Author's response to reviews

Title: HIV Risk and Associations of HIV Infection among men who have sex with men in Peri-Urban Cape Town, South Africa

Authors:

Stefan Baral (sbaral@jhsph.edu)
Earl Burrell (earl.burrell@ucla.edu)
Andrew Scheibe (andrew.scheibe@hiv-research.org.za)
Ben Brown (ben.brown@hiv-research.org.za)
Chris Beyrer (cbeyrer@jhsph.edu)
Linda-Gail Bekker (Linda-gail.bekker@hiv-research.org.za)

Version: 2 Date: 21 March 2011

Author's response to reviews: see over
We want to thank Dr. Jewkes and Dr. Buchbinder for very thorough and critical analyses of our work. We have worked to address the issues raised. I will embed my responses and label with **.

Major essential revisions
The paper has a number of quite large problems that will need to be fixed before it’s really clear what the main findings from the study are. The abstract will then need to be revised to ensure that these are properly summarised and the statistics that have been found need to be included in the abstract. It may be considered unwise to place so much emphasis on bi-sexual practices.

We have made significant changes to the manuscript including a near complete rewrite of the introduction and discussion. And major revisions to methods and results.

The background to the paper is generally adequate but I have a few very specific suggestions that need to be addressed:
Para. 1: HIV prevalence studies have also been conducted in South Africa among MSM as well

** This has now been included in the text.

Para 2: laws criminalising homosexual practices have not been in place for 16 years and I don’t think this is really “recent” and it can’t really account for why there has been limited research on MSM. You need another explantation.

** This has now been addressed in the text.

Para 3: the N on the paper cited here was 1277 not 46 – there were 46 men with MSM experiences out of the 1277. Further the HIV prevalence was 11.5% among men who had had sex with a man and 2% among all men. This is 5.8x higher. The statistic cited in the paragraph is the adjusted OR for the association with HIV and that is a quite different statistic.

** This paper did not provide a sample size for the unadjusted odds ratio calculation. Given that this paper describes prevalent infections, we would be hesitant about comparing rates in this fashion as the outcome as not so rare for the relative risk to approximate the odds ratio. Thus we used the adjusted odds ratio that is reported in the text of Dr. Jewkes’ paper though we now label this as being adjusted.

Lane’s study was much larger but I would caution against the use of ‘definitive’ with respect to a RDS study, especially one conducted in one area as there will be questions about generalisability.

** ** This has now been addressed in the text.

The prevalence of HIV among MSM in the two larger studies cited is actually lower than that among men in the general population. This deserves a comment
and needs to be returned to with respect to the findings of the current study in the discussion.
** This has now been addressed in the text.

Para 4 : introduces the issues of multiple partners and then discussed MSMW. This may be less relevant if the changes I suggest later are pursued – I wonder if concurrency with men as well as women and condom use, transactional sex and other risk practices wouldn’t be better discussed here as a context for a paper primarily on HIV and HIV risk

**We have removed this section
The paper doesn’t have a clear statement of aims and this should come just before the methods. I would suggest you focus on the HIV prevalence and associated factors as you can cover the points you want to raise within this and the current dual focus on HIV and bi-sexuality isn't working well and would needs considerable further work to be able to retain in the paper. I suggest below that biases in your sample may make this a non-ideal population to comment on bi-sexuality in and so perhaps you should focus on HIV.

** we have addressed this. We do think that making the case that concurrency transcends sexuality and that programs addressing concurrency should consider addressing bisexual concurrency.

The paragraph just before methods presently is very unclear and emphasises the importance of studying MSM in Cape Town. As a reader this is puzzling as the authors have not told us why CT should be any different from elsewhere and the MSM population particularly interesting – either do so or de-emphasise the setting.

** This has now been addressed in the text.

Please avoid the word ‘determinants’ as this is cross-sectional data and you do not know about causality. Also if you are going to mention human rights abuses here you need to work them into the results / tables more systematically and I suggest you mention you will describe their prevalence and association with HIV sero-status (not done at present).

** This has now been addressed in the text.

Methods:
The biggest source of bias in this sort of study relates to the venues and sampling. To try and make sense of the results you need to present more about the venues, how they were distributed between townships, how many were used, how many people were recruited per venue and how the interviewers were
instructed about who to recruit, and how many to recruit. How did you avoid interviewing the same person twice?

** This has now been addressed in the text.

How many recruitment staff were used? Can you tell us a bit about them? What efforts were made to reduce bias in recruitment? The sample generated by this sampling method is a form of cluster sample and it is important to have a term for venue and adjust for this appropriately in your multi-variable analyses.

** we do not believe this to be cluster sampling as we did not have full sampling frame from which we chose venues. Formative mapping was done to generate a sampling frame allowing for probability estimates however the venues were too dynamic to allow for the development of a sampling frame. We address more thoroughly the limitations in generalizability.

It’s important to describe the items in the questionnaire – what was measured ?, how? where did the questions come from, what language was it administered in? When was the study done?

** This has now been addressed in the text.

What are the sensitivity and specificity of the OraQuick tests?

** This has now been addressed in the text.

Please reference backward elimination at p=0.1 as commonly a more conservative p is initially recommended (p=0.2). I don’t understand the reference to the use of two levels of statistical significance as actually it didn’t seem to me that associations at p< or =0.1 were being discussed as significant. However, I don’t think they should be and would like the convention of p< or =0.05 to be used. I know the sample size is small so power is limited but there is a risk of spurious conclusions being drawn, especially given the non-random sample.

** I reviewed with our biostats department who were the ones that recommended p=0.1 for stepwise regression. I have addressed this in the text, but the recommendation was that this is a nice balance between inclusion and development of a parsimonious model. We could find no recommendation between setting the regression cutoff at 0.1 or 0.2 in any statistics textbook or in the original descriptions of stepwise regression.

Please ensure the methods used in all analysis are described and as well as information given on candidate variables considered for multi-variable models. Was written informed consent used – please mention.

Results:
There are quite a few problems with the results and a number of errors with the statistical analysis and I suggest that the presentation be somewhat changed and
all statistics be re-checked.

** This has now been addressed in the text. All statistics were checked.

Given that this is not a random sample I would suggest that the authors reduce their emphasis on prevalence (as its quite unclear how their sample related to the broader population of MSM in CT or even at these venues) and focus on associations.

** We do believe that prevalence is useful though further describing the limitations of our sampling methodology. Associations are highlighted.

The authors are primarily interested in HIV prevalence (and previous testing). I would suggest that you start by describing HIV prevalence and then present your tables giving a total column and then a column presenting the proportion (n & % or median) among HIV+ men and then another for HIV- men. A column presenting the chi squared p value for the association between the variables and HIV status would be informative.

** The new table 2 describes associations of HIV with a similar design as described. We want to maintain a table 1 that describes the sociodemographics of our sample.

When a median is used please give range as well as IQR. Given that the age data is apparently not skewed it’s not clear why a median is being used – a mean and 95% CI is an alternative.

We did not include a 95% CI as we know the true value of what was measured and want to avoid the interpretations that this confidence interval would be generalizable to the whole population of MSM. Please revised table 1 according to this. Because of overlap between categories of disclosure please derive a ‘never disclosed’ variable and include this. Please also include a variable for race, marital status (if asked) and for township of recruitment.

** We have included marital status, but not race as this was not asked. I would like to suggest that table 2 be treated in the same way and that all HIV sexual risk behaviours be summarised in this way (so include variables now in table 4). Since you have data on human rights violations I suggest you include this in table 1 or 2 as well.

** We have included descriptions of the association between human rights violations and HIV though have removed excess tables.

I would like to suggest that you present in a table with the multi-variable model of factors associated with being HIV+. This model must include all possibly causal
or confounding candidate variables and given their strong association with HIV in other studies I suggest you retain age and race as well as sampling venue in the model.

The current tables 4 and 5 have some statistical problems and I suggest you drop them with the new approach I recommend. I have checked some of your bi-variable ORs for table 4 and they are not correct – e.g. for not always wearing condoms as well as age, unemployment and partner numbers. I think the ‘disease’ here should be HIV and ‘exposure’ the risk behaviour – which is different from table 5 where it’s the other way round (very confusing – especially as this isn’t explained in the methods section). If table 5 is kept its necessary to describe the other variables adjusted for in each multivariable model.

** The authors are unsure of how the odds ratios were assessed as we only provided two out of the 4 cells needed for this calculation. We have checked the statistics and can provide STATA output logs

I am very confused about table 5 as the one bi-variable OR I checked was actually correct, but the description of the results shows that the multi-variable modelling approach is wrong and bi-sexuality has been modelled as an outcome. Bi-sexually isn’t “caused” by not always wearing condoms – the analysis is the wrong way round – the association of interest is a model where bisexuality is a dependent variable and condom use the outcome. Ditto for the other sexual behaviours etc.

** We did include analysis assessing the associations of bisexual practices. However, since the odds ratio is equivalent in univariate analysis and approximates a chi-squared analysis, this just assesses associations rather than outcomes. We are not suggesting causality in these analysis and are just referring to associations.

The text of the results:
Start by discussing HIV prevalence. The current table is unclear and its best to handle different by race, age etc in a revised table 1 (as suggested above)
At present many of the paragraphs only describe (re-iterate) what is in the tables. This is not generally accepted in journal articles. Please use the text to provide extra information or to highlight statistics from the tables.
p.10 – when discussing condom use by partner sex please make it clear that you present use among men having that type of partner.
Were respondents about to report multiple human rights violations (and if not why not)– its just not clear from the statistics that they did this. The n’s don’t currently add up as 10+21+16=47 not 49.

** We have clarified this in the text. There is overlap between these different estimates and other violations not mentioned individually.

I am not sure about the sentence about testing and being aware of their status – where does the denominator of 50 come from? Please don’t present % by
towship as your n here =3
** This has now been addressed in the text.

Please follow the advice above about reporting and discussing factors associated with HIV as the current paragraphs are very hard to follow. Its not good practice to just present a p value and no effect size/95%CI and also its not good to introduce new stats here that are not in the table – such as association with rape by township status. The sample size is not large enough to describe and compare between townships with confidence. I prefer to focus the reader on the multi-variable model findings and not the bi-variable as these are potentially wrong due to confounding.

**We are not assessing causality here and thus cannot distinguish between confounders, mediators, and determinants. We have made this clear in the manuscript.

Discussion
The discussion needs to focus on the main findings of the paper – all of them and not just a few!

** This has now been addressed in the text. We have basically rewritten the entire discussion to address the points below.

The prevalence is of interest and is higher than the other cited paper, but the reasons need to be discussed. Age and race of the men are the most likely explanations and the possibility that they explain it must be explored. Please don’t test the other study’s prevalence and your’s for significant differences – this isn’t valid.

On p.13 I find the second paragraph bizarre. The whole HIV epidemic in South Africa is massively affected by race because we still have relatively limited racial social (sexual) integration. This needs to be addressed first and foremost as the explanation. The discussion must not speculate, it must use evidence from your own work and that of others to offer explanations. The circumcision discussion is speculative and should be justified with evidence or removed.

p. 14 what do you think was meant by blackmail? what is the hypothesis linking blackmail causally to HIV status? Please offer some suggestions here (preferably informed by interviews with appropriate men)

** We are not assessing causality and are just reviewing associations.

p.14 – the discussion of rape shows exactly why you should not do extra sub-group analyses. You have relatively few men with HIV and rape was quite rare. When you go to different townships your cell size is minute = hence the huge confidence intervals. Please, the pooled rape and HIV association is about 3x higher. Stick to this.

I am concerned that this sampling method was not a good way of accessing MSMW and that the prevalence may have been much lower than would be found
in the general population. It is particular types of MSM who will frequent venues and not others. I suspect we may learn little about bi-sexuality overall from this sampling approach and caution you against a weighty emphasis on these findings in the paper.

The discussion in para 1 of page 15 seems again speculative, please bolster with data or references.

P.15 – the main limitation of this study is the non-random sample. The range of biases introduced potentially by this need discussion with relation to what is known about the venues used. It’s a population based sample (not RDS) that is needed to tell us about MSM in the general population. Venue based sampling may have some value (apart from convenience) as MSM who frequent venues may actually be accessible for MSM-focused interventions in a way that those who don’t perhaps aren’t. This strength could be discussed too.

The discussion needs to conclude with reflections on what the results mean in public health terms and for HIV prevention interventions.

** We have assessed why we feel that population-based surveillance approaches are not well suited for MSM in Africa.

Reviewer 2

** We want to thank Dr. Buchbinder for taking the time to review this manuscript. Responses are embedded.

This is a well-written manuscript that explores the risk practices and HIV prevalence among a cross-sectional sample of 200 MSM recruited from townships near Cape Town, South Africa. Over the past several years, several reports have documented the high HIV prevalence among MSM in sub-Saharan Africa, and risk factors associated with prevalent infection. This study is unique in exploring these issues among peri-Cape Town townships.

Major comments:
1. The investigators conducted HIV-1/2 testing using the OraQuick rapid antibody test on oral specimens; no results were provided to the participants because no confirmatory test was conducted. This raises some concerns, as this population may have limited access to testing. Were other mechanisms for testing offered to study participants? If you were not confident enough of the results to share them with participants, how confident are you in using them as the primary outcome for many of your analyses?
** The sensitivity and specificity of these tests are excellent. However, at the time of this study these tests were not part of the national testing protocols in South Africa and thus we would have not been allowed to use these as diagnostic assessments. We provided significant counseling for all participants to seek VCT and provided referrals to LGBT-friendly clinics.

2. The introduction and discussion could be reworked to set up the major questions in the introduction (without detailed description of published data), and then integrating previous results in interpreting results from this sample. The authors could also make a stronger case for why this particular sample is important to study, as results from several other samples of MSM in Africa have been reported. Do the authors think these results differ from other published data, and if so, how?

** we have revised the introduction and discussion significantly to address this important concern.

Minor revisions:
1. A number of the references are not appropriately formatted.
2. The sentence at the top of page 11 beginning "This trend was more pronounced..." is unclear and could be reworded.
3. In the second paragraph on page 13, there is a statement that a potential explanation for difference in HIV prevalence between Black and Coloured townships may be knowledge of unprotected anal sex as a risk factor. Were these data presented? If not and they were collected, these data should be included in results. If they weren't collected, they either shouldn't be in the discussion or it should be made clear that this is just speculation. Similarly, there is a discussion of male circumcision -- were these data collected?
4. Several other variables should be added to Table 4 including Black vs. Coloured townships, finding partner on the internet, and type of lubricant.

3. The results could be shortened and tightened by highlighting only the most important results from tables, but not repeating all of the results. It would also read more clearly to have results in the text and tables be presented in the same order.

** The above issues have been addressed with near complete rewrites of the introduction, methods, results, and discussion.