SELECTED ISSUES ON HEALTH AND ENVIRONMENT

EDITORS
Narimah Samat
Zuliyadini A. Rahaman
Shariffah Rohayah Syeikh Dawood
Wan Ruslan Ismail
Ruslan Rainis
SELECTED ISSUES ON HEALTH AND ENVIRONMENT: EDITORS:
Narimah Samat, Zuliyadini A. Rahaman, Wan Ruslan Ismail,
Sharifah Rohayah Syeikh Dawood, Ruslan Rainis
1. Environmental health, 2. Environmental management.
I. Narimah Samat. II. Zuliyadini A. Rahaman. III. Wan Ruslan Ismail.
IV. Sharifah Rohayah Syeikh Dawood, V. Ruslan Rainis.
363.7
1

GEOGRAPHIC INFORMATION SYSTEM (GIS) AND PREDICTIVE RISK MAP SOIL-TRANSMITTED HELMINTHIASIS IN PENINSULAR MALAYSIA

Aziz Shafie, Romano Ngui, Chua Kek Heng, Mohd Aidil Roslan, Wan Yusoff Wan Sulaiman & Yvonne Al Lian Lim

INTRODUCTION

In Malaysia, soil-transmitted helminth (STH) infections are considered largely controlled with significant reduction of infection rates particularly among urban populations (Aaron et al, 2011). However, this reduction trend remains significantly unchanged with high prevalence rates and significant morbidity among communities in rural and remote areas (Lim et al, 2009; Aaron et al, 2011). Lim et al. (2009) summarized studies that have been conducted since colonial era in Malaysia demonstrated that foci for high endemcity remain largely unchanged with alarming high prevalent rates, in some cases up to 100% in these rural dwellers. Although Malaysia is still known to have high prevalence of STH infections particularly in rural dwellers (Lim et al, 2009), this disease is recognized as not notifiable by the local public health authorities. Moreover, a precise estimate of the total disease burden has not been fully described as collation of systematic information on STH infections in the country is not currently available. Most of the information or record on the prevalence of STH infections is scattered across the literature and not catalogued systematically. These data are seldom available in an accessible format for policy makers or public health authorities.

In recent years, the geographical information system (GIS) and remote sensing (RS) has been widely used for effective storage, mapping, analysis and development of STH atlas (Brooker & Michael, 2000). Such approach also made data integration and mapping more accessible and reliable. It also offers us the ability for modeling the spatial distribution of STH infections in relation to their ecological factors which are derived from remote sensed (RS) satellite data that are known to influence the distribution pattern, thus deepening our knowledge and understanding in the biology and epidemiology of the infections (Hay, 2000; Brooker et al, 2006). Likewise, it also allows us to predict the spatial distribution of infection and identify endemic areas, thus providing more precise estimates of populations at risk (Brooker & Michael, 2000). By extending such approach to Malaysia, a reliable and accessible GIS database consists of prevalence map and information of environmental and ecological factors that
correlate with their distribution can be explained. The correlations between infection patterns and ecological factors can be used to extrapolate predictive risk maps in areas for which no data are available. In addition, the predictive risk maps can serve as a baseline data to estimate number of population at risk, numbers requiring treatment and cost of delivering anthelmintics. Establishment of such reliable map is essential for the development and implementation of control measures to those populations in greatest need particularly when the resources for control program are finite and limited. Therefore, findings of the current study will be valuable for the public health authorities to justify and facilitate the reassessment of the existing control measures to reduce the prevalence of STH infections.

MATERIALS AND METHODS

Data searches and geo-positioning procedures

Relevant information on the prevalence of soil-transmitted helminth (STH) infections in Malaysia were identified through a combination of (i) an extensive search in electronic bibliographic databases, (ii) manual search of local archives and libraries and (iii) direct contact with local researchers. In addition to these data, we also conducted field investigations from January 2009 until December 2012 in randomly selected locations in Peninsular Malaysia. The geographic coordinates of various surveyed locations were determined using combination of various electronic resources including GeoNet Names Server (http://earth-info.nga.mil), Google Earth (http://www.google.com), Wikimapia (http://www.wikimapia.org), Maplandia (http://www.maplandia.com) and Tago (http://www.tago.com). Each of the identified locations from one source was consequently cross-checked against other sources to ensure consistency of the identified coordinates. For the field investigation survey conducted by our group, the spatial location (i.e., in situ data collection) of the each survey location was recorded using handheld Garmin GPSMAP 60CSx. All the digital data coordinate systems were synchronized using World Geodetic System (WGS 1984) which serve the x (longitude or east-west) and y (latitude or north-south).

Geographic Information System (GIS) and Remote Sensing (RS) database

Peninsular Malaysia comprises 11 states and each state is divided into 81 districts, which are then further divided into the smallest administrative level defined as mukim or sub-district (Peninsular Malaysia comprises 842 mukim). Each of the extracted survey data were then mapped at the sub-district levels. The sub-district in which each survey has been conducted was identified and linked to the Peninsular Malaysia boundary map referred as base map, which was obtained from Department of Surveying and Mapping, Malaysia. A set of environmental variable was gathered from a variety of sources. Monthly averages Land Surface Temperature (LST) at 1 km resolution was derived from the WorldClim (http://www.worldclim.org). Briefly, these data were generated from global weather station temperature records gathered from various sources for
the period of 1950 to 2000. A thin-plate smoothing spline algorithm was then used to interpolate the data, following the approach of Hijmans et al. (2005). Normalized Difference Vegetation Index (NDVI) data was obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) (Scharlemann et al, 2008). Other ecological covariates such as altitude was obtained from the interpolated Digital Elevation Model (DEM) from Department of Surveying and Mapping, Malaysia. For each of these environmental variables, minimum, mean and maximum values were extracted for each pixel that corresponded to the survey locations.

### Statistical analysis, mapping and modeled STH distribution

Geographic information system (GIS) application tool using ArcGIS 9.3 software (ERSI, Redlands, CA, USA) were used to integrate and analyze survey data and environmental variables that were derived from remote sensing (RS) satellite data. The statistical analysis for each test was performed using SPSS software (Statistical Package for the Social Sciences) program for Window version 17 (SPSS Inc, Chicago, USA). Logistic regression model was used to identify significant environmental variables that are known to influence the transmission of infections and for developing statistical risk models (Fielding & Bell, 1997; Pearce & Ferrier, 2000). Following development of statistical risk model, the best fit logistic regression model was then used to generate predictive risk map of the probability of having STH prevalence of more than 50% as recommended by WHO (WHO, 2002).

### RESULTS

Soil-transmitted helminth (STH) has direct life cycles, involving sexual maturation in the human hosts and the free living stages present in the environment. Their development and survival rate are dependent on the surrounding environmental factors such as humidity and temperature. Studies have shown that such environmental factors will indirectly influence their transmission success and spatial patterns of infection (Brooker & Michael, 2000). We have therefore investigated the ecological correlation of STH infections and predict their prevalence in un-sampled areas on the basis of satellite derived environmental data using logistic regression analysis. The results indicated that maximum or mean Land Surface Temperature (LST) and minimum or mean Normalized Difference Vegetation Index (NDVI) were significant explanatory variables for *A. lumbricoides* infection (Table 1.1). The result shown that the odds of *A. lumbricoides* was significantly negatively associated with maximum LST (OR = 0.88; 95% CI = 0.78-0.99) and minimum NDVI (OR = 0.98; 95% CI = 0.96-0.98). Only *A. lumbricoides* model was developed. As for *T. trichiura*, hookworm and estimated combined STH infections, no statistically significant explanatory variables were recorded, thus no model could be developed.
Table 1.1 Regression coefficients used to estimate probability of person being infected in logistic regression model

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Coefficient estimate (B)</th>
<th>Standard error of estimate</th>
<th>OR (95% CI)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ascaris lumbricoides</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>28.992</td>
<td>0.248</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Maximum LST</td>
<td>-0.127</td>
<td>3.942</td>
<td>0.88 (0.78-0.99)</td>
<td>0.047</td>
</tr>
<tr>
<td>Mean LST</td>
<td>0.113</td>
<td>4.019</td>
<td>1.12 (1.00-1.25)</td>
<td>0.045</td>
</tr>
<tr>
<td>Minimum NDVI</td>
<td>-0.020</td>
<td>4.747</td>
<td>0.98 (0.96-0.98)</td>
<td>0.029</td>
</tr>
<tr>
<td>Mean NDVI</td>
<td>0.026</td>
<td>5.619</td>
<td>1.03 (1.01-1.05)</td>
<td>0.016</td>
</tr>
</tbody>
</table>

Following that, these environmental factors (i.e., maximum and mean LST and minimum and mean NDVI) were then used to model and predict the distribution of *A. lumbricoides* infection using best logistic regression model to map probability of infection prevalence being 50% or more. The predictive risk map of *A. lumbricoides* indicated that the prevalence of infection was clustered and higher (i.e., areas within blue color range with prevalence of at least 20% up to 100%) in central and northern plains of Peninsular Malaysia including central Pahang, Kelantan, northern Perak and Kedah particularly in areas which border with southern Thailand. In contrast, predicted prevalence of *A. lumbricoides* was lower (i.e., area within green color range with prevalence of ≤ 20%) along the west coast and southern part of Peninsular Malaysia (Figure 1.1). A visual comparison of our predicted map with observed prevalence map for *A. lumbricoides* showed some similar features and was in agreement. For instance, up to 70% of the surveyed areas with high observed prevalence showed similar levels with our predicted model.
and mean LST and the distribution of sample probability of the A. lumbricoides (i.e., areas within central and northern northern Perak and contrast, predicted a color range with insular Malaysia and prevalence map- ment. For instance, used similar levels

![Image of prevalence map]

**Figure 1.1** Predicted prevalence map of *A. lumbricoides* as derived from logistic regression model of the relation between observed empirical prevalence survey data and remote sensing (RS) - satellite sensor environmental variables.

**DISCUSSION**

Results of the present study indicated significant association between observed prevalence of *A. lumbricoides*, Land Surface Temperature (LST) and Normalized Difference Vegetation Index (NDVI). In addition, the result also indicated negative association between observed prevalence of *A. lumbricoides* with maximum LST and minimum NDVI, in other words as the temperature increases, transmission and prevalence of infection decrease. Such observations are most probably due to the effects of heat and low humidity on the embryonation and survival of ova. The thermal limits of infection as reported in our current work were also in agreement with studies on the mapping of STII distribution in other countries. For instance, finding in Cameroon, Chad and Uganda suggests that *A. lumbricoides* and *T. trichiura* most unlikely to occur in areas where maximum LST exceeds 37°C (Brooker et al, 2002a, 2002b). Similarly, observation in Vietnam also reports low prevalence of *A. lumbricoides* and *T. trichiura*
infections (i.e., less than 10%) in areas where maximum LST above 37°C (Brooker et al., 2003). More recently, mapping of STH infections in Kenya by Pullan et al. (2011) also reported that the odd of *A. lumbricoides* infections was significantly lower with maximum LST. Such thermal limits are also supported by several experimental data. For example, previous studies have reported that the optimal temperature for the embryonation of *A. lumbricoides* ova is 31°C (Seamster, 1950) and do not develop at temperature more than 38°C (WHO, 1967). As for *T. trichiura*, experimental data also shows that the ova requires different days to develop and hatch at different temperature (i.e., takes 28 days to develop at 25°C, 15 days at 30°C and 13 days at 34°C), while do not survive at temperature above 37°C (Beer, 1976). In addition, few studies also suggest that an upper limit of 40°C is considered to be environmental or biological transmission limits for both *A. lumbricoides* and *T. trichiura* infections (Hotez et al., 2003; Pullan et al., 2011). Studies in South Africa showed that the average prevalence of *A. lumbricoides* and *T. trichiura* were below 20% in areas with mean annual temperature below 15°C, postulating that the lower thermal limit might be around 15°C or below for both STH species (Appleton et al., 1999). The presence of vegetation also tends to prevent evaporation and conserve soil moisture, offering favorable condition for the development and transmission of infection (Hotez et al., 2003).

The ability to predict the distribution of infection on the basis of their ecological limits has important consequences for control planning, targeting program at national level. The predictive risk map as illustrated in this study indicated that low prevalence of infection can be found along the west coast and southern part of the country, while high prevalence along the central plain and northern part, suggesting that MDA is most warranted in the central and northern plains of the country, as illustrated by a high probability that infection prevalence exceeds 50%. Moreover, the estimation of uncertainty in unsurveyed area is also another additional advantage of such approach, which takes existing empirical prevalence data to generate continuous map by interpolating prevalence at unsampled location on a grid system. Such approach recognizes error or ambiguity associated with potential biasness in data by generating the possible prevalence value (i.e., probability prevalence for each prediction location). Factors such as lack or low number of survey location, error in survey measurement, data quality and the unavoidable presence of apparently random or variation in prevalence which may result with these uncertainties (Pullan et al., 2011).

CONCLUSION

The present study has successfully mapped and modeled the distribution of soil-transmitted helminthiasis using geographic information system (GIS) and remote sensing (RS) satellite derived environmental data. Such approach offers us the ability to model the spatial distribution of STH infections in relation to the ecological factors which were derived from remote sensed satellite data known to influence their distribution pattern, thus deepening our knowledge and understanding in the biology and epidemiology of infection.
ACKNOWLEDGEMENT

The authors would like to sincerely thank Prof. Simon Brooker and Dr. Rachel Pullan from London School of Hygiene and Tropical Medicine (LSHTM), United Kingdom for their courtesy in providing the Remote Sensing data and also for their input and counsel during the initial stage of this work.

BIBLIOGRAPHY


