Reviewer's report

Title: Evaluation of using composite HPV genotyping assay results to monitor human papillomavirus infection burden through simulation

Version: 3 Date: 29 November 2014

Reviewer: Xianhong Xie

Reviewer's report:

In this manuscript, the author studied the composite HPV test prevalence for different groups of HPV types. Simulations with two prevalence settings and different analytic sensitivities and specificities were done to evaluate the composite test prevalence in relation to the true composite HPV prevalence. The assay specificity stands out as the more important factor to control for compared to the assay sensitivity.

Some comments and questions are:

1. Page 5, lines 97-99. Please discuss when the ratio would be greater than 1. Too many false positives? Also please mention when the ratio would be equal or less than 1. The ratio of estimated-to-true prevalence ratio is a poor measure of accuracy. In the extreme case (especially when the true prevalence is low), those who test positive could be completely different from those who are true positives, and the estimated-to-true ratio could still be equal to 1. The composite test sensitivity, specificity, ROC curve, and AUC are better accuracy measures than the estimated-to-true ratio. Some or all of the latter measures should be presented with the estimated-to-true prevalence ratios.

2. Page 5, line 112, does the author mean that the correlation between each pair of HPV types among the 40 or so HPV types is assumed to be equal to 0.4? It seems unrealistic that all the correlations are equal to a specific value because the HPV types have genetic grouping, some HPV type pairs are more correlated than the others. It would be better that the correlations between the HPV types are estimated from the US data or Canada data, which would make the simulated data more realistic.

3. Page 6, lines 119-123. Please explain how the sensitivity and specificity pair (0.95, 0.95) was chosen in the simulation section instead of in the discussion section. Also, how does the pair compare to the typical setting in the routine HPV testing?

4. Page 6, line 126. Does the author mean in the reduced rows in the Tables 1a-2b, only the prevalence of these 4 HPV types: 6, 11, 16 and 18 were reduced by 50%, while the prevalence of all the other types was unchanged? This seems to contradict with the correlation of the HPV types.

5. Page 7, lines 143-145. Has the author considered increasing the specificity to
more than 0.95 and lower the sensitivity to less than 0.95 and see the effect on the estimation of the composite prevalence? A better way of choosing the sensitivity and specificity pair is to select from the ROC curve.

6. Page 7, line 155-157. It appears that the problem is there are too many false positives. It becomes worse when the “true” prevalence is decreased. The assay sensitivity is too high, but the specificity is not high enough. If one increases the specificity and decreases the sensitivity, the problem of overestimating the composite prevalence could be alleviated.

7. Page 9, lines 203-205. Please explain how the fact that “in the NWT region, the ‘true’ type-specific infection … in the composite measure” causes that “increasing number of HPV types in the composite measures does not always help to ease the overestimating problem in the NWT, Canada” (lines 200-201). It appears that the lower HPV prevalence in NWT, Canada compared to that in US plays a dominant role.

8. Page 15, lines 346-347. Please discuss that the genotype testing on all the HPV types has the same sensitivity $\alpha$ and specificity $\beta$. The HPV types have different true prevalence; for a given test, one usually cannot control both the sensitivity and specificity simultaneously. But by choosing the threshold differently, one can easily achieve higher sensitivity and lower specificity or lower sensitivity and higher specificity.

9. Page 15, line 353. Although the correlation matrix $\Sigma_k$ may have been specified in the reference Lin et al. It would be better to specify the matrix for this manuscript, i.e., giving the actual correlation values.

One minor comment:

10. Page 19, Table 1b, the estimated-to-true ratios are missing for the reduced prevalence row under No. of HPV types = 37.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.