Sains Malaysiana 42(7)(2013): 1003–1010

Spatial Analysis of Infant Mortality in Peninsular Malaysia over Three Decades Using Mixture Models

(Analisis Reruang bagi Kematian Bayi di Semenanjung Malaysia dalam Tempoh Tiga Dekad Menggunakan Model Campuran)

NUZLINDA ABDUL RAHMAN* & ABDUL AZIZ JEMAIN

ABSTRACT

Infant mortality is one of the central public issues in most of the developing countries. In Malaysia, the infant mortality rates have improved at the national level over the last few decades. However, the issue concerned is whether the improvement is uniformly distributed throughout the country. The aim of this study was to investigate the geographical distribution of infant mortality in Peninsular Malaysia from the year 1970 to 2000 using a technique known as disease mapping. It is assumed that the random variable of infant mortality cases comes from Poisson distribution. Mixture models were used to find the number of optimum components/groups for infant mortality data for every district in Peninsular Malaysia. Every component is assumed to have the same distribution, but different parameters. The number of optimum components were obtained by maximum likelihood approach via the EM algorithm. Bayes theorem was used to determine the probability of belonging to each district in every components of the mixture distribution. Each district was assigned to the component that had the highest posterior probability of belonging. The results obtained were visually presented in maps. The analysis showed that in the early year of 1970, the spatial heterogeneity effect was more prominent; however, towards the end of 1990, this pattern tended to disappear. The reduction in the spatial heterogeneity effect in infant mortality data indicated that the provisions of health services throughout the Peninsular Malaysia have improved over the period of the study, particularly towards the year 2000.

Keywords: Disease mapping; infant mortality; mixture model

ABSTRAK

Mortaliti bayi merupakan salah satu isu penting bagi kebanyakan negara membangun. Di Malaysia, kadar kematian bayi pada peringkat kebangsaan telah bertambah baik sejak beberapa dekad yang lalu. Walau bagaimanapun, isu yang akan diketengahkan adalah untuk mengetahui sama ada penambahbaikan ini berlaku secara seragam di seluruh negara atau sebaliknya. Tujuan kajian ini ialah untuk mengkaji taburan geografi bagi kematian bayi di Semenanjung Malaysia dari tahun 1970 sehingga 2000 menggunakan teknik yang dikenali sebagai pemetaan penyakit. Diandaikan pemboleh ubah rawak kes kematian bayi adalah daripada taburan Poisson. Model campuran digunakan untuk mendapatkan bilangan komponen/kumpulan yang optimum bagi data mortaliti bayi bagi setiap daerah di Semenanjung Malaysia. Setiap komponen diandaikan mempunyai taburan yang sama tetapi parameter yang berbeza. Bilangan komponen yang optimum diperoleh menerusi kaedah kebolehjadian maksimum melalui algoritma EM. Teorem Bayes digunakan untuk mengenal pasti kebarangkalian daerah berada dalam setiap komponen bagi taburan campuran yang disuaikan. Setiap daerah diumpukkan kepada komponen yang mempunyai kebarangkalian posterior tertinggi. Keputusan yang diperoleh dipaparkan menerusi peta. Analisis menunjukkan pada awal tahun 1970-an, kesan heterogen reruang adalah lebih ketara, walau bagaimanapun, pada akhir 1990-an, keadaan tersebut telah semakin berkurangan. Pengurangan kesan heterogen reruang dalam data kematian bayi menunjukkan bahawa kemudahan kesihatan yang disediakan di seluruh Semenanjung Malaysia telah semakin baik sepanjang tempoh yang dikaji, terutamanya menjelang tahun 2000.

Kata kunci: Kematian bayi; model campuran; pemetaan penyakit

INTRODUCTION

Child health is a central public issue in most developing countries and one of its components is infant mortality. The most common index for measuring the standard of health and medical services in infant mortality is the rate (number of infant death per 1000 live births). Usually, the infant mortality rate is considered as a good indicator to measure the quality of health, socioeconomic and education level of a country. The improvement and advancement of modern medical technology, especially the development of programs to reduce maternal and infant mortality, have led to a dramatic reduction in infant mortality in many developing countries after World War II. For example, in Japan (Laskar & Harada 2005), the country-level infant mortality rates declined significantly from 1973 to 1998 and in Brazil (Alves & Belluzzo 2004), the infant

mortality rate reduced from 124 to 34 per 1000 live births from 1970 to 2000. Many efforts have been taken through early diagnosis and treatment, vaccination programs and appropriate management to reduce infant mortality and morbidity from diarrhea, pneumonia, measles, malaria, HIV/AIDS and infectious diseases.

Since independence in 1957, Malaysia itself has had a very remarkable decline in infant mortality from the rate of around 100 per 1000 to around 13 per 1000 by the late 1980s and in 2004, the rate has been reduced to 9 per 1000. This achievement is nearly equal to the standard rate of other developed countries, such as United States and United Kingdom, with 7 and 6 deaths per 1000, respectively. Indirectly, this trend is due to the improvement in Malaysia's socioeconomic situation, where the wages has increased and basic facilities, such as water supply, electricity, sewerage, sanitation and health services have improved and have been provided to the wide population. Apart from that, the levels of education and health consciousness have increased among the Malaysians. However, there are many other factors that directly and indirectly influence the infant mortality in Malaysia (Mohamed et al. 1998).

As mentioned earlier, it is fortunate that the infant mortality rate in Malaysia has improved over the last few decades, but the issue concerned is whether the improvement is uniformly distributed throughout the country. Does every district experience the same level of improvement or reduction in the risks? Does the improvement only occur in certain areas while the other areas still remain in the high-risk areas category? If there is a huge gap between the high-risk areas and low-risk areas, then the disease risks can be divided into several categories or considered as heterogeneous. This is the main issue that will be addressed in this paper in the context of infant mortality in Malaysia, using a technique known as disease mapping. It is very useful to produce such maps, especially for government agencies responsible for resource allocation or identifying hazards that contribute to the disease (Lawson & Williams 2001).

In disease mapping, usually two important elements will be addressed. They are obtaining smoothed estimators of relative risk (RR) and categorizing or classifying all districts into different components using shading or coloring to differentiate the level of risks for each component. The most common index that have been used to measure the RR is the Standardized Mortality/Morbidity Ratio (SMR), where the risk is define as the number of observe cases divided by the number of expected cases. Although SMR has been widely used as an index to measure RR, it has some limitations, such as the variance of SMR, which depends on the number of expected cases. The variance will be large when the expected value is small, as contributed by the small population size and the variance will be small when the expected value is large due to the large population size. Furthermore, if the observed value is zero, such as in the case of rare disease, the SMR and the standard deviation will be zero. Another limitation of SMR is the instability of the RR estimation due to the presence of extreme SMR when rare diseases are investigated in small population areas (Rattanasiri et al. 2004).

To overcome the drawbacks of the SMR, a Bayesian approach has been used. Clayton and Kaldor (1987) discussed several models based on empirical Bayes method to smooth the SMR. Marshall (1991) and Meza (2003) used the same model proposed by Clayton and Kaldor (1987), but introduced different method of parameter estimation. Apart from empirical Bayes method, some authors focused on hierarchical Bayesian approach with structured and unstructured spatial random effects (Lawson et al. 2003). Another simple approach of estimating the RR is known as locally weighted averages (Waller & Gotway 2004). This smoothing method is used to reduce noise in disease risks by borrowing information from neighboring districts to obtain better estimate of the RRs. The advantages of using this approach are less tedious derivations or complex computations procedures, not relying on a lot of assumptions and very suitable for an exploratory tool for descriptive statistics. Although the Bayesian and locally weighted averages methods can provide estimate on the RRs for each district, the number of optimum classification for categorizing the districts cannot be obtained based on them. The most common approach that is widely used by many researchers for categorizing areas in disease mapping is classification based on quartiles. However, this method is rather arbitrary and has no guarantee in detecting the classification of high- or low-risk areas. To overcome these drawbacks and also to consider the heterogeneity effect, this paper will focus on a more flexible tool known as the mixture model (Chandrasekaran & Arivarignan 2006; Schlattmann & Bohning 1993; Schlattman et al. 1996). This approach will be applied to examine the geographical distribution of infant mortality data in Peninsular Malaysia from 1970 to 2000.

Materials and Methods

Data

Malaysia is divided into West Malaysia and East Malaysia. In this paper, only the data of West Malaysia will be analyzed. West Malaysia is also known as Peninsular Malaysia and consists of 11 states. Each state is divided into administrative districts. Data on infant mortality and live birth for every district in Peninsular Malaysia were obtained from annual Vital Statistics reports for the period 1970–2000 and analyzed using the mixture model. Vital Statistics is a report published annually by Statistics Department of Malaysia. It presents population estimates and statistics on births and deaths at both state and administrative district levels. The main data source such as birth and death records for this report were provided by National Registration Department. Vital Statistics is the main component used for formulation of policies at national and international levels. Owing to some problems

in obtaining the data for the years 1978, 1980, 1992 and 1998, the analysis for these particular years have been excluded in this study.

DISEASE MAPPING USING MIXTURE MODEL

In disease mapping, the study area to be mapped is divided into M mutually exclusive districts (i=1,2,…,M). Each district has its own observed number of cases, O_i and expected number of cases, J_i. The expected number of cases is calculated as:

$$J_i = N_i \sum_{i=1}^{M} O_i \Big/ \sum_{i=1}^{M} N_i ,$$

where N_i is the population at risk for area i. Here, the standardization is done on the total population at risk which is live birth. The standardization can be done on other factors, such as age and gender (Mantel & Stark 1968; Pollard et al. 1981). It is assumed that the observe number of cases, O_i follows a Poisson distribution with expected value of $J_i\theta$, and the probability density function is given by:

$$f(O_i, \theta) = \frac{\exp(-J_i\theta)(J_i\theta)^{O_i}}{O_i!},$$

where θ is the RR of disease concerned over the study area.

As the assumption of constant risk only considers a homogenous case over the study area, a method that considers a heterogeneous case that can describe spatial variation of risks over the study area, known as the mixture model, will be discussed in this paper. In mixture model, we assume that the entire population throughout the study area comes from K subpopulations known as the components. It is also assumed that the random variables of the observed cases, O_i for every component have the same distribution, but with different parameter, θ_1, θ_2,…, θ_K, and the probability density function for each component is given by $f_k(O_i, \theta_K)$, where k = 1, 2, …, K. Let q_k denote the proportion of regional areas having θ_K risk. Here, q_k is the weight corresponding to θ_K, where $q_1+q_2+\ldots+q_K=1$. This discrete parameter P^* distribution for describing the level of risk can be given as:

$$P^* = \begin{bmatrix} \theta_1,\ldots,\theta_k \\ q_1,\ldots,q_K \end{bmatrix}.$$

Accordingly, we may assume that O_i comes from a nonparametric mixture density identified in the following form:

$$f(O_i; P^*) = \sum_{k=1}^{K} f_k(O_i,\theta_k) q_k .$$

The number of parameters to be estimated in the model with K components considered earlier are $2K-1$, which consists of K unknown RRs $\theta_1,\ldots, \theta_K$ and $K-1$ unknown mixing weights $q_1, q_2, \ldots, q_{K-1}$. A very simple and general introduction to mixture model can be found in Everitt and Hand (1981).

The most widely used method for parameter estimation in mixture model is the maximum likelihood approach by implementing the EM (expectation maximization) algorithm. As mentioned earlier, the study area is divided into M mutually exclusive districts and based on the nonparametric mixture density, $f(O_i;P^*)$ and the likelihood function is given by:

$$L(O_i;P^*) = \prod_{i=1}^{M} \left[\sum_{k=1}^{K} f_k(O_i,\theta_k) q_k \right].$$

To use the mixture model approach for map construction, another important element that should be taken into consideration is the classification aspect, where this step can classify the RR for each district into one of the components of the mixing distribution. Let the probability of region i belong to the kth component given that the observed cases, O_i is known and is denoted as $P(\pi_k | O_i)$. Based on the Bayes Theorem, this probability of belonging can be written as:

$$P(\pi_k | O_i) = \frac{q_k f_k(O_i,\theta_k)}{f(O_i;P^*)}.$$

The suitable algorithm for maximizing the log-likelihood function is known as EM algorithm (Bohning et al. 1992). This algorithm consists of two iterative procedures. By assigning the initial values of θ_k and q_k, the probability of region i belongs to the kth component, given that the observed cases O_i will be obtained in E-step. Based on $P(\pi_k | O_i)$ obtained in E-step, the new values of θ_k and q_k can be calculated in M-step. The new parameter values will again be replaced in E-step and this process will be repeated until the convergence criterion is met. Mathematically, the EM algorithm at s-iteration is given by:

$$\text{E-step:} \quad p^{(s)}(\pi_k | O_i) = \frac{\hat{q}_k^{(s)} f_k\left(O_i,\hat{\theta}_k^{(s)}\right)}{\sum_{k=1}^{K} \hat{q}_k^{(s)} f_k\left(O_i,\hat{\theta}_k^{(s)}\right)}.$$

$$\text{M-step:} \quad \hat{q}_k^{(s+1)} = \frac{1}{M} \sum_{i=1}^{M} P^{(s)}(\pi_k | O_i) \quad \text{and}$$

$$\hat{\theta}_k^{(s+1)} = \frac{\sum_{i=1}^{M} P^{(s)}(\pi_k | O_i) \frac{O_i}{J_i}}{\sum_{i=1}^{M} P^{(s)}(\pi_k | O_i)}.$$

Once the mixture model with K components achieves the convergence criteria, the next important step is to determine the probability of belonging for each district in every components of the mixture distribution. This can be done by calculating the value of $P(\pi_k | O_i)$. The ith district is assigned to component k, for which it has the highest posterior probability of belonging.

An important issue in mixture model is how to determine the optimum number of components compatible to the data. A further step is to determine the most suitable number of components by computing the difference between the log-likelihood for K and K+1 components, which is known as the likelihood ratio statistics (LRS) and is defined as:

$$LRS = -2(I_K - 1_{K+1}) = -2 \log\theta,$$

where I_K and I_{K+1} are log-likelihood values for K and K+1 components, respectively.

The purpose of calculating the LRS is to test the following hypotheses:

H_o: The number of components is K

H_1: The number of components is K+1

A problem arises in determining the number of components when the solution consists of the log-likelihood values that are nearly the same for every component. Conventionally, the LRS test has an asymptotic χ^2 distribution with degrees of freedom equal to the difference between the number of parameters under the alternative and null hypothesis. However, according to Schlattmann and Bohning (1993), these conventional results for LRS do not hold for mixture model and a method proposed to obtain the critical values in determining the number of components is via a simulation technique, for example, by parametric bootstrap.

RESULTS

To illustrate the results, the analysis for the mixture model of infant mortality data for the year 1993 was taken as an example on how to determine the optimum number of components. This result is given in Table 1.

From Table 1, it can be observed that there is no difference between the log-likelihood value of the model with three components and that with four components. For the model with four components, the weight associated with the highest risk category is only 0.007, indicating that only 0.7% of the data for that particular year fall in this category (or may be none of the districts fall in this category). Thus, we decided that the optimum number of components is three. Although it has been suggested by some studies, for example, by Schlattmann and Bohning (1993), that bootstrap methods should be applied in deciding either to choose between three or four components, we based our decision on the previous argument and for the purpose of obtaining parsimonious model. From the fitted model with three components, the first category had the lowest risk with a mean of 0.990 and weight of 0.841 and the highest risk category with a mean of 1.962 and weight of 0.014. The analysis of 1993 data led to the mixture density with three components given by:

$$f(o_i, \hat{P}^*, J_i) = f_1(o_i, 0.990, J_i) \times 0.841 + f_1(o_i, 1.462, J_i) \times 0.145 + f_3(o_i, 1.962, J_i) \times 0.014.$$

The same procedure was used in analyzing the data for all the other years in this study and the final results are summarized in Table 2.

From Table 2, we can conclude that over the three decades, the number of components has reduced from four or five subpopulations in the early 1970 to three or four components, starting in the middle of 1970s till 1987. In the late 1980s until the year 2000, the number of components can be divided into two or three. The reduction in the number of components shows that the heterogeneous effect of the infant mortality data in Peninsular Malaysia has decreased and the districts in the study area tend to be in the same category or component. This situation shows positive improvements from the actions taken in reducing the infant mortality, especially in the resource allocation between the districts. Based on the weight values $\hat{q}_k$ it can be seen that more districts tend to be in the low-risk category, especially towards the end of 1980s and in the 1990s.

To illustrate the results obtained, maps of mixture model for infant mortality data in Peninsular Malaysia are given in Figure 1.

TABLE 1. Mixture-model result of infant mortality data in Peninsular Malaysia for 1993

Number of Components (K)	Mean Relative risks $(\hat{\theta}_k)$	Weight $(\hat{q}_k)$	Log-likelihood (l_K)	*LRS*
K = 4	0.990	0.841	-313.411	0.000
	1.462	0.145		
	1.962	0.007		
	1.962	0.007		
K = 3	0.990	0.841	-313.411	3.584
	1.462	0.145		
	1.962	0.014		
K = 2	0.997	0.863	-315.203	60.520
	1.549	0.137		
K = 1	1.073	1.000	-345.463	

TABLE 2. Mixture-model result of infant mortality data in Peninsular Malaysia from 1970 to 2000

Year	Number of Components (K)	Weight $(\hat{q}_k)$	Mean Relative risks $(\hat{\theta}_k)$	Log-likelihood (l_K)
1970	3	0.599	0.847	-430.428
		0.387	1.312	
		0.014	2.312	
1971	4	0.198	0.710	-342.603
		0.269	0.932	
		0.445	1.145	
		0.088	1.491	
1972	4	0.245	0.667	-353.294
		0.306	0.942	
		0.364	1.183	
		0.085	1.441	
1973	4	0.039	0.610	-341.267
		0.230	0.778	
		0.255	0.909	
		0.476	1.130	
1974	5	0.036	0.487	-329.872
		0.370	0.806	
		0.431	1.026	
		0.145	1.286	
		0.018	1.646	
1975	5	0.049	0.505	-326.679
		0.200	0.760	
		0.573	1.001	
		0.149	1.228	
		0.029	1.643	
1976	3	0.029	0.558	-312.444
		0.578	0.891	
		0.393	1.170	
1977	3	0.231	0.701	-366.038
		0.443	0.925	
		0.326	1.204	
1979	4	0.153	0.641	-337.803
		0.363	0.835	
		0.364	1.122	
		0.120	1.466	
1981	5	0.187	0.647	
		0.349	0.868	
		0.337	1.085	-340.789
		0.114	1.512	
		0.013	2.990	
1982	3	0.162	0.577	
		0.698	1.285	
		0.140	2.084	-475.429
1983	3	0.570	0.867	
		0.406	1.403	
		0.024	2.067	-418.701
1984	3	0.413	0.800	
		0.426	1.200	
		0.161	1.575	-363.070
1985	4	0.257	0.751	
		0.428	1.007	
		0.289	1.397	-355.717
		0.026	2.693	
1986	3	0.521	0.820	
		0.465	1.380	
		0.014	1.885	-379.689

(continue)

Continued TABLE 2

Year	Number of Components (K)	Weight $(\hat{q}_k)$	Mean Relative risks $(\hat{\theta_k})$	Log-likelihood (l_K)
1987	3	0.383	0.799	
		0.555	1.164	
		0.062	1.712	-332.776
1988	2	0.730	0.958	
		0.270	1.424	
1989	2	0.728	0.923	-347.063
		0.272	1.419	
1990	2	0.814	0.998	-333.777
		0.186	1.467	
1991	3	0.140	0.672	-334.668
		0.716	1.022	
		0.144	1.570	
1993	2	0.863	0.997	-315.200
		0.137	1.549	
1994	2	0.861	0.989	-314.365
		0.139	1.525	
1995	3	0.369	0.827	-310.001
		0.471	1.114	
		0.160	1.441	
1996	2	0.484	0.800	-300.969
		0.516	1.242	
1997	2	0.658	0.890	-320.095
		0.342	1.363	
1999	3	0.652	0.928	-313.999
		0.331	1.390	
		0.017	2.951	
2000	3	0.572	0.886	-303.559
		0.405	1.376	
		0.023	2.632	

Figure 1 shows maps for the year 1972, 1981, 1991 and 2000 to describe the spatial changes in the infant mortality data over three decades of the study period. Each map provided the information on mixing weight, q_k and $\theta_{k,}$ and the estimate of RR. Based on the maps, it can be seen that in the 1970s and 1980s, most of the districts in the high-risk category were located in several undeveloped states, such as Kedah, Kelantan, Terengganu and Pahang. In the 1990s, there were positive changes in the infant mortality data in Peninsular Malaysia, where many districts fell in the low-risk category and only few districts remained in the high-risk component. However, the clustering effects were still found to occur throughout the period of study.

DISCUSSION

For quite some time, many researchers have conducted various studies in disease mapping using the traditional methods of classification, such as percentiles method and significant method. However, these methods have some deficiencies and the potential of misrepresenting the graphical distribution and question regarding whether these classifications give a correct interpretation may be raised (Schlattmann & Bohning 1993). An alternative approach suggested is the mixture model, which could produce a smoother map where the random variability has been extracted from the data. The other main advantages of using the mixture distribution are that its discreteness in making the map construction is straightforward and provides the optimum number of components. Based on the maps in Figure 1, it is very clear that the mixture model has removed the random variability from the map and provides a better and clearer picture of classification for high- and low-risk areas.

Based on the results obtained in this study, we can conclude that the map classification using the mixture model reduced the number of components from four or five in the early study period to two or three towards the end of the study period. This indicates that the classification tend to be more homogeneous, implying that the random variability has reduced with time. In general, the mean RRs appeared to reduce over time and the weight for low-risk components were found to increase. This shows that more districts have fallen into the low-risk categories, which

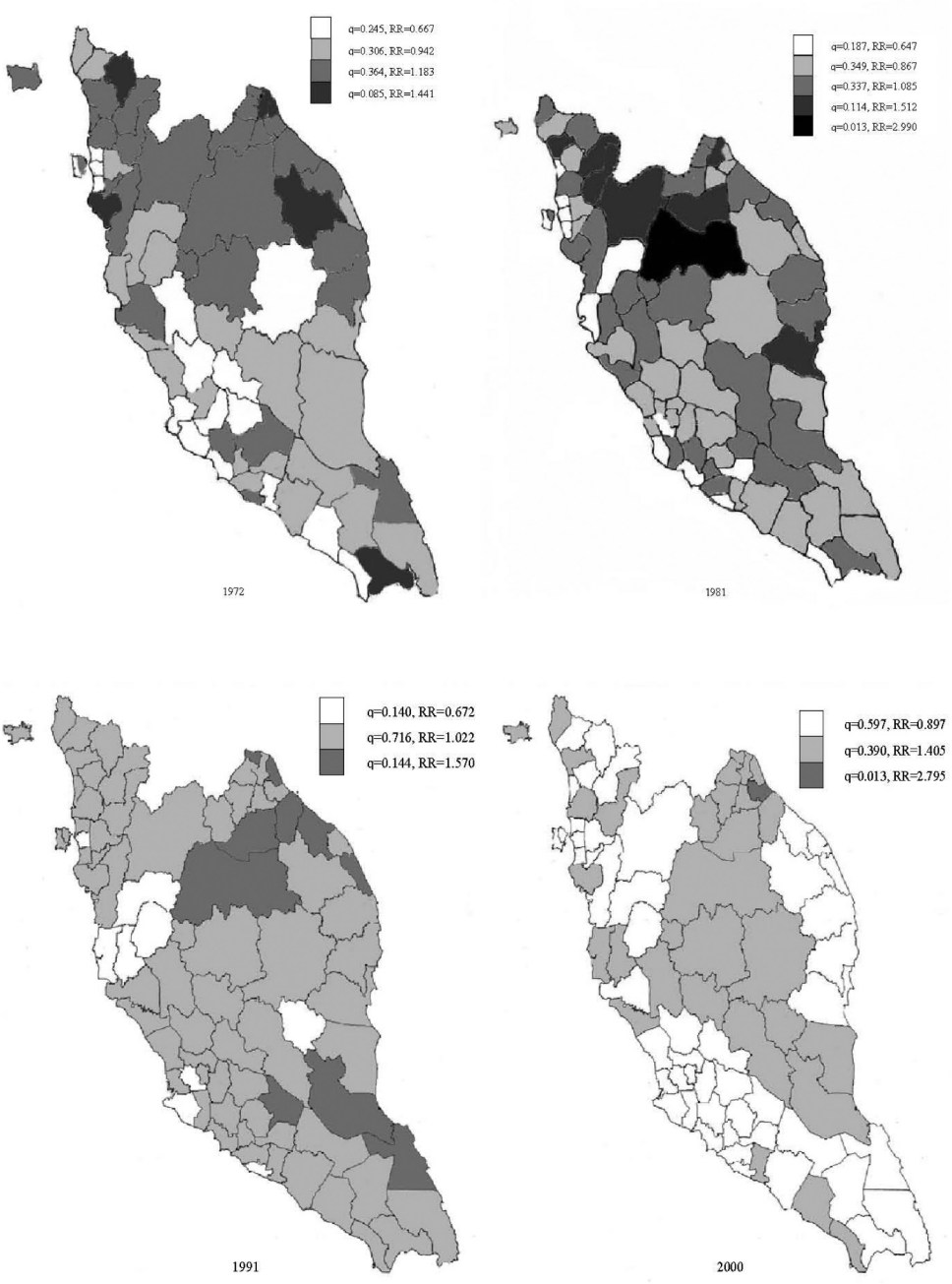

FIGURE 1. Maps of infant mortality in Peninsular Malaysia using mixture model

indicated that the infant mortality in Peninsular Malaysia have improved within the last few decades.

Many studies have stated that infant mortality is more likely to be related to the socioeconomic level, health behavior, quality of antenatal care, support during delivery, postnatal care, nutritional status, educational level, unemployment and birth intervals (Adebayo et al. 2004; Rutstein 2005; Turrell & Mengersen 2000). These factors were addressed in the case of Malaysia, as shown by the increase in health allocation from RM17.30 per capita in 1970 to RM248 per capita in the year 2000 (Estimates of Malaysia Federal Revenue and Expenditure 1970-2000). The number of hospitals were increased from 84 public hospitals in 1965 to 116 public hospitals in 2002, along with many private hospitals, health clinics and rural clinics being built throughout the country to provide better health system in Malaysia (Social Statistics Bulletin Malaysia 1965-2002). As the number of hospitals increased, more facilities were upgraded, such as providing more hospital beds, while at the same time, the number of registered doctors, trained nurses and midwives have also been increased (Social Statistics Bulletin Malaysia 1965-2002).

In conclusion, the improvement of health and medical services in Malaysia in the past four decades contributes to the improvement in infant mortality rates. The authority

in charge has put in a lot of effort, especially in terms of the quality of service, the advancement of medicine and medical technologies, the resolution of the issue of unbalanced distributions of medical resources between rural and urban areas, the establishment of collaborations among government and private hospitals or medical institutions. Numerous campaigns and programs have been carried out by the local government and the Ministry of Health to educate and increase health consciousness among the Malaysians. However, more efforts must be taken, especially by providing more resource and financial allocation to the high-risk areas to reduce the gap or inequality in infant mortality data between the districts in Peninsular Malaysia.

Even though the mixture model has some advantages in estimating the disease risks and providing a better and clearer picture of categorization, this approach still has the disadvantage of the RR for different districts being possibly correlated, i.e. dependent on geographical proximity. The parametric conditional autoregressive model that includes the neighboring factor among the study areas can be considered in future research to overcome this weakness (Clayton & Kaldor 1987).

ACKNOWLEDGEMENT

This research was fully supported by the short-term grant offered by Universiti Sains Malaysia (304/PMATHS/6310041).

REFERENCES

Adebayo, S.B., Fahrmeir, L. & Klasen, S. 2004. Analyzing infant mortality with geoadditive categorical regression model: A case study for Nigeria. *Economics and Human Biology* 2: 229-244.

Alves, D. & Belluzzo, W. 2004. Infant mortality and child health in Brazil. *Economics and Human Biology* 2: 391-410.

Bohning, D., Schlattmann, P. & Lindsay, B. 1992. Computer-assisted analysis of mixtures (C.A.MAN): Statistical algorithms. *Biometrics* 48: 283-303.

Chandrasekaran, S.K. & Arivarignan, G. 2006. Disease mapping using mixture distribution. *Indian Journal of Medical Research* 123(6): 788-798.

Clayton, D. & Kaldor, J. 1987. Empirical Bayes estimates of age-standardized relative risks for use in disease mapping. *Biometrics* 43: 671-681.

Estimates of Malaysia Federal Revenue and Expenditure. Ministry of Finance Malaysia. 1970-2000.

Everitt, B.S. & Hand, D.J. 1981. *Finite Mixture Distributions*. New York: Chapman and Hall.

Laskar, M.S. & Harada, N. 2005. Trends and regional variations in infant mortality rates in Japan, 1973-1998. *Public Health* 119(7): 659-663.

Lawson, A.B. & Williams, F.L.R. 2001. *An Introductory Guide to Disease Mapping*. Chichester: John Wiley & Sons.

Lawson, A.B., Browne, W.J. & Rodeiro, C.L.V. 2003. *Disease Mapping with WinBUGS and MlwiN*. New York: Wiley.

Mantel, N. & Stark, C.R. 1968. Computation of indirect-adjusted rates in the presence of confounding. *Biometrics* 24: 997-1005.

Marshall, R.J. 1991. Mapping disease and mortality rates using empirical Bayes estimators. *Applied Statistics* 40: 283-294.

Meza, J.L. 2003. Empirical Bayes estimation smoothing of relative risks in disease mapping. *Journal of Statistical Planning and Inference* 112: 43-62.

Mohamed, W.N., Diamond, I. & Smith, W.F.P. 1998. The determinants of infant mortality in Malaysia: A graphical chain modeling approach. *Journal of the Royal Statistical Society A* 161(3): 349-366.

Pollard, A.H., Yusuff, F. & Pollard, G.N. 1981. *Demographic Techniques*. Sydney: Pergamon Press.

Rattanasiri, S., Bohning, D., Rojanavipart, P. & Athipanyakom, S. 2004. A mixture model application in disease mapping of malaria. *Southeast Asian Journal Trop. Med. Public Health* 35: 38-47.

Rutstein, S.O. 2005. Effects of preceding birth intervals on neonatal, infant and under-five years mortality and nutritional status in developing countries: Evidence from the demographic and health surveys. *International Journal of Gynecology and Obstetrics* 89: 7-24.

Schlattmann, P. & Bohning, D. 1993. Mixture models and disease mapping. *Statistics in Medicine* 12: 1943-1950.

Schlattmann, P., Dietz, E. & Bohning, D. 1996. Covariate adjusted mixture models and disease mapping ith the program Dismapwin. *Statistics in Medicine* 12: 919-929.

Social Statistics Bulletin Malaysia, Department of Statistics Malaysia. 1965-2002.

Turrell, G. & Mengersen, K. 2000. Socioeconomic status and infant mortality in Australia: A national study of small urban areas, 1985-1989. *Social Science & Medicine* 50: 1209-1225.

Vital Statistics Malaysia, Department of Statistics Malaysia. 1964-2000.

Waller, L.A. & Gotway, C.A. 2004. *Applied Spatial Statistics for Public Health Data*. New Jersey: John Wiley & Sons.

Nuzlinda Abdul Rahman*
School of Mathematical Sciences
Universiti Sains Malaysia
11800 Penang
Malaysia

Abdul Aziz Jemain
School of Mathematical Sciences
Faculty of Science and Technology
Universiti Kebangsaan Malaysia
43000 Bangi, Selangor, D.E.
Malaysia

*Corresponding author; email: nuzlinda@usm.my

Received: 10 November 2011
Accepted: 25 January 2013