Reviewer’s report

Title: Discriminative and predictive properties of disease-specific and generic health status indexes in elderly COPD patients

Version: 1 Date: 18 March 2008

Reviewer: Gordon H Guyatt

Reviewer’s report:

- Discretionary Revisions

General: This is an interesting paper of relevance to the readers of the BMC Pulmonary Medecine. It is generally well written, although improvements remain desirable. The study appears to have been carefully and professionally carried out and interpreted.

Title: Appropriate.

Abstract: Concise with a good structured format. The abbreviations of “SARA” should be definite in the methods. The following statement from the abstract is problematic: “Both scores performed fairly well in identifying people who survived, but not those who died.” The issue is whether the instruments could differentiate those who survived from those who died. The statement should be framed in this way and the results presented accordingly (that the predictive power of both instruments was poor). Indeed, the statement should best be omitted: the hazard ratios tell the story. The authors present no data to substantiate the final statement of the conclusions of the abstract. If this is a statement that could be made irrespective of the results before the study started (just by the nature of the measures) then it doesn’t belong here.

Introduction: This is the weakest component of the paper. After reading it, the reader does not have a clear conviction that this study was worth doing (although the rest of the paper convinces the reader). The listed references are pertinent. The author should have more developed the issue to be addressed in their paper and permit to explain their hypothesis about what they expected to find. In particular, the introduction has a focus on prediction of mortality, and it appears the question is the relative impact of specific and generic instruments on predicting mortality. The authors seem to have emphasized other issues as well (relation between measures) and this is not mentioned in the introduction.

Methods: This section is generally well-done.

1. line 1 : the abbreviation of SARA should be explained as soon as its apparition in the text, and not the second time in line 4

2. paragraphs could be created to separate different parts of this section (study
design, outcome measures, patient classification, statistical analysis…)

3. The analysis is clear up until the last section on Bayesian approaches. This is not comprehensible as written. It needs to be expanded.

4. The authors need to either drop the analysis of the non-obstructed cohort or expand the methods and results to fully articulate the characteristics of this cohort and how they dealt with them. If they keep them, they need to present the results fully (including the hazard ratios associated with the three levels for mortality in this population).

5. It would be very instructive to adjust the results of the survival analysis for FEV1. The issue of relevance is whether the quality of life measures or the other measures have predictive power over and above FEV1.

6. Further, if either model predicts above and beyond FEV1 (if it doesn’t, this is important) then it would be good to run a model with both SGRQ and MDA and see if they predict over and above one another.

Results: Table 2 adds little to the simple statement of the kappa statistic. There is a contradiction in the methods and results that needs clarification. In the methods, the authors state that they abandoned modelling time to death and instead modelled age at death. Yet, the x-axis on figure 1 is not age at death, but follow-up time. This needs to be clarified. The paragraph on probabilities is very confusing. All results should be presented as either the probability of dying or the probability of survival, rather than mixing the two. The reference to different cut-offs is also confusing: surely this can be presented as the probability of survival (presumably at 5 years) in each of the three tertiles.

Discussion: is instructive and well-presented. However, it might change dramatically given the results of the other analyses suggested (whether or not either measure predicts above and beyond FEV1; whether the measures provide predictive power over and above one another).

Conclusion: is clear and effective

References: the 36 listed references appear appropriate and up-to-date.

Tables and Figure: are all appropriate.

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

Level of interest: An article of limited interest

Quality of written English: Acceptable