Reviewer's report

Title: Comparison of retention in observational cohorts and nested Simulated HIV Vaccine Efficacy Trials in the Key populations in Uganda

Version: 0 Date: 09 Jan 2020

Reviewer: Lucy Chimoyi

Reviewer's report:

Abstract

In the methods section - there is inclusion of the sample size of 572 in this section. Consider removing it

In the result section: The last statement talks about the association of drop-out with the duration of living in the community. What is the RR and its corresponding 95%CI?

Introduction

It is important to provide more detail on the key populations in the literature review. For example it may be good to mention that especially for FSW, there are other occupations they are engaged in apart from sex work that may likely lead to dropping out from a study.

Methods

*  Line 10: maintain consistency when using the subscript. Some acronyms have it included and others like in Line 10 are lacking the subscript

*  Line 35-36: Why was the main purpose of collecting follow-up information every 3 months for 12 months? This has not been made clear.

*  Consistency when referring to table numbers in the text is advisable. In some sentences, the numbers are written as (Table 1) where as in other, the "T" is in small caps e.g. line 42.

*  Line: The definition of non-SiVET1 and 2 cohorts are unknown. It is only after looking at the table that I was able to determine that they refer to FF and FSW respectively. Author could make this distinction in parenthesis. E.g. non-SiVET1 (FF).

Results
Table 1:

- Include HIV positive in the SiVET exclusion criteria
- The last three points under the inclusion criteria for SiVET cohorts are specifically for female participants, it may be advisable to indicate that.
- Under exclusion criteria, include not willing to provide written consent as part of the criteria.

Figure 1:

- What does withdrawn (low risk) mean? Include the definition as a footnote under the figure.
- For consistency, the information under the SIVET cohort should be on the left side of the flow diagram as in the manuscript, the author seems to begin statements with the SiVET cohort.
- The flow diagram caption includes the words "FF" and "FSW" but these two key are not distinctly shown in the flow diagram. Consider revising the figure to include the distributions per each key population or drop the reference to the two words from the caption.

Table 2:

- Why is the distribution of the demographic variable (tribes/languages) important in these cohorts relevant? Is this because of the messaging? What role does this play as it's not evident in the discussion.
- Was the category 'restaurant/bar/hair salon' combined during analysis or at the design stage? Findings associated with this category are significant and some background around this would clarify why this occupation is important in this context. Secondly, do the participants work in these businesses or do they own them? If they own them, would they not fall under small-scale businesses?
- What occupations in this study fell under "Other"? include in the foot note.

Table 5:

- New variables have been added to this table (the last seven variables) and are not in previous tables. Is there an explanation for this?
- The last variable in this table is ambiguous. Clarify what "Away" means.

Discussion
Line 21 and 22: what do the authors mean by not educated, since the cohort participants all have some level of education? Did they rather mean, less educated? Even so, the findings in Table 5 do not support that statement since no significant results were reported while looking at the association between education and drop-out.

Line 59: the inclusion of migration in this section when there is no mention of any information recorded on migration is speculative since migration was not measured. Rather exclude this.

Other comments

From the table 5, it seems that quite a number of factors are "protective" against drop-out. E.g. an increasing age, being a Muslim, female, having lived in a place for >1 year, not having genital ulcers and sometimes having sex while drunk and using condoms. The authors should strive to bring out this "positivity" during their discussion of these findings" and explain from previously reported literature why this is the case or not the case. This has been done only for these factors: Muslim and living in a place for >1 year.

For this occupation (restaurant/bar/hair salon), the likelihood of dropping out is very high according to the findings. What are the implications of including or excluding this group from future HIV efficacy trials? If they are to be included, what interventions need consideration to ensure retention?

Are the methods appropriate and well described?
If not, please specify what is required in your comments to the authors.

Yes

Does the work include the necessary controls?
If not, please specify which controls are required in your comments to the authors.

Unable to assess

Are the conclusions drawn adequately supported by the data shown?
If not, please explain in your comments to the authors.

Yes

Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics
Quality of written English
Please indicate the quality of language in the manuscript:

Acceptable

Declaration of competing interests
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

i declare that i have no competing interests

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal