Reviewer’s report

Title: Validation of the Disease Burden Morbidity Assessment by Self-Report in a French-speaking population

Version: 1 Date: 13 October 2011

Reviewer: Andrew Garratt

Reviewer’s report:

Minor essential revisions

1. Translation. In the field of patient-reported outcomes, forward-backwards translation has become the standard method for translating questionnaires. Perhaps this method was not required here given the nature of the questions, which comprise a list of conditions. However, this should be discussed.

2. Why was the additional item of anxiety/depression included? Was this justified according to the results?

3. Cognitive interviews. Please explain how the convenience sample of patients was selected. Was this also in primary care?

4. Cognitive interviews. What was the process of making changes to the questionnaire? How many interviewers were there and what was their professional background? Did the researchers meet up after each interview to discuss potential changes? Please be more explicit.

5. Methods, validation study. Were reminders used in the postal survey?

6. Methods, validation study. In what order were the DBMA-FV based on medical records and the CIRS completed by the research nurse, or were they completed simultaneously?

7. Methods, validation study. One research nurse was used to extract data from patient records and intra- and inter-reliability were not assessed. The limitations of this approach as a means of assessing validity should be discussed.

8. Data analysis. The criterion for incomplete questionnaires is stringent. Does this follow the developers’ recommendations? Is this recommended in practice or are there grounds for taking this approach in the current evaluation of the DBMA? This should also be taken up in the Discussion.

9. Results. The dropout rate due to missing data is very high for T1 compared to T2. Is this because patients lacked time for completion? If so this may be adequate grounds for using the T2 data for validation purposes. However, there may have been a learning effect, with patients having had time to think about the questions between administrations and hence provide data of higher quality that is possibly more reliable and valid at T2.
10. Results. What was the range of missing data? Most patients completed all items but did one or two patients miss out several? Were some items more frequently missed than others? This is mentioned in the Discussion but should also be reported in the Results. Descriptive data for the DBMA items would have been informative including missing data and mean scores (sd) for the activity limitations. How many patients included additional conditions and what were they? How many patients changed areas at T2?

11. Was age the only variable that was statistically significant in the regression model? Please clarify and give the level of significance. Older people have greater comorbidity and hence face a more demanding task when completing the DBMA. How does this relate to the evidence more generally for completion rates and age? This should be taken up in the discussion.

12. The administration of the test-retest study took place in two different settings. The potential implications of this should be taken up in the discussion.

13. Were there differences between the groups who returned (n=85) and did not return (n=12) questionnaires at T2?

14. The test-retest results relate to what might be a select group. This should be taken up in the discussion.

15. Did the source of unreliability relate mostly to the activity limitations ratings or the medical conditions? In the Discussion it is stated that both components of the measure could be used (page 11, para 2) and hence this issue should be taken up.

16. The inclusion of additional areas by patients might increase content validity to the individual patient but are there criteria for determining if an additional area is valid?

17. Discussion. ‘Excellent’ reliability might be taken to imply that reliability could not be improved upon. I would prefer ‘satisfactory’ or ‘high’ which is more appropriate with a test-retest estimate under 0.90.

18. ‘Sensibility’ is used in the Introduction and Table 2 - ‘sensitivity’.

Discretionary revisions

1. Introduction. It is stated that complete measurement properties have not been reported. The other measurement properties that are relevant for such a measure should be briefly summarised here.

2. Methods, The Instrument. ‘Accessible language’ presumably means understandable to patients? Please clarify this. Please also state that items have five-point descriptive scales (as in the original questionnaire). As it stands the reader is left unclear if the scaling is end-point only or all-point descriptive scales.

3. Methods, The Instrument. If patients can add conditions to the list, then in
theory the scores can be higher than 105.

4. The T1 survey data were not used to assess validity because the DBMA was originally developed as a postal questionnaire. However, patient-reported outcome measures are often administered in a clinical setting and hence the DBMA might also be administered in such a manner in future applications. It would also be interesting to compare the two sets of results for T1 and T2.

5. Introduction. “…another index of multimorbidity” can be deleted.

6. Translation. Replace “brought” with “made”.

7. P 9, Reliability. Move “correctly” to before “completed” in the first sentence.

8. P10, Discussion. Replace “relationships” with “associations”.

Level of interest: An article of importance in its field

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.