An exploration of spatial patterns of seasonal diarrhoeal morbidity in Thailand

B. J. J. McCormick*, W. J. Alonso and M. A. Miller

Fogarty International Center, National Institutes of Health, Bethesda, USA

(Accepted 30 August 2011; first published online 19 September 2011)

SUMMARY

Studies of temporal and spatial patterns of diarrhoeal disease can suggest putative aetiological agents and environmental or socioeconomic drivers. Here, the seasonal patterns of monthly acute diarrhoeal morbidity in Thailand, where diarrhoeal morbidity is increasing, are explored. Climatic data (2003–2006) and Thai Ministry of Health annual reports (2003–2009) were used to construct a spatially weighted panel regression model. Seasonal patterns of diarrhoeal disease were generally bimodal with aetiological agents peaking at different times of the year. There is a strong association between daily mean temperature and precipitation and the incidence of hospitalization due to acute diarrhoea in Thailand leading to a distinct spatial pattern in the seasonal pattern of diarrhoea. Model performance varied across the country in relation to per capita GDP and population density. While climatic factors are likely to drive the general pattern of diarrhoeal disease in Thailand, the seasonality of diarrhoeal disease is dampened in affluent urban populations.

Key words: Climate – impact of, diarrhoea.

INTRODUCTION

Despite therapeutic advances and improved social standards the global burden of enteric and diarrhoeal disease continues to be a leading cause of morbidity and mortality in young children worldwide [1–3]. As an illustration of this, the mortality rate due to acute diarrhoea, defined roughly as >2 watery stools within 24 h, has dropped substantially in Thailand during the past three decades. The number of hospitalizations has, however, increased from the late 1970s to the present (Fig. 1).

Different aetiological agents causing diarrhoea have been shown to circulate with distinct seasonal patterns: bacterial diarrhoea tends to occur most frequently in warmer and wetter months when environmental conditions favour the faecal–oral spread of pathogens [4–7]. In contrast, viral diarrhoea exhibits seasonal patterns similar to respiratory infections such as influenza, peaking in colder weather [8–10]. Climatic patterns have been implicated in driving the seasonal pattern of diarrhoeal disease in several temperate countries [11–14] although not in tropical climates where climatic seasons are less distinct [5, 15, 16]. Comparable studies describing patterns of morbidity or mortality in countries with monsoon weather patterns that do not, for example, exhibit a pronounced summer and winter so much as hot and wet seasons remain rare. Previous studies of northern

* Author for correspondence: Dr B. J. J. McCormick, Fogarty International Center, National Institutes of Health, Bethesda, MD 20892, USA.
(Email: ben.mccormick@nih.gov)
Thailand identified strong seasonal patterns associated with monsoon rainfall patterns, but did not comment on the seasonal patterns in the more tropical areas of the same country [17, 18].

In this study, the spatial trends associated with the seasonal dynamics of acute diarrhoeal morbidity are investigated across the whole of Thailand. In addition to describing the spatial pattern of seasonal diarrhoea, the model is tested against a generalized description of the climate providing an opportunity to assess the predictive power of the model. Thailand is particularly interesting as a case study with which to assess the role of climate in driving patterns of diarrhoeal disease since it not only has monsoon weather patterns in the north of the country, but also more stable tropical weather in the south thus representing a diverse range of climatic conditions that might be expected to contribute to a diverse range of diarrhoeal disease incidence.

DATA

Morbidity data

Diarrhoeal morbidity data, defined as the number of hospitalizations, due to acute diarrhoea (ICD-10: A09), enteric fever (typhoid ICD-10: A01.0; paratyphoid ICD-10: A01.1-3, unspecified) and dysentery (bacterial ICD-10: A03 and amoebic ICD-10: A06 and unspecified) were extracted from the Thai Ministry of Health annual reports, 2003–2009 [19]. Data were divided by changwat (the first sub-national administrative unit) and month resulting in 76 geographical units, each with 84 observations per causal agent. Data were input to a GIS and associated with social and physical data. The number of cases per month were standardized into cases/10 000 using the annual population per changwat and log-transformed.

Social data

Social data were taken from a combination of Thai Ministry of Statistics data hosted on the South-East Asian Health and Life Statistics Database [20] (population density/km² and percentage of the population living in municipal areas) and the Thai Household Socioeconomic survey (2000) (per capita GDP).

Environmental data

Climatic data were extracted from the CRU TS 3.0, 0.5° global datasets [21] and the median of each pixel containing the changwat centroid and contiguous pixels were calculated for monthly mean daily temperatures and monthly precipitation for the period 2000–2006. Temperature and rainfall have both been extensively related to diarrhoeal disease [11, 12, 17, 22–24] and are highly correlated with other climatic variables available from the CRU data (minimum, maximum and diurnal range in temperature, the frequency of wet days, and vapour pressure).

Physical data

The median elevation of the changwat was similarly extracted using the centroid pixel and the surrounding pixels. Elevation data was supplied by the CRU.

METHOD

Diarrhoeal disease data was available from 2003 to 2009 inclusive, whereas climatic data was only available until mid-2006. A spatially weighted panel regression [25–27] was therefore fitted for the period 2003–2005 (inclusive) and then used to forecast the diarrhoeal hospitalizations for 2006–2009, assuming a continuation of the climatic pattern and taking the spatial dependence from the first period.

The spatially weighted panel regression was used to compare detrended temperature and precipitation data to detrended diarrhoeal disease data. A temporal lag of 1–11 months was included for both mean daily temperature and mean monthly precipitation.
To assess nonlinear relationships between climate and diarrhoeal disease quadratic terms were included in addition to an interaction term between temperature and precipitation. Following the fitting of a full model, terms were re-assessed and non-significant terms removed.

The spatial panel regression assumes a form of

\[ \ln(\text{inc}_{it}) = \alpha + \sum_{i} \beta_i X_{nit} + u_i + e_{it}, \]

where \( \text{inc}_{it} \) is the incidence of diarrhoea mortality in state \( i \) at time \( t \); \( X \) is a series of independent variables at state \( i \) and time \( t \); \( \alpha \) and \( \beta \) are the regression parameters (global mean and predictor coefficients respectively); \( u_i \) is an error term that captures the state-specific unobserved effects and \( e_{it} \) represents all other unobserved effects (assumed to be normally distributed and independent). Spatial effects were based on a weighting, calculated using the inverse distance between \textit{changwat} centroids.

In order to forecast the remaining acute diarrhoeal hospitalization data and test the fit of the spatial panel regression model, it was necessary to extend the climate data beyond mid-2006. The whole climate time-series (2000–2006) was used to fit Fourier harmonics to the temperature and precipitation data [28, 29]. A combination of an annual and bi-annual harmonics reproduced the stationary (i.e. with no overall trend) seasonal pattern in both climatic variables, generalizing the seasonal pattern as an approximation of an ‘average’ year.

Regression parameters relating the mean daily temperature and mean monthly precipitation along with the spatial coefficients based on the spatial dependence of the monthly mean of the 2003–2005 disease incidence data were applied to the extended Fourier-processed climate data to forecast the acute diarrhoeal hospitalization incidence.

Deficiencies in model fit were compared to social factors that might account for variability between \textit{changwats} that is not captured by climatic variables. Analyses, graphics, and statistical tests were conducted in Matlab (The MathWorks Inc., USA). Spatially weighted panel regressions were performed using the Spatial Econometrics toolbox [27].

**RESULTS**

There was marked seasonality in national diarrhoeal hospitalization resulting from all the different...
aetiologies (Fig. 2). The timing of peak hospitalization was different between the three broad categories; however, for typhoid/paratyphoid there were dominant and unimodal peaks around July; for dysentery, peaks were bimodal, larger in May and smaller in January, although there was a trend of decreasing magnitude, which was more pronounced in the May peaks; for acute diarrhoea there were equally bimodal peaks around July and January.

The number of hospitalizations per month for acute diarrhoea (i.e. unspecified aetiology) substantially outweighed any of the defined causes. Further analysis was only performed for acute diarrhoea, for which there was a sizable number of cases per month per changwat.

Spatial patterning was not visually observed in the mean hospitalizations for different changwats, although there was visual evidence of spatial patterns in the timing of peak incidence. First, it is worth noting that not all changwats had a similar seasonal pattern (Fig. 3), only those changwats that had an early (i.e. around January–February) peak in acute diarrhoeal hospitalizations also showed a peak around June. Changwats with peak mortality in May tended not to have secondary peaks. The unimodal changwats tended to be in the periphery of Thailand, either the extreme north or south of the country. The timing of peaks tended to be earliest in central Thailand, around Bangkok and to the north, which is the area with lowest elevations, then progressively later in the year in areas of higher altitude. By the time that the highlands in the north of Thailand had peaks (around May) the lowlands began a second wave of diarrhoeal hospitalizations.

A spatial panel regression model was made of contemporary acute diarrhoeal disease and climatic data (2003–2005). The full model, using both temperature and precipitation with lags of 1–11 months, achieved an $r^2_{adj} = 0.597$. While a number of both daily mean temperature and mean monthly precipitation lagged months were statistically significant, they tended to cluster around 9–10 months earlier for temperatures and 1–3 months earlier for precipitation relative to the diarrhoeal disease data (Table S1, available online). The reduced model (Table 1) achieved $r^2_{adj} = 0.544$.

The climatic data was projected using annual and bi-annual Fourier harmonics to match the 2006–2009 acute diarrhoeal mortality data. The average climatic profile for each changwat was projected from 2000 to 2006 up to the end of 2009. Assuming that the spatial structure was the same as the mean of the earlier period, and matching to the seasonal profile of acute diarrhoeal disease.
diarrhoeal hospitalization data, derived from Fourier harmonics, achieved a model with an overall $r_{adj}^2 = 0.462$.

The overall model’s goodness of fit could be partitioned to explore where the model deviated from the observed data (Fig. S1, available online). The patterning of goodness of fit was comparable between the models with contemporary and projected climatic data, and performance tended to be good ($r^2 > 0.7$) in the north of Thailand and worst ($r^2 < 0.3$) in the area around the Bangkok metropolis. Where the performance of the fitted and extrapolated model differed there was generally a shift in the seasonal pattern over the time period, for example, the hospitalization incidence in Bangkok shifted from a broadly unimodal peak around January to the presence of a secondary peak in June (Fig. 4). Similarly, Uthai Thani in the south of Thailand demonstrates a shifting seasonal pattern, which not only shows increasing magnitude between minimum and maximum values, but also has a broadening peak (from a peak in July to a peak extending between January and July).

The deficiences of the model to accurately predict the hospitalizations due to acute diarrhoea show a geographical pattern (Fig. 4). None of the social factors proved significant in the spatially weighted regression model and did not improve the fit compared to the climate variables alone (e.g. the fitted climate model with per capita GDP achieved an $r_{adj}^2 = 0.543$). This result is unsurprising given that the social variables provide a single snapshot in time, whereas the climatic variables better capture the temporal dynamics of diarrhoeal disease. The changwat $r^2$ value was, however, significantly correlated with the per capita GDP ($P < 0.001$ for both contemporary and projected climate models, Fig. 5a), where GDP is lower the model accuracy tends to be higher. Additionally, areas of Thailand with higher per capita GDP tended to have reduced diarrhoeal disease seasonal amplitude ($P < 0.001$, Fig. 5b). Per capita GDP correlated positively with both population density and the proportion of the population in urban areas, although it had higher correlation with the acute diarrhoea data. The urban population of Thailand is predominantly found in the central region around and to the north of Bangkok. The highland areas to the north of the country and to the west of Bangkok are more rural and have lower GDP and it is these areas that have greater correspondence between climate and acute diarrhoeal hospitalizations. The south of Thailand can be characterized in this instance by lower population densities and higher per capita GDP than the north and these changwats are also those which tend to have poorer model accuracy.

Peak acute diarrhoeal hospitalizations tended to be associated with the minimum precipitation, lagging by about 1 month (Fig. S2, online). Peak incidence preceded the peak temperatures by about 3 months. In both cases this corresponded to the significant lags in spatial panel regression.

**DISCUSSION**

The seasonality in the annual incidence of diarrhoeal disease in Thailand suggests a role for climatic drivers, as have been implicated in other countries [11, 12, 14, 30]. The degree to which climate plays a role in determining the timing of seasonal diarrhoea is less well characterized in countries with a monsoon weather pattern, however, as much work has concentrated on temperate regions of the world.

There is evidence of differences in the seasonality between the bacterial and protozoan causes of diarrhoeal disease and viral causes; however, in the absence of detailed viral aetiology for Thailand it is not possible to further characterize these differences other than to conjecture that the peak hospitalizations in January categorized as acute diarrhoea do not correspond to either the peaks of typhoid/paratyphoid or of bacterial/amoebic dysentery.

A spatially weighted panel regression model using just the daily mean temperature and monthly mean precipitation performed well in predicting the incidence of hospitalization due to acute diarrhoea in Thailand. The pattern of acute diarrhoea in Thailand varies by changwat, although overall it is strongly

### Table 1. Reduced spatial panel regression model of acute diarrhoeal disease

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>$t$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature ($t_n$)</td>
<td>0.358</td>
<td>13.065</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Temperature ($t_{n-1}$)</td>
<td>-0.007</td>
<td>-12.01</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Temperature ($t_{n-2}$)</td>
<td>0.04</td>
<td>12.846</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Temperature ($t_{n-3}$)</td>
<td>0.014</td>
<td>5.159</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Temperature × precipitation</td>
<td>0.002</td>
<td>3.54</td>
<td>0.176</td>
</tr>
<tr>
<td>Precipitation ($t_n$)</td>
<td>-0.027</td>
<td>-0.86</td>
<td>0.392</td>
</tr>
<tr>
<td>Precipitation ($t_{n-1}$)</td>
<td>-0.009</td>
<td>-2.967</td>
<td>0.003</td>
</tr>
<tr>
<td>Spatial weight</td>
<td>0.01</td>
<td>24.012</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

$t_{n-x}$ is the temporal lag of $n$ months.

Variables in bold are significant ($P < 0.05$).
related to the seasonal patterns in both temperature and precipitation, independently and by interaction. Diarrhoeal hospitalization was highest first around January, towards the end of the dry season when temperatures are comparatively low, and second around July at the beginning of the rainy season and when temperatures drop (after the hot season, roughly March–June). When the transition between the hot and the rainy seasons is less distinct there is merging of the two diarrhoeal peaks around May–June [e.g. in Uthai Thani, Figs 5 and S2 (online)].

Where the population was least dense or urbanized, the same places also tended to have the lowest per capita GDP, the model performance using climatic factors alone was highest. In more affluent and urbanized populations the model predictions diverged from the observed pattern of diarrhoeal disease. A possible parallel, hinted at by the relationship

Fig. 4. Climatic model fitting. (a) Example of seasonal profiles for three changwats (north to south: Chiang Rai, Bangkok, Uthai Thani) showing the model fit from the contemporary climate data (circles) and projected climate (squares) along with the observed acute diarrhoea hospitalizations (black curve). Model $r^2$ for each changwat from (b) the contemporary and (c) projected climate indicated in the colour scale (dark to light indicating $r^2 = 0.0$).

Fig. 5. (a) Comparison of the $r^2$ goodness of fit for the models with contemporary climate data (crosses and solid line) and projected climate (dots and dotted line) and the per capita GDP for each changwat. (b) Comparison of the amplitude of the seasonal pattern (2003–2009) for each changwat with the per capita GDP.
between the amplitude of diarrhoeal disease incidence and GDP, is that urban areas are better able to dampen the seasonal variability of diarrhoeal disease as driven by climate. Urban areas may, for example, have greater investment in sanitation and are thereby better able to dispose of faecal material [17]. The results presented here do suggest that factors other than climate begin to play an important role in determining the seasonal pattern of diarrhoeal disease in urban Thailand.

Previous models of diarrhoeal disease in Thailand have suggested that there is considerable under-reporting of cases [18, 31]. This finding is not surprising as diarrhoeal disease can become routine in the local population such that it is not necessarily reported to national surveillance programmes [32]. The decreasing mortality from diarrhoeal disease in Thailand may also suggest that severe cases have also diminished even though incidence has not, thereby further reducing the reporting of cases. The method of correcting for under-reporting used in those models [18, 31] was to re-fit the model using interpolated values rather than supposedly spurious records. In this study such an iterative procedure was not used, a possible consequence being that the model predictions are ‘less wrong’ than they appear compared to questionable observation data. This is particularly the case with very low observed incidence, for which the model generally predicted higher hospitalization rates. While zero incidences are improbable, the regularly low number of cases between October and December is unlikely to be underreporting as much as less favourable conditions for diarrhoeal disease transmission [31].

There is little evidence to directly relate the climatic conditions at large spatial scales with causal mechanisms of diarrhoeal disease. There is, however, evidence of the effects of micro-climatic temperature and humidity on the replication and survival of bacterial pathogens [33, 34] and thereby the risk of exposure to pathogens [6, 35]. The indirect effects of rainfall have been suggested to be both beneficial and detrimental to bacterial transmission [17]; e.g. a combination of increased bacterial survival during periods of high humidity, and also improved faecal disposal and hygiene practices when water is more abundant [6, 36]. Extrapolating an assumed relationship between climatic factors and diarrhoeal disease burden is subject to the compounding errors of aggregating heterogeneous data [37], although identifying putative determinants at a course scale is likely to increase the targeted research necessary to bridge the gap in resolution.

Extending the understanding of possible causal mechanisms for disease patterns is a useful step to devising appropriately targeted interventions. In Thailand the increasing morbidity associated with diarrhoeal disease is a cause for concern and this study extends the previously observed spatial patterning from identification of ‘season’ as a predictor of disease burden [18, 31] to a possible early warning forecast using meteorological conditions.

NOTE
Supplementary material accompanies this paper on the Journal’s website (http://journals.cambridge.org/hyg).

ACKNOWLEDGEMENTS
We thank Ramya Ambikapathi for her assistance in locating appropriate disease data and Cynthia Schuck-Paim for insightful comments on the manuscript. This study was funded as part of the Malnutrition and Enteric Diseases (Mal-ED) project, sponsored by the Bill and Melinda Gates Foundation and the Fogarty International Center of the National Institutes of Health.

DECLARATION OF INTEREST
None.

REFERENCES
34. Wallis CM, Howard K, Webster AJF. The survival of *Escherichia coli* in an aerosol at air temperatures of 15 °C and 30 °C and a range of humidities. *Epidemiology and Infection* 1986; 97: 489–496.