

# Mapping and modelling the geographical distribution of soil-transmitted helminthiases in Peninsular Malaysia: implications for control approaches

Romano Ngui<sup>1</sup>, Aziz Shafie<sup>2</sup>, Kek H. Chua<sup>3</sup>, Mohd S. Mistam<sup>4</sup>, Hesham M. Al-Mekhlafi<sup>1</sup>, Wan W. W. Sulaiman<sup>1</sup>, Rohela Mahmud<sup>1</sup>, Yvonne A. L. Lim<sup>1</sup>

<sup>1</sup>Department of Parasitology, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia; <sup>2</sup>Department of Geography, Faculty of Arts and Social Sciences, University of Malaya, Kuala Lumpur, Malaysia; <sup>3</sup>Department of Biomedical Science, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia; <sup>4</sup>Department of Orang Asli Development, Ministry of Rural and Regional Development, Kuala Lumpur, Malaysia

**Abstract.** Soil-transmitted helminth (STH) infections in Malaysia are still highly prevalent, especially in rural and remote communities. Complete estimations of the total disease burden in the country has not been performed, since available data are not easily accessible in the public domain. The current study utilised geographical information system (GIS) to collate and map the distribution of STH infections from available empirical survey data in Peninsular Malaysia, highlighting areas where information is lacking. The assembled database, comprising surveys conducted between 1970 and 2012 in 99 different locations, represents one of the most comprehensive compilations of STH infections in the country. It was found that the geographical distribution of STH varies considerably with no clear pattern across the surveyed locations. Our attempt to generate predictive risk maps of STH infections on the basis of ecological limits such as climate and other environmental factors shows that the prevalence of *Ascaris lumbricoides* is low along the western coast and the southern part of the country, whilst the prevalence is high in the central plains and in the North. In the present study, we demonstrate that GIS can play an important role in providing data for the implementation of sustainable and effective STH control programmes to policy-makers and authorities in charge.

**Keywords:** mapping, geostatistical modelling, soil-transmitted helminths, Malaysia.

## Introduction

Soil-transmitted helminth (STH) infections are among the most widespread human diseases, especially among those trapped in the vicious cycle of poverty and destitution in the tropics and subtropics (Hotez et al., 2009). Of particular importance are the two anthropophilic species of hookworm (*Necator americanus* and *Ancylostoma duodenale*), the roundworm *Ascaris lumbricoides* and the whipworm *Trichuris trichiura* (Chan et al., 1994). Although the annual worldwide mortality due to STH infections varies depending on the source (from 12,000 to 135,000) (WHO, 2005), these infections are the leading causes of malnutrition, growth stunting and cognitive deficit among vulnerable groups, such as children and women of reproductive age (WHO, 2005). It can be

estimated that approximately one third of the world's population (~2 billion people) is currently infected with at least one STH species; 1.2 billion people with *A. lumbricoides*, 800 million with *T. trichiura* and 600 million with hookworm (de Silva et al., 2003; Bethony et al., 2006).

Malaysia has witnessed rapid socioeconomic growth and infrastructure development since independence 57 years ago. Nevertheless, the country is still plagued with parasitic diseases, in particular STH infections (Aaron et al., 2011; Ahmed et al., 2011; Ngui et al., 2011; Nasr et al., 2013). Summarising the studies conducted since the colonial era demonstrate that highly endemic foci remain largely unchanged with alarmingly high prevalence rates in poor and rural dwellers (Lim et al., 2009). The STHs are not notifiable diseases, and after the national deworming programme among school-aged children was discontinued in 1983 (Anon, 1985), little effort has been made towards combatting these infections; one of the reasons may be that a precise account of the total STH burden in the country is not available. Most of the information in this field is scattered across the scientific literature and not systematically catalogued.

---

Corresponding author:

Yvonne A. L. Lim

Department of Parasitology, Faculty of Medicine  
University of Malaya, 50603 Kuala Lumpur, Malaysia  
Tel. +603 7967-4746; Fax +603 7967-4754  
E-mail: limailian@um.edu.my

In recent years, however, there has been renewed interest in STH control from international aid organisations leading to increased momentum in attaining more comprehensive data and allowing available control resources to be more rationally and cost-effectively deployed (Brooker and Michael, 2000). As a result of changes in health priorities, tremendous efforts have been made in the development of methods to map the distribution of STH infections, particularly through the use of geographical information systems (GIS) and remote sensing technologies. Most activities with respect to STH mapping have been attempted in African countries (Knopp et al., 2008; Brooker et al., 2009; Snow et al., 2009; Pullan et al., 2011; Tchuem Tchuenté et al., 2012). In addition, this approach has recently been used to model the geographical distribution of STH infections in Bolivia (Chammartin et al., 2013) and Southeast Asia (Brooker, 2002; Brooker et al., 2003).

In the context of a preventive deworming programme, reliable and accurate contemporary maps offering intervention recommendations at a realistic scale are needed to minimise the cost of interventions. The establishment of reliable predictive risk maps for STH infections in Malaysia is imperative for developing and implementing sustainable control measures when resources are finite and limited. We attempt here to generate such maps not only on the basis of empirical survey data, but also by applying logistic regression analysis for the potential effect of climate and other environmental factors. Through investigation of ecological correlations, we planned to identify STH distribution, prioritise target areas and estimate the population at risk and its implications for the STH control programme in Malaysia.

## Materials and methods

### Data searches

Relevant information on the prevalence of 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 brief, an initial systematic search of published articles started in 2008 and was repeated periodically between 2009 and 2012. The online electronic databases PubMed, ISI Web of Knowledge, Google and Yahoo were used to identify relevant studies for STH by using the Medical Subject Headings (MSHs) hookworm, ascariasis, trichuriasis, *Necator*

*americanus*, *Ancylostoma duodenale*, *Ascaris lumbricoides*, *Trichuris trichiura*, intestinal parasites or soil-transmitted helminth (STH) together with Malaysia with no restriction of publication date and language. The second search strategy involved the identification of “grey” literature sources such as university theses, unpublished surveys conducted by government institutions and Ministry of Health (MoH) archives. The third source of information included personal contact with researchers known to have undertaken STH surveys in Malaysia. In addition, cross-sectional surveys were conducted between 2008 and 2012 in order to determine the prevalence of STH infections in several locations in Peninsular Malaysia.

### Geo-positioning procedures

The geographical coordinates of various surveyed locations were determined using combination of various free, on-line 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 Tageo (<http://www.tageo.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 location coordinates were recorded using a global positioning system (GPS) approach with a hand-held Garmin GPSMAP 60CSx instrument. The recorded coordinates were then downloaded from the GPS memory card into a computer using GPS Pathfinder software (<http://www.trimble.com/mappingGIS/PathfinderOffice.aspx>). All the digital data coordinate system were synchronised using World Geodetic System (WGS, 1984) according to longitude and latitude.

### Selection and entry

Pre-determined inclusion and exclusion criteria were applied to information identified through the above search mechanisms. Only cross-sectional surveys were included in the database. Multiple surveys from the same location but surveyed at different times were included as separate entries. Data from hospitals and related, clinical surveys were excluded. Survey data were also excluded if the prevalence was reported without denominator (i.e. sample size and number of positive samples) or if there were inconsistencies or errors in the calculations presented. Surveyed locations that could not be geo-positioned were also

excluded. Each extracted piece of information was entered into a Microsoft Excel database following the standard format by Brooker et al. (2000). Information included in the database consisted of the data source, survey location, survey date, methodology survey, population survey, sample size, examination methods and the number of positive findings for each helminth species.

### *Study area*

Peninsular Malaysia comprises 11 states and each state is divided into 81 districts, which are then further divided into the smallest administrative level, the sub-district or “mukim”, of which there are 842 in total. The extracted survey data were mapped at the sub-district level, each of which identified and linked to the Peninsular Malaysia boundary map obtained from Department of Surveying and Mapping, Malaysia (Anon, 2010a), here referred to as the base map. Places where more than one study had been conducted (or if the survey points were less than 2 km apart) were treated as a single location and the prevalence calculated by taking the weighted mean, according to sample size, of the individual survey prevalence rates.

### *Cumulative prevalence*

The combined STH prevalence (i.e. proportion of infection with any one STH species) is important when making decision about targeting effective treatment. However, the combined prevalence per STH species is seldom reported in epidemiological studies as most survey reports typically only provide proportions of individuals with single-species infections (de Silva and Hall, 2010). The cumulative prevalence of STH (i.e. the combined STH prevalence) was calculated using a simple probabilistic model after incorporating a small correction factor to allow for non-independence between the species in question. Briefly, the combined probability of having infection with any one STH species is the simple probability law for the union between three aspects after assuming the probability of infection with a single species to be independent of each other. The proportion infected with a combination of STH species was then calculated according to the formula:

$$\text{PATH} = A + T + H - (\text{AT}) - (\text{AH}) - (\text{TH}) + (\text{ATH})$$

where PATH is the combined proportion of infections with all three STH species, A the prevalence of *A. lum-*

*bricoides*, T the prevalence of *T. trichiura* and H the prevalence of hookworm infection. However, de Silva and Hall (2010) also reported that there was overestimation in the combined proportion of being simultaneously infected with all STH species (i.e. PATH) as calculated using this probability model. Due to the non-independence effect of each species, it has been demonstrated that the proportion of PATH increases by 0.6% for every 10% increase in the prevalence rate by calculating the true observed combined prevalence of STH through incorporating the over-estimation factor  $p$  as equal to  $\text{PATH} \div 1.06$  (de Silva and Hall, 2010).

### *GIS and remote sensing*

A set of environmental variables was gathered from a variety of sources. Monthly averages of the 1-km resolution land surface temperature (LST) were collected from the WorldClim database (<http://www.worldclim.org>), which comprises global weather station temperature records gathered from various sources for the period of 1950 to 2000. A thin-plate smoothing spline algorithm was used to interpolate the data following the approach of Hijmans et al. (2005). Normalized difference vegetation index (NDVI) data were obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS) as described by Scharlemann et al. (2008). Altitude was gathered from the interpolated digital elevation model (DEM) derived from the Department of Surveying and Mapping, Malaysia (Anon, 2010). For each of the environmental variables, minimum, mean and maximum values were extracted for each pixel corresponding to the various survey locations. These data were then exported and stored into ArcGIS 9.3 software (ESRI; Redlands, USA) for further exploration and analysis.

### *Statistical analysis, mapping and modelled STH distribution*

The statistical analysis for each test was performed using the statistical package for the social sciences (SPSS) software (SPSS Inc; Chicago, USA) for Windows, version 17. Logistic regression was used for developing statistical risk models and to identify significant environmental factors known to influence the transmission of infections (Fielding and Bell, 1997; Pearce and Ferrier, 2000). Since previous studies (Brooker and Michael, 2000; Brooker et al., 2002a, 2002b) have shown that temperature is an important

factor in determining STH distribution because of the effect of heat and low humidity on embryonation, development and survival of free-living infective stages. This variable was entered into the regression model first and then followed by the minimum and mean LST values. The remaining variables, such as NDVI (minimum, maximum and mean) and altitude (minimum, maximum and mean) were added to the model in a stepwise manner (i.e. introducing each variable in turn).

The accuracy of the model was assessed using classification table derived from regression analysis by dividing the probability of infection into two groups with a cut-off value of 0.5 (i.e. an infection probability of  $<0.5$  was categorised as 0) while probabilities  $>0.5$  were categorised as 1 (meaning with infection) (Menard, 2002). The “goodness-of-fit” test advanced by Hosmer and Lemeshow (1989) was used to determine whether our model fitted the observed prevalence data. Moran’s *I* (Moran, 1950) was used for controlling spatial autocorrelation, i.e. to determine whether or not the infection patterns found were spatially correlated. Following the development of the statistical risk model, the best-fit logistic regression model was used to generate a predictive risk map of the distribution of STH prevalence probability exceeding 50%. The observed and predicted prevalence rates were categorised according to prevalence thresholds for treatment established by the World Health Organization (WHO), i.e. 0, 0.1-9.9, 10-19.9, 20-49.9 and 50-100%.

## Results

### *Survey data and STH distribution*

The current database incorporated 99 survey locations conducted between 1970 and 2012 through combined search strategies, of which all locations were successfully geo-positioned to actual coordinates (Table 1). Nineteen surveys had been undertaken in the same locations but at different times and were therefore excluded. Of the included surveys, 80 were from spatially unique locations and represented a study population of 47,118 individuals of all age groups and both genders. Overall, 22,790 (48.4%) individuals were found to have been infected with *T. trichiura*, followed by 15,642 (33.2%) with *A. lumbricoides* and 5,578 (11.8%) with hookworm infection. The estimated combined prevalence of individuals infected with any STH species was 74.5% as calculated using simple probabilistic model following de

Silva and Hall (2010). Data extracted from the published papers accounted for 82.8% surveys and constituted the main data source. Personal communications or direct contact with the researchers was the second most important data source (10.1%) while unpublished reports related to theses accounted for 7.1% (7/99) of the surveys. Based on our communication with the MoH and its relevant agencies such as the public health division or the non-communicable disease division, no studies or data related to STH infections were available in their records. Overall, 75.8% surveys were community-based, while the rest (24.2%) were school-based. The most common way of stool examination with respect to STH diagnosis was combinations of at least two different conventional microscopy techniques (67.7%).

Figs. 1a-1d observed the geographical distribution of each STH species as well as the estimated combined prevalence based on all survey data at the sub-district level. For each STH species and combined STH, the prevalence of infections varied considerably with no clear pattern across the surveyed areas with the exception of hookworm. Fig. 1 shows that the observed prevalence of *A. lumbricoides*, *T. trichiura* and combined STH species was high in most of the surveyed areas. In contrast, most of the areas had low prevalence of hookworm infection. In addition to these geographical variations, there were also marked variations in the observed prevalence of STH infections over time (Table 2). For example, the mean prevalence of *A. lumbricoides* was 61.8% in the period of 1970-1979, but declined to 37.3% for 2010-2012. Similarly, mean prevalence of hookworm also shows distinct variations over time from 31.8% in 1970-1979 to 13.0% for 2000-2012.

### *Statistical and spatial analysis*

The results with regard to ecological correlation in un-sampled areas based on satellite derived environmental data and logistic regression indicated that maximum and mean LST values and minimum and mean NDVI values are significant explanatory variables for *A. lumbricoides* infection (Table 3) with the odds of *A. lumbricoides* negatively associated with maximum LST (odds ratio (OR) = 0.88; 95% confidence interval (CI) = 0.78-0.99) and minimum NDVI (OR = 0.98; 95% CI = 0.96-0.98). Only the *A. lumbricoides* model was developed since no statistically significant explanatory variables were recorded for *T. trichiura*, hookworm and the estimated combined STH infections.

Table 1. Prevalence survey of soil-transmitted helminth (STH) infections between 1970 and 2012 in Peninsular Malaysia: the descriptive database.

Characteristic	Frequency	%
Total number of indentified surveys	99	-
Total number of of individuals examined	47,118	-
Total numbers infected with <i>T. trichiura</i>	22,790	48.4
Total numbers infected with <i>A. lumbricoides</i>	15,642	33.2
Total numbers infected with hookworm	5,578	11.8
Estimated combined prevalence of STH species <sup>a</sup>	-	74.5
Source of survey		
Published papers in scientific journals	82	82.8
Unpublished reports (MOH reports, thesises, symposia)	7	7.1
Personal communications (direct contact with researchers)	10	10.1
Time covered (decades)		
1970-1979	23	23.3
1980-1989	17	17.2
1990-1999	17	17.2
2000-2009	4	4.0
2010-2012	38	38.4
Study populations		
Villages	53	53.5
Schools	24	24.2
Fishing villages	13	13.1
Estates	5	5.1
Squatter areas	4	4.0
Age-range examined		
Children	28	28.3
All age groups	71	71.7
Stool examination methodology		
Kato-Katz	32	32.3
All others <sup>b</sup>	67	67.7

<sup>a</sup>Estimated combined prevalence of STH species was calculated using a simple probabilistic model following de Silva and Hall (2010); <sup>b</sup>Other diagnostic methods included direct smear, formalin ether concentration, zinc-sulphate flotation, merthiolate-iodine-formalin (MIF), thiomersal-iodine-formol (TIF) and/or Harada Mori technique.

The model's predictive ability was generally good and reliable. The accuracy test as derived from classification table of logistic regression showed that the overall percentage of correct predictions of areas with and without infections was 67.7% (18/27). This includes a 63.7% (17/27) capability of exclusively predicting the areas with infection and 70.4% (19/27) indicating only those without infection. The goodness-of-fit (according to Hosmer and Lemeshow, 1989) was not statistically significant ( $\chi^2 = 6.27$ ; degrees of freedom (df) = 8;  $P = 0.508$ ), but we failed to reject the null hypothesis ( $P > 0.05$ ), which states that there is no difference between our model and the collected data, suggesting that our model estimates are indeed well fitted with the data. The Moran's *I* demonstrates that

there was significant positive spatial autocorrelation for *A. lumbricoides* infection within the sub-districts (Moran's *I* index = 0.04; z-score = 6.98;  $P < 0.01$ ).

#### Modelling of helminth distribution

The significant environmental factors such as maximum and mean LST and minimum and mean NDVI were used to model and predict the distribution of *A. lumbricoides* using logistic regression. The predictive risk map (Fig 2) of *A. lumbricoides* shows that the prevalence of infection was clustered and higher than elsewhere (the areas within the blue colour range with prevalence of at least 20% and up to 100%) in the central and northern plains of Peninsular Malaysia

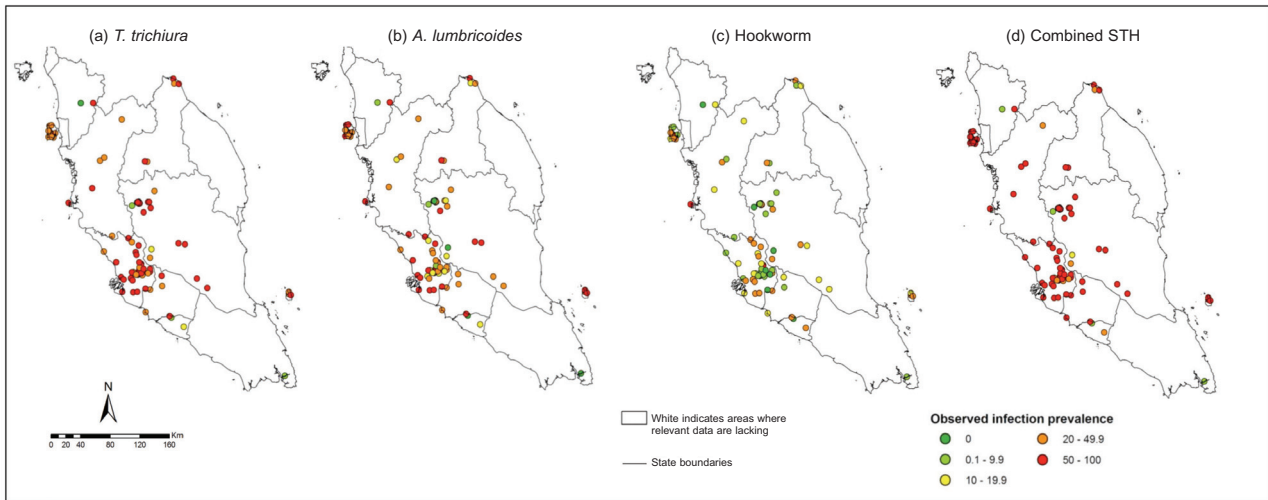


Fig. 1. Distribution maps based on available survey data of STH species at the sub-district levels in Peninsular Malaysia 1970-2012.

including central Pahang, Kelantan, northern Perak and Kedah, particularly in areas bordering southern Thailand. In contrast, the predicted prevalence of *A. lumbricoides* was lower (the areas within the green colour range with prevalence of  $\leq 20\%$ ) along the west coast and southern part of Peninsular Malaysia (Fig. 2). A visual comparison of our prediction map with the observed prevalence map for *A. lumbricoides* was not much different. For example, up to 70% of the surveyed areas with observed high prevalence showed levels similar to our predicted model.

Continuous probability contour map of prevalence exceeding 50% is illustrated in Fig. 3. The areas within the red colour range (i.e. probabilities  $>70\%$ ) are the areas where there is high probability that the WHO mass drug administration (MDA) threshold of 50% is exceeded, while areas within the green colour range (probabilities of at least 30%) are those where there is a low probability of 50% MDA threshold being exceeded. The yellow colour range (probabilities between 30% and 70%) can be assumed as areas of high uncertainty where further surveys would be helpful or continued surveillance programmes are recom-

mended. Figs. 4a and 4b show recommended intervention areas and population density (person per km<sup>2</sup>) in Peninsular Malaysia. A visual comparison between both maps indicates that most of the recommended intervention areas are areas with a population density between 10 and 100 persons per km<sup>2</sup>.

**Discussion**

The present study provides a description of the distribution and prevalence of STH infections in Peninsular Malaysia from available empirical data sources using GIS. Findings of the present study indicate that there is considerable geographical variation in the STH distribution and that geographically targeted control programmes are required to maximise available resources with regard to populations in greatest need. In the absence of an easily accessible database, the identification of priority areas for control has commonly been made using unsystematic data collection. Indeed, absence of appropriate information has sometimes led to deworming being included as a component of public health control programmes in

Table 2. Mean prevalence of individual and estimated combined STH infections between 1970 and 2012.

Decade	<i>Trichuris trichiura</i>		<i>Ascaris lumbricoides</i>		Hookworm		Estimated combined STH	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD*
1970-1979	60.3	26.2	61.8	24.5	31.8	25.3	85.4	18.2
1980-1989	56.3	28.0	36.1	23.0	13.4	12.6	71.3	24.4
1990-1999	44.8	21.3	38.8	17.9	17.9	13.8	70.0	15.7
2000-2009	70.0	32.5	47.2	20.8	25.8	17.8	83.7	21.8
2010-2012	57.6	26.8	37.3	22.9	13.0	9.9	72.3	29.5

\*Standard deviation

Table 3. Regression coefficients used for estimation of the probability of finding a subject infected by *Ascaris lumbricoides*.

Characteristic	Coefficient (B)	Standard error	OR <sup>a</sup> (95% CI <sup>b</sup> )	P-value
Constant	28.992	0.248	-	-
Maximum LST <sup>c</sup>	-0.127	3.942	0.88 (0.78-0.99)	0.047
Mean LST <sup>c</sup>	0.113	4.019	1.12 (1.00-1.25)	0.045
Minimum NDVI <sup>d</sup>	-0.020	4.747	0.98 (0.96-0.98)	0.029
Mean NDVI <sup>d</sup>	0.026	5.819	1.03 (1.01-1.05)	0.016

<sup>a</sup>Odds ratio; <sup>b</sup>CI = confidence interval; <sup>c</sup>land surface temperature; <sup>d</sup>normalized difference vegetation index.

areas with absent or only low prevalence STH infections (Brooker et al., 2009).

Since the STHs have life cycles involving sexual maturation in the human host and free-living stages in the environment, development and survival rate of these parasites depend on environmental factors, such as humidity and temperature. Studies have shown that environmental variables indirectly influence transmission success and thus influence the spatial pattern of these infections (Brooker and Michael, 2000). Therefore, we investigated the ecological correlations of STH infections and predicted their prevalence in unsampled areas based on satellite derived environmental data using logistic regression analysis.

Despite the lack of records on STH infections in

Malaysia at the MoH and its related agencies, and the fact that few or no data are available in many areas of the country, we were able to assemble a database which is not only one of the largest, but also the most detailed on the geographical distribution of STH infections in Malaysia. The situation is not very different in other endemic countries; for example, STH data for some East-African countries were extracted from more than half of the available but unpublished surveys from the Division of Vector Borne Diseases, MoH in Kenya, confirming the importance of this type of exercise (Brooker et al., 2009). Similarly, mapping of filariasis in Vietnam (Meyrowitsch et al., 1998) and malaria in Kenya (Omumbo et al., 1998) also demonstrate the significance of unpublished data.

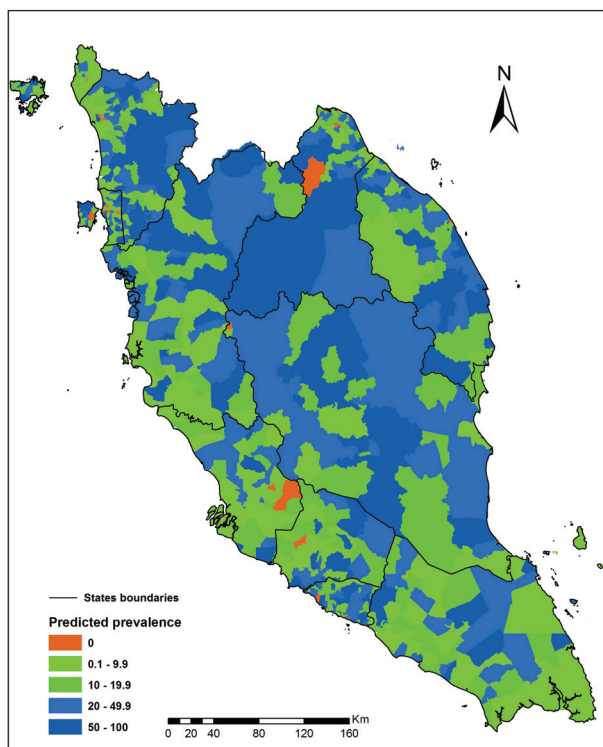


Fig. 2. Predicted prevalence of *A. lumbricoides* in Peninsular Malaysia.

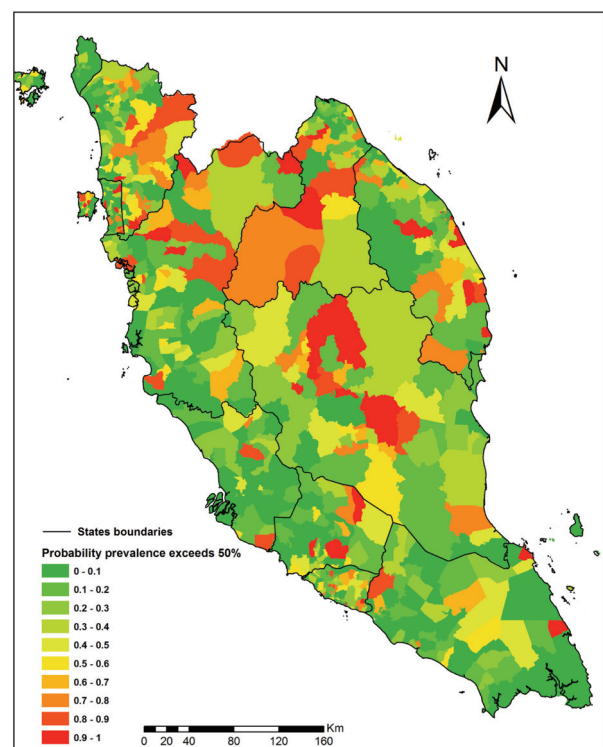


Fig. 3. Spatial distribution probability prevalence of *A. lumbricoides* exceeding 50% in Peninsular Malaysia (continuous probability map).

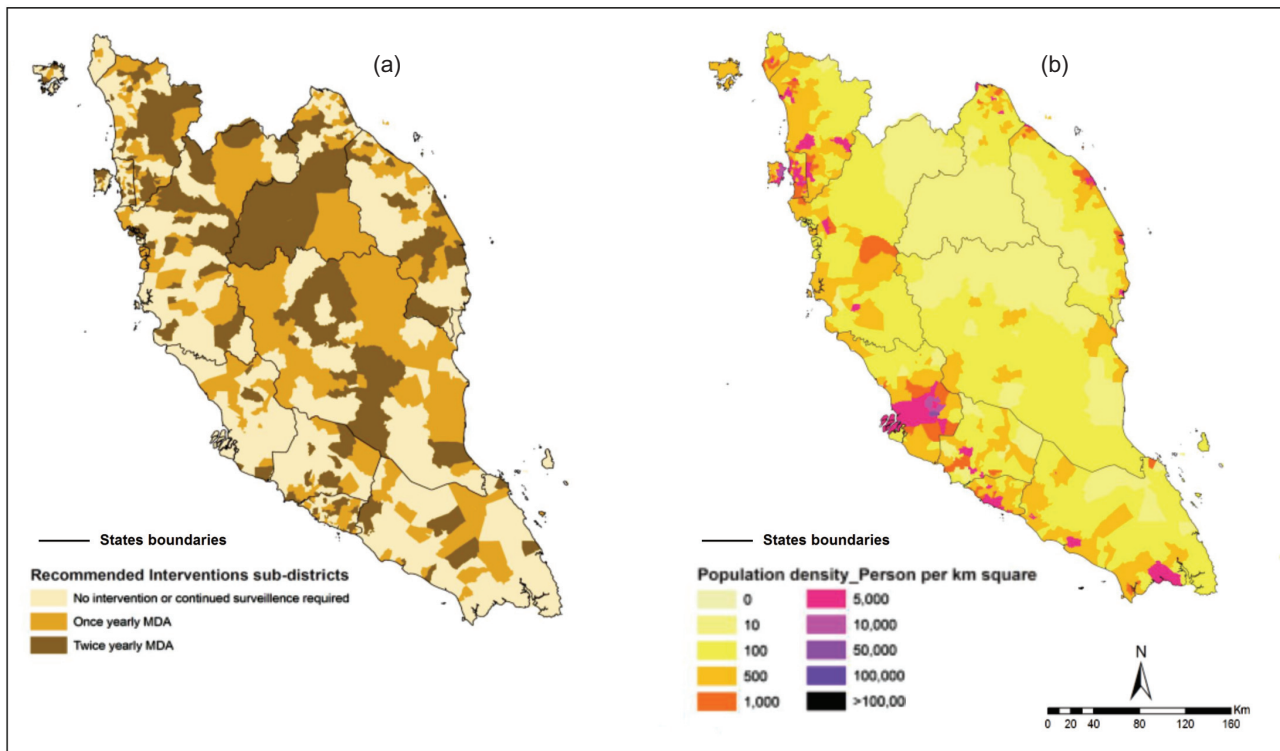


Fig. 4. Maps used for public health control planning in Peninsular Malaysia. Sub-districts recommended for intervention (a) and density (person per  $m^2$ ) of local population (b).

Interestingly, in addition to the wide geographical variation discovered, our analysis indicates that the observed prevalence of STH infections also varied temporally, generally showing a gradual decreasing pattern. For example, mean prevalence of each STH infection was high for the period 1970-1979 but gradually declined over time across the country especially for the period 1980-1999. One possible explanation for this trend may be due to the positive impact of school-based deworming programmes. In 1974, the MoH launched the National Worm Control Programme involving 1,468 schools with more than 3 million children receiving anthelmintic drugs but the programme was discontinued in 1983 (Anon, 1985). Other factors, such as urbanization that usually leads to improvement in the socioeconomic status, may have contributed indirectly through improved living standard. Better personal hygiene and environmental sanitation may also have had a positive impact resulting in the decline of prevalence rates over time.

While considering the value of the mapping approach, it is important to identify the limitations. Different parasitological methods and techniques employed in the surveys might reduce comparability of the data and ability of the map to represent infection prevalence accurately in every survey area (Brooker et al., 2003, 2009). Most of the parasitolog-

ical survey reports were based on single faecal examination. This could lead to potential bias due to several factors such as poor sensitivity in detecting light infections or day-to-day fluctuation, i.e. intermittent egg excretion in excretion by the adult worm (Engels et al., 1996; Booth et al., 2003). Additionally, previous studies indicated that Kato-Katz technique was more sensitive compared to other microscopic method for the examination of faecal samples (Goodman et al., 2007; Knopp et al., 2008). However, only one-third of the reported prevalence survey in our current work used the Kato-Katz technique as the main diagnostic tool. Another potential bias in the observed prevalence surveys was the delay in sample processing after sample collection, which is particularly important with regard to hookworm infection (Dacombe et al., 2007). While this inherent potential bias should be borne in mind when interpreting prevalence data, they are unlikely to lead to any different conclusion at policy decision level. This is because current treatment guidelines are still relying on prevalence data determined by the current diagnostic “gold standard” (i.e. stool examination by microscopy) recommended by WHO (Montresor et al., 1998).

Although it is possible to estimate the national average prevalence of STH infections based on observed



prevalence figures, it is clear that such an exercise is potentially misleading (Brooker et al., 2000). For example, we found very few prevalence studies and these data were obtained mostly from endemic areas known to have high prevalence and national prevalence was actually based on extrapolation of these values. Indeed, such extrapolation more often reflects the number of studies carried out and their locations, rather than being reliable indications of the true prevalence rate. Moreover, a large proportion of these published data were more than a few years old making the present-day interpretation uncertain. Despite these limitations, the current data have provided a crude estimation of the overall picture of STH infections and their distribution in Malaysia. Apart from these limitations, the work presented has generated predictive distribution of STH infections based on environmental variables derived from satellite data. Since ecological factors are known to influence the development and survival of the free-living transmission stage of the STH parasites, they can be used as proxy to prevalence rates (Brooker and Michael, 2000). The correlations between infection patterns and ecological factors can thus serve as tools to extrapolate risk estimates in areas for which no data are available, thus serve as baseline data to identify priority areas and the populations with the greatest needs.

Results of the present study indicate a significant association between prevalence of *A. lumbricoides*, LST and the NDVI. Findings of the current study also suggest a negative association between 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. Temperature and humidity are the density-independent factors influencing STH species transmission and infection patterns as measured by the basic reproductive number ( $R_0$ ). Increases in  $R_0$  will lead to increase in the prevalence and intensity of infection (Anderson and May, 1991). The thermal limits of infection as reported in our current study are also in agreement with other studies. Studies in Africa suggest that *A. lumbricoides* and *T. trichiura* infections were most unlikely to occur in areas where maximum LST exceeds 37 °C (Brooker et al., 2002a,b). Likewise, mapping of STH infections in Kenya (Pullan et al., 2011) and Vietnam (Brooker et al., 2003) also report low prevalence of *A. lumbricoides* and *T. trichiura* infections in areas where maximum LST was above 37 °C. A study in South Africa also showed

that the average prevalence of *A. lumbricoides* and *T. trichiura* was below 20% in areas with mean annual temperature below 15 °C (Appleton et al., 1999). This observation was further supported by field investigations. A national parasitic survey in China demonstrated that prevalence of STH infections was significantly associated with temperature and humidity factors (Lai and His, 1996). In addition, low prevalence of STH infections was observed in northern China where estimated LST was less than 20 °C and in India and Pakistan with the estimated LST as more than 40 °C (Brooker et al., 2004, 2006). Dunn (1972) conducted a field survey to investigate the intestinal parasitism among different rural communities living in different habitats in the Malaysian rain forest. The indication is that communities at higher and cooler elevations suffer less from STH infections.

Although findings of the present study can play important roles and offer basic information for the rapid planning and sustainable control programme of STH in Malaysia, there are some other factors that may result in underestimation of the prevalence and influence the distribution of infection (Brooker et al., 2009; Pullan et al., 2011). Issues such as differences between rural and urban areas, poverty, lack of education, poor hygiene and environmental variables are known to influence the distribution pattern of infection especially at small, spatial scales. We must also bear in mind that this information on socioeconomic and behaviour variables is difficult to collect over large spatial scales (Brooker et al., 2009). Apart from small-scale spatial variations, other factors such as uneven data distribution or the need for a better understanding of the impact on the different life-cycle biological stages with environmental factors are also essential (Brooker and Michael, 2000). These variables could not be incorporated in the current mapping survey, but constitute the aim for a future study. In the present study, however, our method of selecting variables for the geostatistical modelling relied on standard methods such as stepwise logistic regression that is more suitable for non-spatial data. However, ignoring spatial correlation might lead to incorrect estimates of the statistical significance of the predictor included in the model (Raso et al., 2005; Clements et al., 2009). The use of a more rigorous and powerful geostatistical modelling approach such as Bayesian geostatistical approach should be incorporated in future work as it can provide a more robust tool for identifying determinants of the disease distribution and predicting risk and burden at high spatial scale (Diggle et al., 2007).

Despite the fact that predictive model may under-predict true prevalence which may arise as consequences of these small-scale spatial variations, it still provides a reliable indication of whether or not MDA intervention is warranted in prioritised areas and populations with greatest needs in accordance with WHO recommendation.

Reliable maps of infection distribution and risk enhance cost-effectiveness of preventive chemotherapy intervention programmes. GIS and remote sensing facilitate effective storage, mapping, analysis and development of STH distribution maps by making data integration more accessible and reliable than before. This approach also permits the modelling of spatial infection distribution in relation to ecological factors, thereby enhancing understanding of the biology and epidemiology of the infections. The predictive risk maps produced in the present study can be used to identify endemic areas and provide more precise estimation of populations in greatest need, particularly when the resources for control are limited.

## Conclusion

A low-cost approach in surveying, designing and targeting by identifying priority areas and simplifying the monitoring and evaluation processes has been provided and important improvements of the survey and sampling strategy at the national level outlined. Areas that require further surveying in order to accurately estimate disease burden have been highlighted and high-risk population identified so that sustainable strategies and national control measures can be optimally targeted at a realistic scale.

## Acknowledgements

We would like to express our gratitude to Prof. Simon Brooker and Dr. Rachel L. Pullan from the Department of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, London, United Kingdom for their input and counsel during the initial stage of this work. We are also grateful to all those who have assisted and supported the collection of necessary data for this study including MoH, Ministry of Rural Development, headmasters and teachers in the schools and all the villagers who have voluntarily participated in this study. Special thanks also go to Mr. Saidon Ishak and Mr. Muhammad Aidil Roslan for their technical assistance during the fieldtrips. This research project was funded by the UM/MoHE High Impact Research (H-20001-00-E000061), UMRG (RG132/11SUS), PJM-KTP-CITRA FL001-13SBS and E-Science (16-02-03-6034) grant.

## References

- Aaron RJ, Lim YAL, Bethony J, Hotez PJ, Young ND, Gasser RB, 2011. Soil-transmitted helminths of humans in Southeast Asia-towards integrated control. *Adv Parasitol* 74, 231-265.
- Ahmed A, Al-Mekhlafi HM, Choy SH, Ithoi I, Al-Adhroey AH, Abdulsalam AM, Surin J, 2011. The burden of moderate-to-heavy soil-transmitted helminth infections among rural Malaysian aborigines: an urgent need for an integrated control programme. *Parasit Vectors* 4, 242.
- Anderson RM, May RM, 1991. *Infectious diseases of humans: dynamics and control*. Oxford: Oxford University Press.
- Anon, 1985. Annual Report 1983/1984. Ministry of Health (MoH), Malaysia.
- Anon, 2010. Department of Survey and Mapping (*Jabatan Ukur dan Pemetaan, JUPEM*) Malaysia.
- Appleton CC, Maurihungirire M, Gouws E, 1999. The distribution of helminth infections along the coastal plain of Kwazulu-Natal province, South Africa. *Ann Trop Med Parasitol* 93, 859-868.
- Bethony J, Brooker S, Albonico M, Geiger SM, Loukas A, Diemert D, Hotez PJ, 2006. Soil-transmitted helminthes infections: ascariasis, trichuriasis and hookworm. *Lancet* 367, 1521-1532.
- Booth M, Vounatsou P, N'Goran EK, Tanner M, Utzinger J, 2003. The influence of sampling effort and the performance of the Kato-Katz technique in diagnosing *Schistosoma mansoni* and hookworm co-infections in rural Côte d'Ivoire. *Parasitology* 127, 525-531.
- Brooker S, 2002. Human helminth infections in Indonesia, East Timor and the Philippines. Report to UNICEF East Asia and Pacific Region Bangkok, Thailand.
- Brooker S, Beasley NMR, Ndinaromtan M, Madjiouroum EM, Baboguel M, Djenguinabe E, Hay SI, Bundy DAP, 2002a. Use of remote sensing and a geographical information system in a national helminth control programme in Chad. *Bull World Health Organ* 80, 783-789.
- Brooker S, Clements ACA, Bundy DAP, 2006. Global epidemiology, ecology and control of soil-transmitted helminth infections. *Adv Parasitol* 62, 221-261.
- Brooker S, Hay SI, Tchuem Tchuente LA, Ratard R, 2002b. Using NOAA-AVHRR data to model helminth distributions for planning disease control in Cameroon, West Africa. *Photogramm Eng Rem Sens* 68, 175-179.
- Brooker S, Kabatereine NB, Smith JL, Mupfasoni D, Mwanje MT, Ndayishimiye O, Lwambo NJ, Mbotha D, Karanja P, Mwandawiro C et al., 2009. An updated atlas of human helminth infections: the example of East Africa. *Int J Health Geogr* 8, 42.
- Brooker S, Kabatereine NB, Tukahebwa EM, Kazibwe F, 2004. Spatial analysis of the distribution of intestinal nematode infections in Uganda. *Epidemiol Infect* 132, 1065-1071.

- Brooker S, Michael E, 2000. The potential of geographical information systems and remote sensing in the epidemiology and control of human helminth infections. *Adv Parasitol* 47, 245-288.
- Brooker S, Rowlands M, Haller L, Savioli L, Bundy DAP, 2000. Towards an atlas of human helminth infection in sub-Saharan Africa: the use of geographical information systems (GIS). *Parasitol Today* 16, 303-307.
- Brooker S, Singhasivanon P, Waikagul J, Supavej S, Kojima S, Takeuchi T, Luong TV, Looareesuwan S, 2003. Mapping soil-transmitted helminths in Southeast Asia and implications for parasite control. *Southeast Asian J Trop Med Public Health* 34, 24-36.
- Chammartin F, Scholte RGC, Malone JB, Bavia ME, Nieto P, Utzinger J, Vounatsou P, 2013. Modelling the geographical distribution of soil-transmitted helminth infections in Bolivia. *Parasit Vectors* 6, 152.
- Chan MS, Medley GF, Jamison D, Bundy DA, 1994. The evaluation of potential global morbidity attributable to intestinal nematode infections. *Parasitology* 109, 373-387.
- Dacombe RJ, Crampin AC, Floyd S, Randall A, Ndhlovu R, Bickle Q, Fine PE, 2007. Time delays between patient and laboratory selectively affect accuracy of helminth diagnosis. *Trans R Soc Trop Med Hyg* 101, 140-145.
- de Silva NR, Brooker S, Hotez PJ, Montresor A, Engels D, Savioli L, 2003. Soil-transmitted helminth infections: updating the global picture. *Trends Parasitol* 19, 547-551.
- de Silva NR, Hall A, 2010. Using the prevalence of individual species of intestinal nematode worms to estimate the combined prevalence of any species. *PLoS Negl Trop Dis* 4, e655.
- Diggle P, Thomson MC, Christensen OF, Rowlingson B, Obsomer V, Gardon J, Wanji S, Takougang I, Enyong P, Kamgno J et al., 2007. Spatial modelling and the prediction of Loa loa risk: decision making under uncertainty. *Ann Trop Med Parasitol* 101, 499-509.
- Dunn FL, 1972. Intestinal parasites in Malayan aborigines (Orang Asli). *Bull World Health Organ* 46, 99-113.
- Engels D, Sinzinkayo E, Gryseels B, 1996. Day-to-day egg count fluctuation in *Schistosoma mansoni* infection and its operational implications. *Am J Trop Med Hyg* 54, 319-324.
- Fielding AH, Bell JF, 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24, 38-49.
- Goodman D, Haji HJ, Bickle QD, Stoltzfus RJ, Tielsch JM, Ramsan M, Savioli L, Albonico M, 2007. A comparison of methods for detecting the eggs of *Ascaris*, *Trichuris* and hookworm in infant stool, and the epidemiology of infection in Zanzibari infants. *Am J Trop Med Hyg* 76, 725-731.
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A, 2005. Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25, 1965-1978.
- Hosmer D, Lemeshow S, 1989. Applied logistic regression. A much-cited recent treatment utilized in SPSS routines. New York: John Wiley and Sons.
- Hotez PJ, Fenwick A, Savioli L, Molyneux DH, 2009. Rescuing the bottom billion through control of neglected tropical diseases. *Lancet* 373, 1570-1575.
- Knopp S, Mgeni AF, Khamis IS, Steinmann P, Stothard JR, Rollinson D, Marti H, Utzinger J, 2008. Diagnosis of soil-transmitted helminths in the era of preventive chemotherapy: effect of multiple stool sampling and use of different diagnostic techniques. *PLoS Negl Trop Dis* 2, e331.
- Lai D, His BP, 1996. Soil-transmitted helminthiasis in China: a spatial statistical analysis. *Southeast Asian J Trop Med Pub Health* 27, 754-759.
- Lim YAL, Romano N, Colin N, Chow SC, Smith HV, 2009. Intestinal parasitic infections amongst Orang Asli (indigenous) in Malaysia: has socioeconomic development alleviated the problem? *Trop Biomed* 26, 110-122.
- Menard S, 2002. Applied logistic regression analysis. Sage Publications, California.
- Meyrowitsch DW, Toan ND, Hao HT, Dan NT, Michel E, 1998. A review of the present status of lymphatic filariasis in Vietnam. *Acta Trop* 70, 335-347.
- Montresor A, Crompton DWT, Bundy DAP, Hall A, Savioli L, 1998. Guidelines for the evaluation of soil-transmitted helminthiasis and schistosomiasis at community level. *WHO/CTD/SIP* 1, 28-30.
- Moran PAP, 1950. Notes on continuous stochastic phenomena. *Biometrika* 37, 17-23.
- Nasr NA, Al-Mekhlafi HM, Ahmed A, Roslan MA, Bulgiba A, 2013. Towards an effective control programme of soil-transmitted helminth infections among Orang Asli in rural Malaysia. Part 1: prevalence and associated key factors. *Parasit Vectors* 6, 27.
- Ngui R, Saidon I, Chow SC, Rohela M, Lim YAL, 2011. Prevalence and risk factors of intestinal parasitism in rural and remote West Malaysia. *PLoS Negl Trop Dis* 5, e974.
- Omumbo J, Ouma J, Rapuoda B, Craig MH, le Sueur D, Snow RW, 1998. Mapping malaria transmission intensity using geographical information systems (GIS): an example from Kenya. *Ann Trop Med Parasitol* 92, 7-21.
- Pearce J, Ferrier S, 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecol Modell* 133, 225-245.
- Pullan RL, Gething PW, Smith JL, Mwandawiro CS, Sturrock HJW, Gitonga CW, Hay SI, Brooker S, 2011. Spatial modelling of soil-transmitted helminth infections in Kenya: a disease control planning tool. *PLoS Negl Trop Dis* 5, e958.
- Raso G, Matthys B, N'Goran EK, Tanner M, Vounatsou P, Utzinger J, 2005. Spatial risk prediction and mapping of *Schistosoma mansoni* infections among schoolchildren living in western Côte d'Ivoire. *Parasitology* 131, 97-108.
- Scharlemann JPW, Benz D, Hay SI, Purse BV, Tatem AJ, Wint

- GR, Rogers DJ, 2008. Global data for ecology and epidemiology: a novel algorithm for temporal Fourier processing MODIS data. *PLoS One* 3, e1408.
- Snow RW, 2009. An updated atlas of human helminth infections: the example of East Africa. *Int J Health Geogr* 8, 42.
- Tchuem Tchunte LA, Kamwa Ngassam RI, Sumo L, Ngassam P, Dongmo Noumedem C, Nzu DD, Dankoni E, Kenfack CM, Gipwe NF, Akame J, Tarini A, Zhang Y, Angwafo FF, 2012. Mapping of schistosomiasis and soil-transmitted helminthiasis in the regions of centre, east and west Cameroon. *PLoS Negl Trop Dis* 6, e1553.
- WHO, 2005. Deworming for health and development. Report of the third global meeting of the partners for parasite control. Geneva: World Health Organization.