

# Mapping of Pneumonia Disease in Malaysia Using Poisson-Gamma Model

<sup>1</sup>Ijlal Mohd Diah & <sup>1,2</sup>Nazrina Aziz\*

\* Corresponding author:

<sup>1</sup>Department of Mathematics and Statistics, School of Quantitative Sciences, UUM College of Arts and Sciences, Universiti Utara Malaysia, 06010, Sintok, Kedah Darul Aman, Malaysia.

<sup>2</sup>Institute of Strategic Industrial Decision Modelling (ISIDM), UUM.

\*nazrina@uum.edu.my

## Abstract

Bayesian models have been used as one of the estimation techniques for smoothing relative risks estimation for disease mapping. One of the most common uses of Bayesian methodology in disease mapping is Poisson-gamma model. This paper begins with a brief introduction of pneumonia disease. Then followed by a review of methodology used in this paper that is Poisson-gamma model which used to estimate the values of relative risk. Then, the results will be shown into a map to show pneumonia risk areas. The risk areas can be shown clearly through disease map. Thus, the objective of this study is to estimate the value of relative risk for pneumonia based on Poisson-gamma model. Pulau Pinang, Perak, Selangor, Kuala Lumpur & Putrajaya, Sabah, Labuan and Sarawak are categorized as low-risk of pneumonia occurrence, while Terengganu and Negeri Sembilan categorized as high-risk areas.

**Keywords:** disease mapping; pneumonia; relative risk; Poisson-gamma model

## 1. Introduction

This paper demonstrates and discusses the use of the most common and earliest Bayesian method that is Poisson-gamma model in disease mapping. First, we explain about pneumonia. This is followed by description of Poisson-gamma model that was suggested by Lawson et al.<sup>1</sup>. Pneumonia is an infection in one or both lungs which inflames the air sacs with pus or fluid. It can be caused by different types of organisms which include viruses, bacteria and fungi<sup>2</sup>. However, the most common cause for someone to get pneumonia is caused by bacterial pneumonia known as *Streptococcus pneumoniae* and viruses. These viruses include the new virus which is now already became outbreak worldwide that is Coronavirus-2019 (COVID-19), which may trigger pneumonia and can become serious. For children under 5 years old, the most well-known cause of pneumonia are viruses. Even though viral pneumonia usually is just minor, however it can become most terrible in some cases<sup>2</sup>. Both bacterial and viral pneumonia are contagious which mean that they can transmit the disease from a person to another via air when a person inhales airborne droplets from a sneeze or cough of an infected person. It also can be transferred to others through contact with objects or surfaces that are contaminated with bacteria or viruses. When someone is

infected with pneumonia, it will show symptoms such as chest pain when cough or breathe, fever, chills, shortness of breath, and coughing that produce mucus<sup>2</sup>. The symptoms can be varied from minor to serious depend on factors such as age, type of organisms that causes the inflammation, and overall condition of health. For mild symptom, it shows like a normal flu or cold, but it last longer.

The severity of pneumonia can range from mild to life-threatening. Infants and young children especially under-age of five, people aged 65 years and older, and people with weakened immune systems are the higher risk group and can become most serious for them if they are being infected with pneumonia<sup>2</sup>. Pneumonia has been recognized as top five diseases that causes of death globally with 2.6 million in 2017<sup>3</sup>. About 15% from these 2.6 million are deaths among children under-age of five. About 1 child dies of pneumonia every 39 seconds and this made it become number one killer among the leading infectious disease cause of death for children under-age of five<sup>4</sup>. These 2.6 million of deaths concentrated in South Africa, South East Asia, and sub-Saharan Africa. Even though pneumonia is a preventable disease, but it kills more children than other infections. It also recognized as a common cause of death for elderly where in 2017, about 1.13 million of deaths are among adults aged over 65 years<sup>3</sup>. According to Department of Statistics Malaysia, pneumonia has been listed as top three diseases that cause death in Malaysia<sup>5</sup>. In 2019, 145,419 cases had been reported with 7,542 number of deaths<sup>5</sup>. Pneumonia also been reported as top three causes of death for children under-age of five. If this situation persists, it will affect the number of future populations if it is not control where we will be facing ageing population as there are more people aged 70 years and older than there are children under-age of five in the future.

It is necessary to look into this problem as complication for those who infected with influenza and Corona Virus Disease 2019 (COVID-19) can lead someone to get pneumonia disease<sup>5</sup>. Both can contribute to the increase in the number of pneumonia cases. Pneumonia disease might become outbreak if it does not control as the number of pneumonia cases can be increase due to other diseases.

Current approach used in Malaysia to estimate the low and high-risk areas of pneumonia is by monitoring based on the reported overall number of cases in each state. However, this approach only showed general information without consider other important factors such as number of population and geographical areas.

One of the important tools in public health research is disease mapping where it can be help in controlling and also as the stopping strategies for a disease<sup>6</sup>. Disease mapping can be seen as a descriptive picture of pneumonia's burden in some geographical areas. Besides, it can help showing areas that need more attention from government especially in term of health supply and medical treatment. In this study, Empirical Bayes approach specifically Poisson-gamma model are used where this approach are highly recommended in the use of small area estimation as it smooth the relative risk and provides the measures of uncertainty associated with this relative risk estimation and the modeling can take into account the spatial autocorrelation. The approach to smoothing in Bayesian approach is by borrowing strength values from geographically referenced neighboring values.

## 2. Materials and Methods

In this study, the relative risk values are computed using WinBUGS software. This software is a program created to implement Bayesian inference on statistical problem using computations of Markov Chain Monte Carlo (MCMC)<sup>1</sup>. Findings of this study are presented in form of table and graph. A map of the pneumonia risk will be displayed to show the high-low risk areas depend on the values of relative risk which are estimated. ArcGIS software is used to produce the map. For the Poisson-gamma model, the prior parameters values,  $\alpha$  and  $\beta$  are unknown and are assumed to have exponential prior distributions with values of hyperparameter 0.1<sup>1</sup>. The prior expected relative risks in this study using this model is equal to 1.

### 2.1 Poisson-gamma Model

One of the initial Bayesian methods is Poisson-gamma model that has been suggested by other researchers to be used to overcome the weakness of the Standardized Morbidity Ratio (SMR) method<sup>1</sup>. Poisson distribution is used as this is the fundamental model for count data. Here,  $i=1,2,\dots,M$  represent the study areas while  $j=1,2,\dots,T$  represent the time period. Assuming that the number of new infections,  $y_{ij}$  follows the Poisson distribution over a period of time, with mean and variance,  $e_{ij}\theta_{ij}$ . In this model, the expected number of new infectives is expressed as  $e_{ij}$  and the relative risk is expressed as  $\theta_{ij}$ , hence:

$$y_{ij} | e_{ij}, \theta_{ij} \sim \text{Poisson}(e_{ij}\theta_{ij}) \quad (1)$$

The parameter of the relative risk has a gamma prior distribution with  $\alpha$  and  $\beta$  parameters:

$$\theta_{ij} \sim \text{Gamma}(\alpha, \beta) \quad (2)$$

Based on this Poisson-gamma model, the expected posterior relative risk will be one of the outputs of the analysis.

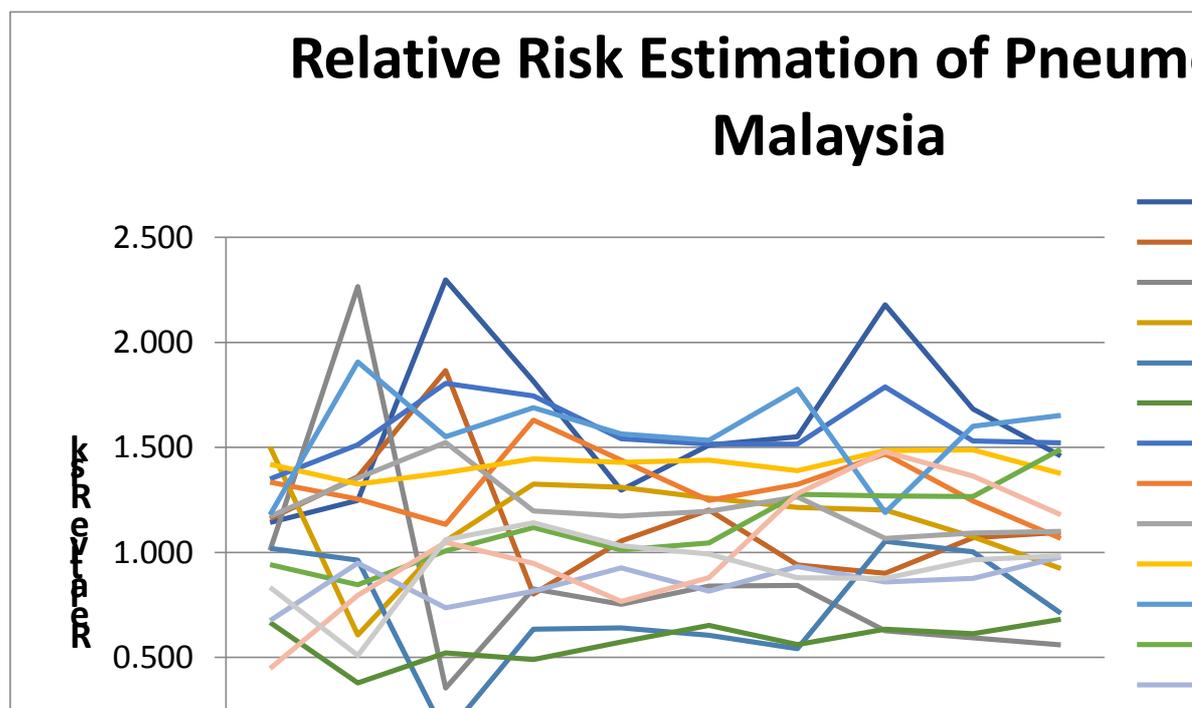
### 2.2 The Data Set

In this study, the data set was given by the Department of Statistics and Ministry of Health in Malaysia. This Poisson-gamma model are implemented to pneumonia data from year 2010 until year 2019 in form of the number of cases for 13 states in Malaysia that are Perlis, Kedah, Penang, Perak, Selangor, Negeri Sembilan, Melaka, Johor, Pahang, Terengganu, Kelantan, Sabah and Sarawak and three federal territories that are Kuala Lumpur, Putrajaya, and Labuan. However, in this study, both territories and states are informed as states for simplicity. In this study, for Putrajaya state pneumonia data are included in the data for the states of Kuala Lumpur where it is written as Kuala Lumpur & Putrajaya.

## 3. Results

Figure 1 shows the estimated results of the relative risk value for 15 states in Malaysia. From the graph, most states show the relative risk values more than one from year 2010 until 2019 which implies that the susceptible people in these states tend to get pneumonia disease compared with people throughout the whole population. The relative risk in this study is defined as the conditional probability that a person inside an area become infected with the disease divided by the conditional

probability that a person in the whole population become infected with the disease. From the graph, Sabah and Sarawak has relative risk close to one for most epidemiology weeks, which means there is no significant difference in terms of the likelihood that the people in these states and within the whole population to contract with pneumonia disease. Conversely, for Selangor, Kuala Lumpur & Putrajaya and Pulau Pinang, these three states have relative risk value less than one for most epidemiology weeks which shows that people within these states are less probable to get pneumonia disease when compared to whole population in Malaysia.



**Fig. 1** Plots of time series for the relative risk estimation using Poisson-gamma model for 15 states in Malaysia.

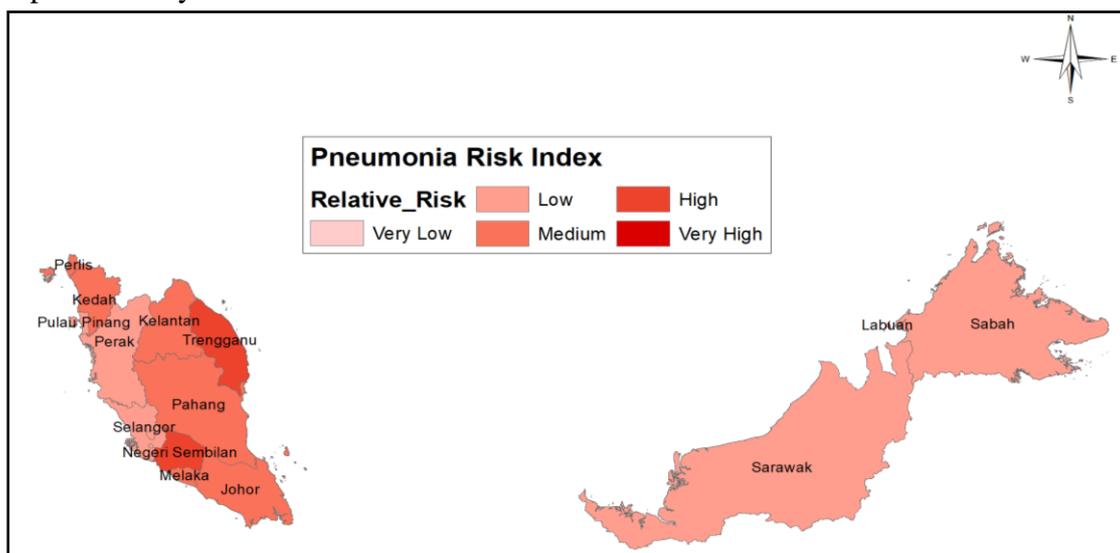
Numerical values for the relative risk are shown in Table 1 for year 2019. From Table 1, susceptible persons in the state of Terengganu has the highest risk of getting infected with pneumonia disease with 1.651 value of relative risk. In contrast with susceptible persons in Pulau Pinang where it recognized to be the lowest risk area with value of 0.559.

States	Relative risk
Perlis	1.458
Kedah	1.094
Pulau Pinang	0.559
Perak	0.922
Kuala Lumpur & Putrajaya	0.710
Selangor	0.682
N. Sembilan	1.521

Melaka	<b>1.066</b>
Johor	<b>1.099</b>
Pahang	<b>1.375</b>
Terengganu	<b>1.651</b>
Kelantan	<b>1.491</b>
Sabah	<b>0.978</b>
Labuan	<b>1.178</b>
Sarawak	<b>0.983</b>

**Table 1. Estimation of Relative Risk for Pneumonia for Year 2019**

Figure 2 shows the choropleth pneumonia map that show high-low risk areas for the occurrence cases of pneumonia for 15 states of Malaysia for year 2019. Each state is allocated with one of five different levels of relative risk, starting from very low to very high risks with intervals of  $[0.0,0.5)$ ,  $[0.5, 1.0)$ ,  $[1.0,1.5)$ ,  $[1.5,2.0)$  and  $[2.0,\infty)$ . Here, very low risk is represented by the brightest shade and for very high risk, it is represented by the darkest shade to differentiate the levels of relative risk.



**Fig. 2. Relative risk estimation map for pneumonia disease using Poisson-gamma model for year 2019**

Figure 2 demonstrates the map of Poisson-gamma. From the map, the state of Negeri Sembilan and Terengganu have high risk areas. Perlis, Kedah, Johor, Pahang and Kelantan have been identified as medium risk areas. The other states are identified as low risks. There is no state identified as very high risk and very low risk area.

#### 4. Conclusion

It is very important to estimate the values of relative risk in order to monitor and control the spread of pneumonia, especially in Malaysia. In this study, the Poisson-gamma model has been used to estimate the relative risk values, which this model is one of the initial approaches of Bayesian in estimating the relative risk. The results for

this study are shown in form of graph, table and also map. From the map, it gives obvious picture of areas with high-low risks. To conclude, there is no state that has been identified as very high-risk and very low-risk area. Terengganu and Negeri Sembilan highest risk area while Pulau Pinang, Perak, Selangor, Kuala Lumpur & Putrajaya, Sabah, Labuan and Sarawak show the lowest risk areas of contracting pneumonia. Based on the study by Lawson et al., they demonstrated the use of Poisson-gamma model in their study and it is shown that a map is smoother with less extreme values of the relative risk compared when using SMR method<sup>1</sup>. This Poisson-gamma model still has its own drawbacks where the covariate adjustment is difficult and there is no likely to deal with spatial correlation between risks in adjacent areas. Hence, this urges researchers to purpose other approaches to estimate the relative risk values. The map should be seen as a tool which can help to inform and direct government strategy for monitoring and controlling pneumonia disease.

### **Acknowledgements**

The authors would like to thank Universiti Utara Malaysia and also family members for the support of this study. The authors also acknowledge support from the Department of Statistics Malaysia and Ministry of Health Malaysia for the data.

### **References**

1. Lawson AB, Browne WJ, and Rodeiro CLV. Disease mapping with WinBUGS and MLwiN. Statistics in Practice. Chichester: John Wiley & Sons, Ltd; 2003.
2. Mayo Clinic. Pneumonia [Internet]. [cited 2020 June 13]. Available from <https://www.mayoclinic.org/diseases-conditions/pneumonia/symptoms-causes/syc-20354204>
3. Greenslade L. The Missing Piece: Why continued neglect of pneumonia threatens the achievement of health goals. Just Actions Report; 2018.
4. UNICEF. Pneumonia: A child dies of pneumonia every 39 seconds. 2019 [cited 2020 April 12]. Available from <https://data.unicef.org/topic/childhealth/pneumonia/>
5. Department of Statistics Malaysia. Statistics on Causes of Death, Malaysia, 2019 [Internet]. 2019 [cited 2020 August 14]. Available from [https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=401&bul\\_id=RUxISDNkcNRVazJnakNCNVN2VGgrdz09&menu\\_id=L0pheU43NWJwRWVVSZklWdzQ4TlhUUT09](https://www.dosm.gov.my/v1/index.php?r=column/cthemeByCat&cat=401&bul_id=RUxISDNkcNRVazJnakNCNVN2VGgrdz09&menu_id=L0pheU43NWJwRWVVSZklWdzQ4TlhUUT09)
6. Samat NA & Percy DF. Dengue disease mapping in Malaysia based on stochastic SIR models in human populations. ICSSBE 2012 – Proceedings, 2012 International Conference on Statistics in Science, Business and Engineering: “Empowering Decision Making with Statistical Sciences”; 623–627. doi:10.1109/ICSSBE.2012.6396640; 2012.