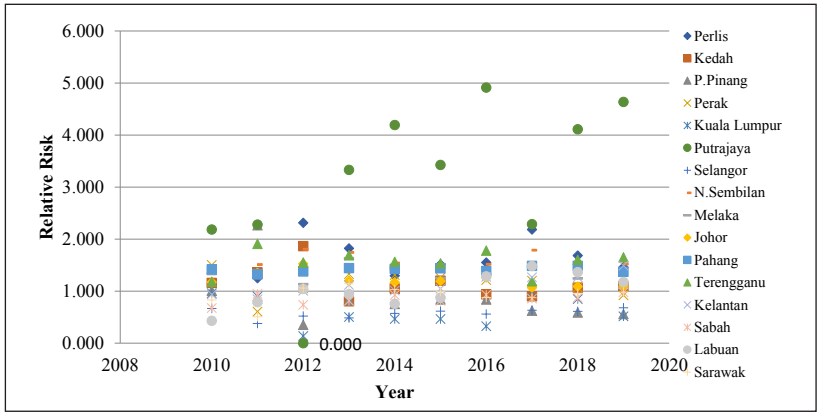


Figure 4

Time Series Plot for Relative Risk Estimation for Different States in Malaysia Using Poisson-Gamma Model

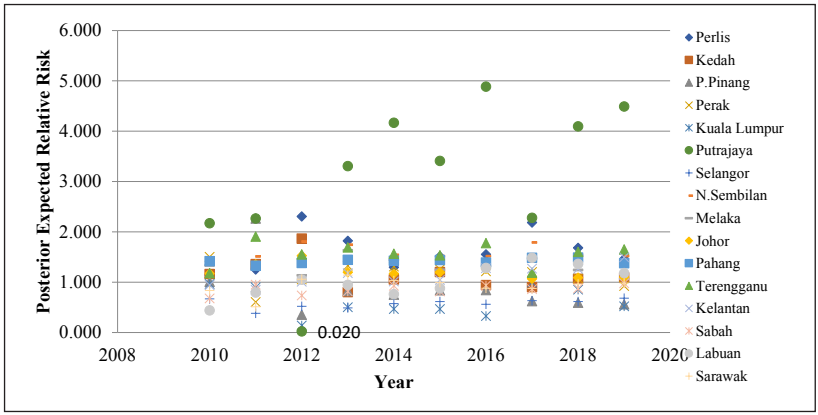


Figure 5

Time Series Plot for Relative Risk Estimation Using Stochastic SIC Model for Malaysian States

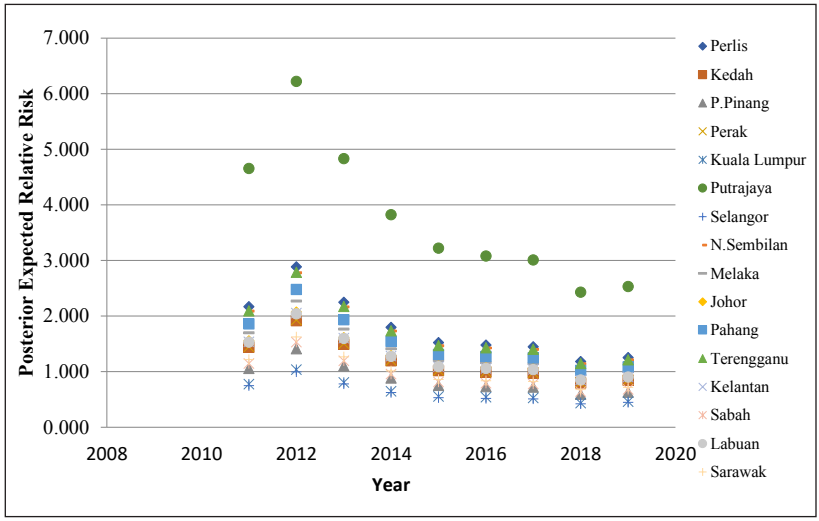


Table 1 displays numerical relative risk values for the year 2019. As seen in Table 1, by using all three models, the most significant

difference was found in the Federal Territory of Putrajaya with the highest value of relative risk for getting pneumonia. This finding revealed that susceptible persons in Putrajaya were more likely to become infected with pneumonia than individuals in the overall population. Conversely, for the lowest risk area, the SMR and Poisson-gamma models both showed that susceptible persons in Kuala Lumpur had the lowest risk of contracting pneumonia, while for the stochastic SIC model, Selangor was recognised to be the lowest risk area with a value of 0.493.

Table 1

A Comparison of Three Different Models for Estimating the Posterior Expected Relative Risk for Pneumonia in 2019

	Relative Risk Estimations of Pneumonia		
State	**SMR**	**Poisson-gamma**	**Stochastic SIC**
Perlis	1.461	1.460	1.530
Kedah	1.094	1.094	0.980
Pulau Pinang	0.559	0.559	0.694
Perak	0.922	0.922	0.993
Kuala Lumpur	0.519	0.519	0.519
Putrajaya	4.638	4.492	3.775
Selangor	0.682	0.682	0.493
N. Sembilan	1.522	1.522	1.463
Melaka	1.065	1.066	1.175
Johor	1.099	1.099	1.062
Pahang	1.375	1.375	1.308
Terengganu	1.652	1.652	1.485
Kelantan	1.491	1.491	1.071
Sabah	0.978	0.978	0.775
Labuan	1.178	1.179	1.045
Sarawak	0.984	0.984	0.827

From the results shown in Table 1, choropleth pneumonia maps in Figures 6, 7, and 8 were constructed to illustrate the low and high-risk areas for pneumonia cases in 16 Malaysian states and federal territories, especially for the year 2019. Every state had been categorised with five distinct levels of risk, ranging from extremely low to very high risk at intervals of (0.0, 0.5), (0.5, 1.0), (1.0, 1.5),

(1.5, 2.0), and $(2.0, \infty)$. This interval division was chosen for a more in-depth analysis based on Samat and Percy (2012). In this case, to distinguish the relative risk levels, the lowest risk was represented by the brightest hue, and the darkest shade signified the very high risk. Interested parties can clearly see from this map which states have the highest risk of pneumonia occurrences and require closer scrutiny or extra consideration.

Figure 6

Relative Risk Estimation Map for Pneumonia based on SMR for the Year 2019

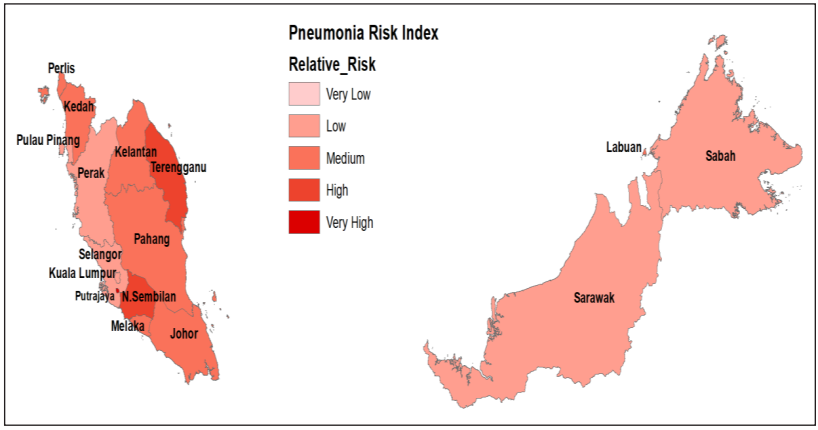


Figure 6 shows the risk map of pneumonia occurrence based on the SMR method. Putrajaya was classified as a very high-risk area with a relative risk value of 4.638, followed by Negeri Sembilan and Terengganu as high-risk areas. Areas with medium risk comprised Kedah, Perlis, Johor, Melaka, Kelantan, Pahang, and Labuan. The other six states were categorised as areas with low risk. No states were identified as very low-risk areas.

Figure 7

Relative Risk Estimation Map for Pneumonia based on Poisson-Gamma Model for the Year 2019

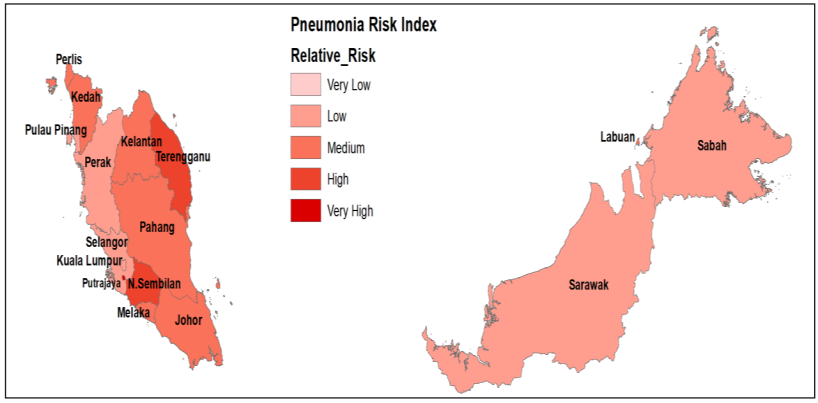


Figure 7 illustrates the relative risk map based on the Poisson-gamma model. The figure showed a similar result with the SMR method for the very high-risk area, which was Putrajaya, and no states categorised as very low-risk areas. As for the Poisson-gamma model, Negeri Sembilan and Terengganu were classified as high-risk areas. The model listed seven states as medium-risk areas: Kedah, Perlis, Johor, Melaka, Kelantan, Pahang, and Labuan. The other six states were identified as low-risk areas.

Figure 8

Relative Risk Estimation Map for Pneumonia using Stochastic SIC Model for the Year 2019

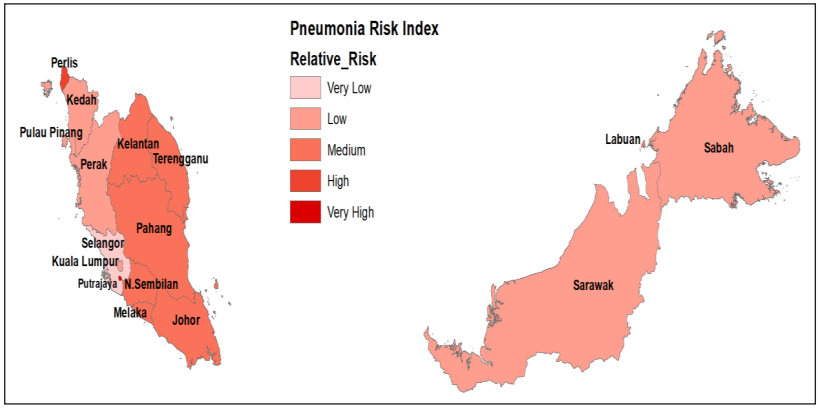


Figure 8 demonstrates the map of pneumonia risk occurrence using the stochastic SIC model. From the map, similar to those results from SMR and Poisson-gamma, Putrajaya was a very high-risk area for pneumonia occurrence. Perlis, Negeri Sembilan, Melaka, Pahang, Johor, Kelantan, and Terengganu were recognised as areas with medium risk. Seven other states were classed as low risk, except for Selangor, which was identified as a very low-risk area for pneumonia occurrence.

From all the maps above, in terms of colour tones, Figure 8 revealed a large gap in risk compared to the other two methods. When using the stochastic SIC model, the relative risk values decreased from high-risk to medium-risk level and from low-risk to very low-risk level. This outcome is because the stochastic SIC model considered extra information such as disease transition and spatial correlation. According to Diah et al. (2017), spatial correlation needs to be considered in estimating relative risks as the states are located next to each other. Therefore, the risk might be transferred to other states.

To identify which model is better among these two models, Poisson-gamma and stochastic SIC, in estimating relative risk values, Deviance Information Criterion (DIC) was used as a goodness-of-fit (GOF) measure. The SMR method was not measured with DIC since there was zero relative risk value for year 2012 when using this method. Therefore, this study only compared Poisson-gamma and stochastic SIC models. Spiegelhalter et al. (2002, 2014) suggested DIC and it has been used by Kristiani et al. (2016) and Alhdiri et al. (2017) in finding the better model. The DIC values for each model were calculated using WinBUGS software, as shown in Table 2. According to Spiegelhalter et al. (2002) and Lawson et al. (2003), the model with the lowest DIC value is believed to be the best model for the data.

Table 2

Deviance Information Criterion (DIC) based on Poisson-gamma and stochastic SIC models

	Poisson-gamma	Stochastic SIC
DIC	1963.32	1692.47

The stochastic SIC model provided a better model for pneumonia as it had the lowest DIC value. In summary, the stochastic SIC model is one of the best models to be used in the future in estimating the relative risk for pneumonia data and other diseases with similar transmission criteria.

DISCUSSION AND CONCLUSION

The discrete time-space stochastic SIC model was proposed in this study to estimate the values of relative risk for pneumonia. This model was developed from the deterministic SIC model proposed by Doura et al. (2000). In this study, the current approaches, SMR and Poisson-gamma models, were compared with the proposed model, the stochastic SIC model. This study's findings were presented in the form of graphs, tables, and maps. The maps provided clear views of states with high to low risks of pneumonia occurrence. About two-thirds of Malaysia's states were more likely to contract pneumonia. As seen in Figure 8, when compared to other states, Putrajaya was a very high-risk area for pneumonia occurrence, whereas Selangor was the lowest risk area of contracting pneumonia. These findings are equivalent to the number of prevalence for pneumonia, which was predicted in 2019 to be 176 per 10,000 persons in Putrajaya and 25 per 10,000 persons in Selangor. According to statistics from the Ministry of Health Malaysia Annual Report 2019, Putrajaya recorded the highest annual population growth rate of 6.55%, while Selangor with 0.82% (Ministry of Health Malaysia, 2019). Moreover, Putrajaya is the smallest territory in Malaysia (Department of Survey and Mapping, 2017). As a result, susceptible people in Putrajaya have the highest risk of getting pneumonia.

It can be concluded that the population's size in the area and the region's size itself can influence the relative risk estimation values. Therefore, this study proposed the stochastic model that considered the spatial elements and the number of populations for each state and federal territory. This study's proposed stochastic SIC model examined the spatial correlation between adjacent regions and allowed covariate adjustment. Consequently, this can help to solve the shortcomings of the SMR and Poisson-gamma models. From the study's findings, the stochastic SIC model can be applied as a fundamental method to estimate relative risk instead of using the number of pneumonia

cases alone. The maps should also be seen as a tool for informing and guiding government strategies for controlling and monitoring pneumonia disease.

ACKNOWLEDGMENT

The authors would like to express their gratitude to Universiti Utara Malaysia and family members for supporting this study. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. The authors also appreciate the Ministry of Health Malaysia and the Department of Statistics Malaysia for the data.

REFERENCES

Awang, A. C. (2017). *The development of stochastic SIR-SI age-structured model for leptospirosis mapping in Malaysia* (Master's thesis, Universiti Pendidikan Sultan Idris, Tanjung Malim. Perak, Malaysia). https://ir.upsi.edu.my/detailsg.php?det=5750

Awang, A. C., & Samat, N. A. (2017, May). Standardized morbidity ratio for leptospirosis mapping in Malaysia. In *AIP Conference Proceedings* (Vol. 1847, No. 1, p. 020006). AIP Publishing LLC. https://doi.org/10.1063/1.4983861

Alhdiri, M. A., Samat, N. A., & Mohamed, Z. (2017). Risk estimation for lung cancer in Libya: Analysis based on Standardized Morbidity Ratio, Poisson-gamma model, BYM model and mixture model. *Asian Pasific Journal of Cancer Prevention*, *18*(3), 673–679. https://doi.org/10.22034/APJCP.2017.18.3.673

Assab, R., Nekkab, N., Crépey, P., Astagneau, P., Guillemot, D., Opatowski, L., & Temime, L. (2017). Mathematical models of infection transmission in healthcare settings: Recent advances from the use of network structured data. *Current Opinion in Infectious Diseases*, *30*(4), 410–418. http://doi: 10.1097/QCO.0000000000000390.

Besag, J., York, J., & Mollie, A. (1991). Bayesian image restoration with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics, 43*, 1–59.

Department of Statistics Malaysia. (2019). *Statistics on causes of death, Malaysia, 2019*. Department of Statistics Malaysia. https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=401&bul_id=RUxlSDNkcnRVazJnakNCNVN2VGgrdz09&menu_id=L0pheU43NWJwRWVSZklWdzQ4TlhUUT09

Department of Survey and Mapping Malaysia. (2017). *Keluasan Malaysia, 2017*. data.gov.my. https://www.data.gov.my/data/ms_MY/dataset/keluasan-malaysia/resource/a04e46de-5044-4081-b6cc-7853a46439b

Diah, I. M., Aziz, N., & Ahmad, N. (2016). Relative risk estimation of tuberculosis with standardized morbidity ratio in Malaysia. *Global Journal of Pure and Applied Mathematics*, *12*(5), 4011–4019.

Diah, I. M., Aziz, N., & Kasim, M. M. (2017). A comparison of four disease mapping techniques as applied to TB diseases in Malaysia. *Journal of Telecommunication, Electronic and Computer Engineering*, *9*(2), 133–137.

Diah, I. M., & Aziz, N. (2021). Mapping of pneumonia disease in Malaysia using Poisson-gamma model. *Annals of the Romanian Society for Cell Biology*, *25*(1), 2062–2067. https://www.annalsofrscb.ro/index.php/journal/article/view/324

Doura, K., Melendez-morales, J. D., Meyer, G. G., & Perez, L. E. (2000). An S-I-S model of streptococcal disease with a class of beta-hemolytic carriers. *Biometric Unit Technical Reports*, 493–518.

Kassa, M., & Murthy, S. N. (2016). Pneumonia control measures under five year children. *IOSR Journal of Mathematics (IOSR-JM)*, *12*(3), 64–70. https://doi.org/10.9790/5728-1203036470

Kristiani, F., Parahyangan, U. K., & Samat, N. A. (2016). Dengue disease mapping in Bandung, Indonesia: An analysis based on Poisson-gamma, Log-normal, BYM and mixture models. *Jurnal Teknologi (Sciences & Engineering)*, *78*(6–5). http://doi.org/10.11113/jt.v78.8991

Lawson, A. B., Browne, W. J., & Rodeiro, C. L. V., (2003). *Disease mapping with WinBUGS and MLwiN*. Statistics in Practise. Chichester: John Wiley & Sons, Ltd.

Lawson, A. B. (2006). *Statistical methods in spatial epidemiology*. Wet Sussex, UK: John Wiley & Sons, Ltd.

Mayo Clinic. (2020). *Pneumonia*. Mayo Clinic. https://www.mayoclinic.org/diseases-conditions/pneumonia/symptoms-causes/syc-20354204

Mbabazi, F. K., Mugisha, J. Y. T., & Kimathi, M. (2019). Hopf-Bifurcation analysis of pneumococcal pneumonia with time delays. *Abstract and Applied Analysis*, *2019*(January). https://doi.org/10.1155/2019/3757036

Mcbryde, E. (2006). Mathematical and statistical modelling of infectious diseases in hospitals. *Sciences-New York*, *10*(November), 777–782. https://doi.org/10.1093/nar/gkr403

Melegaro, A., Gay, N. J., & Medley, G. F. (2004). Estimating the transmission parameters of pneumococcal carriage in households. *Epidemiology and Infection*, *132*(3), 433–441. https://doi.org/10.1017/S0950268804001980

Meza, J. L. (2003) Empirical Bayes estimation smoothing of relative risks in disease mapping. *Journal of Statistical Planning and Inference*, *112*, 43–62. https://doi.org/10.1016/S0378-3758(02)00322-1

Ministry of Health Malaysia. (2019). *Laporan tahunan Kementerian Kesihatan Malaysia 2019*. Ministry of Health Malaysia. https://www.moh.gov.my/moh/resources/Penerbitan/Penerbitan%20Utama/ANNUAL%20REPORT/LAPORAN%20TAHUNAN%20KKM%202019/mobile/index.html#p=1

Molzon, R. (2009, May). Deterministic approximation of stochastic evolutionary dynamics. In *Proceedings of the 2009 International Conference on Game Theory for Networks, GameNets '09* (pp. 323–332). IEEE.

Ndelwa, E. J., Kgosimore, M., Massawe, E. S., & Namkinga, L. (2015). Mathematical modelling and analysis of treatment and screening of pneumonia. *Mathematical Theory and Modeling*, *5*(10), 21–40.

Normandin, B. (2021, November 10). *All you need to know about pneumonia*. healthline. https://www.healthline.com/health/pneumonia

Otieno, O. J., Joseph, M., & John, O. (2012, November). Mathematical model for pneumonia dynamics among children. In *The 2012 Southern Africa Mathematical Sciences Association Conference*, *(SAMSA)* (pp. 26–29). https://su-plus.strathmore.edu/bitstream/handle/11071/3615/Mathematical Model for Pneumonia Dynamics paper.pdf?sequence=1&isAllowed=y

Samat, N. A. (2012). *Mathematical models for vector-borne infectious disease mapping with application to dengue disease in Malaysia*. (Doctoral dissertation, University of Salford, Manchester, UK).

Samat, N. A., & Percy, D. F. (2012, September). Dengue disease mapping in Malaysia based on stochastic SIR models in human populations. In *ICSSBE 2012 - Proceedings, 2012 International Conference on Statistics in Science, Business and Engineering: "Empowering Decision Making with Statistical Sciences"* (pp. 623–627). http://doi:10.1109/ICSSBE.2012.6396640

Samat, N. A., & Imam, S. H. M. (2013). Dengue disease mapping with standardized morbidity ratio and Poisson-gamma model: an analysis of dengue disease in Perak, Malaysia. *International Journal of Mathematical and Computational Sciences*, *7*(8), 640–644. http://doi.org/10.5281/zenodo.1086703

Seramo, R. K., Awol, S. M., Wabe, Y. A., & Ali, M. M. (2022). Determinants of pneumonia among children attending public health facilities in Worabe town. *Scientific Reports*, *12*(1), 6175. https://doi.org/10.1038/s41598-022-10194-z

Smith, T., Lehmann, D., Montgomery, J., Gratten, M., Riley, I. D., & Alpers, M. P. (1993). Acquisition and invasiveness of different serotypes of streptococcus pneumoniae in young children. *Epidemiology and Infection*, *111*(1), 27–39. https://doi.org/10.1017/S0950268800056648

Soliman, B. W. M., & Bueno, A. C. F. (2018). Modelling the spread of pneumonia in the Philippines using Susceptible-Infected-Recovered (SIR) model with demographic changes. *Journal of Technology Management and Business*, *5*(1), 28–35.

Spiegelhalter, D. J., Best N. G., Carlin B. P., & van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal Royal Statistical Society: Series B (Statistical Methodology)*, *64,* 583–616. http://doi:10.1111/1467-9868.00353

Spiegelhalter D. J., Best N. G., Carlin B. P., & van der Linde, A. (2014). The deviance information criterion: 12 years on. *Journal Royal Statistics Society: Series B (Statistical Methodology)*, *76*(3), 485–493. http://doi:10.1111/rssb.12062

Tilahun, G. T., Makinde, O. D., & Malonza, D. (2017). Modelling and optimal control of pneumonia disease with cost-effective strategies. *Journal of Biological Dynamics*, *11*, 400–426. https://doi.org/10.1080/17513758.2017.1337245

UNICEF. (2019). *Pneumonia: A child dies of pneumonia every 39 seconds*. UNICEF Data. https://data.unicef.org/topic/child-health/pneumonia/