Reviewer's report

Title: Factors associated with HIV infection among children born to mothers on the Prevention of Mother to Child Transmission programme at Chitungwiza Hospital, Zimbabwe, 2008

Version: 3 Date: 3 September 2013

Reviewer: Lars Thore Fadnes

Reviewer's report:

The manuscript has improved some during the last revision. However, there is still a need for substantial changes.

Major Compulsory Revisions:

You wrote that “We included the following variables in the stepwise model: Maternal CD4 count, breastfeeding > 3 months, treated for an STI, full immunization status, nevirapine adherence, and exclusive breast feeding for #6 months, hospitalization, mixed feeding.”

Are the variables/categories “breastfeeding > 3 months”, “exclusive breast feeding for #6 months” and “mixed feeding” independent? By definition, they are not, aren’t they? Thus, the structure of the model violates the assumptions of the regression model used. Please re-structure the model (e.g. one variable which includes the different feeding categories: e.g. breastfeeding pattern: category 1: exclusive replacement feeding, category 2: exclusive breastfeeding up to 3 months of age, 3: predominant breastfeeding up to 3 months of age, 4: mixed feeding during first 3 months of age.

You did not answer the following question in the last review, neither can I see that it has been included in the manuscript: From which recall type is exclusive breastfeeding defined in this study? Is it based on 24-hour recall, 1-week recall, ever recall?”

Could you explain why the unadjusted OR of mixed feeding which is 4.3 (1.1-16) according to table 3, while it is 29 (4.2-207) in the adjuste OR in table 4. Is it due to the duration of feeding? Could it be reasons to modify the selection of the included variables? Table 4 does also not specify what the reference values are. Is mixed feeding compared to exclusive breastfeeding or exclusive replacement feeding?

Limitations: What about the case-control design? (Causality vs. associations?)

There are insufficient descriptions of the regression analysis in the methods. What were the criteria for selection of the variables in the stepwise procedure? Which regression assumptions have been checked?

You wrote that “We have reviewed the suggested literature and included it in this
manuscript”. I can only see the reference of WHO and Horvath (the latter is wrongly spelled in the reference list), and I cannot see the essential references of Kuhn 2008, Lehman 2007, Becquet 2009, Coovadia 2008, Thea 2006 and Kumwenda 2008, which all could contribute to improve the discussion. These could lift the discussion.

In my last comments, I wrote that “The presented adjusted odds ratios (aOR) are given with two decimals. This indicates higher precision of the estimates than what is observed. Thus, please have a maximum of two valid digits in the aORs (1 decimal in the aORs<10 and no decimals in the aORs>10).” When you write that “We have revised the aORs as recommended”, it is surprising to still see this unchanged in the abstract, unchanged in the result section, unchanged in table 3 and unchanged in the third column of table 4. The only change I can see that you have done is in the second column of table 4. If you disagree that a change is needed, I would prefer that you would be make this clear in the cover letter and not pretend it has been done in the manuscript.

Minor Essential Revisions:

It would be nice if the percentages in the result section are presented similarly as in the table (without decimals due to accuracy).

In the result section, it is written e.g. “The median age for the mothers was 28 years (Q1=23.5, Q3=31)”. The parenthesis seems to present inter-quartile ranges (IQR). I suggest rather writing “The median age for the mothers was 28 years (IQR 23.5 - 31)”. Please also write out IQR the first time.

The sentence “Assuming exposure in the children who were HIV negative to be 25% and OR to be 3.16 24 the minimum sample size was 60 cases and 60 controls” is unclear and need to be re-phrased.

It would be nice if your definition of mixed feeding is also included.

There is still insufficient description of the study setting (e.g. who interviewed the mothers, in which language, was it done in the clinical setting etc).

Level of interest: An article of limited interest

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