Author’s response to reviews

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

Authors:
  Brian D Rice (briandrice@yahoo.com)
  Jorg Batzing-Feigenbaum (Baetzing-FeigenbaumJ@rki.de)
  Victoria Hosegood (Vicky.Hosegood@lshtm.ac.uk)
  Frank Tanser (tanserf@africacentre.ac.za)
  Caterina Hill (cats_hill@hotmail.com)
  Till Barnighausen (tbarnighausen@africacentre.ac.za)
  Kobus Herbst (kherbst@africacentre.ac.za)
  Tanya Welz (tanyawelz@btinternet.com)
  Marie-Louise Welz (mnewell@africacentre.ac.za)

Version: 2 Date: 29 March 2007

Author’s response to reviews: see over
Dear Sir / Madam,

Resubmit an original article: Population and antenatal-based HIV prevalence estimates in a high contracepting female population in rural South Africa by Brian D Rice, Jörg Bätzing-Feigenbaum et al

Manuscript reference:

May I thank the editor for the opportunity to resubmit the above manuscript and the reviewers for their positive comments and suggestions for amendment. In view of the reviewers comments we have amended the paper. Below are the reviewers’ specific comments and the authors’ response to these.

**Reviewer 1**

**Major Compulsory Revisions 1**

1.1 The first type [of bias due to non-response being differential non-response] 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”.

In response to the reviewers comments we have re-run the analyses and the figures presented in table 4 are correct. The results show the ANC estimate to decrease from 37.9% to 36.2% (by 4.2%) and the population-based survey to increase from 25.2% to 27.5% (by 9.1%). Despite the two estimates becoming more predictive of one another they remained statistically different. Because the word “slightly” is ambiguous we have removed it from the text[A1] and replaced it with the degree of change (paragraph 2, page 14).
1.2 Regarding the second type, [possible differences in infection by responders and non-responders] a more focused discussion on the magnitude and direction of the latter type of possible bias is needed (based on own data and literature).

Due to worries relating to possible differences between responders and non-responders we currently refer to unrepresentative testing and necessary standardisation in the abstract and on pages 4, 5, 6, 9, 11, 12, 13, 16, 17, 18 and 19. At the end of page 17 we state that due to the limitations and aims of this paper “Further analyses are necessary to explore this possible bias”.

**Major Compulsory Revisions 2**

2.1 In this regard [that sample taken in the population survey to be the same as the coverage of the surveillance site] the fact that the authors are age standardizing ANC-based estimates does not make sense at all.

The reviewer raises an important point, normally there is a close match in the two source populations and as such age-standardisation for the ANC sample is redundant. However, we have reason to believe that this may not be the case for the urban-based clinic (which represents much of the sample) in sentinel surveillance that is located close to the southeast border of the study area. Beyond this border lays a relatively high dense population from which we would expect attendees at the urban-based clinic. We discuss this possible bias in the last paragraphs of pages 18 and 19. With place of residence not being collected by sentinel surveillance we thought it wise to account for any anomaly introduced by cross-border flows through age-standardisation.

In response to the reviewers comments we have extended the sentence on page 19 to now read “Furthermore, the urban-based clinic is located at the southeast extremity of the study area and may well attract individuals from urban areas laying outside of the study area, and it is due to this possibility that ANC-based estimates were standardised by age”. 

Brian you need to look at your previous section on this where you adjusted for status, read papers as suggested by second reviewer and discuss with others.
2.2 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 systems……the authors should distinguish clearly between these two levels when analysing/discussion.

In response to the reviewers comments we have amended paragraphs one to five of the discussion accordingly, We have also included the following sentence at the end of paragraph three of the discussion: “It should be noted that comparing estimates at the national level with other national or regional estimates can be problematic due to bias as a result of site selection”.

Major Compulsory Revisions 3
3.1 Methods - Standardising ANC-based estimates does not make sense.

Please refer to answer provided for 2.1

3.2 Methods - Describe the ethical issues (consent etc).

In response to the reviewers comments we have introduced a new paragraph providing further details of our methodology in relation to ethical issues – paragraph 1 of page 8.

Major Compulsory Revisions 4
4.1 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.
In response to the reviewers comments we have removed the sentence in question.

4.2 Discussion - Paragraph 8 “Unrepresentative testing in ….”: Here is a need to discuss magnitude and direction more properly. 

Please refer to our response above to 1.2.

4.3 Discussion - 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.

In response to the reviewers comments we have removed the sentence in question.

4.4 Discussion - 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.

In table 2 we show that the groups under-represented in testing include women from the urban area (area with second highest prevalence) and women amongst the age-groups with the highest prevalence. Age-standardisation increased the crude overall estimate by 9.1%. We also present some evidence that women with a positive test from the first survey are less likely to re-test than women with a negative test. Therefore, in an area with high PMTCT coverage we would expect women to learn of their status through other means and we would expect women already aware of their HIV status to be less likely to consent to test in the population-based survey – this is a likely explanation for lower rates on consent in high prevalence groups. However, we accept that our conclusion on population-based surveys is more strongly linked to wide ANC coverage rather than contraceptive use or fertility and have amended the paragraph accordingly[A4].
Minor Essential Revisions 1

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

In response to the reviewers comments we have amended the whole section under “Population-based survey” and introduced new paragraphs (paragraph 3 on page 7 and paragraph 1 on page 8), including a definition of resident and non-resident individuals.

Minor Essential Revisions 2

Discussion - write properly the 95% CI

We have amended all confidence intervals accordingly.

Minor Essential Revisions 3

Discussion - Paragraph 3: second sentence: include also for Zambia here (ref 18-20).

The paragraph has been amended accordingly.
Reviewer 2

Major Compulsory Revisions 1

The major issue in this paper is the low rate of HIV testing in the population-based survey (36.6% of the total eligible women and 42.9% of the contacted eligible women). The comparative analysis reported in this paper in the background of a low proportion of women consenting for HIV testing in the population-based survey is a serious limitation of this paper. This limitation has not been adequately addressed in this paper. The UNAIDS Report on the Global AIDS Epidemic 2006 and paper by Garcia-Calleja et al in Sex Transm Inf 2006;82:64-70 would be useful readings.

To be same answer as that provided to 1.2 and 4.2 for reviewer 1

Minor Essential Revisions 1

Methods - Paragraph 2 – What is the definition used for resident and non-resident in the population-based HIV survey?

In response to the reviewers comments we have included a new paragraph (paragraph 3 on page 7).

Minor Essential Revisions 2

Methods - Paragraph 2 – Blood was taken from 1,146 women in ANC and this number is shown as 1,111 in Table 1.

The footnotes of table 3 explain that of the 1146 women a test result was not available for 35 women (1111) and age was not available for a further 18 women (1093). However, we agree that the difference in numbers should be better explained and as such under the sub-heading “Antenatal Sentinel Surveillance” bottom of page 11 we have amended the paragraph to better explain the figures. To facilitate age-specific comparisons, table 4 continues to compare the ANC prevalence estimate among the 1093 women for whom there was both a test result and age (rather than that for the
1111 women) to that for the population based survey. The table is footnoted accordingly.

**Minor Essential Revisions 3**

*Results - Paragraph 1 – The repetition of what is shown in Table 1 could be avoided.*

Table 1 presents columnar %s where of all women within a particular clinic catchment area it presents the proportion living within each of the three residency area types. The text in paragraph 1 of the results presents the reverse of this, namely of all women under each residency area type it presents the proportion breakdown by clinic catchment type. The authors believe presenting results by both ways adds meaning to the results.[A5]

**Minor Essential Revisions 4**

*Results - Paragraph 1 (last statement) and Table 1 – In the clinic accessibility model, accessibility was more comparable between the ANC sentinel surveillance and all women and pregnant women eligible for population based HIV testing than with all women in the population reporting pregnancy ever. Please clarify.*

In response to the reviewers comments we have expanded the explanation on page 11, paragraph 1 and also included a figure that highlights the correlation between predicted clinic attendance among all women and reported clinic attendance amongst women in the population-based survey that had recently reported a pregnancy.

**Minor Essential Revisions 5**

*Results - Paragraph 2, line 5 – 12,351 should read 14,476.*

In response to the reviewers comments we have amended the figure as correctly suggested.
Minor Essential Revisions 6

6.1 Results - Table 2 – Please provide sample size for the age categories of both variables from the total eligible women (14,476).

In response to the reviewers comments we have amended table 2. To ensure the numerical breakdown of the overall eligible population is meaningful we have provided a proportional breakdown of the overall sample for each category. To avoid confusion by including these columnar percentages (all other percentages are by row) we have attempted to ensure column headings are self-explanatory and have included footnotes. However, if the reviewers believe this amendment makes the table difficult to interpret we will remove the proportional breakdown of the overall sample.

6.2 Results - Table 2 - Among the contacted women, please mention if there was a significant difference between those who consented to the test and those who did not. The proportion of women who consented for test was lower in 25-34 years age category, and they had a higher HIV prevalence as compared with the other age groups.

To be same answer as that provided to 1.2 and 4.2 for reviewer 1 and first question from reviewer 2. In the last paragraph of page 17 we state “Women in the 25-29 and 30-34 age-groups presented both the highest HIV prevalence estimates and the lowest proportions agreeing to test. Women resident in the urban area were the overall group least likely to consent to test. Prevalence estimates among groups where the proportion of women contacted consenting to test is particularly low should be interpreted with caution”.

BELOW ARE RESULTS OF SIMPLE ANALYSES RUN EARLIER THIS WEEK

1

Is there a difference between the mean of those consenting to test and those refusing?

`ttest age_to_use, by(refused)`

Two-sample t test with equal variances

```
Group |     Obs        Mean    Std. Err.   Std. Dev.   [95% Conf. Interval]
--------- +这么多年                              
0 |    7058    29.47917    .1212362    10.18529    29.24151    29.71683
1 |    5293    27.92745    .1479297    10.76233    27.63745    28.21745
--------- +这么多年                              
combined |   12351    28.81419    .0941581    10.46427    28.62962    28.99875
--------- +这么多年                              
diff |            1.551721    .1897637                1.179755    1.923688
```

Degrees of freedom: 12349
Ho: mean(0) - mean(1) = diff = 0

Ha: diff < 0          Ha: diff != 0          Ha: diff > 0

```
t =   8.1771              t =   8.1771              t =   8.1771
P < t =   1.0000          P > |t| =   0.0000           P > t =   0.0000
```

2

Is there a difference between the medium (due to skew in population distribution) between the two groups?

`. ranksum age_to_use, by(refused) porder`

Two-sample Wilcoxon rank-sum (Mann-Whitney) test

```
refused |     obs    rank sum    expected
--------+----------------------------------
```
Ho: age_to~e(refused==0) = age_to~e(refused==1)  
\[ z = 9.741 \]
\[ \text{Prob} > |z| = 0.0000 \]

\[ P\{\text{age_to~e(refused==0)} > \text{age_to~e(refused==1)}\} = 0.551 \]
<table>
<thead>
<tr>
<th>row</th>
<th>col</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1,030</td>
<td>5,293</td>
</tr>
<tr>
<td>2</td>
<td>1,277</td>
<td>7,058</td>
</tr>
<tr>
<td>Total</td>
<td>2,307</td>
<td>12,351</td>
</tr>
</tbody>
</table>

Pearson chi2(1) = 2.5446  Pr = 0.111

25-29 year olds - . tabi 501 5293\977 7058, chi

<table>
<thead>
<tr>
<th>row</th>
<th>col</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>501</td>
<td>5,293</td>
</tr>
<tr>
<td>2</td>
<td>977</td>
<td>7,058</td>
</tr>
<tr>
<td>Total</td>
<td>1,478</td>
<td>12,351</td>
</tr>
</tbody>
</table>

Pearson chi2(1) = 43.5104  Pr = 0.000

30-34 year olds - . tabi 470 5293\857 7058, chi

<table>
<thead>
<tr>
<th>row</th>
<th>col</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>470</td>
<td>5,293</td>
</tr>
<tr>
<td>2</td>
<td>857</td>
<td>7,058</td>
</tr>
<tr>
<td>Total</td>
<td>1,327</td>
<td>12,351</td>
</tr>
</tbody>
</table>

Pearson chi2(1) = 27.1797  Pr = 0.000

35+ year olds - . tabi 1634 5293\2398 7058, chi

<table>
<thead>
<tr>
<th>row</th>
<th>col</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1,634</td>
<td>5,293</td>
</tr>
<tr>
<td>2</td>
<td>2,398</td>
<td>7,058</td>
</tr>
<tr>
<td>Total</td>
<td>4,032</td>
<td>12,351</td>
</tr>
</tbody>
</table>
4

Is there a difference in the proportion of women who consent to test compare to those in the overall eligible population are aged .......?

Only for the age-group 35+ was there no difference.