The Estrada Real project and endemic diseases: the case of schistosomiasis, geoprocessing and tourism

Omar S Carvalho1,*, Ronald GC Scholte1,4, Ricardo JPS Guimarães1,4, Corina C Freitas2, Sandra C Drummond5, Ronaldo S Amaral6, Luciano V Dutra7, Guilherme Oliveira1, Cristiano L Massara1, Martin J Enk2

1Laboratório de Helminiotologia e Malacologia Médica 2Laboratório de Esquistossomose 3Laboratório de Parasitologia Celular e Molecular, Instituto de Pesquisas René Rachou-Fiocruz, Av. Augusto de Lima 1715, Belo Horizonte, MG, Brasil 4Programa de Pós-Graduação, Santa Casa de Misericórdia, Belo Horizonte, MG, Brasil 5Secretaria de Estado de Saúde de Minas Gerais, Belo Horizonte, MG, Brasil 6Secretaria de Vigilância em Saúde, Ministério da Saúde, Brasília, DF, Brasil 7Instituto Nacional de Pesquisas Espaciais, São José dos Campos, SP, Brasil

Geographical Information System (GIS) is a tool that has recently been applied to better understand spatial disease distributions. Using meteorological, social, sanitation, mollusc distribution data and remote sensing variables, this study aimed to further develop the GIS technology by creating a model for the spatial distribution of schistosomiasis and to apply this model to an area with rural tourism in the Brazilian state of Minas Gerais (MG). The Estrada Real, covering about 1,400 km, is the largest and most important Brazilian tourism project, involving 163 cities in MG with different schistosomiasis prevalence rates. The model with three variables showed a $R^2 = 0.34$, with a standard deviation of risk estimated adequate for public health needs. The main variables selected for modelling were summer vegetation, summer minimal temperature and winter minimal temperature. The results confirmed the importance of Remote Sensing data and the valuable contribution of GIS in identifying priority areas for intervention in tourism regions which are endemic to schistosomiasis.

Key words: Geographic Information System - schistosomiasis - Estrada Real project - control - multiple linear regression

Schistosomiasis mansoni is a major parasitic disease in humans and is endemic in 54 countries in South America, the Caribbean, Africa and the Middle East (Chitsulo et al. 2000, Steinmann et al. 2006, Bruun & Aagaard-Hansen 2008). The trematode Schistosoma mansoni (Sambon 1907), the etiological agent, causes acute, chronic and other clinical forms of disease and/or complications. The treatment is a single oral dose of the safe drug praziquantel (Katz et al. 1989). The control of schistosomiasis includes chemotherapeutic treatment, adequate sanitation, safe drinking water, sewage draining, treatment and health education. In Brazil, the Biomphalaria glabrata, Biomphalaria tenagophila and Biomphalaria straminea, three species of Biomphalaria water snails, are intermediate hosts for the transmission of S. mansoni. B. glabrata is the most important species due to its wide geographic distribution and high susceptibility to infection (Carvalho et al. 2008).

In Brazil, during the last few decades, the distribution of schistosomiasis has been characterised by the spread from rural areas into the periphery of large urban centres. Because of the lack of basic sanitation, sewage disposal in natura into hydric collections and the use of contaminated water contribute to the continuing disease and its geographic expansion (Graeff-Teixeira et al. 1999). As recently described, rural tourism and leisure activities in endemic areas by the middle-class urban residents have been driving the exposure and infection in this population subset (Enk et al. 2004, Massara et al. 2008). Although small outbreaks of acute schistosomiasis have occurred among weekend tourists, this mode of transmission has not been well-publicised. These infections may have a negative impact on the tourism business in endemic areas, so new strategies to prevent infection under these circumstances must be developed.

The Estrada Real project - The Estrada Real (Supplementary data), covering about 1,400 km, is one of the largest and most ambitious Brazilian tourism projects; it involves 178 municipalities in the states of Minas Gerais (MG) (163), Rio de Janeiro (8) and São Paulo (7). In MG, the Estrada Real encompasses municipalities with a strong appeal of rural tourism, but these municipalities have prevalence levels of schistosomiasis mansoni that range from 0.06-28.2%. This is a worrisome situation because tourists come to this region from all over Brazil, including areas without schistosomiasis and from other countries that are free of the disease.

Utilization of the Geographical Information System (GIS) in the study of schistosomiasis - Because schistosomiasis is a disease determined in space and time by certain risk factors, the GIS is a very powerful tool for better understanding disease prevalence and the distri-
bution of multiple risk factors. The use of both GIS and statistical techniques allows for the determination of risk factors and geolocalization of risk areas and may lead to the optimization of resources and choosing the best strategies for disease control (Beck et al. 1997, 2000, Bavia et al. 2001).

The first uses of GIS for schistosomiasis prediction occurred in the Philippines and the Caribbean (Cross et al. 1984). Brown (1994) and Appleton (1978) documented the climate and environmental variables that influenced the distribution of schistosomiasis. GIS was used in a larger scale schistosomiasis study in Asia (Cross et al. 1996, Zhou et al. 2001, Seto et al. 2002, Yang et al. 2005, Zhang et al. 2008) and in several African countries, including Ethiopia (Kristensen et al. 2001, Malone et al. 2001), Egypt (Abdel-Rahman et al. 2001, Malone et al. 1994, 1997), Uganda (Kabatereine et al. 2004), Tanzania (Brooker et al. 2001) and Chad (Beasley et al. 2002, Brooker 2002). In the state of Bahia, Brazil, Bavia et al. (2001) conducted one of the first studies that attempted to correlate disease distribution with environmental variables. Other studies in Brazil using GIS were conducted in Pernambuco (Barbosa et al. 2004) and MG (Freitas et al. 2006, Guimarães et al. 2006, 2008).

This study aims to further develop the GIS technology by creating a model for the spatial distribution of schistosomiasis and to apply this model to the region of the Estrada Real project in the Brazilian state of MG (Supplementary data). To create this model, we used epidemiologic, meteorological, socio-demographic, sanitation and Remote Sensing data, particularly those produced by recently launched orbital sensors such as the Moderate Resolution Imaging Spectroradiometer (MODIS) and the Shuttle Radar Topography Mission (SRTM).

MATERIALS AND METHODS

Schistosomiasis prevalence and Biomphalaria data - Schistosomiasis prevalence data from 35 municipalities of MG (dependent variable) were gathered from the Secretary of Public Health Surveillance and the Secretary of Health of Minas Gerais. The spatial distribution of these data is presented in Fig. 1.

Biomphalaria data - Data on the distribution of Biomphalaria molluscs, important for the presence or absence of schistosomiasis in Brazil, were provided by the Laboratory of Helminthology and Medical Malacology from Instituto de Pesquisas René Rachou-Fiocruz, MG.

Remote sensing products

MODIS - The MODIS instrument is operating on both the Terra and Aqua spacecrafts. The data products derived from MODIS observations describe features of the land, oceans and the atmosphere that can be used for studies of processes and trends on local and global scales (Justice et al. 1998). The variables used in this study comprised the Blue, Red, Near Infrared and the Middle Infrared bands, the Enhanced Vegetation Index and the Normalised Difference Vegetation Index (Freitas et al. 2006, Guimarães et al. 2008).

Spectral Linear Mixing Model (SLMM) - The SLMM was the image processing algorithm that generated the fraction images; the proportion of each component (vegetation, soil and shade) inside the pixel was estimated by minimising the sum of square of the errors (Shimabukuro & Smith 1991).

In this paper, the so-called vegetation, soil and shade fraction images were generated using the MODIS data and the estimated values for the spectral reflectance components were also used as input to the model (Guimarães et al. 2008).

SRTM data - A near-global scale Digital Elevation Model (DEM) was obtained by the SRTM, generating the most complete high-resolution digital topographic database of Earth. SRTM consisted of a specially modified radar system that flew onboard the Space Shuttle Endeavour during an 11-day mission in February of 2000. The ground altitude used in this experiment was given by the value of SRTM DEM for each ground pair of coordinates. The local declivity was derived from SRTM DEM by an appropriate filtering approximation of the first order derivative.

Meteorological variables - The meteorological variables consisted of total precipitation and the minimum and maximum temperatures for summer and winter seasons;
these were obtained from the Brazilian weather forecast centre Centro de Previsão de Tempo e Estudos Climáticos of the Instituto Nacional de Pesquisas Espaciais.

Socio-demographic and sanitation variables - Sanitation (water supply and sewage disposal) and human development index data (for the years of 1991 and 2000) were provided by the National System for Urban Indicators from the Brazilian Minister of the Cities and the Brazilian Institute of Statistic and Geography.

Applied model - In this study, a relationship between schistosomiasis prevalence (dependent variable) and the above-mentioned independent variables from 35 municipalities of the Estrada Real project was established by using multiple regression models.

Mean values for each municipality were used in the models. The relationship among the dependent variable (prevalence, denoted by \( P_v \)) and the 42 independent variables, consisting of social, sanitation, meteorological, remote sensing and mollusc distribution data, were analysed in terms of correlation and multicollinearity. Variables having non-significant correlation with \( P_v \) at the 95% confidence level were initially eliminated. The remaining variables were used as input variables in the multiple regression model and a selection procedure was performed by the \( R^2 \) criterion. This selection technique identified a subset of variables by using a coefficient of determination \( R^2 \) sufficiently close to that when all significant variables were used in the model. In regression analysis, the \( R^2 \) is a statistical measure of how well the regression line approximates the real data points.

Because the prevalence data was given at the municipality level, all input variables were integrated inside the municipality boundaries using GIS systems (ArcGis, ENVI) and exported to a standard spreadsheet for statistical analysis and modelling.

The final statistical model was built by using the data of 35 municipalities from where prevalence information was available. By applying the model to the remaining municipalities, a risk map was built for the entire area of the Estrada Real project.

RESULTS

The final model revealing a \( R^2 \) of 0.34 included three variables: summer vegetation (VEG_S), summer minimal temperature (Tmin_S) and winter minimal temperature (Tmin_W). Tmin_W was negatively correlated with schistosomiasis prevalence, whereas VEG_S and Tmin_S were positively correlated with prevalence. After determining the \( R^2 \), the equation \( P_v = e^{-9.25 + 0.04 (\text{VEG}_S) + 0.81 (\text{Tmin}_S) - 0.4 (\text{Tmin}_W)} \) was built to allow estimation of the prevalence for each municipality; 33 (20.3%) municipalities had a prevalence greater than 10% and 130 (7%) municipalities had a prevalence less than 10% (Fig. 2A). Fig. 2B shows the plot of the residuals resulting from the difference between observed...
and estimated schistosomiasis prevalence (Fig. 2A). In Fig. 2B, dark orange represents overestimated values, light orange underestimated values and yellow the municipalities with an accurate estimative. According to Supplementary data, over 80% of the municipalities of the Estrada Real have untreated sewage and the final destination of the sewage is the river.

**DISCUSSION**

Variables from different sources (social, sanitation, meteorological and remote sensing) were used in this study. The equation of estimated $P_v$ showed that $T_{min\_S}, T_{min\_W}$ (meteorological) and vegetation (remote sensing) were useful variables for the study of schistosomiasis. These same variables were used by Guimarães et al. (2006, 2008) to develop a forecast map for schistosomiasis mansoni in MG.

First, the results of this study led to the determination of significant factors related to disease transmission and reconfirmed the importance of Remote Sensing data. These results, compared to other work, suggest a successful role for Remote Sensing data applications in investigating the spatial epidemiology and control of major human helminths (Hassan et al. 1998a, b, Brooker & Michael 2000, Bavia et al. 2001, Beck et al. 2002, Brooker 2002, Guimarães et al. 2006, 2008). Second and more importantly, our approach facilitated the geographical identification of priority areas for intervention, providing the possibility for more efficient control or surveillance efforts. In this context, the introduced methodology may also serve to direct preventive measures such as environmental education or health awareness programs, especially in regions like the Estrada Real project where tourism is a considerable economic factor.

Therefore, even with a relatively low coefficient of determination and low number of municipalities with information about prevalence, we conclude that the use of GIS, remote sensing and statistical techniques, combined with socio-demographic, epidemiologic and environmental data, together allows for the identification of regional factors associated with the transmission of schistosomiasis and facilitates the delimitation and classification of areas based on their probability for infection with the diseases.

As shown in Supplementary data, in the case of Estrada Real, the environmental and Remote Sensing data were more important to the model construction possibly because more than 80% of the municipalities have no sewage treatment and waste is destined for the river.

Nevertheless, to improve the model’s accuracy for more precise estimates and its application to identify priority areas for intervention on a large scale, it is very important to conduct geo-referenced malacological and prevalence surveys in those municipalities where data are insufficient or not yet available.

**REFERENCES**


Fig. 1: map showing the municipalities of the Estrada Real project in Minas Gerais, Rio de Janeiro and São Paulo, Brazil and sewage destination (Brazilian Ministry of Environment).