Reviewer's report

Title: An initial application of computerized adaptive testing (CAT) for measuring disability in patients with low back pain

Version: 1 Date: 30 June 2008

Reviewer: Alan Jette

Reviewer's report:

Title: An initial application of computerized adaptive testing (CAT) for measuring disability in patients with low back pain

Authors: Atilla H Elhan, Derya Öztuna, Kutlay, Küçükdeveci, Tennant

1. Is the question posed by the authors well defined?
Core sets of items based upon the International Classification of Functioning, Disability and Health (ICF), have been proposed to measure outcome for LBP. Recent approaches to outcome measurement involving Computerized Adaptive Testing (CAT) were employed to measure ICF domains in a way that can reduce the burden upon patient and professional. This paper introduces such an approach for LBP.

The question of whether one can develop better outcome instruments for LBP using the ICF framework is very important and IRT/CAT methodology holds considerable promise in this regard.

2. Are the methods appropriate and well described?
In part 1 of the study 266 patients with low back pain answered questions from a range of widely used questionnaires in LBP research. An exploratory factor analysis (EFA) was used to identify domains which were then subjected to Rasch analysis. Reliability was tested by internal consistency and person separation index (PSI). Validity of disability levels were evaluated by Spearman correlation coefficient (r), intraclass correlation coefficient (ICC) and the Bland-Altman approach. A CAT was developed for each domain, and the results checked against simulated and real applications in a second sample of 133 patients.

Three of the 4 instruments used in this study (Nottinham Health Profile, Roland Morris Questionnaire, and Oswestry) were not developed to measure the ICF framework. Only the WHODAS instrument was. Therefore it is not clear if these choices were the best for a study that aimed to develop measurement indicators and scales that correspond to the ICF framework.

The statistical methods used for the most part were clearly described and appropriate. Although the sample size is small for the Rasch analyses, the authors noted this limitation and explained it in appropriate detail.

My one methodological concern is the correlation analyses between the real CAT
with the entire item pool. The authors describe this analysis as ‘convergent validity’. My understanding of convergent validity is that the new measure under evaluation is correlated with another concept hypothesized to be related to the concept being measured with the new instrument. In this case the CAT (the new measure) is being correlated with all items from the pool from which the CAT items are being selected. This in my view is not convergent validation but rather an assessment of the CAT’s accuracy for estimating the scores generated from the entire item pool.

3. Are the data sound?
The data are sound although the sample sizes are small.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?
Yes, very clear and appropriate.

5. Are the discussion and conclusions well balanced and adequately supported by the data?
This is the one area of major concern I have with the paper. The authors interpret their analyses as evidence of two domains that correspond to two domains of the ICF: Body function and Activity/Participation. I for one cannot see the basis for this interpretation and conclusion. The items that load on the two domains that emerged from their analyses seem to be a mixture of elements that cut across the body function and activity/participation domains. In fact most seem to be drawn FROM THE Activity/Participation domains as outlined in the WHO ICF manual (2002). For example, items related to emotional function, social isolation, sleep, understanding and communicating, and energy all of which have been associated with the activity/participation domains of the ICF appear to be loading on the body function domain in these analyses. If anything the results of the analyses seem to question the existence of a distinct body function domain at least as measured by the items selected from these 4 outcome instruments. One reason for this apparent failure to clearly identify ICF specific domains may lie in the fact that the instruments from which the items were chosen were not developed within an ICF context. The manuscript makes reference to ICF core sets for LBP but these in fact are not used as a source of the items used in the analysis.

6. Are limitations of the work clearly stated?
The methodological limitations are clearly and fairly described. I am not sure the conceptual limitations of what they have done is appreciated nor explained for the reader. The ICF might not be an appropriate framework for this work.

7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?
Yes, the authors make reference to many of the major measurement efforts underway in Europe and within the United States. I am not convinced the measures used are the best if the intent is to work within an ICF framework as
the authors purport to do.

8. Do the title and abstract accurately convey what has been found? The title and abstract are very appropriate.

9. Is the writing acceptable? The writing is extremely clear and difficult and unfamiliar concepts for many are well defined for the reader.