นิพนธ์ต้นฉบับ

Mapping acute febrile illness incidence in Yala province

เสาวรส แก้วสมพักตร์[1]
สุจรรยา บุญประดิษฐ์[2]
จำเนียร ฉุ้นประดับ[3]
ยุพดี ชัยสุขสันต์[4]

Abstract:

Mapping acute febrile illness incidence in Yala province

Kaewsompak S, Boonpradit S, Choonpradub C, Chaisuksant Y.

Department of Mathematics and Computer Science, Department of Science,

Faculty of Science and Technology, Prince of Songkla University, Pattani, 94000, Thailand

Songkla Med J 2005;23(6):455-462

The objectives of this study were to study the epidemic patterns of dengue hemorrhagic fever and other acute febrile illnesses in Yala province in southern Thailand, to investigate relations between the their incidence rates in terms of their geographical distributions, and to develop a methodology that may be applied routinely to geographical epidemiologic research for the spatio-temporal mapping of disease. Schematic range maps and statistical models were used to investigate their distribution by year and location. The analytic methods used poisson and negative binomial distribution models. The concept of a "risk alert" is suggested as a method for highlighting subdistricts with unexpectedly high incidence rates in a given year.

---

[1]M.Sc. (Research Methodology) [2]M.Sc. (Applied Statistics) [3]Ph.D. (Statistics), Department of Mathematics and Computer Science
[4]Ph.D. (Environmental Science), Asst. Prof., Department of Science, Faculty of Science and Technology, Prince of Songkla University, Pattani, 94000, Thailand
รับต้นฉบับวันที่ 11 สิงหาคม 2548 รับลงตีพิมพ์วันที่ 11 มกราคม 2549

We found high correlations between the incidence rates in 2002 for diarrhoea and conjunctivitis (r = 0.805), and between diarrhoea and pneumonia (r = 0.798). For all illnesses in the two years, five risk alerts occurred. For dengue fever, in 2002 there were two subdistricts with moderate risk alerts, namely KaYuBoKo (p = 0.030) and KuTaBaRu (p = 0.042), whereas in 2003 there was just one subdistrict with a high risk alert, KaYuBoKo (p = 0.003). For pyrexia, in 2003 there were two subdistricts with moderate risk alerts, namely Kerikat (p = 0.017) and ThanTo (p = 0.036).

Further studies are needed, over longer periods, involving more provinces, and taking into account known risk factors including age, season, environmental factors, and time series autocorrelations.

Key words: *acute febrile, illnesses, GIS mapping, Poisson distribution, negative binomial distribution*

บทคัดย่อ:

การวิจัยครั้งนี้มีวัตถุประสงค์เพื่อศึกษารูปแบบการแพร่กระจายของโรคไข้เลือดออกและโรคอื่น ๆ ในจังหวัดยะลา, เพื่อศึกษาความสัมพันธ์ระหว่างความชุกของโรคกับลักษณะการกระจายทางภูมิศาสตร์ และเพื่อพัฒนาวิธีการที่อาจจะนำมาประยุกต์ใช้กับงานวิจัยทางด้านระบาดวิทยาในเชิงภูมิศาสตร์ ได้นำแผนที่และโมเดลทางสถิติวิเคราะห์มาใช้ในการศึกษาการกระจายของโรคจำแนกตามเวลาและพื้นที่ รวมทั้งการหาค่าสัมประสิทธิ์สหสัมพันธ์เพื่อวัดความสัมพันธ์ระหว่างอัตราการเกิดโรค และสร้างโมเดลลักษณะการเกิดโรคโดยใช้การแจกแจงปัวส์ซองและการแจกแจงทวินามวิเสธ ลักษณะเกี่ยวกับ "พื้นที่เสี่ยง" คือ วิธีการสำหรับตำบลที่มีอัตราการเกิดโรคสูงในแต่ละปี ผลการศึกษาพบว่า ในปี พ.ศ. 2545 มีความสัมพันธ์กันสูงระหว่างอัตราการเกิดโรคท้องร่วงและโรคตาแดง (r = 0.805), โรคท้องร่วงและโรคปอดบวม (r = 0.798) และโรคตาแดงและโรคปอดบวม (r = 0.580) ในระยะเวลา 2 ปีที่ศึกษามีพื้นที่เสี่ยงต่อการระบาด 5 ตำบล โดยในปี พ.ศ. 2545 โรคไข้เด็งกี่มี 2 ตำบลที่มีความสี่ยงต่อการระบาดในระดับปานกลาง คือ ตำบลกอยูบอเกาะ (p = 0.0296) และโกตาบารู (p = 0.0422) และในปี พ.ศ. 2546 ตำบลกายูบอเกาะ (p = 0.0032) เกิดการระบาดสูง สำหรับโรคไข้ไม่ทราบสาเหตุในปี พ.ศ. 2546 มี 2 ตำบล คือ ตำบลคีรีเขต (p = 0.0172) และตำบลธารโต (p = 0.0358) ที่มีพื้นที่เสี่ยงระบาดระดับปานกลาง

คำสำคัญ: *โรคที่มีอาการไข้, แผนที่, การแจกแจงปัวซ์ซอง, การแจกแจงทวินามวิเสธ*

Introduction

In Thailand, the incidences of dengue haemorrhagic fever (DHF) have increased substantially over the last 40 years[1], with the rates peaking between November and April. However, while the Thai Ministry of Public Health has projected decreased morbidity in the provinces where the disease is most prevalent, the current DHF incidence remains a major public health problem. The Division of Epidemiology, Ministry of Public Health, reported the number of cases for the years 1998-2001 to be 129,954, 24,826, 18,617 and 132,082, respectively, with corresponding deaths 424, 56, 32, and 238. In 2002, the DHF incidence was estimated to be 108,905 cases and 172 deaths, with the highest incidences occurring in Yala, Suratthani, Krabi, Phatthalung, and Nakhon Sri Thammarat provinces, all in the southern part of Thailand. In 2003, however, these numbers fell to 62,526 cases and only two deaths[2], underlining the volatility of DHF disease incidence from year to year.

Yala is the southernmost province of Thailand with an area of 4,521 square kilometers. It lies between latitude 6°54′N and longitude 101°28′E, and is 1,055 kilometers from Bangkok by rail and 1,440 kilometers by road. It is the only one of the 14 southern provinces that does not have a sea coast. Yala is mostly mountainous and covered with forests. The Pattani River flows from the south and forms a basin down the middle of the province. Yala province is divided administratively into 8 districts, 58 subdistricts and 354 villages (Figure 1).

The epidemic pattern of DHF in Yala province has alternated by year from 1999 to 2003. The highest number of cases was recorded in 2002, particularly in Betong, Yala city, and Yaha districts, with morbidity rates of 434.0, 310.4 and 281.0 per 100,000 residents, respectively. In 2003, the number of cases was highest in Yala city, Betong, and Raman districts, with respective morbidity rates of 201.3, 185.8 and 129.6.[3]

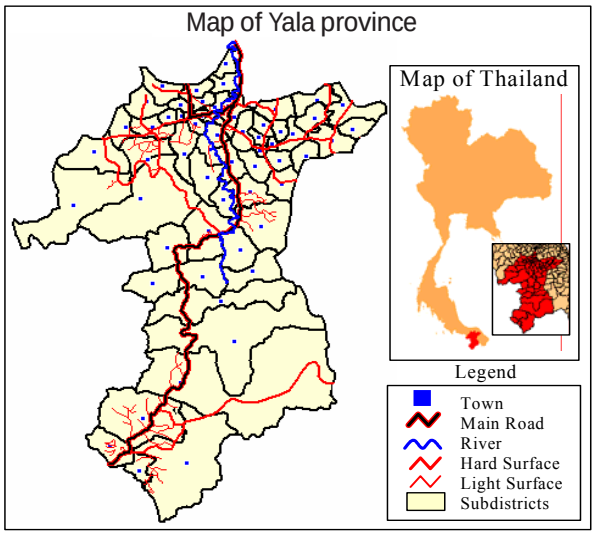

Figure 1 Map of the Yala province

Objectives

The objectives of this study were to study the epidemic patterns of dengue hemorrhagic fever and other acute febrile illnesses in Yala province in southern Thailand, to investigate relations between the their incidence rates in terms of their geographical distributions, and to develop a methodology that may be applied routinely to geographical epidemiologic research for the spatio-temporal mapping of disease.

DHF is only one of several acute febrile illness that account for most hospital admissions in Yala province. Table 1 shows the numbers of admissions for the seven most common acute febrile illnesses from 1999 to 2003. While diarrhea accounted for over 70,000 admissions, and pneumonia, pyrexia, haemorrhagic conjunctivitis all had higher case incidence rates than DHF, diarrhea, pyrexia and conjunctivitis have relatively low case fatality rates compared to DHF, malaria, and pneumonia. The volatility of the DHF and dengue fever incidences is very apparent from this table, showing substantial increases in 2001 and 2002. The number of cases of malaria also shows substantial short-term variation, doubling from 2000 to 2001 and almost doubling from 2002 to 2003.

Table 1 Disease cases in Yala province, 1999-2003

Disease	Year				
	1999	2000	2001	2002	2003
Dengue fever	68	58	1,213	1,548	317
DHF	99	121	1,273	1,217	326
Malaria	376	350	719	797	1,471
Conjunctivitis	2,025	1,859	2,295	3,719	1,571
Diarrhea	12,112	13,081	16,212	14,499	16,071
Pneumonia	3,174	3,368	3,385	2,770	2,845
Pyrexia	1,794	1,499	4,098	4,329	3,924
Total	**22,455**	**23,235**	**23,741**	**28,879**	**28,528**

Materials and methods

Each provincial health office in Thailand maintains a database comprising records based on public health surveillance of all hospital admissions for 78 diseases. Each record comprises the disease, age, gender, subdistrict of residence, date of admission, and various subsequent health details of each person admitted.

The study design was a secondary analysis of hospital-based surveillance system records of all the cases of the most common acute febrile illnesses. The sample comprised data for seven illnesses (dengue fever, dengue hemorrhagic fever, malaria, hemorrhagic conjunctivitis, diarrhea, pneumonia and pyrexia of unknown origin), collected in 2002 and 2003 by the Yala Provincial Public Health Office. These data were classified by disease type and subdistrict, but not by age and gender. Although age is an important determinant of incidence rate for most diseases, it has substantially different patterns for different diseases (for example, DHF occurs mainly in children whereas conjunctivitis mainly afflicts older persons), and for simplicity it is not included as a determinant in this study. The denominator for computing incidence rates was taken to be the number of persons resident in the subdistrict according to the 2000 Population Census of Thailand.

Statistical methods

Poisson distribution

The Poisson distribution for a random variable X is defined as

$$P(X=k)=\frac{\lambda^k}{k!}e^{-\lambda},\quad k=0,\ 1,\ 2,\ldots \qquad (2.1)$$

where λ is density or mean number of event occurrences per unit of time and/or space. In our study of disease occurrence, this unit is taken as a subdistrict of a province during a calendar year. The Poisson distribution arises from observations that are assumed to be statistically independent. It has mean λ and variance λ.

Negative binomial distribution

The negative binomial distribution for a random variable X is defined as

$$P(X=k)=\binom{r-1+k}{r-1}\rho^r(1-\rho)^{r-1+k},\ k=0,\ 1,\ 2,\ldots \qquad (2.2)$$

where $0<\rho<1$ and $r>0$. It has mean $r(1-\rho)/\rho$ and variance $r(1-\rho)/\rho^2$.

If $\lambda = r(1-\rho)/\rho$ and $\alpha = (1-\rho/r)/(1-\rho)$, so that $\rho = 1-\lambda/(1+\lambda\alpha)$ and $r = 1-\lambda+\lambda\alpha$, it follows that X has mean λ and variance $\lambda(1+\lambda\alpha)$, and the poisson distribution is the special case of the negative binomial distribution when $\alpha = 0$. Thus the negative binomial distribution is a generalisation of the Poisson distribution that has greater variability.

It can be shown using statistical theory that as λ tends to infinity the distribution of the standardised random variable $Z = (X-\lambda)/\sqrt{(\lambda+\lambda^2\alpha)}$ has a left-shifted Gamma distribution with mean 0, scale parameter $1/\sqrt{\alpha}$, and shape parameter $1/\alpha$. This distribution has probability density function

$$, z > -1/\sqrt{\alpha}. \qquad (2.3)$$

Generalised linear models were used to fit these distributions to the number of cases reported in subdistricts for a specified year, using the population in the subdistrict as the denominator.

P-values

The p-value is the probability of observing an event at least as extreme as the specific event observed, based on a specified statistical assumption about the distribution of events. In our study, a p-value was associated with the incidence rate for each disease in each subdistrict. If a p-value is sufficiently small, it indicates that an unlikely event has occurred. This concept may be useful for the control of diseases, because a sufficiently small p-value should trigger an alarm, resulting in intervention such as emergency relief.

However, given that a separate p-value may be calculated for each subdistrict, these p-values need to be adjusted to ensure that the false alarm rate (the statistical type I error rate, conventionally 0.05) remains small. Using Bonferroni's rule, a p-value then needs to be smaller than 0.05/n to be statistically significant, to ensure that the overall type I error rate is close to 0.05, where n is the number of subdistricts.

If β denotes the case incidence rate per 1,000 persons at risk for a specified disease in a specified year, assuming that cases are independent, it follows that the number reported in subdistrict i has a poisson distribution with parameter $\lambda_i = \beta N_i/1000$, where N_i is the population of the subdistrict. The (unadjusted) p-value is then given by the formula

$$p_i = \sum_{j=n_i}^{\infty}\frac{\lambda_i^j}{j!}\exp(-\lambda_i), \qquad (2.4)$$

where n_i is the number of infected cases observed in subdistrict i.

If the independence assumption is not satisfied (due to geographical clustering of cases, say) the distribution of the number of observed events in a region over a period of time will be over-dispersed compared to the poisson distribution, so the negative binomial distribution is preferable, and the p-value corresponding to n_i observed events is then given by the formula

$$p_i = \sum_{j=n_i}^{\infty}\binom{r_i-1+j}{r_i-1}\rho_i^{r_i}(1-p_i)^{r_i-1+j}, \qquad (2.5)$$

where $\rho = 1-\lambda_i/(1+\lambda_i\alpha)$ and $r_i = 1-\lambda_i+\lambda_i\alpha$. If n_i is sufficiently large, the asymptotic distribution based on (2.3) may be used as an approximation.

Results

Table 2 shows the average incidence rates of diseases in Yala subdistricts in 2002-2003.

Figure 2 shows the correlations between disease rate incidences in subdistricts in Yala in 2002. Since the distri-

butions of the incidence rates are substantially skewed, they are log-transformed before computing the correlation coefficients. The highest correlations are between conjunctivitis and diarrhoea (0.80), pneumonia and diarrhoea (0.80), and pneumonia and conjunctivitis (0.58). The other correlations are all relatively small, with the exception of that relating DHF to haemorrhagic conjunctivitis (0.46). These co-morbidities are purely statistical and have not yet been supported by biological evidence.[4]

Based on the findings summarized in Table 2 and Figure 2, we constructed some statistical methods for mapping disease data over subdistricts. These methods could be applied to any data of a continuous data type recorded at specified geographical locations in a region. We used *p*-values based on (2.4) or (2.5) to calculate probabilities associated with locations having unexpectedly high morbidity.

Figure 3 shows plots of standardised residuals against scores based on the model distribution after fitting the poisson and negative binomial models. The closeness of the residuals to the line with unit slope indicates how well the model fits the data. The poisson model clearly does not fit the data, but the negative binomial model does.

Table 2 Average incidence rates/1,000 of diseases in Yala subdistricts

Disease	2002	2003
Dengue haemorrhagic fever	2.160	0.642
Dengue fever	3.295	0.619
Haemorrhagic conjunctivitis	10.282	4.070
Diarrhoea	31.965	35.368
Pneumonia	5.928	6.375
Pyrexia of unknown origin	9.530	9.081
Malaria	1.941	3.773

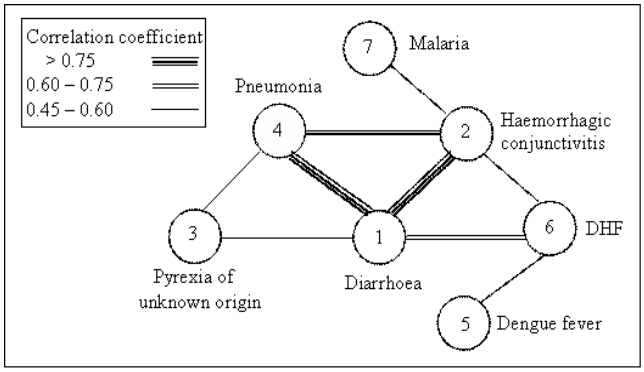

Figure 2 Correlation structure in the disease rates for 2002

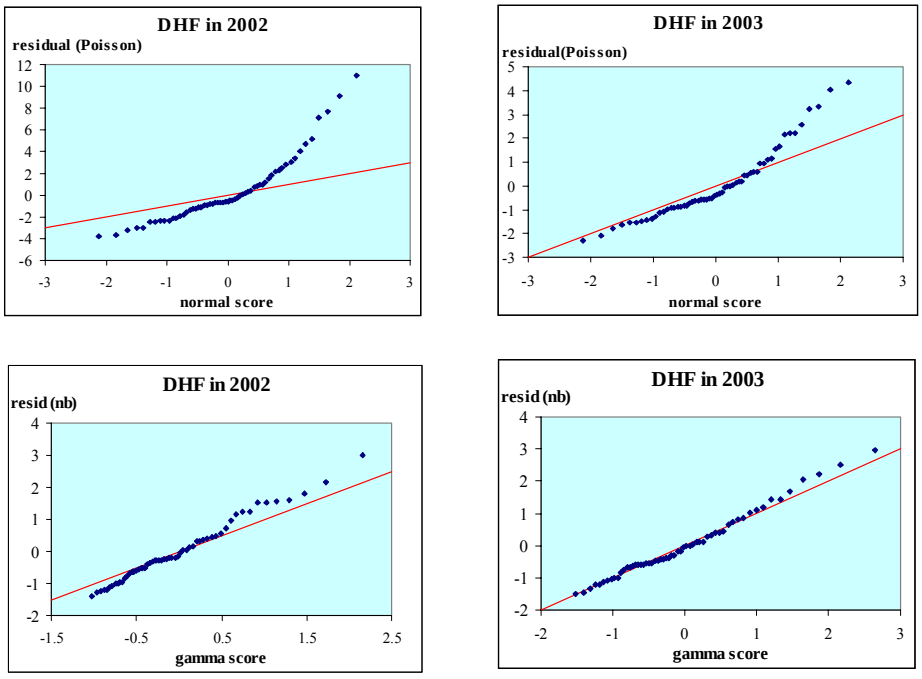

Figure 3 Standardised residuals plots for DHF in Yala subdistricts

Table 3 Estimated parameters in negative binomial model for disease incidences

Disease	2002		2003	
	β	α	β	α
DHF	2.183	0.422	0.644	0.245
	(1.82-2.62)	(0.27-0.65)	(0.54-0.77)	(0.11-0.56)
Dengue	3.224	0.434	0.617	0.522
	(2.69-3.87)	(0.29-0.65)	(0.49-0.78)	(0.30-0.92)
Conjunctivitis	10.234	1.604	4.046	0.917
	(7.34-14.22)	(1.14-2.26)	(3.14-5.21)	(0.63-1.34)
Diarrhoea	31.937	0.704	35.331	0.276
	(25.71-39.68)	(0.50-0.99)	(30.82-40.50)	(0.19-0.39)
Pneumonia	5.929	0.509	6.371	0.215
	(4.91-7.17)	(0.34-0.75)	(5.61-7.24)	(0.14-0.33)
Pyrexia	9.495	0.922	9.050	0.753
	(7.39-12.20)	(0.66-1.30)	(7.21-11.35)	(0.54-1.06)
Malaria	1.944	2.942	3.773	3.700
	(1.24-3.05)	(2.02-4.28)	(2.29-6.21)	(2.57-5.32)

Table 3 shows the parameter estimates and 95% confidence intervals after fitting the negative binomial model to the DHF incidences. Note that the estimates of the parameter β are similar to the values given in Table 2, which occurs because this parameter is the mean of the distribution. The slight differences are due to the fact that the estimates in Table 3 are obtained using maximum likelihood instead of the method of moments.

Alert maps of disease incidence

We used three alert levels to code the subdistricts, based on p-values calculated from the tail of the distribution. For simplicity we used the limiting gamma distribution to compute the p-value associated with a subdistrict. Small p-values were then adjusted for multiplicity by multiplying by n (the number of subregions) as suggested above. In this case n = 58. If the adjusted p-value was greater than 0.05 for any subdistrict the alert level was "None" (blue), if between 0.01 and 0.05 the alert level was "Moderate" (orange) and if less than 0.01 it was "High" (red). Applying this definition, we found that only five alerts were recorded for any disease in 2002 or 2003. There were orange alerts for KaYuBoKo (adjusted p = 0.030) and KuTaBaRu (p = 0.042) for dengue fever in 2002, and for Kerikat (p = 0.017) and ThanTo (p = 0.036) for pyrexia in 2003. There was just one red alert, for KaYuBoKo (p = 0.003) for dengue fever in 2003. Figures 4 and 5 show the corresponding maps for dengue fever and pyrexia.

Discussion

Our basic map region was the subdistrict, so we could not compare disease rates in villages. For example, there was high dengue haemorrhagic fever incidence in BaLa subdistrict in 2002. This could have been due to a single village having many cases of DHF or several villages each having a smaller number of cases. Also, the volatility of incidence rates with time has led some investigators to conclude that there might be a cycle of disease incidence, or possibly environmental factors that wax and wane in the subdistricts. However, Thammapalo, et al.[4] found no evidence of any such effect for DHF.

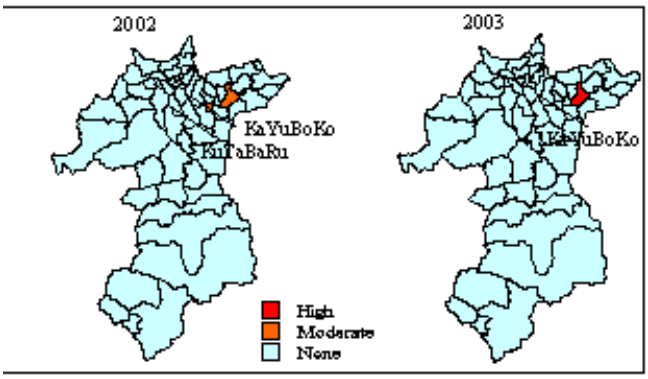

Figure 4 Alert maps of Dengue fever in 2002 and 2003

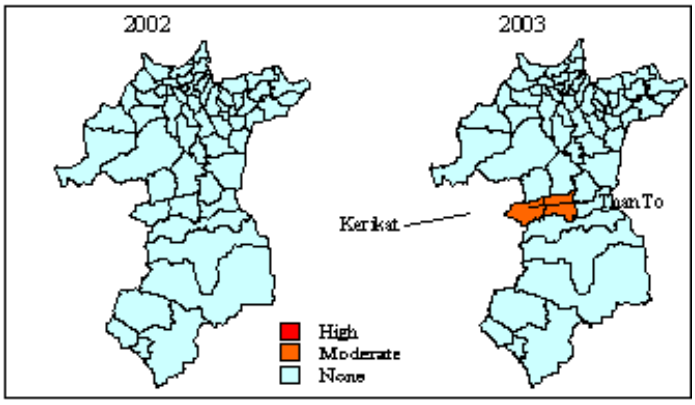

Figure 5 Alert maps of pyrexia in 2002 and 2003

A study of Fenn, et al.[5] found that diarrhoea and pneumonia co-existed among children less than five years of age, and increased severity of disease was associated with higher co-morbidity. While our study provides some further evidence for such correlations, biological explanations have yet to be elucidated. Adenovirus infections may link respiratory illness, pneumonia and diarrhoea.[6-7]

Spatial geographical patterns are associated with the distribution of the incidence of dengue haemorrhagic fever and other diseases. This is consistent with results reported by Luemoh, et al.[8]

Small area maps of disease incidence and growth in incidence are important because they facilitate possible explanations and hypotheses of spatially varying risk factors, including environmental factors such as swamps, chicken farms, industrial waste dumps, and population density and other socio-economic and demographic factors associated with the residents.[9]

Strickman, et al.[10] also found that the incidence of dengue fever increased in small areas. Gubler[11] felt that dengue epidemics might be caused by secondary infections.

However, using statistical methods alone to analyse epidemics and their possible risk factors is not sufficient since important geographic information could be missed. GIS methods are also needed to help researchers to visualize spatial pattens.[12] It is important to note that the risk alert concept we have introduced cannot be related to the overall incidence rate in any given year, because it focuses only on the statistical

significance of unusually high disease incidences relative to the overall mean. Future studies need to develop similar models for data over periods of longer duration than one or two years. The statistical modelling also needs to take other important disease risk factors, such as age and season, into account.

Conclusion

In this study, we used the negative binomial distribution to model the incidence rate of commonly occurring acute febrile illnesses (dengue fever, dengue hemorrhagic fever, malaria, hemorrhagic conjunctivitis, diarrhea, pneumonia and pyrexia of unknown origin), in subdistricts of Yala province in 2002 and 2003, and we computed *p*-values for risk alerts based on this model. The results gave five risk alerts- dengue fever had moderate risk alerts in two subdistricts in 2002 and a high risk alert in one subdistrict in 2003, while pyrexia had moderate risk alerts in two subdistricts in 2003. The method could be applied more generally, and further studies are needed, over longer periods, involving more provinces, and taking into account known risk factors including age, season, environmental factors, and time-series autocorrelations.

Acknowledgements

We would like express our gratitude to Don McNeil, Phattrawan Tongkumchum and Apiradee Saravisupra for their invaluable assistance, encouragement and helpful guidance. Thanks are also due to the Yala Public Health Provincial Office for allowing us to use the disease data from their office.

References

1. Tiprat K, Prasongtang R. Situation of DHF 2003. Weekly Epidemiological Surveillance Report 2004. 2003;35: 230-2.
2. Report of DHF situation (In Thai). Bangkok: Division of Communicable Diseases Control, Thailand; 2003.
3. The situation of dengue haemorrhagic fever in Yala province (In Thai). Yala: Yala Public Health Provincial Office; 2003.
4. Fenn B, Morris SS, Black RE. Comorbidity in childhood in northern Ghana: magnitude, associated factors, and impact on mortality. Int J Epidemiol 2005;34:368-75.
5. Thammapalo S, Chongsuwiwatwong V, McNeil D, Geater A. The climatic factors influencing the Occurrence of dengue hemorrhagic fever in Thailand. J Trop Med Public Health 2005;36:191-6.
6. Jong JC. Epidemiology of enteric adenoviruses 40 and 41 and other adenoviruses in immunocompetent and immunodeficient individuals. Perspect Med Virol 2003; 9:407-45.
7. Crabtree KD, Gerba CP, Rose JB, Haas CN. Waterborne adenovirus: a risk assessment. Water Sci Technol 1997; 35:1-6.
8. Luemoh A, McNeil D, Kuning M. Water consumption and distribution of dengue larvae in Pattani villages. Songkla Med J 2003;21:209-16.
9. Kleinschmidt I, Sharp B, Mueller I, Vounatsou P. Rise in malaria incidence rates in south Africa: a small area spatial analysis of variation in time trends. Am J Epidemiol 2002;155:257-64.
10. Strickman D, Sithiprasasna R, Kittayapong P, Innis BL. Distribution of dengue and Japanese encephalitis among children in rural and suburban Thai villages. Am J Trop Med Hyg 2000;63:27-35.
11. Gubler DJ. Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. Trends Microbiol 2002;10:100-3.
12. Keola S, Tokunaga M, Tripathi NK. Spatial surveillance of epidemiological disease: a case study in Ayutthaya province, Thailand. GIS Development 2002;6:41-4.