Tropical Medicine and International Health

VOLUME 6 NO 12 PP 998–1007 DECEMBER 2001

Predicting the distribution of urinary schistosomiasis in Tanzania using satellite sensor data

Simon Brooker[1], Simon I. Hay[2], Wahab Issae[3], Andrew Hall[4], Charles M. Kihamia[3], Nicholas J. S. Lwambo[5], William Wint[6], David J. Rogers[2] and Don A. P. Bundy[7]

1 *Department of Infectious Disease Epidemiology, Imperial College School of Medicine, London, UK*
2 *Trypanosomiasis and Land Use in Africa (TALA) Research Group, Department of Zoology, University of Oxford, Oxford, UK*
3 *Tanzania Partnership for Child Development, Dar es Salaam, Tanzania*
4 *Helen Keller International, Dhaka, Bangladesh*
5 *National Institute for Medical Research, Mwanza Research Centre, Mwanza, Tanzania*
6 *Environmental Research Group Oxford, Oxford, UK*
7 *Human Development Division, The World Bank, Washington DC, USA*

Summary In this paper, remotely sensed (RS) satellite sensor environmental data, using logistic regression, are used to develop prediction maps of the probability of having infection prevalence exceeding 50%, and warranting mass treatment according to World Health Organization (WHO) guidelines. The model was developed using data from one area of coastal Tanzania and validated with independent data from different areas of the country. Receiver operating characteristic (ROC) analysis was used to evaluate the model's predictive performance. The model allows reasonable discrimination between high and low prevalence schools, at least within those geographical areas in which they were originally developed, and performs reasonably well in other coastal areas, but performs poorly by comparison in the Great Lakes area of Tanzania. These results may be explained by reference to an ecological zone map based on RS-derived environmental data. This map suggests that areas where the model reliably predicts a high prevalence of schistosomiasis fall within the same ecological zone, which has common intermediate-host snail species responsible for transmission. By contrast, the model's performance is poor near Lake Victoria, which is in a different ecological zone with different snail species. The ecological map can potentially define a template for those areas where existing models can be applied, and highlight areas where further data and models are required. The developed model was then used to provide estimates of the number of schoolchildren at risk of high prevalence and associated programme costs.

keywords urinary schistosomiasis, *Schistosoma haematobium*, prediction, remote sensing, receiver operating characteristic analysis, Tanzania

correspondence Simon Brooker, Department of Infectious Disease Epidemiology, Imperial College School of Medicine, Norfolk Place, London W2 1PG, UK. Fax: +44(0)20 7262 7912; E-mail: s.brooker@ic.ac.uk

Introduction

For *Schistosoma haematobium*, an effective approach to help locate high-risk communities or schools requiring mass treatment has been the use of questionnaires (Red Urine Study Group 1995; Partnership for Child Development 1999a). Before undertaking a questionnaire survey, a first step in national planning is to identify those areas in which surveys are needed (WHO 1995b). Yet, in many African countries a paucity of epidemiological data hinders the targeting of a questionnaire approach (Brooker *et al.* 2000). In an effort to overcome this problem, remotely sensed (RS) satellite sensor data and interpolated meteorological surfaces are increasingly used to predict infection risk in unsampled areas (Hay *et al.* 2000; Lindsay & Thomas 2000; Rogers 2000; Hendricks *et al.* 2001). RS data has also been used to estimate the likely distribution of *S. mansoni* in Egypt (Malone *et al.* 1994, 1997) and Ethiopia (Malone *et al.* 2001).

However, if reliable maps of infectious diseases are to be constructed, there is a need to investigate whether prediction models developed for one place can be applied to another, as environmental factors that influence disease transmission are unlikely to be uniform over large geographical areas (Rogers 2000). Political boundaries are the most obvious geographical divisions and are routinely used to define the spatial extent of risk maps. Alternatively, RS-derived environmental data can be used to develop ecological zone maps (Rogers & Wint 1996) that identify areas of ecological similarity.

In this paper, we use environmental data derived from meteorological satellite sensors and interpolated meteorological data to model the distribution of *S. haematobium* in Tanzania. Further, we show how such ecologically based criteria are better able to define where existing predictive models can and cannot be applied.

Methods

Infection data

Prevalence data originate from school questionnaire surveys conducted in Tanzania (Table 1 and Figure 1). Infection prevalence was estimated from carefully validated questionnaire surveys in which schoolchildren were asked whether they have urinary schistosomiasis or blood in urine (termed locally *kichocho*) (Lengeler *et al.* 1991; Guyatt *et al.* 1999; Partnership for Child Development 1999a). Several studies show that prevalence in schools of self-reported *kichocho* underestimates the parasitological prevalence of infection, but by a consistent amount (Table 1). This means that for each school the prevalence of reported *kichocho* can be reliably calibrated and used to exclude areas of low transmission from control efforts (Red Urine Study Group 1995). Consequently, these data from Tanzania are used to define the extrapolated risk of having infection prevalence $\geqslant$ 50%, WHO's criterion for mass treatment (WHO 1995a).

Remotely sensed and other environmental data

Land surface temperature (LST) and the normalized difference vegetation index (NDVI) information were derived from the Advanced Very High Resolution Radiometer (AVHRR) on-board the national oceanic and atmospheric administration's (NOAA) polar-orbiting meteorological satellites (Cracknell 1997) using standard procedures, reviewed in Hay (2000). Daily data at 8×8 km spatial resolution data were first processed for the period 1985–1998 to exclude unreliable pixels due to extreme sun and sensor viewing angles and cloud contamination (see Hay & Lennon 1999). Single monthly images were then maximum value composited (Holben 1986). Minimum, mean and maximum values of these data were extracted for each pixel that corresponded to the location of the parasitological surveys. Image processing was performed using the Earth Resources Data Analysis System (ERDAS) Imagine 8.4^{TM} (ERDAS Inc. Atlanta, GA, USA).

Interpolated rainfall surfaces were taken from the spatial characterization tool (Corbett & O'Brien 1997) and an interpolated digital elevation model (DEM) of Africa was obtained from the Global Land Information System (GLIS) of the United States Geological Survey (EROS Data Center 1996).

The location of schools was obtained by transcribing co-ordinates from 1 : 25 000 scale maps used in the original survey or collected in the field using a Magellan Global Positioning System (GPS) (Magellan Systems Corporation, San Dimas, CA, USA). Geographical data were displayed and analysed in ArcView (Version 3.0, ESRI, CA, USA, 1996).

Data analysis and model validation

To examine the relationship between environmental variables and the need for mass treatment, schools were classified as having estimated prevalence above or below 50%, WHO's treatment threshold (WHO recommends

Table 1 Summary of data on reported prevalence of schistosomiasis used in the analysis

Location	Number of schools	Number of children	Threshold taken to be equivalent to a prevalence of 50% or more (%)	Schools with prevalence $\geqslant$ 50% or equivalent (%)	Reference
Tanga Region	591	134 924	25	26.1	PCD (1999a)
Kilosa District	164	15 073	35	31.1	Lengeler *et al.* (1991)
Magu District	49	9800	30	81.6	Guyatt *et al.* (1999)
Mtwara Region	176	6302	25	30.8	PCD (pers. comm.)
Total	980	166 099		43.2	

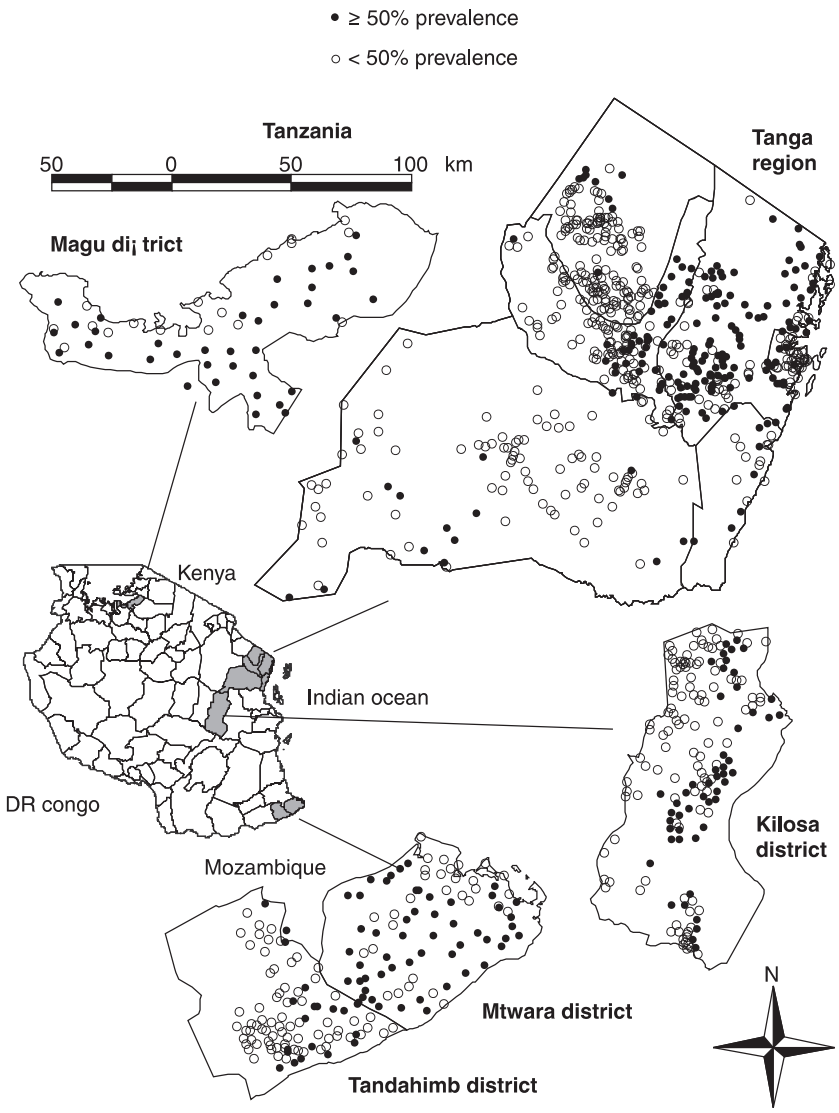


Figure 1 The spatial distribution of reported urinary schistosomiasis in Tanzania. Data are available for 166 099 children from 591 schools in Tanga Region (Partnership for Child Development 1999a), 164 schools in Kilosa district (Lengeler *et al.* 1991; WHO 1995b), 49 schools in Magu district (Guyatt *et al.* 1999) and 176 schools in Mtwara Region (Partnership for Child Development, unpublished data). ● ⩾ 50% prevalence, ○ < 50% prevalence.

that mass treatment is warranted if the prevalence in a school exceeds 50% infection). Logistic regression models were developed to identify significant environmental variables associated with infection patterns. A potential problem in developing models using environmental variables is that many are highly intercorrelated so that it is difficult to separate the effects of the independent variables statistically (Morgenstern 1998). To reduce the dimensionality of these colinear variables, we first selected those variables likely to have greater biological significance on infection transmission (Brooker & Michael 2000). Second, the remaining variables were added to the models in a stepwise fashion, and comparing the statistical fits of alternative models using the residual deviance of models including and excluding correlated variables using a χ^2 distribution (Venables & Ripley 1999). Analysis was performed using S-Plus 4.5 Professional Release 2 (Math Soft, Seattle, WA, USA).

To test the predictive performance of the final model, within the area for which it was developed, the data were divided into two randomly selected sub-samples: one to develop the model (the 'training' data); and the other to assess the accuracy of model predictions (the 'validation' data). Data from the training set in Tanga Region were used to develop a local model of the probability of having an infection prevalence > 50%. The accuracy of the model was then assessed using data from the validation set. However, the real test of accuracy and usefulness of a

model lies in applying it to different locations (Fielding & Bell 1997). Here, we validated the model for Tanga Region using data from elsewhere in Tanzania (Kilosa, Magu, Mtwara and Tandahimba districts).

The predictive accuracy of the developed model was assessed in terms of sensitivity (the percentage of locations with disease/infection present correctly predicted) and specificity (the percentage of locations with disease/infection absent correctly predicted). These measures rely on a single probability cut-off point to classify a school as having infection prevalence 50% or more. A more complete description of the classification accuracy is given by the area under a receiver-operating characteristic (ROC) curve (Fielding & Bell 1997; Greiner *et al.* 2000; Pearce & Ferrier 2000). This curve plots the sensitivity and specificity for an entire range of possible cut-off points. A model with perfect discrimination between occurrence and absence of disease/infection has a ROC curve that passes through the upper left corner (100% sensitivity and 100% specificity). This model will have an area under curve (AUC) of 1.0. As a general rule, an AUC between 0.5 and 0.7 indicates a poor discriminative capacity; 0.7–0.9 indicates a reasonable capacity; and > 0.9 indicates a very good capacity.

Estimates of population at risk and programme costs

Logistic regression equations were then used to map the probability of infection prevalence being 50% or greater using Idrisi Version 2 (The Idrisi Project, Worcester, MA, USA). For the purposes of classification, we have used a probability threshold that maximizes the accurate exclusion of low prevalence schools to develop predictive maps to identify priority areas for a blood in urine questionnaire approach. On this basis, the number of school-aged children at risk of significant schistosomiasis transmission, and the target of a questionnaire approach can be quantified by overlaying the predictive maps of infection prevalence on population data. Data on the school-aged population for every district came from 1990 national population forecasts (Deichmann 1996), and projected to 2001 assuming a country and year specific inter-census growth rates (US Census Bureau. 2001). With the increasing decentralization of health systems in Africa, disease control activities are likely to be undertaken at the district level. Consequently, we have estimated populations at risk for those districts where high prevalence is predicted in 75% of the district's area.

Detailed prospective cost analyses have been conducted for school based anthelmintic programmes in Ghana and Tanzania (Partnership for Child Development 1999b). The cost of delivering praziquantel for schistosomes – which required targeting schools by a questionnaire, and a calculation to determine the dose based on the height of the child – was US$ 0.67 in Ghana and US$ 0.21 in Tanzania. The figures for Tanzania and Ghana were used to provide lower and upper estimates of programme costs.

Results

A number of logistic regression models were fitted to a 50% random sub-sample of schools in Tanga Region. Table 2 presents the final model results and shows that altitude has a negative effect on the probability of a school having prevalence > 50%, whereas minimum LST and mean NDVI both have a positive effect.

The remaining 50% of schools in Tanga Region not selected to develop the model were used to assess the accuracy of the model. Figure 2a shows that the model for Tanga Region allows reasonable discrimination between high and low prevalence schools, within those geographical areas in which surveys were conducted. The plots further indicate that within Tanzania, the model developed for Tanga Region also performs reasonably well in neighbouring Kilosa District and further south in the similar coastal area of Mtwara Region (Figures 2b and c), but performs poorly in comparison in Magu District, near Lake Victoria (Fig. 2d).

These results may usefully be explained by reference to an ecological zone map (Figure 3). This map shows that Tanzania comprises three ecological zones (see Table in Figure 3). Those areas where the Tanga Region model reasonably predicts a high prevalence of schistosomiasis all fall within the same ecological zone (Zone 1). By contrast, the model's performance is poor near Lake Victoria – an area that represents a different ecological zone (Zone 2).

Examination of the ROC curves allows the probability threshold to be identified that optimizes the preferences for maximal sensitivity or specificity. Here, we have used a probability threshold that maximizes the accurate exclusion of low prevalence schools (i.e. $P = 0.2$, see Fig. 2b–d).

Table 2 Regression coefficients describing the logistic regression model for Tanga Region, Tanzania*. LST = land surface temperature; NDVI = normalized difference vegetation index

	B	Residual deviance	*P* (<)
Constant	−543.23	327.6	
Minimum LST	0.052	244.8	0.0003
Mean NDVI	0.188	257.5	0.0005
Altitude	−0.001	269.4	0.0001

* The variables available to the regression analysis was mean, minimum and maximum LST and NDVI, rainfall and altitude.

Ideally, we would want detailed survey data from each of the ecological zones in Tanzania in order to develop a risk model for the whole of Tanzania. However, because it is useful to have a 'national' model for country-level planning purposes and in the absence of detail data, by way of example, we use the Tanga model to develop a *preliminary* risk map for country (Figure 4) – recognizing the limitations it presents, as we help highlight. Comparing this map with a (admittedly outdated) map of the suggested distribution of endemic *S. haematobium* in East Africa (Figure 5) shows a broad correspondence between the suggested distribution of *S. haematobium* and the model's prediction – the slight difference being a broader distribution of transmission south of Lake Victoria.

Thus, using the Tanga Region model, we have estimated populations at risk for those districts in Tanzania where high prevalence is predicted in 75% of the district's area. On this basis, it is estimated that 4.9 million children (in 37 of 97 districts) in Tanzania would be the target for a

Figure 3 Map of ecological zones in Africa based on RS-derived ecological variables. This map is based on the mean annual summaries (1982–2000 inclusive) of multitemporal RS data from the Advanced Very High Resolution Radiometer (AVHRR) processed using standard procedures, reviewed in Hay (2000), to provide, middle infrared brightness temperatures, land surface temperature (LST) and photosynthetic activity estimates [expressed as the Normalized Difference Vegetation Index (NDVI)] for the African continent. These data in combination with a digital elevation model (DEM) of Africa were used to generate 20 'ecological zones' using the unsupervised classification procedures of Earth Resources Data Analysis System (ERDAS) Imagine 8.4™ software. ERDAS implements the Iterative Self-Organizing Data Analysis Technique (ISODATA), an iterative method that uses the Euclidean distance as a similarity measure to cluster data into different classes. A complete synopsis of the environmental criteria used to define and hence separate each zone is not appropriate here but the table shows the mean values of those clusters used to define the zones mentioned in the text.

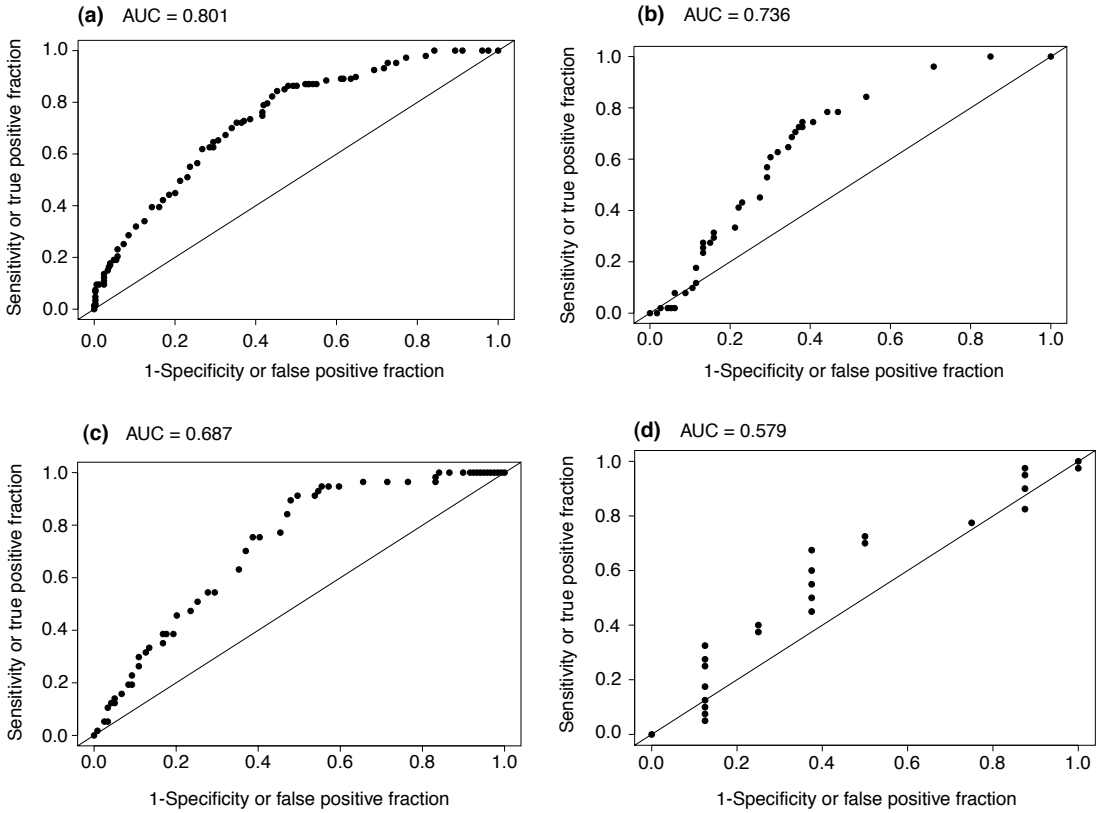
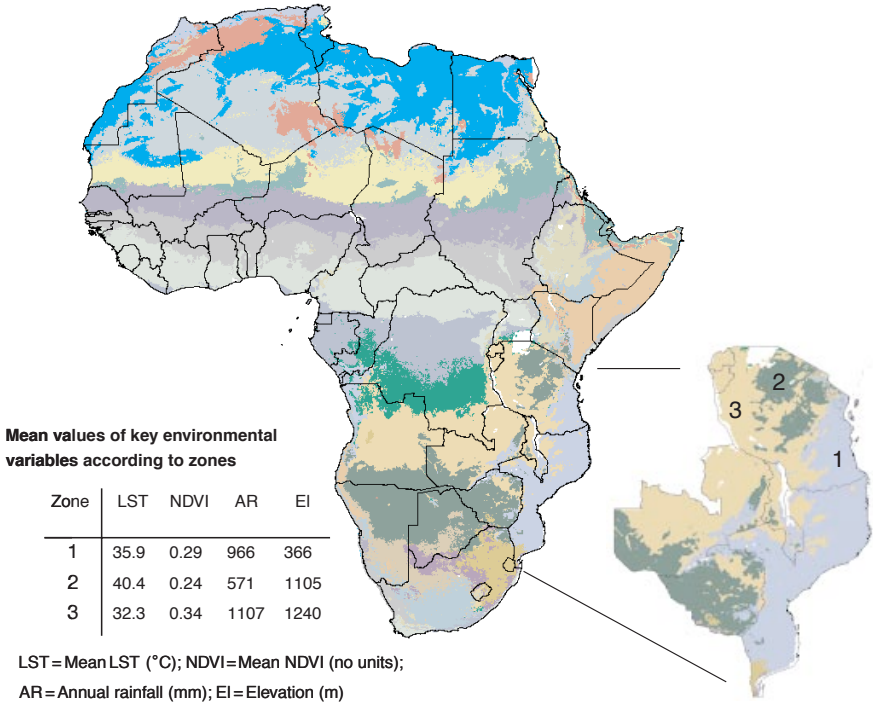


Figure 2 Receiver operator characteristic (ROC) plots for schistosomiasis ecological models. (a) Tanga Region model applied to validation data in Tanga Region ($n = 290$); Tanga Region model applied to (b) Kilosa District ($n = 164$), (c) Mtwara and Tandahimba districts ($n = 176$) and (d) Magu District ($n = 49$).



Mean values of key environmental variables according to zones

Zone	LST	NDVI	AR	EI
1	35.9	0.29	966	366
2	40.4	0.24	571	1105
3	32.3	0.34	1107	1240

LST = Mean LST (°C); NDVI = Mean NDVI (no units);
AR = Annual rainfall (mm); EI = Elevation (m)

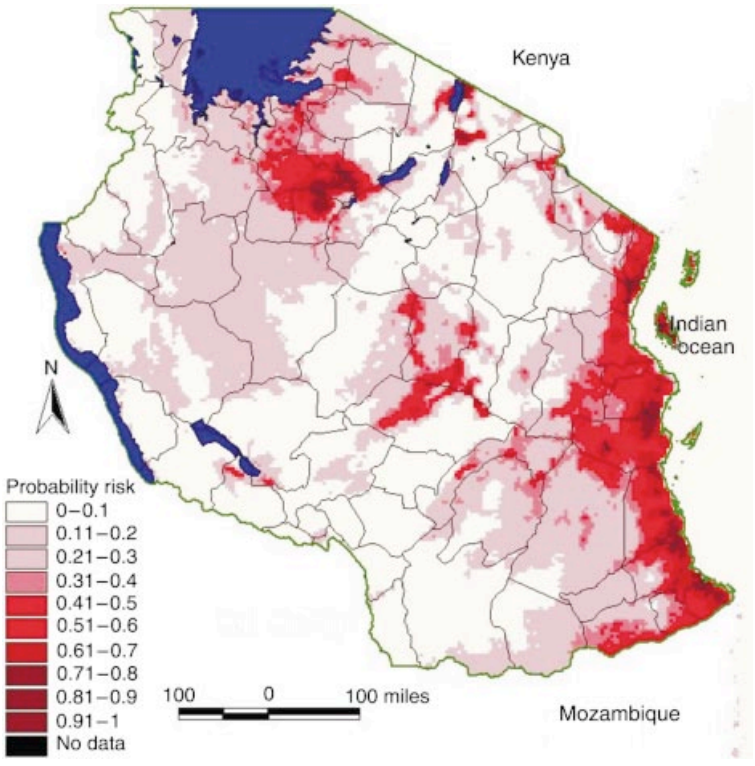


Figure 4 Prediction models for *S. haematobium* transmission in Tanzania. The map shows the probability of an area having an infection prevalence > 50%.

school-based national schistosomiasis control programme. Using the Tanzania and Ghana costs as lower and upper estimates, it is envisaged that the cost of control in Tanzania would be US$ 1–3.2 million. This represents a maximum estimate as the questionnaire survey would further exclude schools from mass drug administration, although the cost of the questionnaire survey, which represented 44% of total costs in Ghana and 19% in Tanzania, would remain.

Discussion

Although unable to capture the well-known small-scale focality of schistosomiasis, low-resolution (5–10 km) RS/GIS models can usefully stratify areas for planning national control activities. In particular, they can help exclude areas where urinary schistosomiasis is unlikely to be a public health problem, and so help focus on priority areas where questionnaire surveys should be undertaken to more precisely target control.

In the interpretation of the current results it is useful to consider the distribution of the snail species involved in local transmission (Sturrock 1993; Brown 1994). In Tanzania, the main snail species are *B. africanus*, *B. globosus* and *B. nasutus* (Webbe & Msangi 1958; McCullough 1972; Zumstein 1983; Marti *et al.* 1985). In coastal areas, *B. globosus* is the principal snail host responsible for transmission, although *B. africanus* may be locally important (Brown 1994). By contrast, *B. nasutus* is the main host in north-western Tanzania and is found in temporary water bodies (Webbe 1962; McCullough 1972; Lwambo *et al.* 1999). *B. nasutus* does occur in coastal East Africa, but appears to be incompatible with the *S. haematobium* parasite strain found in *B. globosus* in that area (Zumstein 1983; Stothard *et al.* 2000). Such differences in the distribution of snail species may explain why the model developed for Tanga Region had a reasonable performance in coastal areas with a common snail species, but had a poor performance in Magu District where a different intermediate host occurs.

The finding that the areas where the model's performance is reasonable are within the same ecological zone may suggest that the model for Tanga Region can be extended elsewhere within the same ecological zone. This zone extends down to Mozambique and parts of southern Malawi (Figure 3). For Mozambique unfortunately, much of the existing information on schistosomiasis is now outdated (Morais 1957), making a detailed validation of the model across the country different. However, crude validations are possible with more recently published data. For example, Traquinho *et al.* (1998) report a parasitological survey among schoolchildren in 12 schools in three districts (Montepuez, Balama and Namuno) in the northern province of Cabo Delgado. The overall prevalence of *S. haematobium* was 84.4%; the lowest prevalence recorded was 77.5%. Applying our model to the three districts, the mean predicted probability for infection prevalence being 50% or greater was 0.21 (minimum 0.07, max, 0.47). Thus, based on our threshold of $P = 0.2$, these three districts would correctly be defined as areas where schistosomiasis is likely to be a public health problem. Further south, in Xai-Xai and Bilene districts in Gaza Region, the prevalence of *S. haematobium* among schoolchildren was 22.1 and 40.2%, respectively (Bobrow & Zacher 1999). The mean predicted probabilities for Xai-Xai and Bilebe districts were 0.07 and 0.11, respectively; thus on the basis of our model they would not be defined as an area where schistosomiasis is a public health problem requiring mass treatment. Although a simplification, the above comparison does support the application of the model developed for coastal Tanzania to other areas within the same ecological zone. More extensive validation of the approach is the subject of ongoing research.

The finding that the areas where the model's performance is poor in different ecological zones suggests that further data and different models would ideally be needed for other areas of Tanzania. In planning future surveys, the use of ecological zone maps derived from RS satellite sensor data can usefully guide sample protocols (Brooker *et al.* 2001). At present however, we are unaware of further detailed data for western Tanzania, and we believe that the current model represents the best we have, despite its limitations, which we highlight here. At present, additional financial resources are being mobilized for the control of disease due to schistosomiasis in Africa. Thus, there is an urgent need for an evidence-based framework in which to support political decision-makers with data on which international priorities for schistosomiasis control can be set.

To conclude, we have demonstrated that it is possible to predict the distribution of schistosomiasis in Tanzania using RS data. The work further suggests where developed prediction models can be applied and where separate or modified development of models needs to be constructed. The development of such an approach in different ecological zones and for separate snail species would therefore appear to offer an important tool for identifying areas for targeted control programmes. Such models can usefully stratify areas for planning national control activities. In particular, they can help exclude areas where schistosomiasis is unlikely to be a public health problem, and so help focus on priority areas where local targeting of treatment using specific procedures should be undertaken. It is within this objective, evidence-based framework that disease

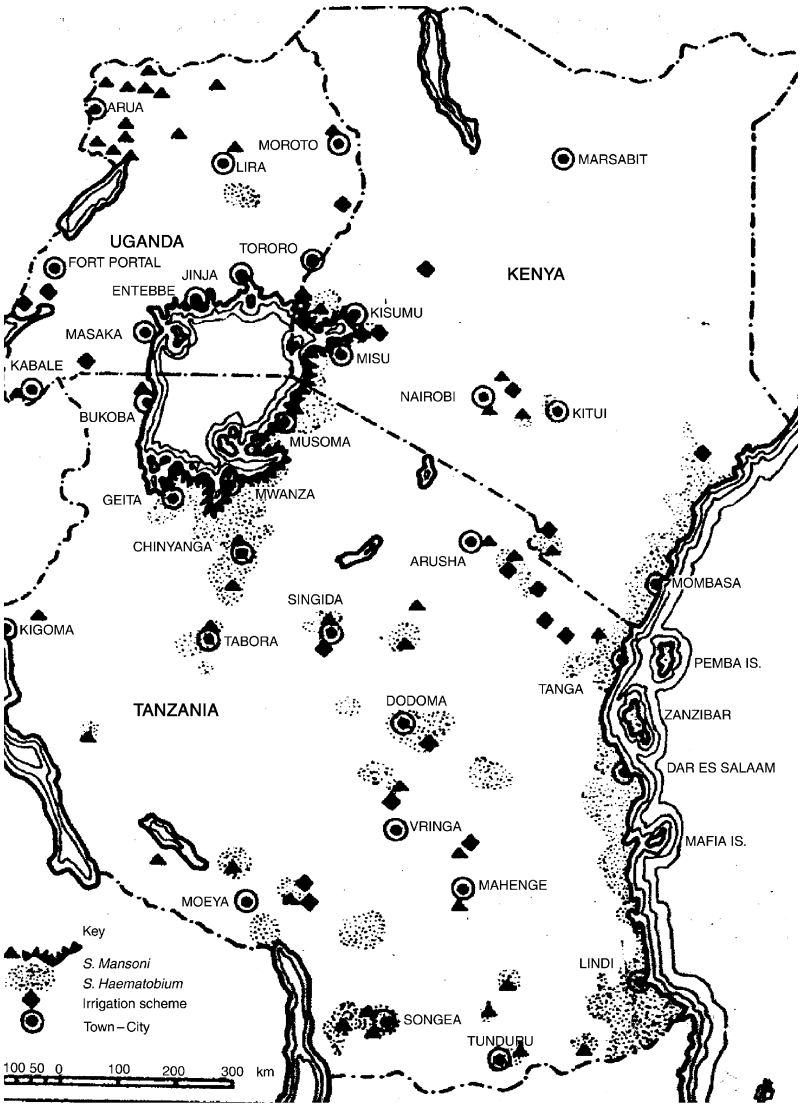


Figure 5 Known distribution of endemic *S. haematobium* in East Africa (Taken from McCullough 1972).

control initiatives should be undertaken. While more research is required, we believe we have provided the first *preliminary* risk map based on RS satellite sensor and meteorological data for *S. haematobium* for sub-Saharan Africa. Further use and validation of the approach is underway.

Acknowledgements

The survey work in Tanzania was funded by the Partnership for Child Development, the Edna McConnell Clark Foundation, and The Wellcome Trust. SB and SIH are supported by a Wellcome Trust Prize Fellowship (#062 692) and Wellcome Trust Advanced Training Fellowship (#056 642), respectively. The following formal acknowledgement is requested of those who use the AVHRR data: 'data used in this study include data produced through funding from the Earth Observing System Pathfinder Program of NASA's Mission to Planet Earth in co-operation with National Oceanic and Atmospheric Administration. The data were provided by the Earth Observing System Data and Information System (EOSDIS), Distributed Active Archive Center (DAAC) at Goddard Space Flight Center which archives, manages and distributes this dataset'. We thank the RIPS project for providing school location data in Mtwara Region.

We thank Joanne Webster, Vaughan Southgate, Andrew Roddam, and members of the Tanzania Partnership for Child Development for their input and helpful comments. We also thank two anonymous reviewers for their comments, which greatly improved the manuscript.

References

Bobrow EA & Zacher AM (1999) *School Health and Micronutrient Initiative. A Baseline Report for Xai-Xai and Bilene Districts in Gaza Province, Mozambique*. Save the Child US/ Mozambique Field Office, May 1999.

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.

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). *Parasitology Today* **16,** 303–307.

Brooker S, Beasley M, Ndinaromtan M *et al.* (2001) A practical application of remote sensing and geographical information systems: targeting a national helminth control programme in Chad. *Bulletin of the World Health Organization* (in press).

Brown DS (1994) *Freshwater Snails of Africa and Their Importance*. Taylor & Francis, London.

Corbett JD & O'Brien RF (1997) *The Spatial Characterization Tool –Africa V1.0*. Texas Agricultural Experiment Station. Texas A&M University, Blackland Research Center, Report no. 97–03, Documentation and CD ROM.

Cracknell AP (1997) *The Advanced Very High Resolution Radiometer*. Taylor & Francis, London.

Deichmann U (1996) *African Population Database Documentation*. National Center for Geographic Information and Analysis, Santa Barbara, USA, URL: http://grid2.cr.usgs.gov/globalpop/ index/html.

EROS Data Center (1996) *GTOPO30 Documentation*. Universal Resourse Locator, Global Land Information System, Sioux. Falls, URL: http://edcwww.cr.usgs.gov/landdaac/gtopo30.

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.

Greiner M, Pfeiffer D & Smith RD (2000) Principles and practical application of the receiver-operating characteristic analysis for diagnostic tests. *Preventive Veterinary Medicine* **45,** 23–41.

Guyatt HL, Brooker S, Lwambo NJS, Siza JE & Bundy DAP (1999) The performance of school-based questionnaires of reported blood in urine in diagnosing *S. haematobium* infection: patterns by age and sex. *Tropical Medicine and International Health* **4,** 751–757.

Hay SI (2000) An overview of remote sensing and geodesy for epidemiology and public health applications. *Advances in Parasitology* **47,** 2–27.

Hay SI & Lennon JJ (1999) Deriving meteorological variables across Africa for the study and control of vector-borne disease: a comparison of remote sensing and spatial interpolation of climate. *Tropical Medicine and International Health* **4,** 58–71.

Hay SI, Omumbo JA, Craig MH & Snow RW (2000) Earth observation, geographic information systems and *Plasmodium falciparum* malaria in sub-Saharan Africa. *Advances in Parasitology* **47,** 174–206.

Hendricks S, de La Rocque S, Reid R & Wint W (2001) Spatial trypanosomosis management: from data-layers to decision making. *Trends in Parasitology* **17,** 35–41.

Holben BN (1986) Characteristics of maximum-value composite images from AVHRR data. *International Journal of Remote Sensing* **7,** 1417–1434.

Lengeler C, Kilima P, Mshinda H, Morona D, Hatz C & Tanner M (1991) Rapid, low-cost, two-step method to screen for urinary schistosomiasis at the district level: the Kilosa experience. *Bulletin of the World Health Organization* **69,** 179–189.

Lindsay SW & Thomas CJ (2000) Mapping and estimating the population at risk from lymphatic filariasis in Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **94,** 37–44.

Lwambo NJS, Siza JE, Brooker S, Bundy DAP & Guyatt H (1999) Patterns of concurrent infection with hookworm and schistosomiasis in school children in Tanzania. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **93,** 497–502.

Malone JB, Abdel Rahman MS, El Bahy MM, Huh OK, Shafik M & Bavia M (1997) Geographic information systems and the distribution of *Schistosoma mansoni* in the Nile Delta. *Parasitology Today* **13,** 112–119.

Malone JB, Huh OK, Fehler DP *et al.* (1994) Temperature data from satellite imagery and the distribution of schistosomiasis in Egypt. *American Journal of Tropical Medicine and Hygiene* **50,** 714–722.

Malone JB, Yilma JM, McCarroll JC, Erko B, Mukaratirwa S & Xinyu Zhou (2001) Satellite climatology and the environmental risk of *Schistosoma mansoni*. Ethiopia and east Africa. *Acta Tropica* **79,** 59–72.

Marti HP, Tanner HP, Degrémont AA & Freyvogel TA (1985) Studies on the ecology of *Bulinus globosus*, the intermediate host of *Schistosoma haematobium*. the Ifakara area, Tanzania. *Acta Tropica* **42,** 171–187.

McCullough FS (1972) The Distribution of *Schistosoma mansoni* and *S. haematobium*. East Africa. *Tropical and Geographical Medicine* **24,** 199–207.

Morais AT (1957) The incidence and geographical distribution of the human bilharziasis in Moçambique. In: *C.R. 3^e Congrès de la P.I.O.S.A.*, Tananarive, section G, pp. 93–96.

Morgenstern H (1998) Ecologic studies. In: *Modern Epidemiology* (eds KJ Rothman & S Greenland), Lippin-Raven, Philadelphia, PA, pp. 469–471.

Partnership for Child Development (1999a) Self-diagnosis as a possible basis for treating urinary schistosomiasis: a study of schoolchildren in a rural area of the United Republic of Tanzania. *Bulletin of the World Health Organization* **77,** 477–483.

Partnership for Child Development (1999b) The cost of large-scale school health programmes which deliver anthelmintics in Ghana and Tanzania. *Acta Tropica* **73,** 183–204.

Pearce J & Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling* **133,** 225–245.

Red Urine Study Group (1995) *Identification of High Risk Communities for Schistosomiasis in Africa: a Multi-Country Study. Social and Economic Research Project Reports, No. 15.* World Health Organisation, Geneva.

Rogers DJ (2000) Satellites, space, time and the African trypanosomiases. *Advances in Parasitology* **47,** 130–165.

Rogers DJ & Wint W (1996) *Towards Identifying Priority Areas for Tsetse Control in East Africa.* Consultants' Report to the Food and Agriculture Organisation, TALA, Oxford.

Stothard JR, Loxton N, Rollinson D *et al.* (2000) The transmission status of *Bulinus* on Zanzibar Island (Unguja) with implications for control of urinary schistosomiasis. *Annals of Tropical Medicine and Parasitology* **94,** 87–94.

Sturrock RF (1993) The intermediate hosts and host-parasite relationships. In: *Human Schistosomiasis* (eds P Jordan, G Webbe & RF Sturrock), pp. 33–85. CAB International, Wallingford.

Traquinho GA, Qunitó LI, Nalá RM, Gama Vaz R & Corachan M (1998) Schistosomiasis in northern Mozambique. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **92,** 279–281.

US Census Bureau (2001) *International Data Base.* http://www.census.gov.

Verables, WN & Ripley BD (1999) *Modern Applied Statistics with S-Plus.* Springer-Verlag, NY, USA.

Webbe G (1962) The transmission of *Schistosoma haematobium* in an area of Lake Province, Tanganyika. *Bulletin of the World Health Organization* **27,** 59–85.

Webbe G & Msangi AS (1958) Observations on three species of *Bulinus* on the east coast of Africa. *Annals of Tropical Medicine and Parasitology* **52,** 302–312.

WHO (1995a) *Health of School Age Children. Treatment of Intestinal Helminths and Schistosomiasis.* WHO/SCHISTO/95.112. WHO/CDS/95.1, WHO, Geneva, 1995.

WHO (1995b) *The Schistosomiasis Manual. Social and Economic Research Project Reports, No. 3.* World Health Organisation, Geneva.

Zumstein A (1983) A study of some factors influencing the epidemiology of urinary schistosomiasis at Ifakara (Kilombero District, Morogoro Region, Tanzania). *Acta Tropica* **40,** 187–204.