Reviewer’s report

Title: Isolation of non-tuberculous mycobacteria (NTM) from pastoral ecosystems of Uganda: Public Health significance

Version: 1 Date: 4 January 2011

Reviewer: Rebecca R Prevots

Reviewer’s report:

The authors have conducted an interesting environmental study of potential sources of exposure to NTM in a sample of households in pastoral areas of Uganda, based on sampling of sources of drinking water as well as soil. They additionally ascertain the degree of exposure to certain water sources and practices related to these sources (eg sharing with wild animals) based on administration of standardized questionnaires. However, because the inferences from this study depend greatly on the selection of the study population and the sampling strategy, additional clarification of the methods is needed.

Major Compulsory Revisions

1. Methods:
   a. Study design and sampling: “…selection of study households from a list of households was done based on a systematic approach…” . Authors should clarify that this is a systematic sampling approach (ie every xth household on the list).
   b. Collection of samples: It is not clear how the sources of water ascertained in the questionnaire correlate with the samples taken for mycobacterial isolation. This should be clarified. Is there any way to describe distance from households, frequency of use, etc.
   c. Seasonality: were the authors interested in ascertaining the seasonality of NTM concentrations? This should be clarified in the introduction; that is, why did they choose to sample across different months, rather than perhaps focusing on just the wet seasons when concentrations would be highest.

2. Data analysis:
   a. The outcome and exposure variables are unclear. Are the authors trying to correlate mycobacterial presence in a given household water supply with the use of that water supply by certain types of animals? The specifics of the data analysis are confusing, and this part needs more detail. In the methods related to logistic regression, the authors state that the outcome of interest is “isolation of mycobacteria in the household environment”, but it appears that the Spiderplot relates to “potential exposures”. Are the authors not relating the actual presence of mycobacteria (yes\no) to the documented exposures ascertained in the questionnaire? This section is critical, and needs clarification. Similarly, the authors state “summary statistics of the explanatory variables with respect to
occurrence of exposure routes to NTM in the pastoral community environment were carried out using the tab commend of Stata”- this part needs more clarification. Explanatory variables are being related to exposure routes rather than prevalence of exposure to NTM? Do the authors mean related to water or soil sources with high prevalence of NTM?

b. The authors state “the validity of the models in explaining usefulness of the explanatory variables”… do the authors mean they are trying to assess the fit of the model? Assessing validity is beyond the application of model fit statistics. Suggest rephrasing.

3. Results:

a. If the main unit of analysis is households, then at least the main results should be presented in terms of the number and proportion of households that had mycobacteria (either proportion that had any, proportion that had a certain concentration, or proportion that a certain type known to be pathogenic). Currently most of the information at the beginning of the results relates to the environmental samples without relating these to the households.

b. The use of the “radar\spiderweb” method should be clarified; further explanation of this method of analysis would be helpful to understanding the results. What does the numeric scale in the middle refer to? Is that the proportion of water sources or of household with any mycobacteria in the sample?

c. Table 4: the title states “multivariable logistic regression analysis showing the household health concern...”. Do the authors mean health concerns or are they relating specific exposures to presence of mycobacteria? This point is not clear.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: No, the manuscript does not need to be seen by a statistician.

Declaration of competing interests:

I have no competing interests