**Response to examiners report**

Project Title: The Spatial Distribution of S. *haematobium* Infection Among School Children in a Rural Sub-District of South Africa: An Application of Geographical Information System, 2007.

**Literary style**
The whole report has been edited and proof read. Grammatical and spellings errors have been corrected.

**Abstract:**
- The examiners suggested that I report other significant factors associated with *S. haematobium* infection in the results section of the abstract. Hence, proximity to water bodies, altitude, access to toilet and portable water and water contact behavior has been included.

- The method used to examine the urine samples has also been included in the methods section of the abstract.

**Chapter one: Introduction**
- Page 1, $2^{nd}$ paragraph, $2^{nd}$ sentence: High rates of morbidity and mortality have been reported in southern Americas, specifically in Brazil. This has been included in the report.

- In page 9 and 12, behavioural has been included in the list of risk factors.
- Page 12, specific objectives, bullet 1, year of testing (2007) included.

**Chapter two: Materials and Methods**
- Page 15, the examiners suggested I include a detail map with the distribution of permanent water bodies. This was to allow comparison with the map showing the spatial distribution of *S. haematobium* infection in the DSA. However, I have not been able to generate a detail map due lack of the geo-reference data and GIS software.

- Page 16, the date of testing and the sample strategy: The study was conducted in 2007 academic year. Screening and testing started in April 2007 and was postpone due the nation wide strike action by teachers. The study resumes again in October 2007 and ended in November 2007. On the sampling strategy, there was no sampling done. There are 33 primary schools in the DSA and all grade five and six school children in all the 33 primary schools were eligible for participation in the study. This has been included in the report (refer to page17, $1^{st}$ paragraph, $1^{st}$ line).

- Page 19, stata transfer corrected to stat transfer.

- Page 20, the response rate description in the method section was deleted since it was described in results section. The data set contains no demographic information on non-responders. It was therefore not possible to compare the baseline characteristics of responders to non-responders to check for possible selection bias. To the best of my knowledge, there was a little possibility for any selection bias since all children were given equal opportunity to participate in the study.

- One Examiner questioned the choice of Principal Component Analysis (PCA) against Multiple Correspondence Analysis (MCA) in constructing the wealth quintals. MCA is suitable for exploring the structure of categorical variables. PCA has been used extensively in previous studies using the same data set from the African Centre. I was comfortable using PCA to construct the wealth quintal than MCA which I have not applied before. However, I don't think the choice of PCA against MCA has in any away compromised the quality of results reported in the study.

- Page 21 and page 22 (section 2.9 and 2.9.1). The sentence has been restructured and repeated sentences deleted.

- I used water collection as an indicator for water contact behavior. Swimming as an indicator of water contact behavior was not appropriate because the study was conducted in both the summer and winter seasons. School children water contact behavior will definitely vary in these two distinct climatic conditions. Therefore including swimming will introduce some bias into the study.

- Page 23, on the cut off point for including a variable into the multivariable model. I included all variable that have been reported in the literature to have association with the outcome variable in the final model. However, a P-value of 0.05 or less was considered a statistically significant result. The techniques used to assess the final model fit have been included. Check page 23nd paragraph last line.

- Page 24, an elaborated description of thiessen polygon is given.


**Chapter three: Results**

- Page 28, table 3.2: number of children who tested positive and the percentage included.

- Page 30: I reported on the isolated high prevalence area found in the Northern part of the DSA.

- Page 31, the 95% CI has been corrected from (0.79, 0.40) to (0.79, 1.40).
- Page 32, the OR that was wrongly reported has been corrected

- Page 33, table 3.4, the reported OR and CI has been corrected from (0.69: 1.01, 2.82) in both table and text to AOR 1.69 95% CI: 1.01, 2.82
- List of variables names in both bivariate and multivariate tables are consistent.

**Chapter four: Discussion**

The wrong report in the discussion that a substantial number of infected individuals have light infections compared to those who had heavy infection has been corrected. See page 40, second paragraph.

On the spatial distribution of infection (fig 3.2) even though I did not integrate the observed spatial prevalence with the various risk factors in the discussion, I offered possible explanation for the isolated peak in prevalence in certain parts of the DSA. See page 44, second paragraph.

**Chapter five: Conclusion**

Page 50, the last paragraph on page 50 which belongs to the last paragraph on page 49 has been deleted.

The examiners recommended the inclusion of the combined approach (both schistosomiasis and geohelminth control) using a single school-based delivery system. This proposal was included as one of the main recommendation to control schistosomiasis in the DSA, see page 48, 1st paragraph.

**Reference:**

The reference has been reformatted and is now uniform, coherent and consistent.