Reviewer's report

Title: An easy and effective way to classify prognostic comorbidity in candidates for radical prostatectomy.

Version: 3  Date: 11 January 2014

Reviewer: Bimal Bhindi

Reviewer's report:

Manuscript review

Title: An easy and effective way to classify prognostic comorbidity in candidates for radical prostatectomy.

Authors: Michael Froehner, Anna-Elisa Kellner, Rainer Koch, Gustavo B Baretton and Manfred P Wirth

Thank you for inviting me to review the revised manuscript.

The authors submitted a manuscript developing a combined score (Charlson Score + ASA risk classification + BMI) to predict 10-year overall survival in men undergoing radical prostatectomy. They have re-submitted their manuscript following with several revisions. While the manuscript has been improved, the following issues remain:

Major Compulsory revisions:

1) I would recommend revising the title. Deriving two separate scores, measuring BMI, and then calculating a combined score, may not be considered “easy” in the clinical setting. Also, it should somehow be mentioned in the title that a combined score is being created – the present title may unintentionally suggest that an entirely new comorbidity classification system is being proposed.

2) The statement of the study objective remains difficult to decipher. Please revise. Also, please avoid the use of the term “interaction” since this has certain statistical connotations (i.e. interaction terms in multivariable models), and that is not what is meant to be implied here from my understanding. Perhaps, for example: “In this study, we investigated several comorbidity classifications as predictors of overall mortality after radical prostatectomy, searching for measures providing complementary prognostic information which could be combined into a single score.” – or something along those lines.

3) Caution should be used in suggesting that this new score can help decision-making. First, the study population consists of solely of men in whom the decision has already been made (i.e. the decision to undergo radical prostatectomy). The primary urologist has already decided that the risk of PC-mortality exceeds the risk of non-PC mortality, and has gone ahead to alter this risk profile. Moreover, one cannot tell whether undergoing RP is the right or wrong decision based on the proposed co-morbidity score derived in this study.
population.

In the revision letter, the authors state that “… in our opinion, it appears most promising to apply comorbidity classifications to patient populations in which they were developed.” I certainly agree with this. The current study reports on a co-morbidity classification that estimates overall survival in a population undergoing radical prostatectomy. Although this new way classifying co-morbidity may be helpful in the pre-treatment setting, caution should be used in generalizing study findings in this way because this was not the population studied (i.e. patients undergoing radiation or watchful waiting, who may have a completely different comorbidity profile, were not studied). Indeed, as pointed out by the authors in the revision letter, a patient with a Charlson score of 0 undergoing RP vs. radiation have inherently different survival profiles (84% vs. 55% at 10 years), because of inherent differences in these populations (not all of which are easily measured). As previously stated, there should be an explicit statement in this regard about generalizability to the pre-treatment setting (where a decision of RP vs. radiation vs. AS/WW is being made) in the limitations section.

4) The study is looking at overall survival. Therefore dying of prostate cancer and dying of something else are considered the same in the analysis. Predicting a median 10-year overall survival of x% does not help to decide whether or not to undergo surgery. Given that cause of death is available, it would be worthwhile also showing analyses with non-PC-related mortality as the outcome measure, since this may be of particular interest when assessing the predictive ability of a comorbidity measure.

5) The authors state that parameters significantly associated with mortality in univariate analyses were included in multivariable models. Looking at table 1 and 2, it seems that almost all variables should have been included, yet they do not appear in table 3 (the multivariable analysis). Can the authors explain this more clearly? Furthermore, there is evidence to suggest that using p<0.05 on univariate analyses as a criterion for inclusion into multivariable models is too restrictive (if this is the variable selection method chosen). A cutoff of p<0.20 or p<0.25 would be a better initial screening cutoff for variable inclusion.

6) Did the authors formally test for multicollinearity? There is likely significant overlap between co-morbidity scores, particularly Charlson score vs. disease count. While they state the “correlations” were not very strong between the co-morbidity measures in the revision letter, I would nonetheless recommend testing this formally for the multivariable models (ex. using tolerance or variance inflation factor) and mentioning it in the manuscript. To my understanding, SAS does not automatically remove collinear variables as suggested in the revision letter.

7) Although the final figure makes more sense in the revised manuscript, it is still confusing and may be difficult for readers to readily interpret. The sloped dotted lines seem to imply a linear trend, but the point estimates do not imply the same. There is also a break in the y-axis, which really makes this graph even harder to
interpret. This is much more easily interpreted in a table (as the authors have shown in the revision letter, but not in the manuscript), and therefore thus the figure can be removed.

8) As previously mentioned, the difference in survival between lowest and highest categories is not a commonly accepted metric of performance of a co-morbidity score. This difference is inflated by small numbers of patients at the extreme tail ends (in this study, it is based off 53 patients with a score of 0 and 35 patients with a score of 5). Of the 1079 patients in this analysis, 82% fall between a score of 1-3. The survival corresponding difference in mean overall 10-yr survival is 88% vs. 78% (not to mention the confidence intervals and uncertainty of the point estimate). Therefore in the vast majority of patients, this combined score provides very little prognostic information. Perhaps the authors should consider using ROC curve analyses (commonly used in assessing discrimination in prediction models), to see if co-morbidity scores add discrimination to a base model (as was done in the bladder cancer paper cited in the discussion) predicting 5 or 10 year survival. Although challenging to use this in survival analysis (due to censoring), given that there is nearly complete follow up, this is certainly possible to do.

9) It seems that the analysis shown in the last figure is looking at the new proposed score was only performed in the subset of patients who were 65 or older (<50% of the cohort). What were the findings when the entire cohort was analyzed? And why was it chosen to only display findings in the >65 years old subset, and not the whole cohort? If this new proposed score only applies to men 65 or older, then it should be more clearly stated in the abstract and conclusion sections. Specifically, if the score did not have predictive value in men < 65 years old then this is certainly possible to do.

10) Can the authors comment in the discussion section on the use of ASA by anaesthesiologists versus urologists, and the inter-observer reliability in the discussion? This is a score designed for the assessment of peri-operative risk (which they are trained to do), and is not commonly assessed by urologists. This would be analogous to asking an anesthesiologist to assess feasibility of partial nephrectomy for a given renal mass. Moreover, there are studies suggesting that inter-observer agreement even among anaesthesiologists is not great.

11) The conventional way to classify BMI into 3 categories is normal (<25) vs. overweight (25-30) vs. obese (>30). Can the authors explain why they chose to use non-obese [<30] vs. obese class 1 [30-35] vs. obese class 2-3 [>35], particularly since there were only 29 patients in the last category?

12) Can the authors comment in the discussion on the precise intended use of the new proposed score? As commented above regarding clinical use, there are concerns about generalizing it to the pre-treatment setting. It is however perhaps useful in men already selected for RP. In the research setting, precise explanations are needed. Provided that BMI, ASA class and Charlson Score are available, what advantage is there of putting this new proposed score into the model versus simply putting BMI, ASA and Charlson Score as separate variables
in their multivariable models? Please comment in the discussion.

Minor essential revisions:

1) The term “supplementary prognostic information/ significance” is used throughout the text and is somewhat confusing. Perhaps it would be better to say “additional” or “complementary” instead, or rephrase in another way to improve clarity.

2) Figure 1 and 2: The y-axis is not labelled and there are no units.

3) Results, first paragraph: States follow up to “now”. Please mention an exact date of maximum follow up to be more precise.

4) It would be clearer if Table 3 more clearly stated the reference category being used. (For example, the model with two stratifications says “ASA 3”. Although one can figure it out by careful examination of other tables, it would be better if the table was made more user-friendly in this regard.

Typos:

Investigated variables section in methods, line 6: Age was treated as a continuous variable.

Discussion line 8: analysis is misspelled.

**Level of interest:** An article of limited interest

**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