Spatial risk prediction and mapping of *Schistosoma mansoni* infections among schoolchildren living in western Côte d'Ivoire

G. RASO[1,2], B. MATTHYS[1,2], E. K. N'GORAN[2,3], M. TANNER[1], P. VOUNATSOU[1] *and* J. UTZINGER[1]*

[1] *Department of Public Health and Epidemiology, Swiss Tropical Institute, P.O. Box, CH-4002 Basel, Switzerland*
[2] *Centre Suisse de Recherches Scientifiques, 01 BP 1303, Abidjan 01, Côte d'Ivoire*
[3] *UFR Biosciences, Université d'Abidjan-Cocody, 22 BP 770, Abidjan 22, Côte d'Ivoire*

(*Received 24 November 2004; revised 16 January 2005; accepted 16 January 2005*)

SUMMARY

The objectives of this study were (1) to examine risk factors for *Schistosoma mansoni* infection among schoolchildren living in western Côte d'Ivoire, and (2) to carry forward spatial risk prediction and mapping at non-sampled locations. First, demographic and socio-economic data were obtained from 3818 children, aged 6–16 years, from 55 schools. Second, a single stool sample was examined from each child by the Kato-Katz technique to assess infection status of *S. mansoni* and its intensity. Third, remotely sensed environmental data were derived from satellite imagery and digitized ground maps. With these databases a comprehensive geographical information system was established. Bayesian variogram models were applied for spatial risk modelling and prediction. The infection prevalence of *S. mansoni* was 38·9%, ranging from 0% to 89·3% among schools. Results showed that age, sex, the richest wealth quintile, elevation and rainfall explained the geographical variation of the school prevalences of *S. mansoni* infection. The goodness of fit of different spatial models revealed that age, sex and socio-economic status had a stronger influence on infection prevalence than environmental covariates. The generated risk map can be used by decision-makers for the design and implementation of schistosomiasis control in this setting. If successfully validated elsewhere, this approach can guide control programmes quite generally.

Key words: Bayesian geostatistics, Côte d'Ivoire, geographical information system, kriging, prediction, remote sensing, risk mapping, *Schistosoma mansoni*, spatial analysis.

INTRODUCTION

Schistosomiasis is a parasitic disease caused by trematode worms belonging to the genus of *Schistosoma*. In sub-Saharan Africa, schistosomiasis remains of major public health and economic significance. Recent estimates for this region of the world suggest that 436 million people live at risk of infection with *Schistosoma haematobium*, and 393 million people are at risk of *S. mansoni*. The estimated annual mortality rate might exceed 200 000 (Chitsulo *et al*. 2000; WHO, 2002; van der Werf *et al*. 2003). Transmission occurs in freshwater bodies, where specific aquatic snails act as intermediate host. Infected snails release cercariae that can penetrate the skin of humans during occupational and/or recreational activities. Cercariae rapidly develop into schistosomula which, after several weeks, become adult schistosome worms that constantly produce eggs. Some of the eggs are released into the environment by human excreta. The remainder are trapped in the tissues of the host organs causing inflammation, which can lead to severe morbidity. Suitable climatic and environmental conditions for both the parasite and intermediate host snail, coupled with inadequate water supply and sanitation and low hygiene conditions, are the root causes for the persistence of schistosomiasis (Utzinger *et al*. 2003).

Progress has been made over the past 10–15 years with the application of geographical information system (GIS) and remote sensing (RS) for risk mapping of parasitic diseases, including schistosomiasis. Hence, these techniques have become important tools for the design and implementation of control programmes (Brooker, Hay & Bundy, 2002*a*). In addition, developments in Bayesian methods and Markov chain Monte Carlo (MCMC) inference (Gelfand & Smith, 1990) have advanced spatial modelling. GIS allows identification and visualization of demographic, environmental and socio-economic covariates for infection risk. Linked to Bayesian spatial statistics, GIS can save scarce resources of otherwise expensive parasitological surveys for detection and monitoring of high-risk areas

* Corresponding author: Department of Public Health and Epidemiology, Swiss Tropical Institute, P.O. Box, CH-4002 Basel, Switzerland. Tel: +41 61 284 8129. Fax: +41 61 284 8105. E-mail: juerg.utzinger@unibas.ch

Parasitology (2005), **131**, 97–108.
doi:10.1017/S0031182005007432 Printed in the United Kingdom

(Brooker & Michael, 2000; Robinson, 2000). Several climate and environmental factors have been linked with schistosome infections on a broad scale (Brooker *et al*. 2001; Malone *et al*. 2001; Zhou *et al*. 2001; Yang *et al*. 2005). For example, a model had been developed for the distribution of *Biomphalaria pfeifferi*, the intermediate host snail of *S. mansoni*, in Ethiopia. The model, based on the normalized difference vegetation index (NDVI) and maximum land surface temperature (LST), predicted *B. pfeifferi* distribution, and hence infection prevalence of *S. mansoni* (Kristensen, Malone & McCarroll, 2001). Using RS to derive rainfall and maximum LST data, a model had been produced for prediction of *S. haematobium* infections in Cameroon (Brooker *et al*. 2002*b*). This model, however, only showed good prediction within the boundaries of a given ecozone (Brooker *et al*. 2002*a*).

Previous research has shown that analyses on a broad scale render it difficult to capture the small-scale focality, which is a typical epidemiological feature of schistosomiasis (Bavia *et al*. 2001; Brooker, 2002; Lengeler, Utzinger & Tanner, 2002). Several factors, primarily acting at a local scale, may be at the origin of such heterogeneity (Husting, 1983; Kloos *et al*. 1997; Watts *et al*. 1998). Consequently, appropriate models are required which can capture potential risk areas at different scales (Bavia *et al*. 2001; Brooker, 2002). The need for broad-scale analyses derives from decision-making often taking place at the district level, while locally distinct needs can drive small-scale analyses.

The objectives of this study were (1) to identify risk factors explaining the geographical distribution of *S. mansoni* infections in a mountainous region of western Côte d'Ivoire, and (2) to make predictions at non-sampled locations. A cross-sectional survey was done among several thousand schoolchildren. They were screened for *S. mansoni*, and interviewed for demographic and socio-economic indicators. Environmental factors were obtained and a GIS was established. Finally, Bayesian geostatistics were employed for prediction of *S. mansoni* infection. This work contributes to an ongoing parasitic disease research and integrated control programme with emphasis on schistosomiasis, soil-transmitted helminthiasis and malaria in the western part of Côte d'Ivoire.

MATERIALS AND METHODS

Study area and population

The study region is a 40 × 60 km area situated in the mountainous region of Man, western Côte d'Ivoire. This setting has been known to be a *S. mansoni* focus for over 3 decades (Doumenge *et al*. 1987). There are 2 wet seasons; the main one between April and June and a shorter one in September. Mountains, inselbergs (remnants of erosion processes forming isolated, typically rounded mountains which can range in elevation from a few to several hundreds of meters) and small valleys dominate the northern part of the study area, while the southern part is a plain that acts as a drain for the numerous small rivers. Recent studies confirmed that this setting is indeed endemic for *S. mansoni* with all age-groups concerned (Utzinger *et al*. 2000; Keiser *et al*. 2002; Raso *et al*. 2005).

The study protocol was approved by the internal review boards of the Swiss Tropical Institute (Basel, Switzerland) and the Centre Suisse de Recherches Scientifiques (Abidjan, Côte d'Ivoire). It received ethical clearance from the Ministry of Health in Côte d'Ivoire. Data presented here are derived from a cross-sectional epidemiological survey carried out between October, 2001 and February, 2002. All schools located in the town of Man and those schools in rural areas with less than 100 pupils on the education registries were excluded from the survey. In the remaining 57 rural schools, all schoolchildren attending grades 3–5 were enrolled for parasitological screening and an interview by the teachers.

Parasitological data

Details of the parasitological surveys have been described elsewhere (Raso *et al*. 2005). In brief, after explaining the objectives and procedures of the study to the education officers, class lists for the school year 2001/2002 were obtained, containing the name, age and sex of each pupil. Unique identification numbers were assigned to all schoolchildren attending grades 3–5. Schoolchildren were examined for *S. mansoni*, soil-transmitted helminths, intestinal protozoa and *Plasmodium* infections. Emphasis here is placed on the field and laboratory procedures pertaining to *S. mansoni*.

The research team visited one school after another. Small containers were distributed to all study participants. They were invited to return the containers with a small portion of their own morning stool. The stool specimens were collected and transferred to the central laboratory in the nearby town of Man. A single 42 mg Kato-Katz thick smear was prepared from each stool specimen on microscope slides (Katz, Chaves & Pellegrino, 1972) and allowed to clear for 30–45 min. The slides were examined by experienced laboratory technologists under a light microscope at low magnification. All *S. mansoni* eggs were counted and recorded.

Treatment

At the end of the parasitological survey, schoolchildren who were egg-positive for *S. mansoni* were treated according to the existing treatment schedule recently developed by the regional health authorities,

https://doi.org/10.1017/S0031182005007432 Published online by Cambridge University Press

as described before (Raso *et al*. 2005). Praziquantel was administered at a single oral dose of 40 mg/kg (WHO, 2002).

Socio-economic data

Socio-economic data were obtained from a pre-tested and validated questionnaire, which was administered by teachers to all schoolchildren who were previously examined for parasite infections. Neither teachers nor schoolchildren had prior knowledge on *S. mansoni* infection status. The questionnaire included a list of 17 morbidity indicators, 12 household assets and the question ''do you live within the main village or in settlements outside?'' A simple asset-based approach was adopted to stratify schoolchildren into socio-economic groups (Filmer & Pritchett, 2001). Household asset variables were weighed using principal component analysis. Schoolchildren were ranked according to their total sum of asset scores. Finally, schoolchildren were attached to wealth quintiles, namely (1) most poor, (2) very poor, (3) poor, (4) less poor, and (5) least poor.

Environmental data

The geographical locations of schools (longitude, latitude and elevation) were recorded in the field using a hand-held global positioning system (GPS; Thales Navigation, Santa Clara, CA, USA).

In the absence of recent digital maps, available maps from the 1960s (scales: 1 : 200 000 and 1 : 50 000) were georeferenced. Streets, village boundaries, rivers and elevation lines were digitized. Satellite imagery data from Digital Enhanced Landsat Thematic Mapper (ETM+) (image dates: 27 January 2002; 11 November 2002) and GPS-referenced control points taken in the field served as validation of the digitized maps. Maps of soil types (scale: 1 : 500 000) were georeferenced and major soil types of the study area digitized. These were either ferrallitic soils slightly regenerated (modal, complex and fairly desaturated) or hydromorphic mineral soils. Land cover types were obtained from satellite image (Advanced Very High Resolution Radiometer (AVHRR) satellite, US Geological Survey (USGS) Africa Land Cover Characteristics Database v. 2: Africa Seasonal Cover Regions, USGS Earth Resources Observation System (EROS) Data Centre) at a 1×1 km spatial resolution (http://edcdaac.usgs.gov/glcc/glcc.asp). Land cover types were field-validated by taking GPS control points of relevant vegetation classes. An interpolated digital elevation model (DEM) was obtained from the USGS EROS Data Center (http://lpdaac.usgs.gov/gtopo30/gtopo?links.asp). LST and NDVI data were downloaded from Moderate Resolution Imaging Spectroradiometer (MODIS) from USGS EROS Data Center (http://edcdaac.usgs.gov/dataproducts.asp). Rainfall estimate (RFE) data with an 8×8 km spatial resolution from Meteosat 7 satellite were obtained from the Africa Data Dissemination Service (ADDS) (http://edcw2ks21.cr.usgs.gov/adds/). LST, NDVI and RFE were downloaded for the period of September 2001 to August 2002 and processed as suggested by Hay (2000). LST day and night data were available as 8-day-maximum value composites at 1×1 km spatial resolution. Monthly minimum, mean and maximum composites were calculated and values extracted for each pixel corresponding to the school locations. Monthly maximum value composites were also calculated from NDVI 16-day-maximum value composites at 1×1 km spatial resolution and values were extracted for each school location. The same procedure was applied for the RFE data. Annual LST and NDVI were obtained as average of the monthly mean values. For LST, annual estimates from monthly minimum and maximum values were also obtained. RFE was calculated as the total amount of rainfall over the 1-year period. Distance from schools to the nearest permanent water body was computed.

Georeferencing of maps, processing of the environmental data and distance calculation were done in IDRISI 32 (Clark Labs, Clarks University, Worcester, MA, USA). Data were displayed in ArcView GIS v. 3.2 (Environmental Systems Research Institute, Inc., Redlands, CA, USA).

Statistical analysis

Parasitological and questionnaire data were double-entered and validated with EpiInfo v. 6.04 (Centers for Disease Control and Prevention, Atlanta, GA, USA). Schoolchildren were subdivided into 2 age groups; namely (1) 6–10 years, and (2) 11–16 years.

Elevation at each school location was defined as the mean of the GPS value and the elevation derived from the digitized maps. Environmental covariates with continuous values, which showed a non-linear relationship to *S. mansoni* infection prevalence, were categorized. Logistic regression models were fitted to *S. mansoni* infection prevalence to identify significant demographic (age and sex), socio-economic and environmental covariates in STATA v. 8.0 (Stata Corporation, College Station, TX, USA). Those covariates with a significance level below 0·15 were fitted into logistic geostatistical models using WinBUGS v. 1.4 (Imperial College & Medical Research Council, London, UK). To take into account the spatial heterogeneity, location-specific random effects were integrated in the logistic models, assuming that they are distributed according to a multivariate normal distribution with variance-covariance matrix related to the variogram of the spatial process (Gelfand, Ravishanker & Ecker,

https://doi.org/10.1017/S0031182005007432 Published online by Cambridge University Press

1999). MCMC simulation was employed to estimate the model parameters (Gelfand & Smith, 1990). Significant covariates from the final model were selected to generate a smooth map of *S. mansoni* infection risk using Bayesian kriging (Diggle, Tawn & Moyeed, 1998). For appraisal of the best fitting model, the deviance information criterion (DIC) was applied (Spiegelhalter *et al*. 2002). A smaller DIC indicates a better model. Further details on the spatial models are given in the Appendix.

RESULTS

Study compliance and school elevation

Fig. 1 shows that in the school year of 2001/2002, 5448 schoolchildren were listed on the education registries for grades 3–5 in the 57 rural schools in the region of Man included in this study. During the parasitological surveys 5019 children were actually at school, but 264 failed to provide sufficient quantities of stool for diagnosis of *S. mansoni*. Hence, 4755 schoolchildren had a single Kato-Katz thick smear examined by light microscopy.

Questionnaire data were obtained from 55 schools. One school returned uncompleted forms, and 1 school failed to send back the questionnaires. Overall, the teachers interviewed 4376 schoolchildren. For spatial risk prediction and mapping, the final cohort consisted of 3818 schoolchildren, who had both parasitological and questionnaire data. There were 1528 girls (40·0%) and 2290 boys, with 2093 children aged 6–10 years (54·8%) and 1725 children aged 11–16 years.

School elevations ranged between 291 m and 842 m. Forty schools (73%) were located at elevations below 400 m, while the remaining 15 schools were located at higher elevations. The highest annual rainfall value (1128 mm) was observed at the school with the highest elevation. Rainfall and elevation were found to be positively correlated ($P<0·001$), however there was no co-linearity.

S. mansoni *infection*

The overall *S. mansoni* infection prevalence was 38·9%. Infection prevalences of girls and boys were 36·3% and 40·6%, respectively. Infection prevalences among the 6–10 year and the 11–16 year old children were 36·4% and 41·9%, respectively.

The cumulative infection prevalences in the schools ranged from 0% to 89·3%. Fig. 2 displays mean school *S. mansoni* infection prevalences at the 55 sampled locations. The highest prevalences were found in the south-western part of the study area. Some distinct foci of high prevalence were also found in the central part. In contrast, mean school prevalences of *S. mansoni* infections were below 20% in the north-eastern part of the study area.

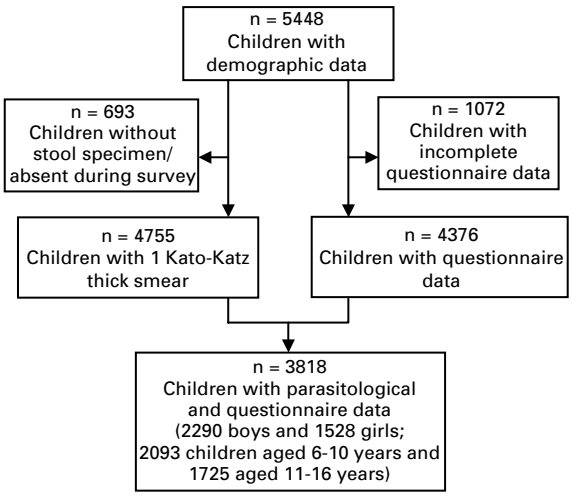

Fig. 1. Study profile and schoolchildren's compliance.

Associations between S. mansoni *and demographic, socio-economic and environmental covariates*

Table 1 shows the results of the non-spatial bivariate logistic regression analyses. Schoolchildrens' age and sex were significantly associated with an *S. mansoni* infection. Children from the older age group were significantly more likely to be infected with *S. mansoni* than their younger counterparts (odds ratio (OR) = 1·22, $P=0·001$). Boys had significantly higher infection prevalence than girls (OR = 1·19, $P=0·008$). There was no significant association between *S. mansoni* infection and socio-economic status.

Children living within the boundaries of the main villages showed significantly lower *S. mansoni* infection prevalences than those living in settlements outside (OR = 0·72, $P=0·001$). Children living 1–4·9 km away from the nearest health facility were more likely to be infected with *S. mansoni* than those living in close proximity (OR = 1·27, $P=0·002$).

The results from the non-spatial bivariate logistic regression between *S. mansoni* and the environmental covariates are summarized in Table 2. The annual mean NDVI showed the strongest association with *S. mansoni* infection among the set of covariates investigated. The association was highly negative; hence high *S. mansoni* infection prevalences were characterized by low NDVI values (OR = 0·001, $P<0·001$). Elevation was also negatively associated with the prevalence of *S. mansoni* infections at the school level (OR = 0·21, $P<0·001$). Children living above 400 m elevation were less likely to be infected with *S. mansoni*.

Spatial analysis of S. mansoni *infection*

The results of the bivariate spatial models between *S. mansoni* and demographic, socio-economic and physical environment covariates are shown in

https://doi.org/10.1017/S0031182005007432 Published online by Cambridge University Press

Table 1. Bivariate associations between *Schistosoma mansoni* infection prevalence and demographic, socio-economic and physical environment-related indicators arising from non-spatial and spatial logistic models

(Odds ratios (OR) are displayed with their respective 95% confidence intervals (CI).)

	Non-spatial models*		Spatial models†	
Indicator	OR	95% CI	OR	95% CI
Demography				
Age group				
6–10 years	1·00		1·00	
11–16 years	1·22	1·09, 1·37	1·48	1·26, 1·73
Sex				
Female	1·00		1·00	
Male	1·19	1·06, 1·34	1·24	1·06, 1·45
Socio-economic status				
Most poor	1·00		1·00	
Very poor	1·08	0·88, 1·33	1·12	0·86, 1·44
Poor	1·10	0·89, 1·35	1·12	0·86, 1·43
Less poor	0·99	0·80, 1·22	0·93	0·70, 1·19
Least poor	0·89	0·72, 1·10	0·73	0·55, 0·95
Physical environment				
Household within main village	0·72	0·59, 0·88	0·84	0·65, 1·08
Travel distance to nearest health facility				
<1 km	1·00		1·00	
1–4·9 km	1·27	1·09, 1·47	1·84	0·88, 3·44
⩾5 km	0·88	0·76, 1·01	1·85	0·76, 3·78

* Non-spatial models were fitted in STATA.
† Spatial Bayesian models were fitted in WinBUGS.

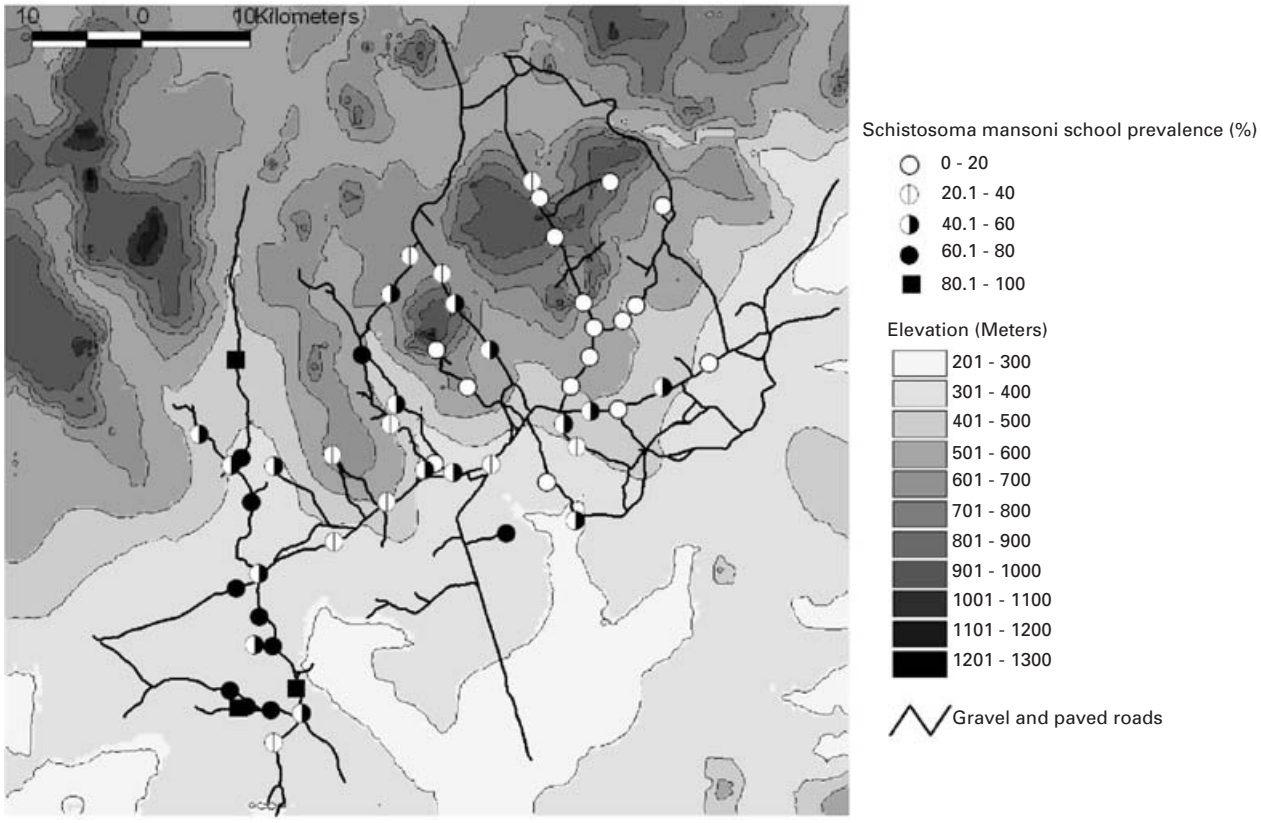

Fig. 2. Distribution of mean *S. mansoni* infection prevalences in 55 rural schools in the mountainous region of Man, western Côte d'Ivoire.

https://doi.org/10.1017/S0031182005007432 Published online by Cambridge University Press

Table 2. Bivariate associations between *Schistosoma mansoni* infection prevalence and environmental indicators arising from non-spatial and spatial logistic models

(Odds ratios (OR) are displayed with their respective 95% confidence intervals (CI).)

	Non-spatial models*		Spatial models†	
Environmental indicator	OR	95% CI	OR	95% CI
Minimum LST				
<25·0 °C	1·00		1·00	
25·0–26·4 °C	0·81	0·71, 0·91	0·99	0·43, 1·91
⩾26·5 °C	0·34	0·27, 0·42	0·43	0·10, 1·21
Mean LST				
<25·0 °C	1·00		1·00	
25·0–26·4 °C	0·57	0·50, 0·66	1·06	0·48, 2·02
⩾26·5 °C	0·27	0·22, 0·33	1·88	0·33, 5·89
Maximum LST				
<25·0 °C	1·00		1·00	
25·0–26·4 °C	1·48	1·03, 2·11	3·05	0·41, 10·88
⩾26·5 °C	1·02	0·71, 1·45	4·83	0·64, 17·31
Day-night temperature difference				
<6·0 °C	1·00		1·00	
6·0–7·9 °C	0·50	0·44, 0·58	1·85	0·63, 4·21
⩾8·0 °C	0·33	0·26, 0·42	1·90	0·27, 6·41
Rainfall	0·99	0·99, 0·99	0·99	0·99, 0·99
NDVI	0·001	0·0003, 0·006	6·35	0·00, 183·6
Land cover				
Forest	1·00		1·00	
Savannah	0·99	0·87, 1·12	1·12	0·59, 1·96
Cropland	0·75	0·61, 0·90	0·74	0·30, 1·55
Soil types				
Ferrallitic soils	1·00		1·00	
Hydromorphic mineral soils	1·96	1·63, 2·36	1·08	0·38, 2·50
Elevation				
<400 m	1·00		1·00	
⩾400 m	0·21	0·17, 0·25	0·22	0·08, 0·54
Slope of the landscape				
<20°	1·00		1·00	
20–39°	0·03	0·01, 0·08	0·23	0·01, 1·21
40–59°	0·32	0·24, 0·43	1·22	0·19, 4·10
⩾60°	0·18	0·07, 0·43	0·34	0·02, 1·70
Distance to permanent water bodies				
<500 m	1·00		1·00	
500–999 m	1·33	1·14, 1·55	1·02	0·45, 2·03
⩾1000 m	1·21	1·05, 1·40	0·58	0·24, 1·19

* Non-spatial models were fitted in STATA.
† Spatial Bayesian models were fitted in WinBUGS.

Table 1, and the results between *S. mansoni* and environmental covariates are summarized in Table 2. Age group, sex, annual rainfall and elevation remained significant covariates when spatial correlation was taken into account. The 5th wealth quintile was significant in this bivariate spatial model. NDVI, which was highly significant in the non-spatial model, was not related to *S. mansoni* infection prevalence after accounting for spatial correlation.

Results of a multivariate non-spatial model (Model 1) and 3 different multivariate spatial models (Models 2–4) are summarized in Table 3. The lower DICs of the spatial models underscore the importance of taking into account spatial correlation when analysing data in space. To assess whether environmental factors or socio-economic status had a more pronounced effect on *S. mansoni* infection, the goodness of fit of Models 2–4 were compared with each other. Model 2 included schoolchildren's demographic covariates and their socio-economic status. Model 3 was built on demographic covariates and elevation alone, whereas Model 4 additionally had schoolchildren's socio-economic status included. Model 4 showed the best fit. Schoolchildren's socio-economic status was more important than

https://doi.org/10.1017/S0031182005007432 Published online by Cambridge University Press

Table 3. Comparison of 4 Bayesian models fitted in WinBUGS

(Values represent mean odds ratios (OR) with 95% confidence intervals (CI).)

	Model 1 (non-spatial)		Model 2 (spatial)		Model 3 (spatial)		Model 4 (spatial)	
Covariate	OR	95% CI	OR	95% CI	OR	95% CI	OR	95% CI
Demography								
Age group								
6–10 years	1·00		1·00		1·00		1·00	
11–16 years	1·27	1·11, 1·46	1·45	1·24, 1·71	1·47	1·25, 1·72	1·46	1·24, 1·71
Sex								
Female	1·00		1·00		1·00		1·00	
Male	1·18	1·03, 1·36	1·21	1·03, 1·42	1·22	1·04, 1·43	1·21	1·03, 1·41
Socio-economic status								
Most poor	1·00		1·00				1·00	
Very poor	1·06	0·84, 1·31	1·12	0·85, 1·44			1·12	0·85, 1·44
Poor	1·07	0·86, 1·33	1·09	0·83, 1·40			1·08	0·82, 1·39
Less poor	0·97	0·78, 1·20	0·92	0·70, 1·84			0·91	0·69, 1·18
Least poor	0·72	0·58, 0·89	0·74	0·56, 0·96			0·73	0·55, 0·95
Environment								
Elevation								
<400 m	1·00				1·00		1·00	
⩾400 m	0·20	0·16, 0·24			0·22	0·07, 0·53	0·20	0·07, 0·50
Sigma*			2·80	1·37, 6·44	1·73	0·92, 3·71	1·72	0·93, 3·53
u†			0·0001	0·00004, 0·0003	0·0003	0·00007, 0·002	0·0004	0·00008, 0·002
DIC‡	**4717·09**		**4025·43**		**4029·26**		**4025·03**	

* Sigma is the estimate of the geographical variability.
† *u* is the smoothing parameter (correlation decay), measuring the range of the geographical dependency. 3/*u* indicates the minimum distance at which spatial correlation between 2 locations becomes less than 5%.
‡ DIC is the measure for the model fit. A smaller DIC indicates a better fit of the model data.

elevation in explaining the small-scale distribution of *S. mansoni* infection. Adding the rainfall covariate to Model 4 failed to improve the fit of this model.

In the current epidemiological setting, the minimum distance at which spatial correlation between 2 locations was below 5% was 7·5 km.

Prediction of S. mansoni *infection*

The predicted prevalence of *S. mansoni* infection among school-age children in the region of Man, western Côte d'Ivoire, is shown in Fig. 3. Low prevalences were predicted for the north-eastern part, whereas high prevalences were predicted for the south-western part of the study area, thus confirming the point estimates derived from the school-based parasitological surveys. The map shows surface predictions of *S. mansoni* infection prevalence at non-sampled locations.

The standard deviation of the predicted *S. mansoni* infection prevalences is given in Fig. 4. This map exhibits the precision of the prediction. It shows that with increasing distance from surveyed locations the error of the prediction increased.

Fig. 5 shows the random effects of the predictive model at non-sampled locations. The random effects basically constitute a residual term that indicates how much is explained by other factors, which are not included in Model 4. Large absolute random effect values correspond to a larger deviation from the average prevalence. The random effects with the highest absolute values were observed in areas of either low or high predicted *S. mansoni* infection prevalence. Hence, the uncertainty of the model is higher in those areas.

DISCUSSION

Microscopical examination of a single Kato-Katz thick smear from a population sample of 3818 schoolchildren in rural western Côte d'Ivoire revealed an overall *S. mansoni* infection prevalence of 38·9%. These results confirm the high endemicity of this parasitic infection in this epidemiological setting (Utzinger *et al.* 2000; Raso *et al.* 2005). Analysis at the school level showed that the infection prevalence of *S. mansoni* varied considerably from one location to another, often within short distances, confirming small-scale focality of this parasite (Greer *et al.* 1990; Ratard *et al.* 1990; Gryseels, 1991; Lengeler *et al.* 2002). Bayesian spatial statistical analysis revealed that the covariates age, sex, 5th wealth quintile, elevation and rainfall explained the geographical variation of *S. mansoni* infection prevalence. Interestingly, age, sex and socio-economic status showed stronger influence on the geographical

https://doi.org/10.1017/S0031182005007432 Published online by Cambridge University Press

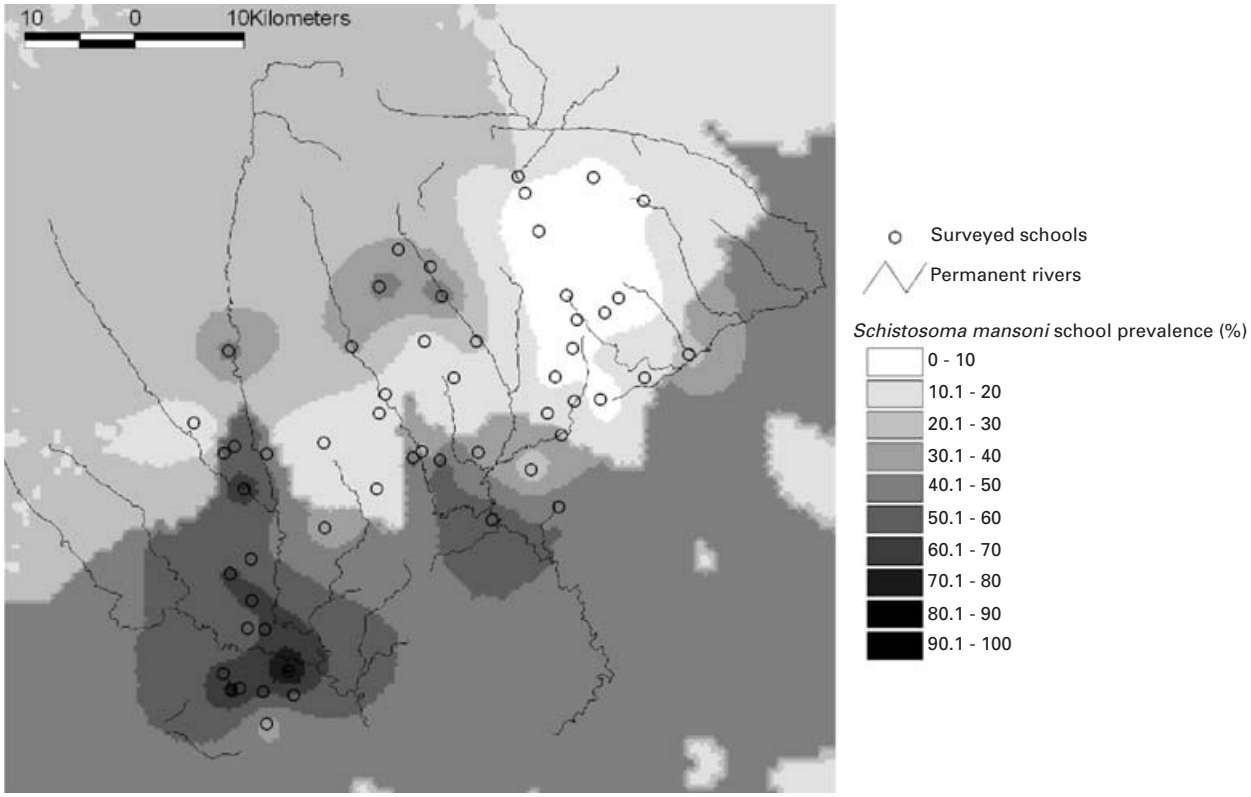

Fig. 3. Smoothed map of the predicted *Schistosoma mansoni* infection prevalence based on Model 4 and produced with Bayesian kriging for the region of Man, western Côte d'Ivoire.

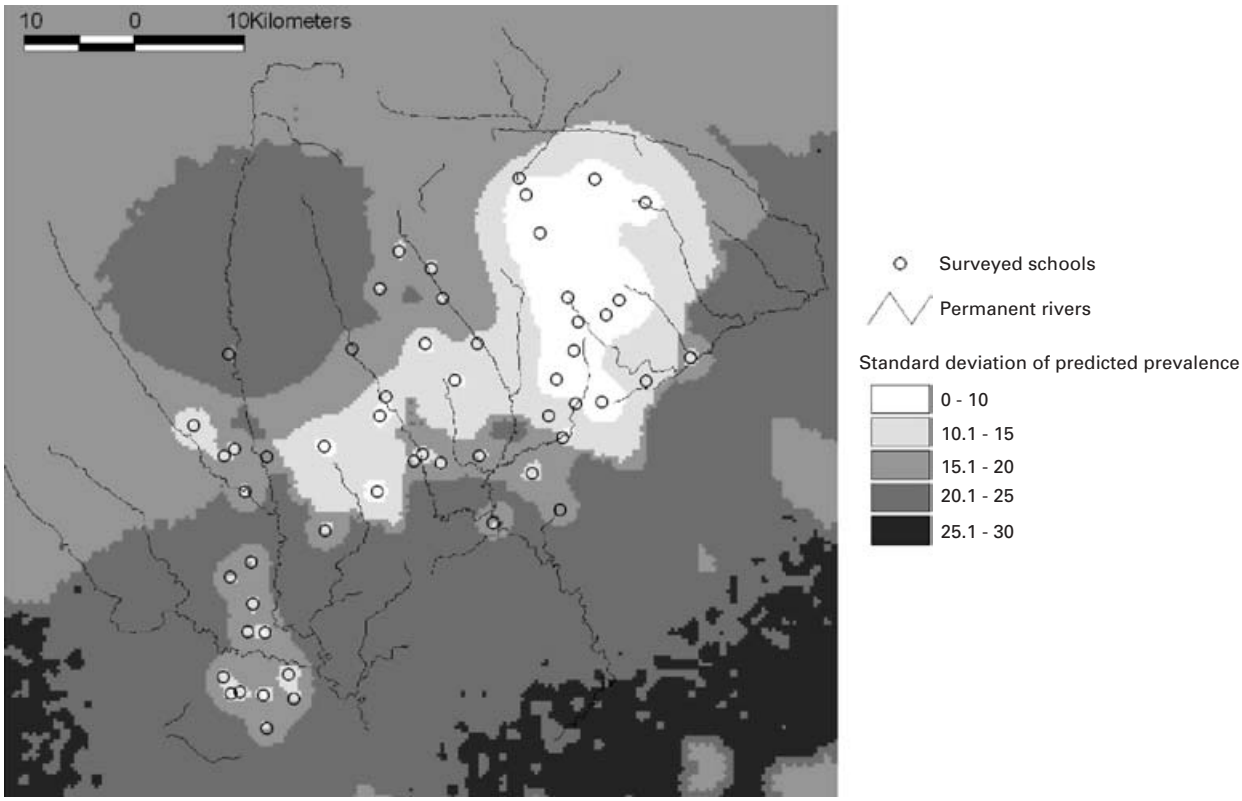

Fig. 4. Smoothed map with standard errors of the predicted *S. mansoni* infection prevalence based on Model 4 and produced with Bayesian kriging for the region of Man, western Côte d'Ivoire.

variation of this parasite at small-scale when compared to the environmental covariates investigated.

Three methodological shortcomings are important to note. First, diagnostic sensitivity of a single Kato-Katz thick smear is low due to the significant

https://doi.org/10.1017/S0031182005007432 Published online by Cambridge University Press

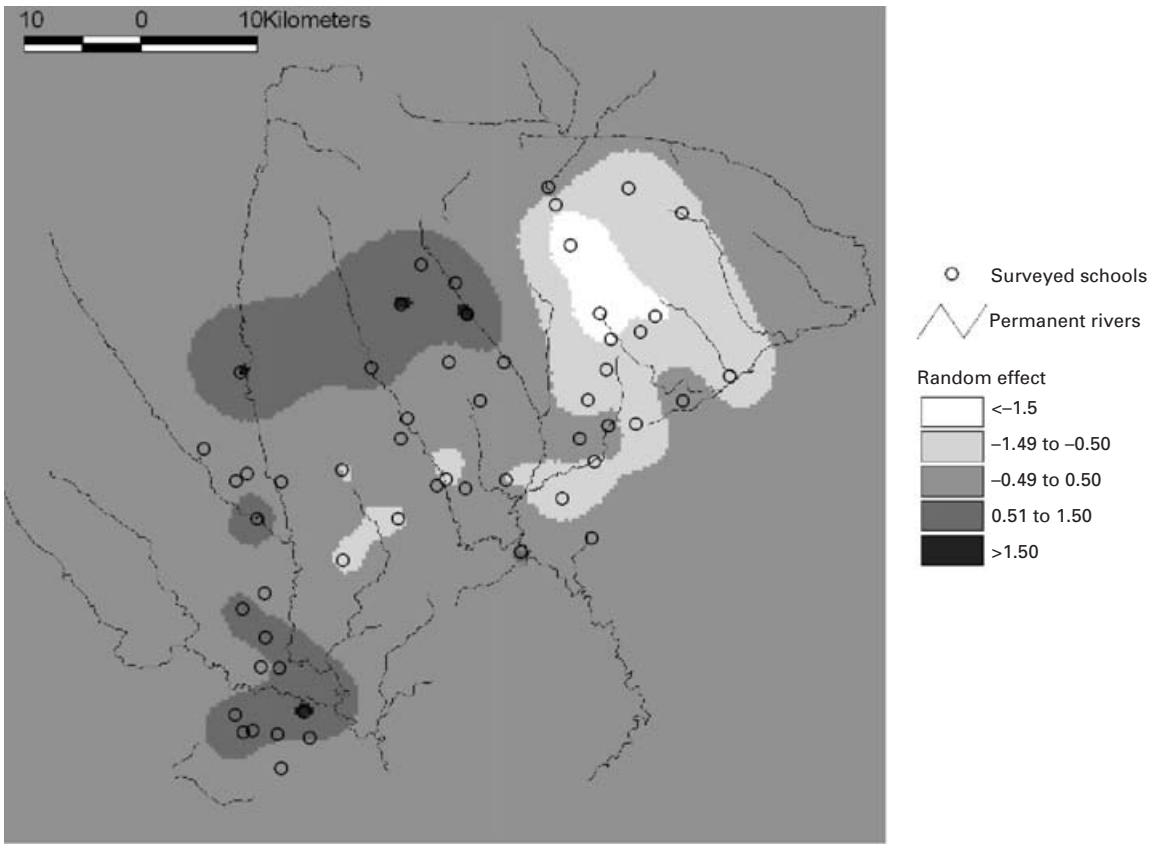

Fig. 5. Smoothed map with location-specific random effects based on Model 4 and produced with Bayesian kriging for the region of Man, western Côte d'Ivoire.

day-to-day and intra-specimen variation of *S. mansoni* egg counts in stool samples (Engels, Sinzinkayo & Gryseels, 1996; Utzinger *et al*. 2001; Booth *et al*. 2003; Berhe *et al*. 2004; Raso *et al*. 2004). However, mainly light infections are being missed, while moderate and heavy infections, which are particularly important from a public health point of view, are more likely to be detected. Second, we have employed environmental factors with different spatial resolutions. Using data at different spatial resolutions may affect the model results by introducing bias (Levin, 1992). Third, the alignment between the environmental covariates and the locations where the *S. mansoni* prevalence data have been collected (i.e. school) introduces a bias, since the environmental features correspond to a wider geographical area than single point measures (Curran *et al*. 2000).

Previous work has been done to investigate how exposure affects *S. mansoni* infection. Distance to transmission sites, human water contact patterns, and household characteristics were associated with individuals or groups at highest risk of *S. mansoni* infections (Husting, 1983; Bethony *et al*. 2001, 2004; Gazzinelli *et al*. 2001). For example, micro-geographical studies carried out in Brazil and Kenya showed that distance from households to streams was negatively associated with *S. mansoni* infection (Kloos *et al*. 1997; Kloos, Gazzinelli & van Zuyle, 1998). With a view to approximate exposure related to water contact, we have used the distance to permanent rivers. In the non-spatial logistic regression analysis, this feature was significantly associated to an *S. mansoni* infection. However, in the spatially-explicit models, distance to permanent water bodies showed no significant association. It follows that water contact activities approximated by distance to rivers may be more complex than previously proposed. Consequently, it seems difficult to capture water contact activities adequately with this single variable at a meso-geographical scale, because the relationship may be biased by uneven distribution of intermediate host snails.

The results of the best fitting spatial model (Model 4) showed that the 5th wealth quintile explained a significant part of the geographical variation of *S. mansoni*, as shown by the improved goodness of fit of the model. One possible explanation arises from distance to health-care delivery centres. However, as previously shown and confirmed in this study, there is no significant association with the school prevalence of *S. mansoni* infection and physical access, as measured by travel distance, to the nearest health facility (Raso *et al*. 2005). Household decision-making processes may play an equal or even more important role in this case. Underlying reasons include health seeking behaviour, transport costs to a dispensary, treatment costs, and knowledge and perception of the disease by household members, as recently shown in

https://doi.org/10.1017/S0031182005007432 Published online by Cambridge University Press

studies from Tanzania (Armstrong Schellenberg *et al*. 2003). On the other hand, socio-economic status may also be an indicator for the presence of a latrine in the richer wealth quintiles, which in turn is a protective factor for *S. mansoni* transmission. Taken together, socio-economic status plays a role in the small-scale distribution of *S. mansoni*, but it acts as an aggregate measure for multiple risk factors that are connected to well-being and equity.

Elevation and rainfall were the 2 most important environmental covariates to capture the spatial distribution of *S. mansoni* infection prevalence. These 2 covariates were highly correlated. It is plausible that the mountains in the northern part of the study area are responsible for the higher rainfall observed there, generating distinct regional climate conditions. It is speculated that both covariates have an effect on the flow velocity of rivers, which in turn is likely to influence the presence of *B. pfeifferi*, the intermediate host snail of *S. mansoni*. Malacological studies have shown that *B. pfeifferi* tolerate a maximum flow speed of up to 0·3 m/s (Appleton, 1978; Kloos *et al*. 2001). Recently, the slope of the landscape has been proposed as a proxy for water flow (Brooker & Michael, 2000). Although this covariate showed no significant association to *S. mansoni* infection prevalence in the spatial models presented here, the elevation, which is related to the slope, was significant. Elevation thus is a useful covariate to indicate higher flow speed in the mountainous part of the present study area. In addition to elevation, rainfall is another covariate that is likely to increase the flow speed of rivers (Brooker & Michael, 2000). It follows that *B. pfeifferi* are less likely to populate rivers in mountainous parts. Snails are more likely to proliferate in the plain, which would explain why children living at locations below 400 m were at a 5-fold higher risk of an *S. mansoni* infection when compared to those living at altitudes above this threshold.

In contrast to previous GIS/RS applications with an emphasis on schistosomiasis, the study presented here is, to our knowledge, one of the first attempts to carry forward Bayesian geostatistics for assessment of environmental risk factors. The results underscore the importance of using appropriate geostatistical approaches for the analysis of spatially-explicit data. Different environmental covariates were tested for their association with the school prevalence of *S. mansoni* infection. The majority of the environmental factors were significantly associated to *S. mansoni* infection in the non-spatial logistic regressions. However, when adding a spatial component to the logistic models, NDVI, LST, and day-night temperature difference were not significant anymore. This finding is expected since omission of spatial correlation underestimates the standard errors of the covariate coefficient (Cressie, 1993).

The results show that demographic covariates and socio-economic status play a more important role in the small-scale variation of *S. mansoni* infection than any of the environmental covariates investigated, as assessed by the DIC. However, in areas of either low or high prevalences, model predictions were less certain, suggesting that other risk factors (e.g. of behavioural or genetic nature, which were not included in the analysis) might also play important roles in these areas. Furthermore, ecological details at small-scale that are important for *B. pfeifferi* habitat formation (Genner & Michel, 2003), and which are suppressed in RS studies at broad or regional scale, may also be at the origin of the uncertainty expressed in our model. Although the random effects and standard errors of the predictions teach us about the model uncertainty, our model needs to be validated in order to assess its reliability, e.g. by random sampling at other school locations that have not been currently included.

Concluding, smoothed maps of *S. mansoni* infection risk, together with the standard error of the predictions and random effects displaying uncertainties, were generated for the region of Man. These maps can be used by decision-makers for the design and implementation of *S. mansoni* control interventions reaching those at highest risk. An important next step is to validate the models in similar ecozones. If they perform well, predictive modelling can guide schistosomiasis control programmes in similar eco-epidemiological settings of Côte d'Ivoire and elsewhere in sub-Saharan Africa.

Thanks are addressed to the education officers, the directors and teachers of the schools surveyed, and the field and laboratory team (A. Allangba, A. Fondio, K. L. Lohourignon, F. Sangaré, B. Sosthène and M. Traoré) for their commitment in this study. We are indebted to Dr C. Châtelain for help with the establishment of the GIS for the region of Man. This investigation received financial support from the Claire Sturzenegger-Jean Favre Foundation, the Roche Research Foundation through a fellowship to G. Raso, the Swiss National Science Foundation (SNF) supporting P. Vounatsou (Project No. 3252B0-102136) and J. Utzinger through an 'SNF-Förderungsprofessur' (Project No. PP00B-102883), and the Individual Project 4 (IP4) 'Health and Well-being' of the NCCR North-South: 'Research Partnerships for Mitigating Syndromes of Global Change', funded by SNF, with partial support to B. Matthys and M. Tanner.

REFERENCES

APPLETON, C. C. (1978). Review of literature on abiotic factors influencing the distribution and life-cycles of bilharziasis intermediate host snails. *Malacological Review* **11**, 1–25.

ARMSTRONG SCHELLENBERG, J., VICTORA, C. G., MUSHI, A., DE SAVIGNY, D., SCHELLENBERG, D., MSHINDA, H. & BRYCE, J. (2003). Inequities among the very poor: health care for children in rural southern Tanzania. *Lancet* **361**, 561–566.

BAVIA, M. E., MALONE, J. B., HALE, L., DANTAS, A., MARRONI, L. & REIS, R. (2001). Use of thermal and vegetation index

https://doi.org/10.1017/S0031182005007432 Published online by Cambridge University Press

data from earth observing satellites to evaluate the risk of schistosomiasis in Bahia, Brazil. *Acta Tropica* **79**, 79–85.

BERHE, N., MEDHIN, G., ERKO, B., SMITH, T., GEDAMU, S., BEREDED, D., MOORE, R., HABTE, E., REDDA, A., GEBRE-MICHAEL, T. & GUNDERSEN, S. G. (2004). Variations in helminth faecal egg counts in Kato-Katz thick smears and their implications in assessing infection status with *Schistosoma mansoni*. *Acta Tropica* **92**, 205–212.

BETHONY, J., WILLIAMS, J. T., BROOKER, S., GAZZINELLI, A., GAZZINELLI, M. F., LO VERDE, P. T., CORREA-OLIVEIRA, R. & KLOOS, H. (2004). Exposure to *Schistosoma mansoni* infection in a rural area in Brazil. Part III: Household aggregation of water-contact behaviour. *Tropical Medicine and International Health* **9**, 381–389.

BETHONY, J., WILLIAMS, J. T., KLOOS, H., BLANGERO, J., ALVES-FRAGA, L., BUCK, G., MICHALEK, A., WILLIAMS-BLANGERO, S., LO VERDE, P. T., CORREA-OLIVEIRA, R. & GAZZINELLI, A. (2001). Exposure to *Schistosoma mansoni* infection in a rural area in Brazil. II: Household risk factors. *Tropical Medicine and International Health* **6**, 136–145.

BOOTH, M., VOUNATSOU, P., N'GORAN, E. K., 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). Schistosomes, snails and satellites. *Acta Tropica* **82**, 207–214.

BROOKER, S., HAY, S. I. & BUNDY, D. A. P. (2002*a*). Tools from ecology: useful for evaluating infection risk models? *Trends in Parasitology* **18**, 70–74.

BROOKER, S., HAY, S. I., ISSAE, W., HALL, A., KIHAMIA, C. M., LWAMBO, N. J., WINT, W., ROGERS, D. J. & BUNDY, D. A. P. (2001). Predicting the distribution of urinary schistosomiasis in Tanzania using satellite sensor data. *Tropical Medicine and International Health* **6**, 998–1007.

BROOKER, S., HAY, S. I., TCHUEM TCHUENTÉ, L. A. & RATARD, R. (2002*b*). Using NOAA-AVHRR data to model human helminth distributions on planning disease control in Cameroon, West Africa. *Photogrammetric Engineering and Remote Sensing* **68**, 175–179.

BROOKER, S. & MICHAEL, E. (2000). The potential of geographical information systems and remote sensing in the epidemiology and control of human helminth infections. *Advances in Parasitology* **47**, 245–288.

CHITSULO, L., ENGELS, D., MONTRESOR, A. & SAVIOLI, L. (2000). The global status of schistosomiasis and its control. *Acta Tropica* **77**, 41–51.

CRESSIE, N. A. C. (1993). *Statistics for Spatial Data*. Wiley, New York.

CURRAN, P. J., ATKINSON, P. M., FOODY, G. M. & MILTON, E. J. (2000). Linking remote sensing, land cover and disease. *Advances in Parasitology* **47**, 37–80.

DIGGLE, P. J., TAWN, J. A. & MOYEED, R. A. (1998). Model-based geostatistics. *Applied Statistics* **47**, 299–350.

DOUMENGE, J. P., MOTT, K. E., CHEUNG, C., VILLENAVE, D., CHAPUIS, O., PERRIN, M. F. & REAUD-THOMAS, G. (1987). *Atlas of the Global Distribution of Schistosomiasis*. CEGET/CNRS – Word Health Organization, Talence/Geneva.

ENGELS, D., SINZINKAYO, E. & GRYSEELS, B. (1996). Day-to-day egg count fluctuation in *Schistosoma mansoni* infection and its operational implications. *American Journal of Tropical Medicine and Hygiene* **54**, 319–324.

FILMER, D. & PRITCHETT, L. H. (2001). Estimating wealth effects without expenditure data – or tears: an application to educational enrollments in states of India. *Demography* **38**, 115–132.

GAZZINELLI, A., BETHONY, J., FRAGA, L. A., LO VERDE, P. T., CORREA-OLIVEIRA, R. & KLOOS, H. (2001). Exposure to *Schistosoma mansoni* infection in a rural area of Brazil. I: Water contact. *Tropical Medicine and International Health* **6**, 126–135.

GELFAND, A. E., RAVISHANKER, N. & ECKER, M. (1999). Modeling and inference for point-referenced binary spatial data. In *Generalized Linear Models* (ed. Dey, D., Ghosh, S. & Mallick, B.), pp. 373–386. Marcel Dekker Inc., New York.

GELFAND, A. E. & SMITH, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**, 398–409.

GENNER, M. J. & MICHEL, E. (2003). Fine-scale habitat associations of soft-sediment gastropods at Cape Maclear, Lake Malawi. *Journal of Molluscan Studies* **69**, 325–328.

GREER, G. J., MIMPFOUNDI, R., MALEK, E. A., JOKY, A., NGONSEU, E. & RATARD, R. C. (1990). Human schistosomiasis in Cameroon. II. Distribution of the snail hosts. *American Journal of Tropical Medicine and Hygiene* **42**, 573–580.

GRYSEELS, B. (1991). The epidemiology of schistosomiasis in Burundi and its consequences for control. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **85**, 626–633.

HAY, S. I. (2000). An overview of remote sensing and geodesy for epidemiology and public health application. *Advances in Parasitology* **47**, 1–35.

HUSTING, E. L. (1983). Human water contact activities related to the transmission of bilharziasis (schistosomiasis). *Journal of Tropical Medicine and Hygiene* **86**, 23–35.

KATZ, N., CHAVES, A. & PELLEGRINO, J. (1972). A simple device for quantitative stool thick-smear technique in schistosomiasis mansoni. *Revista do Instituto Medicina Tropical São Paulo* **14**, 397–400.

KEISER, J., N'GORAN, E. K., TRAORÉ, M., LOHOURIGNON, K. L., SINGER, B. H., LENGELER, C., TANNER, M. & UTZINGER, J. (2002). Polyparasitism with *Schistosoma mansoni*, geohelminths, and intestinal protozoa in rural Côte d'Ivoire. *Journal of Parasitology* **88**, 461–466.

KLOOS, H., DE SOUZA, C., GAZZINELLI, A., SOARES FILHO, B. S., DA COSTA, T., BETHONY, J., PAGE, K., GRZYWACZ, C., LEWIS, F., MINCHELLA, D., LO VERDE, P. & OLIVEIRA, R. C. (2001). The distribution of *Biomphalaria* spp. in different habitats in relation to physical, biological, water contact and cognitive factors in a rural area in Minas Gerais, Brazil. *Memorias do Instituto Oswaldo Cruz* **96**, 57–66.

KLOOS, H., FULFORD, A. J., BUTTERWORTH, A. E., STURROCK, R. F., OUMA, J. H., KARIUKI, H. C., THIONGO, F. W., DALTON, P. R. & KLUMPP, R. K. (1997). Spatial patterns of human water contact and *Schistosoma mansoni* transmission and infection in four rural areas in Machakos district, Kenya. *Social Science and Medicine* **44**, 949–968.

KLOOS, H., GAZZINELLI, A. & VAN ZUYLE, P. (1998). Microgeographical patterns of schistosomiasis and water contact behavior; examples from Africa and Brazil. *Memorias do Instituto Oswaldo Cruz* **93** (Suppl. 1), 37–50.

KRISTENSEN, T. K., MALONE, J. B. & McCARROLL, J. C. (2001). Use of satellite remote sensing and geographic information systems to model the distribution and abundance of snail intermediate hosts in Africa: a preliminary model for *Biomphalaria pfeifferi* in Ethiopia. *Acta Tropica* **79**, 73–78.

LENGELER, C., UTZINGER, J. & TANNER, M. (2002). Questionnaires for rapid screening of schistosomiasis in sub-Saharan Africa. *Bulletin of the World Health Organization* **80**, 235–242.

LEVIN, S. A. (1992). The problem of pattern and scale in ecology. *Ecology* **73**, 1943–1967.

MALONE, J. B., YILMA, J. M., McCARROLL, J. C., ERKO, B., MUKARATIRWA, S. & ZHOU, X. N. (2001). Satellite climatology and the environmental risk of *Schistosoma mansoni* in Ethiopia and East Africa. *Acta Tropica* **79**, 59–72.

RASO, G., N'GORAN, E. K., TOTY, A., LUGINBÜHL, A., ADJOUA, C. A., TIAN-BI, N. T., BOGOCH, I. I., VOUNATSOU, P., TANNER, M. & UTZINGER, J. (2004). Efficacy and side effects of praziquantel against *Schistosoma mansoni* in a community of western Côte d'Ivoire. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **98**, 18–27.

RASO, G., UTZINGER, J., SILUÉ, K. D., OUATTARA, M., YAPI, A., TOTY, A., MATTHYS, B., VOUNATSOU, P., TANNER, M. & N'GORAN, E. K. (2005). Disparities in parasitic infections, perceived ill health and access to health care among poorer and less poor schoolchildren of rural Côte d'Ivoire. *Tropical Medicine and International Health* **10**, 42–57.

RATARD, R. C., KOUEMENI, L. E., BESSALA, M. M., NDAMKOU, C. N., GREER, G. J., SPILSBURY, J. & CLINE, B. L. (1990). Human schistosomiasis in Cameroon. I. Distribution of schistosomiasis. *American Journal of Tropical Medicine and Hygiene* **42**, 561–572.

ROBINSON, T. P. (2000). Spatial statistics and geographical information systems in epidemiology and public health. *Advances in Parasitology* **47**, 81–128.

SPIEGELHALTER, D. J., BEST, N., CHARLIN, B. P. & VAN DER LINDE, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B* **64**, 583–639.

UTZINGER, J., BERGQUIST, R., XIAO, S. H., SINGER, B. H. & TANNER, M. (2003). Sustainable schistosomiasis control – the way forward. *Lancet* **362**, 1932–1934.

UTZINGER, J., BOOTH, M., N'GORAN, E. K., MÜLLER, I., TANNER, M. & LENGELER, C. (2001). Relative contribution of day-to-day and intra-specimen variation in faecal egg counts of *Schistosoma mansoni* before and after treatment with praziquantel. *Parasitology* **122**, 537–544.

UTZINGER, J., N'GORAN, E. K., OSSEY, Y. A., BOOTH, M., TRAORÉ, M., LOHOURIGNON, K. L., ALLANGBA, A., AHIBA, L. A., TANNER, M. & LENGELER, C. (2000). Rapid screening for *Schistosoma mansoni* in western Côte d'Ivoire using a simple school questionnaire. *Bulletin of the World Health Organization* **78**, 389–398.

VAN DER WERF, M. J., DE VLAS, S. J., BROOKER, S., LOOMAN, C. W. N., NAGELKERKE, N. J. D., HABBEMA, J. D. F. & ENGELS, D. (2003). Quantification of clinical morbidity associated with schistosome infection in sub-Saharan Africa. *Acta Tropica* **86**, 125–139.

WATTS, S., KHALLAAYOUNE, K., BENSEFIA, R., LAAMRANI, H. & GRYSEELS, B. (1998). The study of human behavior and schistosomiasis transmission in an irrigated area in Morocco. *Social Science and Medicine* **46**, 755–765.

WORLD HEALTH ORGANIZATION (2002). *Prevention and Control of Schistosomiasis and Soil-Transmitted Helminthiasis: Report of a WHO Expert Committee*. WHO Technical Report Series No. 912. WHO, Geneva.

YANG, G. J., VOUNATSOU, P., ZHOU, X. N., TANNER, M. & UTZINGER, J. (2005). A Bayesian-based approach for spatio-temporal modeling of county level prevalence of *Schistosoma japonicum* infection in Jiangsu province, China. *International Journal for Parasitology*, **35**, 155–162.

ZHOU, X. N., MALONE, J. B., KRISTENSEN, T. K. & BERGQUIST, N. R. (2001). Application of geographic information systems and remote sensing to schistosomiasis control in China. *Acta Tropica* **79**, 97–106.

APPENDIX

Let Y_{ij} and p_{ij} be the status and probability of an *S. mansoni* infection, respectively, of schoolchild j in village i. We assume that Y_{ij} arises from a Bernoulli distribution, $Y_{ij} \sim Be(p_{ij})$. We model covariates X_{ij} and village-specific random effect ϕ_i on the log $it(p_{ij})$, that is log $it(p_{ij}) = X_{ij}^T\beta + \phi_i$, where β is the vector of regression coefficients. We introduce the spatial correlation on the ϕ_i's by assuming that $\phi = (\phi_1, \phi_1, \ldots, \phi_N)^T$ has a multivariate normal distribution, $\phi \sim MVN(0, \Sigma)$, with variance-covariance matrix Σ. We also assume an isotropic stationary spatial process, where $\Sigma_{kl} = \sigma^2 \exp(-ud_{kl})$, d_{kl} is the Euclidean distance between villages k and l, σ^2 is the geographical variability known as the sill, u is a smoothing parameter that controls the rate of correlation decay with increasing distance and measures the range of geographical dependency. The range is defined as the minimum distance at which spatial correlation between locations is below 5%. This distance can be calculated as $3/u$ meters.

Following a Bayesian model specification, we adopt prior distributions for the model parameters. We choose vague Normal distributions for the β parameters with large variances (i.e. 10 000), an inverse gamma prior for σ^2 and a uniform prior for u. MCMC simulation was applied to fit the models. We run a single chain sampler with a burn-in of 5000 iterations. Convergence was assessed by inspection of ergodic averages of selected model parameters. The chain thereafter sampled every single iteration, until a sample size of 5000 had been attained.

https://doi.org/10.1017/S0031182005007432 Published online by Cambridge University Press