Reviewer's report

Title: Population and antenatal-based HIV prevalence estimates in a high contracepting female population in rural South Africa

Version: Date: 5 February 2007

Reviewer: Knut Fylkesnes

Reviewer's report:

General
The authors are comparing HIV prevalence amongst women in the general population (consenting to test during the 2005 annual population-based serological survey) with ANC-based. The data are from the rural sub-district of Hlabisa, northern KwaZulu-Natal. The authors intend to "explore differences… with particular emphasis on unrepresentative selection of ANC clinics and unrepresentative testing in the population".

HIV prevalence trend data from African populations are still primarily ANC-based, and validity efforts from a variety of communities are needed (including by variation in use of modern contraceptives) in order to make proper interpretation of such trends. The assumption that the representativeness of ANC-based versus population-based data will differ during the various stages of an epidemic is highly interesting in this regard.

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

The problem of non-response
In terms of method for HIV prevalence estimation, population-based surveys will be the potential gold standard. As the authors conclude, their findings highlight the importance of not assuming population-based estimates equate to a gold standard. In this conclusion they also highlight the major weakness of their own validation effort – the high non-response rate in the population-based survey (57% of women contacted refused to be tested, and a substantial variation by age will give a bias in the overall crude estimate).

We might distinguish between two types of bias due to non-response:

i) differential non-response, e.g. as mentioned above such as variation by age/place/other groups, and by

ii) the possible differences in infection by responders and non-responders.

The first type seems to have been taken into account by standardisation (age, place), but statement “... , age-standardisation slightly reduced differences between the two sources of prevalence estimates, .." is rather surprising. This reviewer would expect it to be more than “slightly”. Regarding the second type, a more focused discussion on the magnitude and direction of the latter type of possible bias is needed (based on own data and literature).

Analytical approach
Studies in the past (sub-Saharan Africa mostly) have showed close match between ANC-based estimates and population-based prevalence estimates when men and women are pooled together. These comparisons have been made mostly in local settings with the sample taken in the population survey to be the same as the coverage of the surveillance site (examples of references are 18-20, 36). In this regard the fact that the authors are age standardizing ANC-based estimates does not make sense at al.

A second level of comparison is the national level (National ANC data versus a nationally representative sample used in the population survey).
Here there are very often great differences between the two estimates – indicating some problems with the selection of sites for the national system, examples are Botswana and Kenya – whereas reference 38 is among those very few examples showing very close match also at this level.

The authors should distinguish clearly between these two levels when analysing/discussion.

Details
Methods
- Standardising ANC-based estimates does not make sense
- Describe the ethical issues (consent etc)

Discussion
Paragraph 4: Seems to be some kind of misunderstanding in this text: The well known pattern is an overestimation by ANC-based data in young age-groups and underestimation in older age-groups (as compared with population-based findings (seen in most local based studies – and also in the national level in Zambia (38). In this paragraph there is a mix between the local versus national level of comparison – and should be avoided.
Paragraph 8 ("Unrepresentative testing in ...."): Here is a need to discuss magnitude and direction more properly.
Paragraph 9: In Zambia analyses of the national population-based ...........": this is what the paper is suggesting – it is actually not providing any suggestion in this regard.
Paragraph 15: First sentence: The findings of this study suggest that ...
This sentence is rather surprising – and to this reviewer not plausible, i.e. that the population-based estimate underestimates – when ANC coverage is high, contraceptive use is high and fertility low. The best conclusion might be that there are great uncertainties regarding the magnitude of non-response bias in the population survey restricting us to conclude.

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Methods
- Describe the sampling methods better (the references given are not readily available)

Discussion
- Write properly the 95% CI
-- Paragraph 3: second sentence: include also for Zambia here (ref 18-20)

Discretionary Revisions (which the author can choose to ignore)

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

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