Reviewer's report

Title: Chlamydia trachomatis infection rates among a cohort of mobile soldiers stationed at Fort Bragg, North Carolina, 2005-2010.

Version: 2 Date: 9 September 2013

Reviewer: Sharon Kuhlmann

Reviewer's report:

Comments are given by page number, paragraph (par xx) and line within paragraph (line xx); line number is given as +xx meaning number of lines from first line of the paragraph, or as –xx, meaning lines from last line of paragraph.

Major Compulsory Revisions

1. Page 5, par 2, line 6: Where only one recurrent infection per person considered, i.e. the first one that occurred at least 30 days after the incident infection?

2. Page 7, par 1, line -4: If test of cure was being conducted, why was it necessary to remove Ct infections within 30 days of another infection? Does this mean that that data was not used?

3. Page 7, par 2, line -5: How were non-white race-ethnicity categories combined and why? It is not clear in the text nor in Table 1. Also age seems to have been combined in Table 1/Men but it is not mentioned or motivated in the Statistical analysis section.

4. Page 7, par 2, line -5: It is not clear how recurrent infections were modeled, specifically what the outcome was since one person might have had several recurrent infections. Was only the first recurrent infection included? If so, how does this imply limitations to the study? Alternatively, were all infections included? If so, how was time calculated for the second, third, ... infections? And how was the repeated measurements aspect taken into account?

5. Page 7, par 2, line -3: The term “censoring events” is currently confusing. The event being studied, here Ct infection, is usually called event in survival analysis, while the conditions for being excluded from the cohort are called censoring (as in “censoring was defined as ....”).

6. Page 8, par 1, line -2: What type of test was used to compare? This should be mentioned in the Statistical analysis section.

7. Page 8, par 2, line 2: It is not clear in this paragraph if only incident infections are included or all Ct infections. For example the sentence “The final number of infections excluded 72 that occurred within 30 days of a prior Ct diagnosis.” seems to indicate that incident and recurrent infections are included in the total, similarly the following sentence “The number of individuals found to have had at least one Ct infection during the study period was 2,198, or 3.3% of the cohort.”

8. Page 8, par 2, line 4: If the paragraph only refers to incident infections, and
2493 persons reported an incident infection with 72 lost to follow up, then 2493-72=2421 persons should be included in the study. However, the next sentence states “The number of individuals found to have had at least one Ct infection during the study period was 2,198, or 3.3% of the cohort.” It is not clear where does 2198 come from.

9. Page 8, par 2, line 7: “50% received their Ct diagnosis within 6 months of return to Fort Bragg (…)”: do these numbers in the sentence refer only to incident infections or any kind of Ct infection?

10. Page 8, par 2, line -4: Does the crude incidence rate refer to incident infections only?

11. Page 8, par 2, line -3: Is the number of person-years of follow up mentioned when comparing incidence rates in women and men the same as the denominator used when calculating the crude incidence-rate in the previous sentence?

12. Page 9, par 1, line 6: What type of test was used to compare? This should be mentioned in the Statistical analysis section.

13. Page 10, par 1, line 4: there is not mention to the incidence rates previously calculated for active duty personal nor a reference is given either.

14. Where the assumptions of proportional hazard checked?

15. It is not clear if covariates were updated (e.g. age, marital status, deployment) when analyzing recurrent infections.

Minor Essential Revisions
1. Page 4, par 2, line 6: “Ct” abbreviation has not been defined previously in the article.

2. Page 5, par 1, line 2: The size of the active duty population should be added to provide context.

3. Page 7, par 1, line 1: What type of screening is carried out at Fort Brags? Is it target to any special group?

4. Page 8, par 1, line 4: when reporting mean age, it would be preferable to report median age as usually age has a skewed distribution (which is apparent here when comparing range and mean).

5. Page 8, par 1, line 6: Also when reporting follow-up time, I would like to see the median time.

6. The software used for statistical analysis is not mentioned.

Discretionary Revisions
1. Page 5, par 2, line 3: For clarity of presentation, I suggest that the text starting “Ct infection were ascertained … reliable protocol” should be included under the heading of “Case definition” or “Outcome” as it is not “Study location and population”.

2. Page 6, par 2, line 2: The sentence on the protocol review should be moved to the end as it breaks the flow as it is now.
3. Data sources section could be shortened by removing information that is not relevant for the study, such as serological testing of other diseases.

4. Table 1 and 2: I suggest breaking up the table in two: the first should include Total and Unadjusted estimates for women and men, and a second table with include adjusted results for women and men.

5. Page 7, par 3, line +1: The first sentence of the paragraph belongs to the previous paragraph as it refers to censoring definition.

6. Page 5, par 4, line -2: The sentence "The population at ..." describes more clearly the objective of the study than what is currently stated in the background section. You might consider moving this sentence earlier for the benefit of the reader.

**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