Reviewer’s report

Title: Socioeconomic factors and other sources of variation in the prevalence of genital chlamydia infections: A systematic review and meta-analysis

Version: 2 Date: 23 April 2015

Reviewer: Ingrid Viola Francine van den Broek

Reviewer’s report:

The topic and research questions of this study are highly relevant and interesting. The study is well-performed and, as far as I can judge, uses proper methods for systematic review, statistical measures and meta analysis.

Major compulsory revisions (2)

The readability of the manuscript should however be improved. Especially in the Results section the reader looses the line of the story due to listing of all different comparisons, including the reporting of ORs and the level of heterogeneity as I², as such, while the interpretation of the numbers seems to be left to the reader. Is it possible to show the results of the actual (pooled) prevalence in the results or in a table, so that the reader can see what the actual difference is in reported prevalence. Now, only the range of estimates in men and women (p10, line 230) is given, but not per study or subgroup.

- The heterogeneity is moderate or high for almost all comparisons, and the number of studies relatively low, which results in a high uncertainty of the associations found. It is not clear to me and could be explained better in the discussion why the association between socio-economic status and chlamydia prevalence is quite strong, whereas the association with e.g. gender and age is weaker, due to a higher heterogeneity between studies. I would expect the latter to be more obvious than the first. Does this depend on the number of studies or type of studies included in either comparison? Can there be a bias due to e.g. participation (lower participation in low ses-classes, bias towards higher-risk participants?). Or is the effect of ses-class so strong that it is consistent among studies?

Minor revisions:

- Background line 84-85: risk of infection is higher among men aged 19-24 than in younger men. Is this only true for men younger that 19 or also for men older than 24? Or are data only available of <25 years old?

- The risk of bias was assessed, scoring Low, medium, high. It is not explained on the basis of which criteria this classification is done. Furthermore, the reporting of this is limited to the additional files, but it is not clear whether this was taken into account and how in the analysis (also not when reading results paragraph on this, line 181..)

- Under statistical analysis it is said (line 146..) Heterogeneity in prevalence
estimates was investigated by stratifying meta analyses by gender and type of socio-economic positions. Further down on the same page we read (line 159...) that it was not possible to include socio economic position in the meta regression analyses due to missing or incompatible dummy variable data. This seems contradictory, and it also confuses to what is reported on socio economic status in the results (not meta analyses?).

- the term ‘dummy’ can be omitted in line 152, 161, in my opinion these are not dummy but additional variables at study level.

- the timing of data collection is analysed as a binary variable: before or after 2006 (where is 2006?); why not use year as a continuous variable? The choice for 2006 is based on the implementation of screening policy in England, which is not really relevant to the majority of studies in the systematic review who took place outside England, UK, or even Europe.

- The number of studies reported in the results (eg 10 studies, line 181), is not the same as what is shown in the Figure; in the Table it would be more clear to indicate which results are from the same study (both genders) and put them in adjacent lines (or is there a specific order of studies in Figure 2 that I have missed?)

- The results (line 193) give the combined OR of all studies for men and women together, and a bit further ORs are given for men, women and both sexes combined. I assume the latter is then based on studies where results of both sexes are combined, and not a pooled OR from all studies together?

- The measure 'adjusted risk difference' is not clear to me. In the discussion it is mentioned as a 1.1% difference (RD 0.011) between men and women. Is this the difference between the actual prevalence (eg 2.0% and 3.1%) or a relative measure (eg 2.0% and 2.022%). Here it would be good to see the actual pooled estimates of the prevalence for men and women to clarify this.

- The discussion is too long, it could be shortened especially in the section Strength and limitations. It appears more a methodological discussion with only in the end (from line 337) a discussion of actual findings/explanations for the effect of socio economic status. As a reader of BMC Public health I would like to see this more in the forefront of the discussion.

- the last line on population-based serological studies: I am not sure that is the answer, because serological studies still have quite large limitations and uncertainties as well. It is good as food for thouht, but I would recommend to at least mention this has its limitations as well.

Level of interest: An article of importance in its field

Quality of written English: Acceptable

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
I have no competing interests to declare