Reviewer's report

Title: Challenges of Self-Reported Medical Conditions and Electronic Medical Records among Members of a Large Military Cohort

Version: 1 Date: 30 January 2008

Reviewer: Liset van Dijk

Reviewer's report:

General
This paper covers an interesting methodological problem: the comparison of self-reported data with electronic medical records for a wide range of diseases. However, I have the idea that the authors do not fully use the fact that they have data on such a wide range of diseases. Moreover, I think that the question that was posed in the questionnaire limits the possibilities to compare the self-reported data with the electronic records. I will explain this in more detail below.

Major Compulsory Revisions

1. The way the question on self-reported medical conditions is posed (â## has your doctor or other health professional EVER told you that you have any of the following conditionsâ## ) makes it very hard to compare data of all diseases with the electronic medical records. For example, what if I had sinusitis when I was a child. I will report I have had a sinusitis. However, it will not show up in the medical records. I think the comparison between the self-reported conditions and the electronic medical records can only be made for chronic diseases. For acute complaints it is hard, even for those who are in the military service for over 16 years (so stratifying does not solve all problems, especially not for acute problems). I would suggest to leave out all acute complaints and focus on the chronic complaints. At least the authors should pay much more attention to this issue in the discussion.

2. The authors discuss the results methodologically. The discussion would gain strength if they would also look at the results for the different diseases and reason why the results for certain diseases were found and why this may differ for different diseases. The strength of their study is that they are able to do that. Asthma for example: it is possible that a person has been diagnosed asthma as a child, but that s/he does not have asthma as an adult. For â## kidney failure requiring dialysisâ## one can imagine that a person really know when s/he has that. That would imply that there is underreporting in the electronic medical data. Methodological problems could as such be illustrated with the results for certain diseases. For example for some diseased the translation from self-reported diagnosed to ICD-9 code is easier than for others: did that affect the results? How? Could that be illustrated with figures for these diseases?
Minor Compulsory Revisions

1. For a non-American non-military reader some parts of the paper are hard to read. Some information could be left out of the paper, for example the deployment experience to Southwest Asia, Bosnia, since it is not used in the paper.

2. How complete are the medical record data from the three sources used? And is all care that military people use included in the military databases?

3. As I read the paper, the survey was held before the idea to perform this study (to link self-reported data with medical records) evolved. The paper would benefit from mentioning this explicitly.

4. The methods section is long and contains much information on what was not done. I understand why the authors chose to do that (to explain their choices). However, it would be more clear if only the method chosen would be described. The authors could put the other information in an appendix, for those who are interested to read it.

5. Table 3 is much easier to read than table 2. Why not including all diseases in table 3? Also because the high values for negative agreement are mainly caused by the low prevalence of the diseased included in the study.

6. The discussion section is very long. I would focus more on the main problem (the lack of concordance between the time interval asked for in the questionnaire and the time interval for which medical records are available). And as stated before, the discussion would gain from a better connection between methodological arguments and examples for specific diseases.

Discretionary Revisions

1. In the methods section more attention could be paid to the question how the representative the included cohort is compared to the first panel. That is mentioned in the Discussion section, but it would be better in the Methods section.

**What next?:** Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

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

**Quality of written English:** Needs some language corrections before being published

**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