Reviewer's report

**Title:** Healthcare provision for HIV co-infected tuberculosis patients in rural Zambia: an observational cohort study at primary care centers

**Version:** 8 **Date:** 7 September 2013

**Reviewer:** Jennifer Harris

**Reviewer's report:**

**Minor essential revisions:**

Page 6, last 2 lines: You say that patients are matched to ART clinics using 3 data sources (National HIV registers, HIV case records, and the electronic system) but in the limitations you say that enrollment was determined only by matching people in the electronic register. Please make these statements consistent; did you or didn’t you check for enrollment by reviewing paper registers?

Page 7, lines 2-3: You state that you interviewed patients when data was not recorded properly…but the limitations say that you assessed enrollment only by looking at the electronic register. Please clarify who you interviewed- did you contact who did not have an ART enrollment record? Or just people who met specific criteria? Collecting data by interview is very different than relying on secondary data sources- and might have identified more enrollees than you were able to do by matching data in the electronic medical record. Thus it is important to be clear about what data sources you used to assess enrollment.

Page 7, lines 6-7: I’m confused by your statement on data analysis. From Table 2, it looks like your p-values came from logistic regression. Is this true? If so, the p-values did not come from a chi-squared test (though your SPSS output may say something like “Wald Chi-square test” which is commonly used in logistic regression – but is different from a Pearson’s chi-square test and when people say “chi-square test” it implies Pearson’s test). If you initially ran chi-square tests before do logistic regression, these results are not presented. Please revise your methods section so that it matches the analysis you did as shown by your tables. Please consult a biostatistician or textbook to make sure your description is correct. I apologize for not catching this on the first round of reviews – I think I misunderstood what you did.

Table 2: For treatment outcomes, it appears that you made dummy variables for each outcome and included only “died” in your multivariable model- you did not include “default” or “transfer out.” This is considered incorrect statistical technique. With a categorical variable, you need to include all levels in your multivariable model, even if some levels were not statistically significant in univariate analysis. Your multivariable model should be re-run with all levels of the treatment outcomes. Please consult a biostatistician or text on logistic regression for further explanation.
Discretionary revisions

Page 7, line 1: Do you mean SmartCare when you say “district electronic information system for HIV patients”? If yes, it is a national system and I’d recommend calling it by its name (SmartCare) as is done in other publications.

Page 6 & 7: I think your use of quotes around the names of your data sources is not correct; typically quotes are used for proper nouns (official names) of things.

Page 12, line 4: I would change the word “analysis” in the sentence “Our analysis suggested that….” Your ‘analysis’ is quantitative and only tells you that there were gaps in enrollment and ART initiation. All of the things mentioned in the sentence are hypothesized reasons for the gaps based on your observations of activities in Chongwe- you didn’t actually analyze any of these reasons.

Page 18: the limitation about only having data from 2009 and 2010 is confusing because you could have collected data from 2011 & 2012. I’d remove this (I don’t really think it’s a limitation) or say that there’s no data from before 2008 (but that’s not a limitation- people want to know what the recent situation is in Zambia; not what it was before 2008).

Page 19, line 1: I don’t think selection of ART sites by the MOH was done “randomly” – I think it was done “strategically” to provide as much geographic coverage as possible. I would check on this and consider changing it.

Page 19, lines 6-7: Your last clause about the prevalence of newly diagnosed HIV is confusing. I would re-phrase it as “the percentage of HIV-infected patients who were newly diagnosed by the TB clinic.” However, I think a more important comparison is the overall percentage of TB patients who were HIV infected and would consider replacing the current comparison with the % of TB patients who were HIV-infected in each group of clinics.

Table 2: I would put the p-values on the same row as the OR & CI. It will make the table easier to read.

Finally, the English is OK in this paper (and for someone who is not a native speaker, it is truly excellent- I apologize for sounding critical!) but at times it’s a bit confusing and doesn’t sound as professional as it could. If you could ask a native English speaker to edit this, I think it would strengthen the paper overall.

I’m also concerned that the tone may not be what you are aiming for in some parts; one specific example is that your description of the health center staff sounds very negative- yet your acknowledgements say “authors are grateful to the healthcare workers…who provide quality care and great support for TB/HIV co-infected patients.” So I don’t know if you meant to portray health workers so negatively in the paper.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable
Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:

I declare that I have no competing interests