Southeast-Asian J. of Sciences: Vol. 2, No 1 (2013) pp. 108-116

MALARIA MORTALITY MODELING IN THAILAND

Wattanavadee Sriwattanapongse
and
Sukon Prasitwattanaseree

Biostatistics and Applied Statistics Lab., Dept. of Statistics
Chiang Mai University, Chiang Mai 50200 Thailand
e-mail: wattanavadee.s@cmu.ac.th

Abstract

Malaria has been a leading cause of mortality in Thailand for many decades. The objective was to model of malaria mortality rate in Thailand by gender, age, year and region. A retrospective analysis of the malaria death was conducted by using the national vital registration database for the 10-year period from 2000 to 2009, provided by the Ministry of Interior and coded as cause-of-death using ICD-10. The linear, Poisson regression and negative binomial models were used for modeling and forecasting age-specific malaria mortality rates in Thailand. We use these models to forecast the malaria mortality which is likely to occur in the near future in order to prevent the malaria mortality through the use of suitable measures. Among the models fitted, the best were chosen based on the analysis of deviance and the negative binomial generalized linear model was clearly appropriate fit. The model contains additive effects associated with the gender, age group, year and regions. There is need of malaria control measures to remain on a sustained and long-term basis for the high malaria burden rate of Thailand.

1 Introduction

Malaria is a potentially fatal tropical disease that's caused by a parasite known as Plasmodium. It's spread through the bite of an infected female mosquito. The malaria is ranked 5th on the list in the 10 leading causes of death in low-income countries [1].There were 216 million cases of malaria and an estimated

Key words: mortality, malaria, modeling, analysis of deviance.

655, 000 deaths in 2010. Malaria mortality rates have fallen by more than 25% globally since 2000 and by 33% in the WHO African Region. Most children living in Africa was deaths by malaria where a child dies every minute and the disease accounts for approximately 22% of all childhood deaths [2].

Malaria has been a leading cause of mortality in Thailand. Especially, the border areas close to Myanmar and Cambodia are affected. Non-immune migrant workers occupied with gem mining in forests, logging, agriculture and construction are the most vulnerable and most affected. Malaria epidemics occurred periodically in high risk areas, especially along the international borders of Thailand and Myanmar, and Thailand and Cambodia. The death rate from malaria for males is currently more than for females. Thailand are marked differences regional in the level of incidence of malaria [3].

Thailand is a country located at the centre of the Indochina peninsula in Southeast Asia. It is divided into 77 provinces, which are gathered into 4 groups of regions by provinces [4]. There are four regions: Central, North, North-East and South. It is bordered to the North by Myanmar and Laos, to the East by Laos and Cambodia, to the South by the Gulf of Thailand and Malaysia, and to the West by the Andaman Sea and the Southern extremity of Myanmar. The estimating Thai population by department of provincial administration is of 65,479,453 [5]. Sriwattanapongse et al.[6] studied model the patterns of hospital-diagnosed malaria incidences by month, district and age-group for the two North-Western border provinces in Thailand. The model used linear regression, Poisson regression and negative binomial regression to forecast the districts and age groups. Among the models fitted, the best were the negative binomial generalized linear model. Additional, Sriwattanapongse and Kuning [7] studied the patterns of hospital diagnosed malaria incidences in districts and quarterly periods in the North-Western region of Thailand which was described by regression models based on principal components. The results show that malaria incidence rates decreased substantially in most districts during the study period, but remained very high in border districts with Myanmar. Investigating regional and temporal patterns is commonly used to detect areas with malaria problems and to evaluate periods of likely epidemics for a variety of disease. The forecasting of mortality and disease burden are essential for setting current and future health system priorities. The objective of our study was to model and forecast malaria mortality rates in Thailand. Moreover, Sriwattanapongse et al [8] studied to model and forecast malaria mortality rate in Thailand using death certificate reports. Multivariate regression was used for modeling and forecasting age-specific malaria mortality rates in Thailand. The trends of malaria mortality remained stable in most age groups with decreases in others and decreases during ten-year period (2000 to 2009). Malaria mortality was higher in males and increase with increasing age .

2 Materials and methods

Data for registered deaths due to malaria (ICD10: B50-B64) were obtained from the national vital registration database for the 10-year period from 2000 to 2009.The database is provided by the Ministry of Interior and coded as cause-of-death using ICD-10 by the Bureau of Policy and Strategy, Ministry of Public Health.

Age, gender, residential area by region in Thailand and year were selected as the explanatory variables in studying the mortality rates of malaria. Age was divided into nine groups (0-9,10-19,20-29,30-39,40-49,50-59,60-69,70-79 and above 80 yrs). For each region and gender combination; linear regression, Poisson and negative binomial model was used to investigate and forecast malaria mortality by gender, age group year and regions.

Various approaches have been developed to improve for forecasting morbidity and mortality rates. This paper focuses upon the model proposed by Lee and Carter [9] and Lee and Miller [10] that used initially for projections of the age-specific mortality rates in the United States. The Lee-Carter-based modeling frameworks is viewed in the current literature as among the most efficient and transparent methods of modeling and projecting mortality improvements [11]. This method is also regarded as the state-of-the-art in mortality forecasting and became more and more popular for long-run forecasts of age-specific mortality rates.

Statistical methods

The present study aims to find a suitable statistical model for predicting malaria mortality rates of reported death cases of malaria in Thailand with high risk of disease based on routinely collected data available from the national vital registration database for the 10-year period from 2000 to 2009, provided by the Ministry of Interior and coded as cause-of-death using ICD-10. It is interest to compare the risk of malaria mortality in infants, school children, young adults and older adults and regions.

Linear regression

The simplest model is based on linear regression with the outcome variable defined as the mortality incidence rate in a cell indexed by gender, age group, year, and region. Such incidence rates generally have positively skewed distributions so it is conventional to transform them by taking logarithms. And since monthly disease counts based on small regions are often zero, it is necessary to make some adjustment to avoid taking logarithms of 0: the method we use is to define the outcome as

$$y = \ln\left(1 + \frac{n}{P} \times K\right), \tag{1}$$

where n is the number of disease cases in the cell, P is the population at risk, and K is a specified constant(here, K =10,000). Such an observation-driven

model with k regions could take the form

$$Y_{ijt} = \mu + \alpha_i + \beta_j + \eta_t + \gamma_k + \varepsilon_{ijt}, \tag{2}$$

where N_{ijt} is a random variable denoting the population reported number of disease cases in gender i, age group j ,year t for the region of interest k and n_{ijt} is the sample corresponding number observed, Y_{ijt} is the outcome variable specified in Equation (1) and y_{ijt} the corresponding number observed, ε_{ijt} comprises a set of independent normally distributed random variables with mean 0.In this model we assume that the first of each set of gender, age group, year and region, and seasonal parameters is 0, that is, $\alpha_1 = 0$, $\beta_1 = 0$, $\eta_1 = 0$ and $\gamma_k = 0$. While linear time trends could be included in the model, they are less useful for short-term forecasting purposes in the presence of high serial correlations, and are not considered in the present study.

Generalized linear models

Davis et al [12] suggested observation-driven models for time series counts N_t based on the Poisson distribution with mean λ_t where $\ln(\lambda_t)$ is expressed as an additive function of determinants. Thus a suitable generalized linear model based on the Poisson distribution could take the form:

$$\ln(\lambda_{ijt}) = \mu + \alpha_i + \beta_j + \eta_t + \gamma_k, \tag{3}$$

where λ_{ijt} is the mean of N_{ijt}.

Poisson models for disease counts are often over-dispersed due to spatial or temporal clustering of cases, in which case the negative binomial distribution may be more appropriate. This distribution has an additional parameter and takes the form

$$\text{Prob}(N_t = n) = \frac{\Gamma(n+\gamma)}{\Gamma(n+1)\Gamma(\gamma)} \left(\frac{\gamma}{\gamma+\lambda_t}\right)^{\gamma} \left(\frac{\lambda_t}{\gamma+\lambda_t}\right)^{n}. \tag{4}$$

As for the Poisson model λ_t is the conditional expected value of N_t, but the conditional variance is $\lambda_t + \lambda_t^2/\gamma$ (see for example, Jansakul and Hinde,2004[13]). The parameter γ is actually inversely related to the over-dispersion, so that the Poisson model arises as the special case in the limit as $\gamma \to \infty$.

3 Results

Distributions of malaria mortality rates

During the study period from January 2000 to December 2009, numbering of 2,436 malaria cause-of-death were reported in the national vital registration database, about 72% were male. The number of cases in a year for a particular gender, age group, year and region varied from zero to 51 and the corresponding mortality rates were 1.51 cases per 10,000 which is male, age group 30-39, year 2000 and Northern region. The highest mortality was 1.57 cases per 10,000.

From model, the estimates of the regression coefficients, standard errors were estimates and tests for whether each regression coefficient can be assumed to be zero as in Table 1. While all components in the model are statistically significant exception age group 10-19 , the linear regression model was r-square statistics (71.74%). The highest residual of 3.06, corresponding to 9 cases that occurred among infants below 10 years of age in Central region in 2000, heralded a small epidemic comprising 4, 4, 3, 4 and 1 cases in the following five year.

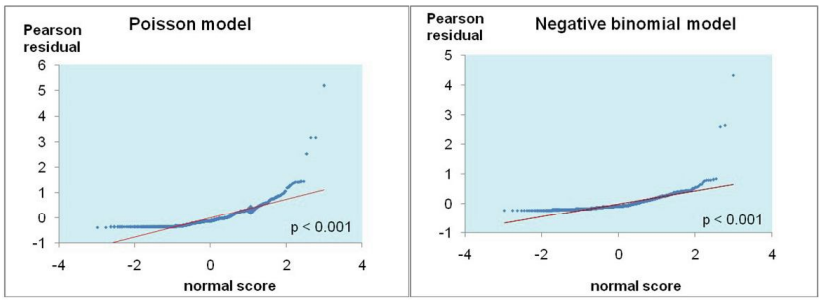

Figure 1: Plots of Pearson residuals versus asymptotic scores after fitting Poisson and negative binomial models for malaria mortality disease counts in Thailand.

Table 1: Regression coefficient and standard errors based on the linear, Poisson GLM and negative binomial GLM fitted to malaria mortality rates in Thailand.

Determinant		Linear model		Poisson GLM		Negative Binomial GLM	
		Coeff	St.Error	Coeff	St.Error	Coeff	St.Error
Constant		-3.000029	0.0987681*	-3.02435	0.09656*	-2.99103	0.11940*
Gender :	Male	0	1	0	1	0	1
	Female	-0.5980028	0.0421127*	-0.99530	0.04512*	-0.93487	0.05618*
Age Group:	0-9	0	1	0	1	0	1
	10-19	***-0.0009125***	***0.0893345***	***0.05492***	***0.11420***	***0.03910***	***0.13120***
	20-29	0.4779750	0.0893345*	0.95496	0.09705*	0.85445	0.11701*
	30-39	0.8345250	0.0893345*	1.47499	0.09162*	1.34943	0.11195*
	40-49	0.8544500	0.0893345*	1.24640	0.09520*	1.21232	0.11433*
	50-59	0.8322625	0.0893345*	1.14545	0.10335*	1.09728	0.12153*
	60-69	0.9752625	0.0893345*	0.98982	0.11712*	0.99645	0.13274*
	70-79	1.3845903	0.0920559*	1.26298	0.13232*	1.23957	0.14701*
	80+	2.1304375	0.0874731*	1.42815	0.18106*	1.39707	0.19056*
Year:	2543	0	1	0	1	0	1
	2544	-0.4011806	0.0941668*	-0.40056	0.06292*	-0.42575	0.09371*
	2545	-0.4488889	0.0941668*	-0.57388	0.06611*	-0.54762	0.09542*
	2546	-0.8280694	0.0941668*	-1.15673	0.08064*	-1.10221	0.10637*
	2547	-0.6533889	0.0941668*	-1.04867	0.07714*	-0.94981	0.10253*
	2548	-0.9964861	0.0941668*	-1.41717	0.08840*	-1.35820	0.11241*
	2549	-0.9362917	0.0941668*	-1.41466	0.08798*	-1.34870	0.11180*
	2550	-1.2801111	0.0941668*	-1.94377	0.10916*	-1.88081	0.12945*
	2551	-1.2646806	0.0941668*	-1.91336	0.10729*	-1.86352	0.12829*
	2552	-1.5122330	0.0947046*	-2.29848	0.12648*	-2.26165	0.14555*
Region:	Central	0	1	0	1	0	1
	North East	-0.2387778	0.0595563*	-0.12013	0.05488*	-0.22106	0.07357*
	North	0.5790556	0.0595563*	0.72386	0.05127*	0.67651	0.06972*
	South	0.4225278	0.0595563*	***0.06779***	***0.06984***	***0.09206***	***0.08413***
df:698		R-squared: 0.7174		Deviance: 1,050.0		Deviance: 822.1	
				AIC: 2,646.3		AIC: 2,605.5	

*Significant at 5%

Turning to the Poisson and negative binomial regression models given by Equations (3) and (4), Figure 1 shows plots of Pearson residuals versus corresponding (normal or gamma) scores. Since, the deviance was used to test model specification, the Poisson model gives residual deviances of 1,050.0 and the negative binomial model gives residual deviances of 822.1 respectively. The AIC (Akaike information criterion) was a measure of goodness of fit, the Poisson model gives AIC of 2,646.3 and the negative binomial model gives AIC of

2,605.5 respectively, so the negative binomial model is clearly appropriate.

For the Poisson model the sum of the predicted disease counts equals the sum of the observed counts. However, when the negative binomial model is fitted using maximum likelihood rather than moment estimators this constraint is not necessarily satisfied, and satisfying this requirement could govern the choice of K. We found that choosing K = 10,000 gives sums of predicted disease counts that are reasonably close to the observed sums.

Table 1 gives the results obtained from fitting the Poisson and negative regression model given by Equations (4) to the malaria mortality disease counts. The largest Pearson residual for Poisson obtained is 8.89, corresponding to 6 cases reported female among 60-69in South region in 2004.Since four further cases were reported in the same region and age group in the following four months (1, 0, 0, and 0 respectively). The largest Pearson residual for negative binomial obtained is 8.71, it was same as group of Poisson model. The dispersion parameter of negative binomial estimates is 9.22.

4 Conclusion

We applied linear, Poisson and negative regression to model and forecast the malaria mortality in Thailand. Among the models fitted, the best were chosen based on the analysis of deviance and the negative binomial generalized linear model was clearly appropriate fit.

Mortality was highest in age groups above 30-39 years in Northern area and year 2000 in male (1.51)and over 80+ in Central year 2002 in female(1.57). Co-morbidity and decrease immune function are important factors in the increasing malaria mortality among the elderly. In the ten-year period (2000 to 2009), the trends of malaria mortality fluctuation in most age groups with decrease in other and trend to decrease. In the Central and Northern, there was pronounced bulge in mortality among males between 30 and 39 years of age. Except the Southern, there was pronounced bulge in mortality among male between 70 and 79 years of age. In addition, there was increase of malaria mortality fluctuation in most age groups in between 20 and 80+ years of age in female.

Although the Lee-Carter model is often used for forecasting, this non-linear model cannot be fitted by ordinary regression methods, and thus does not routinely provide standard errors for estimated parameters. Booth et al.[14] use of the Lee-Carter method with Australian data is compromised by significant departures from linearity in the time component and changes over time in the age component. The model is also expanded to take account of age-time interactions by incorporating additional terms, but these are not readily incorporated into forecasts. Delwarde et al.[15] studied model to forecast future mortality rates.The result shows that it is possible to smooth the estimated

x's in the LeeCarter and Poisson log-bilinear models for mortality projection. Finally, penalized least-squares or maximum likelihood analysis is performed. The optimal value of the smoothing parameter is selected with the help of cross validation.

For each region and gender combination; Wattanavadee et al [11] used multivariate linear regression model to investigate and forecast malaria mortality by age group and year. The multivariate linear regression has the additional advantage is that it takes account of correlations between data in different age groups. The graphical method provides an informative display of the variation in mortality by gender, age group and region.

Whereas this study proposed three models; linear, Poisson and negative binomial regression mode were used to model malaria mortality. Among the models fitted, the best were chosen based on the analysis of deviance and the negative binomial generalized linear model was clearly appropriate fit. The model contains additive effects associated with the gender, age group, year and regions. The negative binomial regression is another method was to forecast disease mortality.

Acknowledgements

We thank the Ministry of Public Health and Ministry of Interior for providing the data. We are grateful to Prof. Don McNeil who supervised our research.

References

[1] WHO, Malaria Bulletin 2008, Available at: http://www.who.int/mediacentre/factsheets/fs310/en/index.html.

[2] World malaria report 2011, Available at: http://www.who.int/malaria/world_malaria_report_2011/en/.

[3] S. Wibulpolprasert, *Thailand Health Profile 2005-2007. In:* The War Veterans Organization of Thailand, Printing Press, Bangkok, 2007.

[4] Online Resources, Available at: http://en.wikipedia.org/wiki/Provinces_of_Thailand

[5] National Statistics Office. 100th anniversary of population censuses in Thailand: Population and housing census, 2010.

[6] W. Sriwattanapongse, M. Kuning and N. Jansakul, *Malaria in North-Western Thailand*, Songklanakarin J. Sci. Technol. 30 (2008), 207–14

[7] W. Sriwattanapongse and M. Kuning, *Modeling malaria incidence in North-Western*, Chiang Mai J. Sci. 36 (2009), 403–10.

[8] W. Sriwattanapongs, S. Prasitwattanaseree and S. Khanabsakdi, *Mortality Rate due to Malaria in Thailand*, Walailak J Sci & Tech. 9(2) (2012), 135–39.

[9] R.D. Lee and L.R. Carter, *Modeling and forecasting U.S.*, Mortality Am. Stat. Assoc. 87 (1992), 659–71.

[10] R.D. Lee and T. Miller, *Evaluating the performance of the Lee-Carter method for forecasting mortality*, Demography 38 (2001), 537–49.

[11] Z. Butt and S. Haberman, *A Collection of R Functions for Fitting a Class of Lee-Carter Mortality Models using Iterative Fitting Algorithms*, Sir John Cass Business School, London (2009).

[12] R.A. Davis, W.T.M. Dunsmuir and S.B. Streett, *Observation-driven models for Poisson count*, Biometrika 90(4) (2003), 777–90.

[13] N. Jansakul and J.P. Hinde, *Linear mean-variance negative binomial models for analysis of orange tissue-culture data*, Songklanakarin Journal of Science and Technology 26(5) (2004), 683–96.

[14] H. Booth, J. Maindonald and L. Smith, *Applying Lee-Carter under conditions of variable mortality decline*, Population Studies 56 (2002), 325-36.

[15] A. Delwarde, M. Denuit and .P Eirlers, *Smoothing the Lee-Carter and Poisson log-bilinear models for mortality forecasting*, Stat. Model. 7 (2007), 29–48.