Reviewer's report

Title: Geographical information system and predictive risk maps of urinary schistosomiasis in Ogun State, Nigeria

Version: 1 Date: 2 January 2008

Reviewer: Archie Clements

Reviewer's report:

General
Unfortunately this manuscript is not suitable for publication in its current form. There is insufficient detail to assess the quality of the analysis --- the methods section is particularly sparse. I think that, with some assistance from a statistician, a spatial analysis of the data could lead to useful results for disease control and I would encourage the authors to give it another go. If they do so, applicability of the resultant maps needs to be considered in the context of real-world disease control, so that appropriate methods can be used (I see no evidence of this type of thought process in the manuscript). I suggest that the authors read, and reference, some recent papers on spatial analysis in the context of schistosomiasis control, by Raso et al and Clements et al, as the ideas presented in these papers might help to structure their approach. I have many problems with the reporting of this study (which are too numerous to list) and the following is a less than comprehensive description of these problems.

Abstract
Grammar needs to be improved throughout (examples include "data" being incorrectly used in the singular, and overall poor sentence structure). Much of the "background" in the abstract is better placed in the introductory section of the manuscript and should be replaced with statements on the hypothesis and aims, including operational objectives (i.e. why the analysis was done, how the maps would be used in practice). The methods section is inadequate. Needs more detail on the questionnaire (how it was administered, to whom) and the analytical methods (what type of "GIS" analysis, how the logistic regression was performed). The results are also inadequate. Put in descriptive stats on the questionnaire and key model coefficients (plus 95% CI's). The conclusions should relate to the aims of the analysis and should not be vague, banal statements about interventions.

Introduction
The general ideas are there, but (probably due to overlooking subtleties in English grammar) most statements miss the point. For example, the focal nature of schistosomiasis doesn't make the distribution difficult to predict as such, it just means that, in the absence of good surveillance data, other methods (such as statistical prediction) need to be applied to plan resource allocation; and the fact that FAO and other international agencies are making GIS data freely
available is not due to the availability of software, but a general commitment to improving the availability and use of spatial datasets (following increasing recognition that these datasets facilitate the development of powerful tools for disease control). There is nothing in the introduction on the aims of the study.

Methods

More information on the questionnaire is needed. How was it validated? What was the context of the survey? Is there an ongoing control programme in the state? If so, was this survey done as part of the planning phase? What was the size of the sampling frame (how many schools and school children are there in the state)? Did the authors do the image processing? If not, where did they get the RS data from? What was the resolution of the RS data? How was the interpolation done of the rainfall, elevation and soil data (or at least, please provide a reference)? Did the authors do the population projections? Were school locations collected prospectively or were they available from previous sources?

What is meant by “ArcView GIS” was used to perform spatial analysis? What kind of spatial analysis? I only see logistic regression being performed (in SPSS). As far as I can see the models don’t look at transmission, but infection prevalence. More details need to be provided on the logistic regression, including how variables were selected, how model fit was assessed, etc. From the results, it appears that the analysis was done at the location level, after dichotomising the results into presence-absence and above-below a 50% prevalence threshold (using Bernoulli models??). I would suggest using a binomial logistic regression approach where the proportion of infected individuals is used as the outcome variable. I would make all of this explicit in the manuscript. (Note, further reading of the results section suggests that they might actually have done a binomial model, but not understood how to interpret it not all clear to me what they have done).

Results

Why present a choropleth map when actual locations of schools are known? Much of the detail on the logistic models should be moved to the methods section. Did the authors do any model validation? While I commend the authors for using the prediction maps to calculate the burden of infection (note, this output should be included in the aims), I don’t trust the approach they took to deriving the prediction maps and would appreciate some more detail on the population density maps.

Discussion

I don’t think that the fact that risk maps have not been developed for schisto in Nigeria (as opposed to the many other places where such maps have been used) is adequate justification for the study. It is not until these maps (which need to be recreated using robust methods) have been placed in the context of disease control in Nigeria, that the true justification for doing the analysis becomes clear. I would refocus the discussion along these lines. Discussion of the estimate of population at risk is good, but again, I think the results are
meaningless unless I can be convinced of the methods used to derive the risk maps. Conclusions are banal and pointless. Best to rewrite these after careful consideration of the aims and objectives of the study. There needs to be some discussion of the fact that the questionnaire is an indirect measure of infection and has imperfect sensitivity and specificity. There are many papers that have looked at the performance of the questionnaire in different populations, age and gender groups and at different levels of endemicity.

**What next?:** Reject because scientifically unsound

**Level of interest:** An article of insufficient interest to warrant publication in a scientific/medical journal

**Quality of written English:** Not suitable for publication unless extensively edited

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests