Reviewer’s report

Title: A cohort study of HPV type prevalence according to age.

Version: 1 Date: 2 July 2012

Reviewer: Rebecca Howell-Jones

Reviewer’s report:

The majority of the comments below are major compulsory revisions. In summary, the paper needs cutting down considerably, 95% CIs need to be presented around estimates, further details needs to be given on the population of women included in the study and the discussion and conclusions need to reflect the results of the study. A few minor comments are listed at the end.

Major Compulsory Revisions

Title: This is not a cohort study, but rather a cross-sectional study and this should be reflected in the title. The title does not reflect that the authors have looked at age and cytology grade.

Abstract: See comments below on results and conclusions. Changes suggested for the main body of the report should be made in the abstract too. This is not a cohort study. The results section contains inaccuracies: HPV42 was not the most common type in all age-groups. The main message from the study (i.e. what this study adds and what the results can be used for) is not clear.

Methods:

Study population: Further information needs to be provided on the population of women included in the study, as there is insufficient information to know whether this sample of women is unusual in any way or from which population they were sampled. Were women attending the Alexandra Hospital referred due to previous abnormal cervical screen or symptoms, or is this a screening population? What was the participation rate for inclusion in the study?

Age-groups: The authors state in the methods that they have used three age-groups but it is not clear why, when the title of the study is to look at HPV by age, that such wide age-ranges have been used.

The statistical tests used for comparisons are given in detail in the methods. In the results however, the statistics are given just by p values and it is frequently unclear which comparison and statistic the p-value is referring to. It would be useful for readers to give the test statistic alongside p values in the results, so that readers can see at a glance which test the authors have used, or at the very least make it clearer which comparison or trend the p value refers to.

Results:

This section is too long with too many tables and figures. It would be preferable
to present the findings in a more succinct and informative manner. Tables 3 and 4 could be given as supplementary material, and a summary figure or table given. The graphs of HPV 16,18,6,11 etc would be more informative with 95% CIs and narrower age-bands. Although there might not be sufficient positive samples to look at all types with 5-year age bands, there are sufficient HPV16 positive samples to look at this trend in greater detail.

Confidence intervals around estimates should be presented, in both the text for key figures and on the graphs. As data are presented for the most part in 3 age-groups it is not possible to know how many very young and how many very old participants are included, and therefore how precise the estimates are anticipated to be at these extremes. If this was an opportunistic study of women undergoing cervical screening (see methods comments), then the youngest and oldest women are likely to be unusual.

Much of the text describing the graphs in detail should be deleted. This is difficult to read, unnecessary and frequently exaggerates what is seen in the figures.

Discussion:
The discussion is extremely long and frequently repeats the results.

It is not informative to compare the prevalence identified in this study with other studies without understanding from which population these women were sampled.

3rd paragraph, the study from the University of Thessaloniki should be referenced. It is not possible to compare the incidence reported in that study with the prevalence identified in this study (presume the term incidence used in error, or further explanation is required).

The discussion on HPV prevalence by age needs further work and consideration. This needs to be informed by an understanding of the population of women included in the study. The authors state that HPV prevalence plateaus around age 40 years – in fact the authors have identified a similar prevalence in two 5-year age bands which could be entirely due to chance fluctuation around the point estimate. The suggested explanation for this plateau is conjecture and should be deleted. The sentence on the susceptibility of young women to HPV infection needs rephrasing (contraception-use is not necessarily the same as condom-use). There is evidence that younger women have a higher rate of partner change and more partners than older women, but the authors should not use the phrase “more liberal sexual relations” without clarifying what they mean by this.

4th paragraph: This study does not provide evidence on risk of progression, and it should be clear that this paragraph refers to information from the literature rather than this study. I don’t think it is accurate to say that young women don’t progress to high-grade lesion due to a better immune response, Maucort-Boulch D et al, Int J Cancer 2010 have shown that older women clear incident infections as quickly as younger women. In populations where new infections are less common in older women, a greater proportion of infections identified in these women are likely to be persistent infections. Reference 13 is not the correct
reference for the statement on immunological response.

5th paragraph: The data presented identifies HPV 51 as the 2nd most common type in all age-groups. The authors however state that HPV 51 significantly drops below HPV 16 in the older age-groups.

7th paragraph: It is incorrect to state that a new vaccine is required due to the presence of types other than 16/18 in low-grade lesions – the significance and impact of these types lies in their contribution to cervical cancer, not to low grade lesions.

8th paragraph: This paragraph should be deleted.

9th paragraph: This study is not representative of the Greek population. The last sentence of the discussion should be deleted or clearly revised.

Conclusion:
HPV42 is found in one more infection than HPV 16 – it is therefore misleading to state as the main conclusion that HPV42 was the most common type. It remains unclear to me exactly how these results differ greatly from northern European or US data.

Tables 2 and 3: It needs to be made clear whether the %s and n’s represent the % and n of infections or samples.

Minor/Discretionary comments:
There are several typos and errors in the text:
Conclusion heading mis-spelt.
There should be consistency in the use of commas or decimal points in numbers.

Discretionary
The data from Tables 1 could be incorporated into Table 2.
Unnecessary use of 2 decimal places when reporting the mean age of participants.

**Level of interest:** An article of limited interest

**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