Reviewer’s report

Title: Socio-Economic determinants of HIV disease progression in Kenya.

Version: 3 Date: 21 August 2014

Reviewer: Sara Lodi

Reviewer’s report:

The article consists of a secondary analysis of data collected within a prevention clinical trial. The article explores if socio-economic characteristics are predictive of fast HIV disease progression.

I have the following comments and suggestions to improve the article. I hope you find them helpful.

Major comments

Introduction

1) The introduction opens with the role of treatment in serodiscordant couples, which is not the topic of the manuscript. The introduction should be rewritten putting more focus on i) the role of socio-economic characteristics on HIV disease progression and on ii) how this could be helpful to prioritize vulnerable groups in the HAART scale-up.

2) The objectives should be explicitly stated.

Methods

1) The inclusion criteria are unclear. My understanding is that the included patients are the HIV-positive individuals in serodiscordant couples, who had a CD4 count >350 and did not qualify for HAART initiation at baseline.

2) Baseline is not defined. Is it the time of randomization in the clinical trial, enrolment into the clinic

3) Were individuals initially HIV negative included later on if they seroconverted? Or you just included individuals who were HIV-positive at the beginning of follow-up?

4) The outcome definition is unclear. Was it evaluated at 6 or 24 months after baseline? From the methods it appears the outcome is WHO clinical disease progression according to staging criteria OR CD4<350, whatever occurred earlier (a composite endpoint). However, in results the outcome seems to be CD4<350 only. The argument that only few individuals had disease progression according to WHO staging criteria is not enough to change the definition of the outcome at the analysis stage.

Results

1) The subsection “Cohort selection” and the first sentence of “Population
characteristics” are redundant, because they were already presented in Methods. Please consider excluding them.

2) It is surprising that in 80% of serodiscordant couples the HIV-positive person was the woman. Maybe I misinterpreted the definition of baseline?

3) The paragraph “Disease progression” is unclear and I suggest excluding the sentences from “By decrease… multivariate analysis”. It is not surprising that individuals with lower CD4 count at baseline are more likely to have disease progression during follow-up and less focus should be put on this finding.

4) The sentence “Approximately 49.1% …” is unclear. Does it mean that the remaining 50% had no household income or that this information was not available?

5) Starting from line 12 page 6 it should be made clear that you are now presenting the results from an adjusted/multivariable model.

6) What is the difference between “Daily income expenditure” and “household income”? Please be consistent with the terminology.

7) Line 18 page 6. Why are you now mentioning a survival analysis if this was not mentioned in the Methods? In case a survival analysis approach is used in addition to the logistic regression approach, appropriate definition of censoring and follow-up should be carefully described in methods.

8) I am concerned about Figure 2: the time has been discretized at times 6, 12, 18 and 24 months and the jumps of the curves appear only at censoring times. I suggest excluding this Figure.

Table 1
1) Daily household income available: the categories in the table are not ordered correctly

2) Baseline WHO staging: the total number 330 is wrong (it is 312 according to the heading of the table)

Table 2
1) It should be made clear in the Table heading that these are univariate analyses and the odds ratios with confidence should be reported together with the p-value. Following my previous comment on the definition of the outcome, these analyses should be performed using the composite outcome WHO stage OR CD4 count <350.

2) There is a trend of increased proportions of disease progression with higher level of education. In literature the trend goes in the other direction (lower educational level associated with more HIV disease progression).

3) Sex has a p-value of 0.06 and I think it should be included in the final model too.

Table 3
1) It should be clear that these are results from a multivariable model including the variables that were considered significant in the univariate analysis.
2) What does COR and AOR mean?
3) It is established in the literature that the relationship between and HIV disease progression is not linear. I suggest you refit the model including a transformation of CD4 count or including CD4 count as a categorical variable.
4) CD4 count at entry. Is entry the same as baseline??

Discussion

1) Since the objective of the study is to explore of socio-economic variables on disease progression, I think that the paragraphs Disease progression and Age and Disease progression should be summarized in a few sentences after the discussion of the main results on socio-economic variables.

2) The sentence “These associations cannot be explained by downwards…” is unclear. Please rephrase.

3) The statement “We recommend that roll out for HAART during scale up …” is too strong and is not supported by the data

Kind regards

Sara Lodi, MSc PhD

Level of interest: An article of limited interest

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.

Sara Lodi