Reviewer's report

Title: The promise of record linkage for assessing the uptake of health services in resource constrained settings: a pilot study from South Africa

Version: Date: 25 March 2014

Reviewer: Katie Harron

Reviewer's report:

This manuscript describes an interesting study that evaluates different linkage methods for the purpose of enhanced surveillance. Generally the manuscript is well written and clear, and the limitations of the gold standard fingerprint data are discussed. I only have a couple of minor essential revisions and discretionary revisions.

Minor essential revisions:

1. page 9: "using the best of the 15 probabilistic linkage scenarios". It is not clear how "best" is defined in this context. Please clarify.

2. It is not clear why adding another household member's first name would improve the matching - can this be explained?

Discretionary revisions:

1. Although sensitivity and PPV are useful for comparing different linkage approaches, it would also be very useful if some actual results were compared. This would give the reader a sense of how results of analysis might differ from those where no linkage error was present (i.e. based on the gold-standard data). The results that you might compare depend on the purpose of the linkage - i.e. you could look at rates of a particular condition, or the association between an outcome and a covariate. It would be particularly interesting to focus on a covariate or group that might be affected by the linkage, e.g. women / former refugees / poorly educated / older respondents. This would enable us to see the real benefit (or not) of investing the extra time and effort in manual review. I think this would add greatly to the interest of the paper and wouldn't require too much additional work to include.

2. It would be helpful to give a sense of the additional time associated with manual review. The numbers of records in this study are fairly small - but manual review can quickly become unfeasible for large numbers of records in population-based data sources. This point could be brought out more in the discussion.

Level of interest: An article whose findings are important to those with closely related research interests
**Quality of written English:** Needs some language corrections before being published

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

**Declaration of competing interests:**

I declare that I have no competing interests