Leptospirosis disease mapping with standardized morbidity ratio and Poisson-Gamma model: An analysis of Leptospirosis disease in Kelantan, Malaysia

Aznida Che Awang and Nor Azah Samat
Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, 35900 Tanjung Malim, Perak, Malaysia.
E-mail:nieda_wanimira@yahoo.com, norazah@fsmt.upsi.edu.my

Abstract. Leptospirosis is a disease caused by the infection of pathogenic species from the genus of Leptospira. Human can be infected by the leptospirosis from direct or indirect exposure to the urine of infected animals. The excretion of urine from the animal host that carries pathogenic Leptospira causes the soil or water to be contaminated. Therefore, people can become infected when they are exposed to contaminated soil and water by cut on the skin as well as open wound. It also can enter the human body by mucous membrane such nose, eyes and mouth, for example by splashing contaminated water or urine into the eyes or swallowing contaminated water or food. Currently, there is no vaccine available for the prevention or treatment of leptospirosis disease but this disease can be treated if it is diagnosed early to avoid any complication. The disease risk mapping is important in a way to control and prevention of disease. Using a good choice of statistical model will produce a good disease risk map. Therefore, the aim of this study is to estimate the relative risk for leptospirosis disease based initially on the most common statistic used in disease mapping called Standardized Morbidity Ratio (SMR) and Poisson-gamma model. This paper begins by providing a review of the SMR method and Poisson-gamma model, which we then applied to leptospirosis data of Kelantan, Malaysia. Both results are displayed and compared using graph, tables and maps. The result shows that the second method Poisson-gamma model produces better relative risk estimates compared to the SMR method. This is because the Poisson-gamma model can overcome the drawback of SMR where the relative risk will become zero when there is no observed leptospirosis case in certain regions. However, the Poisson-gamma model also faced problems where the covariate adjustment for this model is difficult and no possibility for allowing spatial correlation between risks in neighbouring areas. The problems of this model have motivated many researchers to introduce other alternative methods for estimating the risk.

1. Introduction
The most common methods used in the study of disease mapping are Standardized Morbidity Ratio (SMR) and Poisson-gamma model and we have applied these two methods to map leptospirosis data in Kelantan, Malaysia. Firstly, this paper discusses the definition and the drawback of classical methods in estimating relative risk using Standardized Morbidity Ratio. Next, the second method that has been discussed was the earliest application of Bayesian methodology called Poisson-gamma model. Lastly, these two methods are applied to leptospirosis data of Kelantan in Malaysia for the purpose of displaying and to discover a better method for estimating leptospirosis risk.
Leptospirosis is an infectious disease caused by a type of pathogenic species from the genus of Leptospira. This disease can be transmitted directly or indirectly to human by rats. Human can be infected when they are exposed to contaminated soil and water that contain pathogenic Leptospira through cut on skin such as open wound. In addition, it also can enter human body by mucous membrane such nose, eyes and mouth, for example by splashing contaminated water or urine into the eyes or swallowing contaminated water or food [1]. The symptoms for this disease include sudden fever of 39°C and above, sudden headache, vomiting and nausea, kidney or liver problem, anemia and rash [2]. According to [2], leptospirosis in Malaysia is reemerging disease because of the high humidity and warm temperatures allowing Leptospira to survive for a long term in the environment.

1.1. Leptospirosis Scenario in Malaysia
In Malaysia, the number of leptospirosis cases reported is high and keep increasing every year. Ministry of Health Malaysia (2016) recorded that there are about 8291 leptospirosis cases reported in 2015 as shown in Figure 1. The state of Kelantan has the highest number of cases reported with 1956 cases, then followed by the state of Selangor with 1233 cases. The other 13 states recorded between 16 to 800 cases in 2015. This serious predicament should not be underestimated. Therefore, the study of leptospirosis is important for the country to reduce the possible risk of the disease and also to control its occurrences.

![Figure 1. Number of leptospirosis cases reported for the 15 states in Malaysia in 2015.](image)

Currently, the study of risk maps for leptospirosis in Malaysia is rare. In Malaysia, the current methods used to estimate the high-low risk areas of leptospirosis are based on the total numbers of leptospirosis cases in each region that have been reported. However, this method can only show the numbers of leptospirosis reported cases without considering important features such as geographical area, number of population and other parameters. This is important so that the areas can be accurately identified as high or low-risk areas for the disease occurrences.

Disease mapping used the total number of individuals that were infected by the disease within small areas and also can be defined as the estimation and presentation of summary measures of health outcomes. Disease mapping is classified as an important tool in controlling and reducing the number of leptospirosis occurrences. However, good disease mapping relies on the accurate value of relative risk estimation [3]. Currently, the study of relative risk estimation for disease mapping is still ongoing because of the important of investigating geographical distribution. In this study of leptospirosis, SMR and Poisson-gamma model have been used as conventional ways to estimate relative risk.
2. Standardized Morbidity Ratio
SMR is a traditional method used in disease mapping to estimate the relative risk. SMR is defined as Standardized Morbidity Ratio where morbidity is referring to the incidence of the disease and this method basically compares the observed cases with expected cases. Consequently, the SMR is defined as

\[ SMR_i = \hat{\Theta}_i = \frac{O_i}{E_i} \]  

where \( O_i \) represented the observed number of cases and \( E_i \) represented the expected number of cases.

In disease mapping, supposed that the study region to be mapped is divided into \( R \) mutually exclusive regions \((i=1,2,\ldots,R)\). Each region has its own observed cases \( O_i \) where this value could be found from resources such as from the Ministry of Health Malaysia and expected number of cases \( E_i \) could be calculated by using a particular formula as below:

\[ E_i = N_i \frac{\sum O_j}{\sum N_j} \]  

where \( N_i \) is the population of state \( i \) and time \( j \) and the summations (\( \Sigma \)) are for \( i=1,2,\ldots,R \) and \( j=1,2,\ldots,T \). Here, the standardization is completed by the total population at risk, assuming everybody is equal at risk. The formula above can be interpreted as the probability that a person within a specified region contracts the disease divided by the probability that a person in the population contracts the disease.

Even though the SMR is commonly used to estimate relative risk, it has several drawbacks. According to [4], the mean and variance of SMR depend on \( E_i \) since SMR is based on a ratio estimator. This can cause problems where the SMR will become very large if the expected numbers of cases are small, while if the expected number of cases is large, the SMR will become small and the SMR will automatically zero in areas where there are no cases or count data.

3. Poisson-Gamma Model
Since the SMR have many drawbacks, many researchers have investigated using a different method in estimating the relative risk. This is including the use of earliest example of Bayesian method that is called Poisson-gamma model [5]. In Poisson-gamma model, the study regions are represented by \( i=1,2,\ldots,R \) and time periods are represented by \( j=1,2,\ldots,T \). The number of new infective \( y_{ij} \) is assumed to follow a Poisson distribution within a given time period, with mean and variance \( e_{ij} \hat{\Theta}_{ij} \) where \( e_{ij} \) is the expected number of a new infective and \( \hat{\Theta}_{ij} \) is the relative risk:

\[ y_{ij} \mid e_{ij}, \hat{\Theta}_{ij} \sim \text{Poisson}(e_{ij} \hat{\Theta}_{ij}) \]  

The relative risk parameter has a gamma prior distribution with parameter \( \alpha \) and \( \beta \):

\[ \hat{\Theta}_{ij} \sim \text{Gamma}(\alpha, \beta) \]  

The output of the analysis includes the posterior expected relative risk for all regions and time periods, based on this Poisson-gamma model.
4. Application of Standardized Morbidity Ratio and Poisson-Gamma Model to Leptospirosis Disease Mapping

This section demonstrates and displays the results of the applications of relative risk estimation method based on the Standardized Morbidity Ratio (SMR) and Poisson-Gamma Model using observed leptospirosis data in Kelantan, Malaysia. The data set were analyzed using Win BUGS software, which is a package designed to carry out wide variety of Bayesian models. Based on the relative risk results, map of the leptospirosis risk for the state of Kelantan, Malaysia is constructed.

4.1 The Data Set.
Data used in this analysis were provided by the Ministry of Health (MOH) Malaysia and Kelantan State Department of Statistics Malaysia. These SMR method and Poisson-gamma model were applied to leptospirosis data in the form of count of cases within ten districts in Kelantan which are Bachok, Gua Musang, Jeli, Kota Bharu, Kuala Krai, Machang, Pasir Mas, Pasir Puteh, Tanah Merah and Tumpat. In this analysis, leptospirosis data referred to the observed infective leptospirosis cases of humans, which were observed monthly from January to December in 2015.

4.2 The Results.
The outcomes of relative risk estimation for SMR model and Poisson-Gamma model in ten districts of Kelantan, Malaysia are displayed in Figure 2, Figure 3, Table 1, Figure 4 and Figure 5.

Figure 2 shows the time series plot of the SMR value for ten districts of Kelantan during 2015. It can be observed that the district of Jeli has the highest number of relative risk in August with relative risk value equal to 7.67. This is followed by the district of Machang with relative risk value equals to 3.22, for the month of November. The other 8 districts recorded relative risk of less than 3 for all months investigated. The graph shows that most of the states have relative risk less than 1 for most of the months. Based on the definition of relative risk explained in [6], a relative risk less than one indicates that susceptible people within this district is generally less likely to catch leptospirosis compared to people in the overall population in Kelantan, Malaysia.

![Figure 2](image-url)

**Figure 2.** Time Series Plot for the Estimated Relative Risk of Leptospirosis based on the SMR method for 10 districts in Kelantan, Malaysia.

Figure 3 shows the time series plot of the estimated relative risk of leptospirosis using Poisson-gamma model for the ten districts of Kelantan in 2015. Similarly, it can be seen from Figure 3 that most districts also have relative risk less than one for most of the months and this leads to the same conclusion as above. Figure 3 also demonstrates that district of Jeli has the highest number of relative risk value in August which is 5.74. These are followed by the districts of Gua Musang and Machang with relative risk value of 2.70 in January and November 2015 respectively. The other seven districts recorded relative risk of less than 2.50 for all months.
Figure 3. Time Series Plot for the Estimated Relative Risk of Leptospirosis based on Poisson-gamma model for 10 districts in Kelantan, Malaysia.

Both graphs display that there are five districts recorded relative risk larger than one for most of the month which are the districts of Gua Musang, Jeli, Machang, Pasir Mas and Tanah Merah. The relative risk larger than one means that susceptible people within these districts are more likely to contract leptospirosis compared with people in the overall population of Kelantan.

Table 1. Comparison of the Relative Risk Estimation Based on SMR Method and Poisson-Gamma Model in November 2015.

| Districts       | Relative Risk Estimation based on SMR Method | Relative Risk Estimation based on Poisson-Gamma Model |
|-----------------|---------------------------------------------|------------------------------------------------------|
| Bachok          | 0.47                                        | 0.65                                                 |
| Gua Musang      | 3.05                                        | 2.58                                                 |
| Jeli            | 2.28                                        | 1.67                                                 |
| Kota Bharu      | 0.64                                        | 0.68                                                 |
| Kuala Krai      | 0.57                                        | 0.73                                                 |
| Machang         | 3.22                                        | 2.70                                                 |
| Pasir Mas       | 1.16                                        | 1.17                                                 |
| Pasir Puteh     | 0                                           | 0.24                                                 |
| Tanah Merah     | 0.52                                        | 0.65                                                 |
| Tumpat          | 0.92                                        | 0.95                                                 |
| **Range**       | **3.22**                                    | **2.46**                                             |
| **Mean**        | **1.28**                                    | **1.20**                                             |
| **Standard Deviation** | **1.09**           | **0.80**                                             |

Table 1 shows the numerical values of the relative risk based on SMR method and Poisson-gamma model for the month of November 2015. This time period is chosen as an example for demonstration purposes only. From Table 1, it can be seen that by using SMR model, susceptible people within the district of Machang has the highest risk of contracting leptospirosis, while susceptible people within the district of Pasir Puteh has the lowest risk of contracting leptospirosis compared with people in the overall population of Kelantan. The corresponding values of relative risk are approximately 3.22 and 0, respectively.

Estimating relative risk using Poisson-gamma model also displays the susceptible people within the district of Machang has the highest risk of contracting leptospirosis with relative risk value of 2.70. Meanwhile susceptible people within Pasir Puteh district has the lowest risk of contracting leptospirosis compared with people in the overall population of Kelantan with relative risk of 0.24. Besides that,
Table 1 also shows that the relative risk for the district of Pasir Puteh is zero when there is no observed leptospirosis case in that area. This situation displays the drawback of SMR method where the relative risks become zero when there are no observed cases in certain regions. However, the Poisson-gamma model can overcome this drawback since it can generate positive estimates of relative risk in districts that have no observed cases such as Pasir Puteh. In addition, it can be seen clearly from Table 1 that the range, mean and standard deviation for relative risk based on SMR is higher than the range, mean and standard deviation for relative risk based on Poisson-gamma model. This shows that relative risk estimation based on Poisson-Gamma model produced a smoother map with less extreme values for relative risk estimated compared to the use of SMR.

In order to show a clearer picture of the risk area, choropleth map with single-hue progression colours is used to display and differentiate between high and low risk areas of leptospirosis cases occurrence for each districts in Kelantan, Malaysia. Figure 4 and Figure 5 show the risk maps for leptospirosis based on the SMR method and Poisson-gamma model for ten districts of Kelantan, Malaysia, respectively. The maps demonstrate the results for the month of November in 2015. The relative risk is categorized into five different levels which are very low, low, medium, high and very high with respective intervals of [0,0.5), [0.5,1.0), [1.0,1.5), [1.5,2.0) and [2.0,∞), respectively. The darkest shade in the map represents very high risk area while the lightest shade represents very low risk area.

The SMR map in Figure 4 shows that the districts that have very high risk are Gua Musang, Jeli and Machang. There is no district that has high risk. Next, the district with medium risk is Pasir Mas followed by the district of Kota Bharu, Kuala Krai, Tanah Merah and Tumpat with low risk of leptospirosis. Lastly, the districts with very low risk are Bachok and Pasir Puteh. The Poisson-gamma model in Figure 5 demonstrates that the districts with very high risk are Gua Musang and Machang. This is followed by the district of Jeli with high risk and district of Pasir Mas with medium risk of leptospirosis occurrence. While the district of Bachok, Kota Bharu, Kuala Krai, Tanah Merah and Tumpat have low risk. Lastly, the district with very low risk is Pasir Puteh. Based on the comparison using these two maps, there is no obvious difference between estimated relative risks using these two methods. However, these disease maps can be utilized to visualize and identify the districts with high and low risk of leptospirosis occurrences, hence further action could be made to this priority regions.
5. Conclusion

It is important to estimate the relative risk of the spread of leptospirosis, so that its risk map can be constructed. This relative risk was estimated using the most common statistic method which is SMR method and the earliest application of Bayesian method called Poisson-gamma model. The disease map of estimated relative risks for leptospirosis based on the SMR Method and Poisson-gamma model will show a clear visual impression of disease’s incidence in regions of interest. By referring to the relative risk maps, the states that deserve serious attention and disease prevention can be identified. The Poisson-gamma model has been demonstrated as a better model in estimating relative risk compared to SMR method because it can overcome the drawback of SMR. For example the SMR will automatically zeroed in areas where there are no cases or count data. In addition, the SMR will become very large if the expected numbers of cases are small, while if the expected numbers of cases are large, the SMR will become small. However, the Poisson-gamma model also has some problems where covariate adjustment in this model is difficult and there is no possibility for allowing spatial correlation between risks in adjacent areas. These disadvantages have motivated many researchers to propose other alternative methods for estimating the risk.

References

[1] Pongsumpun P 2012 *J. World Acad. Sci., Eng. and Technol.* **72** 266-271

[2] Benacer D, Thong K L, Verasahib K, Galloway R L, Hartskeerl R A, Lewis J W and Mohd Zain S N 2016 *Asia Pacific J. Publ. Health* **28** 290-302

[3] Diah I M 2016 *Res. J. Pure and Appl. Math*. **11** 4011-4019

[4] Meza J L 2003 *J. Stat. Planning and Inference* **112** 43-62

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

[6] Samat N A and Iman S H 2013 *Internat. J. Math. Comput. Sci. and Eng.* **7** 46-50