Reviewer’s report

Title: Human papillomavirus genotypes detected in clinician-collected and self-collected specimens from women living in the Mississippi Delta

Version: 1 Date: 10 September 2012

Reviewer: Diane DMH Harper

Reviewer’s report:

1. Is the question posed by the authors well defined?
Yes, the authors are statistically testing whether age and site of collection are associated with different prevalence levels of HPV infection.

2. Are the methods appropriate and well described?
No, there is no power analysis to document whether the prevalence of HPV is sufficiently high from the original study designed for other purposes to answer the stated aim for this paper.

No, the methods of specimen transport, while referenced as appropriate because of another study showing possible use of mouthwash as an inexpensive transport medium, is not appropriate. SurePath medium is being faulted throughout the literature as a medium for HPV testing because the sensitivity and specificity in studies of over 40,000 women show values in the mid to high 90%s, but not 100%. Scope mouthwash has no FDA approval as a medium from which HPV testing can be taken, and has no widespread validation as an accurate method with equivalence to STM as a transport medium for HPV testing. While this is a research study and not a screening used for clinical management, this method of HPV transport medium is critical to the study aim and hence cannot be compromised if the study question is to be answered appropriately.

No, the reclassification of HPV types into phylogenetic trees by mucosotropic favoritism has not been recognized by virology experts. While one of these authors has previously published mucosotropic favoritism by specific HPV types, his supporting article used only 332 women. Perhaps the data from this study should be used to validate in the Mississippi Delta population the mucosotropic predilection of HPV types to garner or deny support for this classification scheme.

No, there is no biological hypothesis to speculate that HPV prevalence would be different between screened and under-screened women, so a primary analysis method of logistic regression including these two categories of screening history in a multivariate model is inappropriate. Proving that the screening history was immaterial is pertinent to your manuscript, but instead of a univariate logistic regression, simple chi square is an effective methodology.

No, you do not state in your methods when you would resort to using a p for trend which you present in your abstract.
No, the primary aim as you have stated should be analyzed with McNemar's statistics, not logistic regression. The kappa score for agreement should be listed with the McNemar results. Multivariate logistic regression in this manuscript is to explore relationships you find with McNemar.

No, IARC defines oncogenic HPV as 15 types, not 13 as listed in the paper.

3. Are the data sound?

There is no presentation of complete or partial concordance of HPV types by collection method. The only concordance presented was for one or more types within the generic categories you created (Table 2 and pulling out HPV 16/18 in Table 3).

Table 2 presents prevalence, yet the text indicates that the prevalences were in some way adjusted. This is unclear.

The supplemental table should be your primary results.

Figure 1 relies on a weak statistical test to make inferences about the ability of age to modify the study results. Remove this weak statistic and just present the graphs.

Line 197 is not supported by the data: age was not a statistically strong influencer of HPV type detected.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?

Yes

5. Are the discussion and conclusions well balanced and adequately supported by the data?

The discussion section lacks inclusion of the non-neoplastic activity of the higher prevalence of oncogenic HPV types in the vagina. The discussion section lacks inclusion of the concept of HPV reservoir as the incubator for anal HPV infections in women without penetrative anal sex. The discussion lacks inclusion of speculation about the high prevalence of vaginal HPV and the women’s sexual partners’ HPV load for penile cancer in this under-screened and high morbidity area of the US.

6. Are limitations of the work clearly stated?

No.

7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?

The discussion lacks reference to many of the prior researchers who have authored and championed work in this area.

8. Do the title and abstract accurately convey what has been found?

The title is accurate. The abstract should indicate that only 122 women had specimens positive for HPV, not 443 originally recruited. The statistics of p for trend should be removed from the abstract.

9. Is the writing acceptable?
Yes, the use of English is acceptable.

• Major Compulsory Revisions
The author must respond to these before a decision on publication can be reached.

Restrict the dataset to those women 30 and over for whom self-collection for oncogenic HPV testing is a clinical possibility.

Conduct a power analysis to determine the number of specimens needed to be collected in order to determine a difference in prevalence of the 15 oncogenic HPV types detected by self-methods vs. clinician methods.

Remove the p for trend statistics.

Clearly describe the self-collection instrument in the methods section including why it was chosen as the detection instrument.

Expand the limitations section to include the unknown performance of the collection medium for clinical testing; the likelihood of an increase in false negatives among women who self-sample for HPV; the fact that the device used for self-collection was not standardized, FDA approved or fully tested for clinical performance; the categories of HPV classification are created by the author and are not recognized by the HPV field as categories of condyloma, neoplastic or vaginal; HPV prevalence does not indicate the lifespan of the infection and is likely not indicative of a persistent infection.

Expand the citations section to include those who have set the ground work for self-sampling options for women who do not or cannot tolerate a speculum exam. For example:


• Minor Essential Revisions

Figure 1 typography is not clear on the x or y-axis labels, nor are the legends clearly presented. Please reconstruct the graphs with clear labels, legends, titles and explanatory text.
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.