Reviewer's report

**Title:** Developing an algorithm to identify people with Chronic Obstructive Pulmonary Disease (COPD) using administrative data

**Version:** 1  **Date:** 9 February 2012

**Reviewer:** Ann-Britt Wiréhn

**Reviewer's report:**

I believe that the three-step method used in this study is a reasonable method to evaluate the aim. My opinion is, however, that the presentation of data and the result has to be improved.

**Major compulsory revisions:**

#1 As a guidance for the reader to understand the calculations in this study, which design I would say is similar to a serial testing with its different steps, I think it is necessary to present all data and not only the calculated measures, i.e. sensitivity, specificity, PPV and NPV. I suggest that you create a comprehensive table by age groups (and perhaps gender) over the data capture and physicians’ verifications in the different steps. This will clear things out.

#2 My opinion is that the current conclusion “…the algorithm should mostly be regarded as a tool for identifying chronic lung disease…” requires a presentation of data capture on chronic lung disease or perhaps asthma particularly.

#3 Since the aim in this study is to identify as much patients as possible, the most interesting measure ought to be the sensitivity. I believe that this is also the author’s opinion and the reason why the conclusion was formulated as above (italics in #2). The overall sensitivity was 29.7, which is not mentioned in the abstract. I would therefore recommend more focusing on sensitivity in the result part as well as in the abstract. The specificity, PPV and NPV is interesting but not as important as the sensitivity.

**Level of interest:** An article of importance in its field

**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.