ASM Sc. J., 19, 2024
https://doi.org/10.32802/asmscj.2023.1526

COVID-19 Disease Mapping in Malaysia Using Poisson-gamma Model to Safeguard Community

J.E. Tee[1], N. Aziz[1,2*] and I. Mohd Diah[3]

[1]School of Quantitative Sciences, UUM College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah Darul Aman, Malaysia

[2]Institue of Strategic Industrial Decision Modelling (ISIDM), Universiti Utara Malaysia, 06010 Sintok, Kedah Darul Aman, Malaysia

[3]Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, 35900 Tanjung Malim, Perak Darul Ridzuan, Malaysia

Coronavirus disease (COVID-19) is an extremely transmissible disease that causes tremendous human suffering worldwide, with millions of people dying within two years. In Malaysia, the COVID-19 pandemic continues with an increasing number of positive cases. In order to monitor and prevent the disease, disease mapping is proposed for quantifying and describing the disease risk. The primary objective of this research is to estimate the COVID-19 relative risk using one of the most widely employed Bayesian approaches, namely the Poisson-gamma model, and eventually display a disease map to show the COVID-19 risk level in each state. The relative risk results are divided into five different levels to classify the COVID-19 risk level in each Malaysian state. The illustration of five different shades in the disease mapping has distinguished the risk levels within each state, where the most light-coloured area indicates very low risk whilst the most dark-coloured area indicates very high risk. According to the results, the area with the lowest risk of contracting COVID-19 in 2020 is Perlis, and the area with the highest risk is Labuan, whereas in 2021, Perlis remains the lowest risk area while Selangor is the highest risk area.

Keywords: COVID-19; disease mapping; relative risk; Poisson-gamma model; risk level

I. INTRODUCTION

In the process of civilisation development, infectious diseases have caused significant human agony and death, as well as paying huge economic costs. The number of emerging pathogens, such as Zika virus, Ebola virus, and severe acute respiratory syndrome coronavirus (SARS-CoV), has quadrupled in the last 40 years. SARS-CoV-2 is a recently emerged new pathogen and the cause of Coronavirus disease (COVID-19) (Kuppalli & Madad, 2021).

Since the initial emergence of COVID-19 cases in China, the contagious virus has been spreading rapidly worldwide and it has been recognised as a global crisis within four months (Liu *et al.*, 2020). Although the very first COVID-19 epidemic was linked to a wildlife and seafood market in the Wuhan city of China, some early cases were not related to the market. Thus, the original source of the virus remains elusive (Zhang & Holmes, 2020). Numerous investigations are still underway to further explore its exact origin in China.

COVID-19 affects different individuals in various ways. It is generally known that the conditions can be asymptomatic or have symptoms ranging from mild to severe (Elibol, 2021). The disease's most prevalent symptoms are fatigue, dry cough, and fever, whereas the uncommon symptoms are myalgia, sore throat, diarrhoea, pink eye, headache, hyposmia (loss of smell), or hypogeusia (loss of taste), skin rashes and discolouration of toes or fingers. These symptoms might develop complications such as dyspnea (shortness of breath), chest pain, and movement or speech impairments (WHO, 2021b).

*Corresponding author's e-mail: nazrina@uum.edu.my

As a viral disease, treatment of antibiotics will be ineffective. The spread of the virus can be easily transmitted through the secretion of saliva or mucus from coughing or sneezing by an infected individual. To slow down and prevent the spread of COVID-19, it is important to have informative knowledge of the disease and also practice preventive measures such as handwashing, consistently wearing masks, social distancing, and getting vaccinated (WHO, 2021a).

In 2021, COVID-19 is a global pandemic, and the number of positive cases is still increasing rapidly in Malaysia. As of 31 December 2021, 2.7 million confirmed cases of COVID-19 have been recorded nationwide, including over 31,000 deaths (Ministry of Health (MOH) Malaysia and COVID-19 Immunisation Task Force (CITF), 2022). According to statements from MOH Malaysia (2020), the common approach to evaluate the COVID-19 risk areas is based on count data which includes the daily cases number and cumulative cases number of the past 14 days. This method is not appropriate as it does not reflect the real situation and only focuses on the occurrence of the diseases without considering the areas where the disease is transmitted.

To capture the situation of each state in Malaysia, disease mapping is proposed in this research. According to Nieto-Barajas (2008), disease mapping is an appropriate method for describing and quantifying the spatial variation in disease risk, as well as identifying the possible associations between adjacent areas within a specific geographical region. Accurate estimation of the relevant relative risk is important when developing the disease map.

In estimating the relative risk, the most commonly used statistical method is the Standardized Morbidity Ratio (SMR). Nevertheless, it has drawbacks, such as extreme results that might occur by having a large SMR value if the expected case number is small and vice versa, or the SMR value being equal to zero when no cases are observed. To deal with the disadvantages of the SMR approach, the relative risk is measured using the Poisson-gamma model in disease mapping, particularly when no cases are observed in the area. When estimating relative risk, it can yield smoother disease maps that have fewer extreme values and only produce positive estimates (Lawson *et al.*, 2003; Samat & Ma'arof, 2013). Furthermore, it is highly suggested to use the Poisson-gamma model in estimating small areas given that it can provide prior information on the uncertainty measures that are associated with the relative risk estimation in the entire map (Diah *et al.*, 2016; Diah & Aziz, 2021). This study has applied the Poisson-gamma model to the relevant COVID-19 incidence data to accurately estimate the relative risk values and subsequently visualise the risk level of the affected states in Malaysia through disease mapping.

II. MATERIALS AND METHOD

In this study, two tools are applied to conduct the data analysis. WinBUGS software is utilised to compute the relative risk of COVID-19, while ArcGIS software is employed to map estimated relative risk in Malaysia by state. WinBUGS software is a designated programme that uses the computations of Markov Chain Monte Carlo (MCMC) to conduct Bayesian inference for statistical issues (Lawson *et al.*, 2003). Meanwhile, ArcGIS is a software for data storage and mapping where the user is allowed to generate and present the data analysis. Overall results are shown in table and figures, whereas the COVID-19 risk maps are generated using the computed data.

A. Poisson-gamma Model

In view of the SMR method's weaknesses, the Poisson-gamma model as one of the primarily used Bayesian approaches has been suggested by several researchers (Lawson *et al.*, 2003; Yensy, 2021). Poisson distribution is adopted in this research since it provides the fundamental framework for counting data.

It is assumed in this model that the number of new infections, y_{ij} follows a Poisson distribution over a particular period with mean and variance, $e_{ij}\theta_{ij}$. Here, $i = 1,2,\dots,M$ represent the areas of study while $j = 1,2,\dots,T$ represent the period. The expected new infectives number is expressed as e_{ij}, while the relative risk is expressed as θ_{ij}:

$$y_{ij} \mid e_{ij},\, \theta_{ij} \sim Poisson\,(e_{ij}\theta_{ij}) \tag{1}$$

The relative risk parameter follows a gamma prior distribution, with parameters, α and β:

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

According to Lawson *et al.* (2003), the prior parameter values, α and β, of the Poisson-gamma model are unknown and the hyperparameter values are assumed to be 0.1 in the exponential prior distributions. In this study, the model-based prior predicted relative risk is equal to one. The posterior predicted relative risk for each state and for each period of time will be the data analysis output.

B. Data Set

The data sets collected in this study are obtained from an open data community in partnership with the MOH Malaysia and CITF. As to collate the latest necessary data, the official website was made accessible in early 2022. They include the COVID-19 cases instances in Malaysia by 13 states, namely Perak, Penang, Kedah, Perlis, Negeri Sembilan, Malacca, Johor, Pahang, Kelantan, Terengganu, Selangor, Sabah and Sarawak, and also three federal territories: Kuala Lumpur, Putrajaya and Labuan. For simplicity, the states and federal territories are both referred to as states in this research. The Poisson-gamma model is applied to analyse the data of COVID-19 in the form of count cases across 16 Malaysian states from January 2020 to December 2021.

III. RESULT AND DISCUSSION

A. Relative Risk Estimation of COVID-19

According to Samat and Percy (2012), a relative risk value close to 1 implies that individuals in a particular area are not significantly different from the entire population in their probability of contracting the disease. A relative risk value below 1 signifies that the people in the area are less likely to contract the disease than the overall population. On the other hand, an area with a relative risk value larger than 1 indicates that the population there has a higher chance of contracting the disease than the general population. The conditional likelihood of an individual within a specific area getting COVID-19 infection divided by the conditional likelihood of an individual getting the infection in the general population is the relative risk in this data analysis.

Figure 1 displays the estimated relative risk of COVID-19 in each of the Malaysian states in 2020 and 2021. From Figure 1, it can be seen that the relative risks of most states in 2020 are less than 1, which suggests that the tendency of people in most states to be infected with COVID-19 is lower than that of the whole population in Malaysia. In contrast, the relative risks of Negeri Sembilan, Kuala Lumpur, Sabah, and Labuan are more than 1. The values indicate that the people in these states are more susceptible to COVID-19 than the general population in Malaysia. Meanwhile, Selangor is the only state where the relative risk is nearly equal to 1, indicating that the difference in the likelihood of people contracting COVID-19 within this state and the entire Malaysian population is insignificant. Similarly, Figure 1 has also shown that in 2021, more than half of the states have a relative risk value of less than 1, which implies that people in most states are less prone to COVID-19 infection compared to the entire Malaysian population. On the other hand, states with relative risks close to 1 include Penang, Negeri Sembilan, Kelantan, Selangor, Kuala Lumpur, Sarawak, and Labuan. The estimation shows little difference in the likelihood of being infected with COVID-19 between people within these states and the general population in Malaysia.

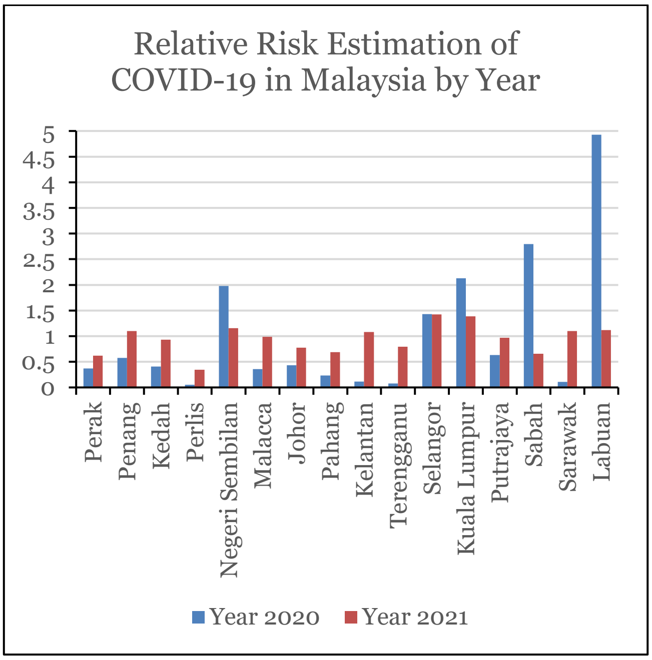

Figure 1. Plots of estimated COVID-19 relative risk using the Poisson-gamma model for 16 Malaysian states for the year 2020 and 2021

The numerical relative risk estimates using the Poisson-gamma model for the years 2020 and 2021 are shown in Table 1. It is observed that Perlis has the smallest relative risk in 2020 and 2021 (0.050 and 0.343, respectively). This finding might be because it is the least populated state, and its population density is in a lower range than other states (Pang *et al.*, 2021). Another possible reason for Perlis having the lowest COVID-19 risk in both years is the limited movement of people across the state. This is due to the fact that the international border remains closed (Flanders Investment and Trade, 2021) and its sole neighbouring state is Kedah. In 2020, Labuan had the greatest risk value of contracting COVID-19, with a relative risk of 4.930. This is because of its relatively high population density among all states, despite its small population (Pang *et al.*, 2021). As can be seen from Table 1, Selangor has recorded the highest COVID-19 relative risk estimate (1.421) in 2021. This result may be attributed to it being a densely populated state with urban areas (Ganasegeran *et al.*, 2021).

Table 1. Relative risk estimation of COVID-19 based on the Poisson-gamma model for the year 2020 and 2021

State	**Relative Risk by Year**	
	2020	**2021**
Perak	0.371	0.619
Penang	0.576	1.098
Kedah	0.405	0.931
Perlis	0.050	0.343
Negeri Sembilan	1.980	1.156
Malacca	0.358	0.988
Johor	0.435	0.778
Pahang	0.236	0.689
Kelantan	0.114	1.078
Terengganu	0.075	0.797
Selangor	1.431	1.421
Kuala Lumpur	2.131	1.384
Putrajaya	0.634	0.967
Sabah	2.798	0.656
Sarawak	0.106	1.100
Labuan	4.930	1.121

B. COVID-19 Disease Mapping in Malaysia

The computed results, which are presented in Table 1, are illustrated on maps to indicate areas of high-low risk of COVID-19 infection. These maps enable one to easily ascertain which areas are at low and high risk of being affected and consequently help with resource allocation and policy development based on the obtained information. According to Diah, Aziz and Kasim (2017), the classification of COVID-19 risk levels consists of five categories: very low, low, medium, high and very high, with intervals of [0.0, 0.5), [0.5, 1.0), [1.0, 1.5), [1.5, 2.0) and [2.0,∞) respectively. Using different shades in the mapping, the risk levels within the 16 states in Malaysia are differentiated, where the most light-coloured indicates a very low-risk area while the most dark-coloured indicates a very high-risk area.

Figure 2 depicts the disease map for relative risk of COVID-19 in 2020 using the Poisson-gamma model. From Figure 2, Kuala Lumpur, Sabah, and Labuan were classified as very high-risk areas in 2020. These states are followed by Negeri Sembilan and Selangor with high and medium risk respectively. The risks of the remaining states are very low except for Penang and Putrajaya with low risk.

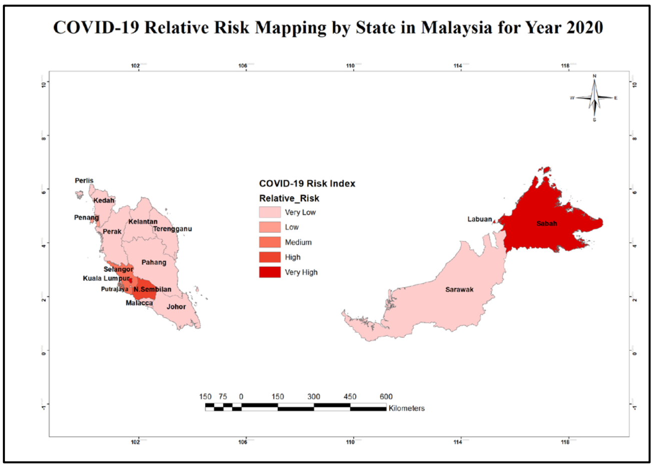

Figure 2. Disease map for COVID-19 relative risk based on the Poisson-gamma model for the year 2020

Figure 3 displays the COVID-19 risk map based on the Poisson-gamma model in 2021. This map shows that the states of Selangor, Kuala Lumpur, Negeri Sembilan, Labuan, Sarawak, Penang, and Kelantan are at medium risk. The rest of the states are rated as low risk except for Perlis, which has a very low risk.

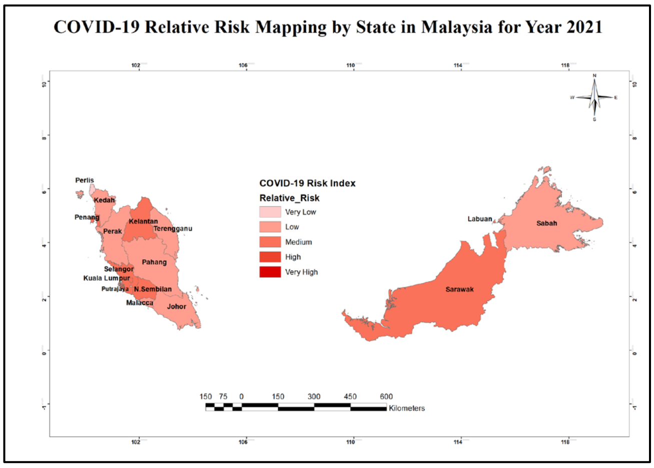

Figure 3. Disease map for COVID-19 relative risk based on the Poisson-gamma model for the year 2021

IV. CONCLUSION

In this research, the Poisson-gamma model is suggested for estimating the relative risk of COVID-19, which can detect the risk in small areas and pinpoint the high-low-risk areas by considering the population density of the area. It can also deal with the problem of the SMR method, especially when COVID-19 cases are not observed in certain areas. In addition to tables and graphs, this research's findings are also displayed as maps.

In conclusion, the areas with the highest risk in 2020 and 2021 are Labuan and Selangor, respectively, whereas Perlis is recognised as the lowest risk area in these two years. It is worth noting that in 2021, no state is considered a very high-risk or high-risk area. The gradual control of the situation is likely to be attributed to the effectiveness of government mitigation measures, such as the phased implementation of Movement Control Orders and vaccination programmes. These findings of the study might be beneficial to the Malaysian government, as they can use the disease mapping information to identify locations that require monitoring and urgent treatment, particularly in terms of health resource planning, and indicate the significance of the risks as well.

It is expected that the Poisson-gamma model utilised in this research will yield a more accurate estimate of relative risk than the Standardized Morbidity Ratio (SMR), a conventional disease mapping methodology. However, this modelling has limitations, including difficulty adjusting covariates and the inability to handle spatial correlations. Therefore, further studies of other methods in estimating more accurate relative risk values are encouraged.

V. ACKNOWLEDGEMENT

The authors are grateful to MOH Malaysia and CITF for their partnership with the open data community and for providing the essential data for this research.

VI. REFERENCES

Diah, IM & Aziz, N 2021, 'Mapping of pneumonia disease in Malaysia using Poisson-gamma model', Annals of the Romanian Society for Cell Biology, vol. 25, no. 1, pp. 2062–2067.

Diah, IM, Aziz, N & Ahmad, N 2016, 'Tuberculosis disease mapping with Poisson-gamma model in Malaysia', Research Journal of Applied Sciences, vol. 11, no. 9, pp. 822–825. doi: 10.3923/rjasci.2016.822.825.

Diah, IM, Aziz, N & Kasim, MM 2017, 'A comparison of four disease mapping techniques as applied to TB diseases in Malaysia', Journal of Telecommunication, Electronic and Computer Engineering, vol. 9, no. 2–11, pp. 133–137.

Elibol, E 2021, 'Otolaryngological symptoms in COVID-19', European Archives of Oto-Rhino-Laryngology, vol. 278, no. 4, pp. 1233–1236. doi: 10.1007/s00405-020-06319-7.

Flanders Investment & Trade 2021, CORONA VIRUS – The situation in Malaysia, viewed 27 December 2021, <https://www.flandersinvestmentandtrade.com/export/nieuw/corona-virus-—-situation-malaysia>.

Ganasegeran, K, Jamil, MFA, Ch'ng, ASH, Irene, L & Peariasamy, KM 2021, 'Influence of population density for covid-19 spread in Malaysia: An ecological study', International Journal of Environmental Research and Public Health, vol. 18, no. 18, p. 9866. doi: 10.3390/ijerph18189866.

Kuppalli, K & Madad, S 2021, 'Emerging infectious diseases during COVID-19', Contagion, vol. 6, no. 1, pp. 18–19, viewed 13 June 2021, <https://cdn.sanity.io/files/0vv8moc6/contagion/93f2d33b34e28089b21b2d387ce8e5f79f4cd060.pdf>.

Lawson, AB, Browne, WJ & Vidal Rodeiro, CL 2003, Disease Mapping with WinBUGS and MLwiN, Disease Mapping with WinBUGS and MLwiN. doi: 10.1002/0470856068.

Liu, YC, Kuo, RL & Shih, SR 2020, 'COVID-19: The first documented coronavirus pandemic in history', Biomedical Journal, vol. 43, no. 4, pp. 328–333. doi: 10.1016/j.bj.2020.04.007.

Ministry of Health Malaysia 2020, Risk assessment and determination of COVID-19 infection and transmission for an area, viewed 6 June 2021, <http://covid-19.moh.gov.my/sorotan/112020/penilaian-dan-penentuan-risiko-risk-assessment-jangkitan-dan-penularan-covid-19-bagi-sesuatu-kawasan>.

Ministry of Health Malaysia & COVID-19 Immunisation Task Force 2022, COVIDNOW in Malaysia, viewed 1 January 2022, <https://covidnow.moh.gov.my/>.

Nieto-Barajas, LE 2008, 'A Markov gamma random field for modelling disease mapping data', Statistical Modelling, vol. 8, no. 1, pp. 97–114. doi: 10.1177/1471082X0700800107.

Pang, NTP, Kamu, A, Mohd Kassim, MA & Ho, CM 2021, 'Monitoring the impact of Movement Control Order (MCO) in flattening the cummulative daily cases curve of Covid-19 in Malaysia: A generalized logistic growth modeling approach', Infectious Disease Modelling, vol. 6, pp. 898–908. doi: 10.1016/j.idm.2021.07.004.

Samat, NA & Ma'arof, SHMI 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, vol. 7, no. 8, pp. 1299–1303. doi: doi.org/10.5281/zenodo.1086703.

Samat, NA & Percy, DF 2012, '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', pp. 623–627. doi: 10.1109/ICSSBE.2012.6396640.

World Health Organization 2021a, Coronavirus: Overview, viewed 27 November 2021, <https://www.who.int/health-topics/coronavirus#tab=tab_1>.

World Health Organization 2021b, Coronavirus: Symptoms, viewed 1 June 2021, <https://www.who.int/health-topics/coronavirus#tab=tab_3>.

Yensy, NA 2021, 'The Small Area Estimation by Using Empirical Bayes Method', Proceedings of the International Conference on Educational Sciences and Teacher Profession (ICETeP 2020), vol. 532, pp. 343-347. doi: 10.2991/assehr.k.210227.058.

Zhang, YZ & Holmes, EC 2020, 'A Genomic Perspective on the Origin and Emergence of SARS-CoV-2', Cell, vol. 181, no. 2, pp. 223–227. doi: 10.1016/j.cell.2020.03.035.