Malaria Disease Mapping in Malaysia based on Besag-York-Mollie (BYM) Model

Nor Azah Samat and Liew Wan Mey

Department of Mathematics, Faculty of Science and Mathematics, Universiti Pendidikan Sultan Idris, 35900 Tanjong Malim, Perak, MALAYSIA.

E-mail: norazah@fsmt.upsi.edu.my, wanmey8859@gmail.com

Abstract. Disease mapping is the visual representation of the geographical distribution which give an overview info about the incidence of disease within a population through spatial epidemiology data. Based on the result of map, it helps in monitoring and planning resource needs at all levels of health care and designing appropriate interventions, tailored towards areas that deserve closer scrutiny or communities that lead to further investigations to identify important risk factors. Therefore, the choice of statistical model used for relative risk estimation is important because production of disease risk map relies on the model used. This paper proposes Besag-York-Mollie (BYM) model to estimate the relative risk for Malaria in Malaysia. The analysis involved using the number of Malaria cases that obtained from the Ministry of Health Malaysia. The outcomes of analysis are displayed through graph and map, including Malaria disease risk map that constructed according to the estimation of relative risk. The distribution of high and low risk areas of Malaria disease occurrences for all states in Malaysia can be identified in the risk map.

1. Introduction

This paper was motivated by a yearly report of the Malaria data collected by the Ministry of Health Malaysia. Seeking a proper model for disease mapping is the aim of this paper. Therefore, Besag-York-Mollie (BYM) model is applied to Malaria data of year 2013 in Malaysia and the estimation of relative risk are used for constructing Malaria maps. Disease mapping defined as a method for illustrating the spatial distribution of disease occurrence for the formulation of aetiological hypotheses, identify areas of unusually high risk so that action may be taken and provide a 'clear' map of disease risk in a defined area of interest to allow better resource allocation and risk assessment. The paper begins to briefly describe the methodology in estimating relative risk using BYM model. This is followed by illustrate the application and the result of the study. Lastly, the discussion and conclusion are drawn.

2. Malaria Disease and Its Situation in Malaysia

According to the World Malaria Report 2014, Malaria transmission occurs in all six WHO regions. Globally, an estimated 3.2 billion people in 97 countries and territories are at risk of being infected with Malaria and developing disease, and 1.2 billion are at high risk (>1 in 1000 chance of getting Malaria in a year). According to the latest estimates, 198 million cases of Malaria occurred globally in 2013 and the disease led to 584 000 deaths, representing a decrease in Malaria case incidence and mortality rates of 30% and 47% since 2000, respectively.
The data used in this paper provided by the Ministry of Health Malaysia which are the observed count Malaria data for 16 states in Malaysia from epidemiology week 1 to epidemiology week 52 in the year of 2013. Figure 1 shows that in 2013, Sabah had the highest total of Malaria cases, 1606 cases, followed by Sarawak with 1004 cases. Putrajaya and Labuan are the only two states had zero case.

![Figure 1. Total Number of Malaria Cases in All States in Malaysia.](image)

3. Besag-York-Mollie (BYM) Model
Clayton and Kaldor (1987) were the first attempt to introduce empirical Bayesian inference for relative risks and developed to a fully Bayesian setting by Besag et al. (1991) which are commonly known as the BYM model. Nowadays, most of the applications still depend on the BYM model which had been considered as the only fully Bayesian spatial model that used in those published applications of disease mapping outside of the statistical literature. These shown that the BYM model is one of the most popular hierarchical Bayesian models.

For relative risks, BYM model incorporates random effects due to unstructured (correlated heterogeneity) and spatially structured heterogeneity (uncorrelated heterogeneity) into the log-linear model. The inclusion of these random effects allows smoothing relative risks at the levels of global and local. For Malaria disease, the observed number of Malaria cases ($y_i$) in area $i$ is assumed to follow a Poisson distribution with mean $e_i \cdot \theta_i$:

$$y_i \sim \text{Poisson}(e_i \cdot \theta_i),$$

where $e_i$ denotes the expected number of cases in the $i^{th}$ geographic unit, and $\theta_i$ is the "true" but unknown relative risk in area $i$ to be estimated. At the next stage of the model, the variability of the log relative risk $\log \theta_i$, is split into three components:

$$\log \theta_i = \alpha + u_i + v_i,$$

where $\alpha$ is an overall level of the relative risk, spatial random effect $u_i$ reflecting the correlated heterogeneity and random effect $v_i$ representing the uncorrelated heterogeneity.
Bayesian modeling needs specification of prior distributions for random effects. The distribution model for the uncorrelated heterogeneity, $v_i$, does not depend on geographic location and is assumed to follow a normal distribution with zero mean and a common variance (precision parameter) $\tau_v^2$:

$$v_i \sim N(0, \tau_v^2).$$

For the clustering component, a spatial correlation structure is used, where estimation of the risk in any area depends on neighbouring areas. The latter is modelled via conditional autoregressive (CAR) model; a value of a parameter in one area is influenced by the average value of its neighbouring areas, with additional variability quantified by a conditional variance depending on the number of neighbours.

$$[u_i | u_j, i \neq j, \tau_u^2] \sim N(\bar{u_i}, \tau_u^2)$$

The mean of the areas bordering area $i$,

$$\bar{u_i} = \frac{1}{\sum j \omega_{ij}} \sum_j u_j \omega_{ij},$$

$$\tau_i^2 = \frac{\tau_u^2}{\sum_j \omega_{ij}},$$

Where the prior mean of each $u_i$ is defined as a weighted average of the other $u_j$, $i \neq j$. The $\omega_{ij}$ define the relationship between the area $i$ and $j$. The common first order binary weighting scheme is $\omega_{ij} = 1$ if $i, j$ are adjacent (or 0 if they are not). The precision parameters $\tau_v^2$ and $\tau_u^2$ control the amount of variability of random effects $v$ and $u$.

In a full Bayesian model analysis, prior distributions must be specified for those precision parameters $\tau_v^2$ and $\tau_u^2$. With no prior estimation for precisions of the random effects, distributions with large variance are recommended. These follow by consideration of gamma distributions (0.5, 0.0005) which yield a probability of 99% for both, as suggested by Bernardinelli et al. (1995b). This prior choice is less informative and allows the likelihood data to dominate the prior information; hence, it will have minimal effect on the inference of relative risks.

The BYM model was then applied on the Malaria disease data in Malaysia to estimate the relative risks. The model for each data set was fitted using WinBUGS version 1.4 that implements the Markov chain Monte Carlo (MCMC) methods. The weights must also be entered as data. The easiest way to define them is to create a loop in the WinBUGS code.

4. Application of Relative Risk Estimation for Malaria Disease

In this section, the result of relative risk estimation based on the application of Besag-York-Mollie (BYM) model which was applied on the observed number cases of Malaria disease transmission in Malaysia is displayed in the form of graphs, table and maps. Figure 2 depicts the time series plot of Malaria disease based on the number of cases for each state in Malaysia from epidemiology week 1 to epidemiology week 52. From the plot, Sabah almost had the
highest number of cases in every epidemiology week in year 2013. Among these 52 epidemiology week, most of the cases were above 10 for Sabah and Sarawak. As for Kelantan, there were sudden dramatically increases cases between epidemiology week 20 and 27 which even achieved the highest number of cases. However, the number of cases in Kelantan went down in epidemiology week 22 with 11 cases and increased again in epidemiology week 23 with 48 cases. While for rest of 13 states, most had less than 10 cases throughout the 52 epidemiology weeks in 2013.

Figure 2 demonstrates the time series plot for the estimated relative risk based on BYM model for Malaria disease in all states of Malaysia in year 2013. According to Figure 3, Sabah and Sarawak almost had relative risks above 2 and highest relative risk for epidemiology week 1 to epidemiology week 52. These indicated that people live in both of these states are more likely to be infected by Malaria disease compared to people in overall population. However, between epidemiology week 19 to epidemiology week 25, relative risk of Kelantan was fluctuates dramatically and even had the highest relative risk which was 8. Moreover, Pahang and Pulau Pinang had relative risks more than 1 but less than 2 for few of epidemiology weeks. These mean that there is no real different in terms of risk between people in these states compared to people in overall population. While, the other 11 states had relative risks below 1 which means that people in these states are less likely to contract with Malaria disease compared to people in overall population.

According to Figures 2 and 3, there was contrasting result for Sabah, for example, in epidemiology week 37 in Figure 2, Sabah reported to had the highest number cases which was 50. But in Figure 3, relative risk of Sabah is only 4.2570. This may be due to the number of population considered in the calculation of expected cases, where the population size for Sabah was 3,496,600 compared to Sabah which was 1,675,100.
5. Risk Map for Malaria Disease Mapping in Malaysia
In this section, the results of relative risk estimation based on the application of Besag-York-Mollie (BYM) model are used to construct the risk map. In this paper, the Malaria risk maps cluster and identify the regions with different levels of risks. There are two type of risk maps which been constructed.

For Malaria risk map which based on the number of cases, the five different levels of risks used to categorize the relative risk values were very low, low, medium, high and very high with the intervals of (<0.5, [0.5-1.0), [1.0-1.5), [1.5-2.0) and >2.0), respectively. However, there were also five different levels of risks were used to categorize the relative risk values for Malaria risk map that based on BYM model to estimate the relative risk, which were very low, low, medium, high and very high with the intervals of (<10, [10-19), [19-28), [28-37) and >37), respectively. The darkest shade in the map showing the area is in very high risk and the lightest shade region is considered as very low risk area.

Figure 4 represents the risk map for Malaria disease based on the number of cases for epidemiology week 29. From this figure, Sabah shown as a very high risk area, followed by Sarawak as a medium risk area and Kelantan as a low risk area. Other states were considered as very low risk areas.
Figure 4. Risk Map for Malaria based on Number of Cases for Epidemiology Week 29.

Figure 4 displays the risk map for Malaria disease based on BYM model for epidemiology week 29. This figure depicts that both states of Sarawak and Sabah were in very high risk for infective Malaria cases. The figure also showed Kelantan as a medium risk area, followed by Labuan and Pahang as a low risk area. Other states were considered as very low risk areas. From this map, interested parties can easily identified which states have a very high risk and need closer scrutiny or further attention. However, Figures 5 and 6 give different levels of risks result, this may due to the consideration of the total number of population in every states in the calculation of expected cases.

Figure 5. Risk Map for Malaria based on BYM Model for Epidemiology Week 29.
6. Conclusion and Future Works

Disease risk maps act as a useful tool for risk communication as it helps to disseminate information in a readily understandable format for the general public and fringe personnel. Based on the information, people in charge in city management or the public who stay in low-risk areas can take immediate action such as prevention and control strategy, rather than a reaction towards an occurred event of outbreak. People should always concern about early diagnosis and prompt treatment as it is a fundamental to Malaria control and need to be available wherever Malaria occurs. Naturally, the risk map does not coincide exactly with the observed number of Malaria cases. This could be due to model used for analysis as the production of disease maps relies on modelling to estimate and predict risks. In this paper, BYM model prove to be one of the methods that could be used to estimate the relative risk of Malaria disease. Information about the performance of different methods of modelling geographical patterns of health events in situations that include an important number of small areas show that BYM models can be easily applied in epidemiology and public health. Nevertheless, further improvement must be done to improve the current model. There are several methodological alternatives for generating estimates with Bayesian hierarchical models and increased accuracy of the risk map could be made for more detailed comparisons by including more confounding factors.

References

[1] Goovaerts, P and Gebreab, S (2008) *International Journal of Health Geographics*, 7:6. DOI: 10.1186/1476-072X-7-6.

[2] Hassan, H, Shohaimi, S, and Hashim, N R (2012) Risk mapping of dengue in Selangor and Kuala Lumpur, Malaysia. *Geospatial Health* 7 (1), pp. 21-25.

[3] Lawson, A B, and Williams, F L R (2001). *An Introductory Guide to Disease Mapping*. John Wiley and Sons, Ltd, Chichester, UK.

[4] Lawson, A B, Browne, W J and Rodeiro, C L V (2003) *Disease Mapping with WinBUGS and MlwiN*. John Wiley and Sons, LTD, Chichester, UK.

[5] Lawson, A B, et al. (2000) Statistics in Medicine, 19: 2217-2241.

[6] Lekdee, K, Sammatat, S, and Boonsit, N (2014) *International Journal of Medical, Health, Pharmaceutical and Biomedical Engineering* 8 (7).

[7] Ministry Of Health Malaysia. (2013). *Management Guidelines of Malaria In Malaysia*.

[8] Mohamad Imam Ma’arof, S H and Samat, N A (2014) *Journal of Science and Mathematics JSM (in press)*, ISSN 1985-7918 (Print).

[9] Mohamad Imam Ma’arof, S H and Samat, N A (2014) *Journal of Science and Mathematics JSM (in press)*, ISSN 1985-7918 (Print).

[10] Samat, N A and Percy, D F (2012) *Journal of Applied Statistics*, 39 (9), pp. 2029-2046. DOI: 10.1080/02664763.2012.700450.

[11] Spiegelhalter, D, et al. (2003) *WinBUGS User Manual Version 1.4*. MRC Biostatistics Unit, Cambridge, UK.