Reviewer's report

Title: Determinant factors of Age at first sexual initiation among youths in South West, Ethiopia: A cross sectional study

Version: 1 Date: 29 March 2013

Reviewer: Mekitie Wondafrash

Reviewer's report:

Major compulsory revision:

1. In the title: the final analysis of this study focuses on factors influencing age at first sex rather than estimation of age at first sexual experience. The authors need to correct the inconsistency observed in the document about these two issues.

2. In 'abstract' section: should be a summary of what has been done and found in general. However, the following issues needs to be addressed by the authors:

   • In 'methods' section: how was the individual study subjects identified after the stratification is done by health facility? This is not also mentioned elsewhere in the methods section. Who are the study subjects in terms of age range?
   • In 'Result' section: what does "actively participated" mean in this section? Survival analysis should be done to estimate the time to first sexual experience; here it seems that the authors divided the time of initiation as early and late. Another model (binary logistic regression) should be appropriate in this case.
   • In "Conclusion' section: The second sentence doesn't seem to emerge from the findings of this study and should be removed. The text should also be reduced significantly.

3. In the 'Background' section of the manuscript:

   • This section focuses mainly on the magnitude, age and impact of initiation of sex by youth. However, there is no introduction about the factors which might influence sexual debut among youths in Ethiopian context and elsewhere.
   • What is the reference for the median age of marriage in Ethiopia? Reference number 6 does not have Ethiopian data.
   • The overall aim of the study should be clearly stated here as well; which is missing from the background section.

4. "Methods" section appears in the document as a subheading. The authors need to create another subtitle; for instance 'study setting and period' under "Methods" section. In the first part of the methods some unnecessary details should be taken out.

5. In the ' Sampling procedure ' subheading:
How was individual study subjects selected after stratification was done based on the client flow to the study health facilities?

6. In the 'Data collection and quality control' subheading:

• What does "...standard questionnaire" refer to?

• What has been attempted to avoid information bias from a client who comes for VCT since you used the same VCT counselors as data collectors on timing of sexual debut? How was the VCT and data collection organized?

• The last two sentences of this section are not clear to your reader. The author need to be clear on how "data coding, cleaning, and verification" can’t ensure data quality.

• How was khat and alcohol use defined for this study? Who are users and non-users?

7. In the 'Data processing and analysis' subheading:

• It is not clear where the proportion 41% comes from for the sample size calculation. Reference needs to be cited.

• The model used to estimate the time to first sex is confusing to your readers and needs to be clear. Moreover, the time variable needs to be adequately described if survival analysis is to be used.

• The Kaplan-Meier curves are recommended to be done for categorical predictors to have an insight about survival function. Moreover, it is mentioned that log-rank test was used and the objective of using it is not mentioned. The purpose of each one of them needs to be clear for the reader. Both are used for exploration and select covariates for multivariate analysis.

• Criteria for entering covariates to the multivariate Cox Proportional Hazards model should be indicated.

• What are the "dependent and independent variables"? At least the dependent variable (time variable) needs to be adequately described or defined.

• If the aim of this study is to estimate the time to first sexual experience, there is no need to categorize the time variable into <18 and >=18 years. Or else linear or logistic regression analysis can be used to identify predictors of sexual initiation (age at first sex as continuous or categorical variable) as suggested in the abstract section.

8. In the 'Result' section:

• The first paragraph needs re-writing. It is not clear grammatically to your reader.

• The first sentence of the second paragraph is not clearly written as well.

• Whose income was asked on the last line of the second paragraph of this section? How was income of VCT clients measured as majority of the study participants are students (48.6%)?

9. Under 'Distribution of the study participants with respect to sexual and sexual behaviors' subheading:
•How was alcohol and khat use measured or defined? It has to be mentioned in the methods section.

10. Under the subheading 'Differences of the study participants with respect to age at first sexual initiation'
•The cut off for early and late sexual initiation should be defined in the methods section.
•It is not recommended to reporting the result of bi-variate analysis separately; the authors then need to combine it with the following subheading. Kaplan-Meier curves or log-rank test are used to select variables to enter into the multivariate Cox Proportional Hazard regression model.

11. Under the subheading 'Predictors of age at first sexual initiation among youths in South West Ethiopia'
•The first sentence in this section should be taken to the method section.
•What variables were entered into the multivariate analysis and what were the criteria?
•If the age has to be categorized as <18 and >=18 years, logistic regression analysis is rather appropriate than Cox proportional Hazards Model (Table 3). Survival analysis should be used if the objective is to estimate the age at first sex as indicated in the title of the manuscript. This discrepancy needs to be corrected.

12. In the 'Discussion' section:

The discussion section needs further work by the authors for the following reasons:
•No evidence has been presented for the explanation given to age at first sex. The last sentence of the first paragraph looks like a recommendation and needs to be rewritten.
•The last two sentences of the second paragraph should be evidenced as well. There is no reference cited for both statements or no part of the result was used to explain the association between khat or alcohol use and early first sexual experience.
•How was employment defined among youth? This has to be clearly stated in the methods section.
•The last sentence of paragraph three does not have reference even if Ministry of Health is mentioned as a source.
•How was "comprehensive knowledge on HIV/AIDS" measured and compared with other studies in relation to timing of first sexual experience?
•What is the "legal issues" in relation to viewing of pornographic materials in Ethiopia? It is not clear for the reader

Minor essential revision:
1. In the title 'Determinant factors of age at first sexual initiation among youths in southwest, Ethiopia: A cross sectional study', the phrase "Determinant factors" is redundant and needs to be rewritten.

2. In the 'ethical consideration' section: how did parents or legal guardians provide written consent for those less than 18 years old?

3. The authors should address the formatting problems throughout the manuscript such as spaces between words, consistency in placing headings and subheadings, etc.

4. The authors need to ensure proper use of active and passive voices.

5. The authors need to be sure that all variables are defined in the 'Methods' section; for example, comprehensive HIV/AIDS knowledge.

6. The authors generally do not need to present HR or other figures with more than 1 decimal when it is greater than 1 and it has to be consistent throughout the document (e.g. 1.1 instead of 1.123). 2 decimals are used if it is less than one (e.g. 0.56). In the article some have 3 and some have less.

7. In the 'conclusion' section:
   • In the first sentence it is mentioned that the prevalence of early sexual initiation is high but it should be emphasized that it is in comparison with other studies in the country.

8. In the 'reference' section, writing of the authors (individuals or organizations) needs to be consistent in terms of style.

9. In 'Table 3' it is recommended to include percentages of youth <18 and >= 18 years old. Nevertheless, the model used here should be checked for its appropriateness to identify 'predictors' of first sexual experience.

10. Kaplan-Meier curves are mostly used for exploration and to select variables for multivariate analysis. The figures (1-4) should be optional, preferably they should removed from the article.

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

**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