Reviewer’s report

Title: Reliability, construct validity and measurement potential of the ICF comprehensive core set for osteoarthritis

Version: 1 Date: 8 September 2011

Reviewer: Thomas Ewert

Reviewer’s report:

GENERAL ISSUES TO NOTE
-------------------------------

1. Is the question posed by the authors well defined?
   Yes

2. Are the methods appropriate and well described?
   Yes

3. Are the data sound?
   Yes

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

5. Are the discussion and conclusions well balanced and adequately supported by the data?
   Predominately, yes

6. Are limitations of the work clearly stated?
   Yes, but in my view not exhaustive

7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?
   No, they missed a validation study for the brief core set for osteoarthritis. This core set is a fraction of the tested core set, therefore it needs to be mentioned despite a different methodology (Xie et al., 2008; Scand J Rheumatol). Additionally an other study (Roe et al., 2009; J Rehabil Med) investigating the dimensionality of the core set for low back pain using Rasch analysis may improve the discussion. This paper used the same way to recode the qualifier for the environmental factors for Rasch analysis.

8. Do the title and abstract accurately convey what has been found?
   Yes
9. Is the writing acceptable?
Yes, it is well written

REPORT TEMPLATE
--------------------
Confidential comments to editors
--------------------------------

Basically a sound literature search must be the starting point of a manuscript. The authors missed a validation study for the brief core set for osteoarthritis (Xie et al., 2008; Scand J Rheumatol). In my view it is necessary to encourage the authors to cover more relevant references. An other point is that the separation of activities and participation does not violate the assumptions of the ICF, but according to publications in this field it would be interesting for the reader to know, on which rationale the separation was based.

Since I make a major compulsory revision comment, according to the guidelines from your journal I have to choose the option “Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions”. Nevertheless, in my view this is a sound and interesting manuscript. It would be easy for the authors to revise the manuscript accordingly.

Reviewer's report
-----------------

General the topic of data based on ICF categories (like ICF Core Sets) is of interest. The Rasch analyses had some advantages in comparison to methods from classical test theory. Some relatively new techniques were used to overcome methodological challenges (like local dependency).

- Minor issues not for publication

1) Abstract, Methods.
   Typo: “besides The Short Form” -> besides the Short Form

2) Results, Patients characteristics
   Typo “…were either retired (%20) or housewives (%70).” -> 20% and 70%

3) RESULTS, “Activities and participation” component, first sentence
   Typo: “… it is also possible to designate some domains as activities and others as participation (ref: ICF book).” –> Ref [4]

- Major Compulsory Revisions

1) There is a validation study for the brief core set for osteoarthritis. The brief
core set is a fraction of the tested core set, therefore I highly recommend to cite and discuss this paper to ensure that all relevant information is included although in the validation study a different methodology have been used (Xie et al., 2008; Scand J Rheumatol).

- Minor Essential Revisions

1) INTRODUCTION, last paragraph

“Thus the ICF classification comprises 1545 categories divided into the four components”

Please provide the reference or cite the ICF. The ICF stated (ICF, p. 23): “At more detailed levels, these codes number up to 1424 items.”

2) Abstract, conclusion, introduction and conclusion section

“…the ICF core set for OA were shown to be valid and reliable through Rasch analysis and classical psychometric methods…” and last sentence of introduction “To accomplish this aim, the scalability of components of this ICF core set was tested by both modern and classical psychometric methods.”

I would expect therefore a direct comparison of both methods. Neither in the method section nor the results or the discussion section the term classical methods is used or reported. I my view the analyses would not cover the whole range of classical methods to examine the aim. Therefore, it would be helpful for the reader to make it more explicit, what is considered as “classical psychometric methods” in the respective sections (or leave it out, in the case it is not a major point).

3) MATERIAL AND METHODS, patients and settings

Please provide more detailed and specific in- and exclusion criteria. I guess there are some missing like age (full of age), language skills to complete questionnaires etc.

4) RESULTS, internal construct validity, “Body functions and body structures” component, first sentence

“In this study, BF and BS categories were grouped and analyzed together since there are only 6 categories in BS component.”

Since you wrote in MATERIAL AND METHODS, internal construct validity, first sentence

“The internal construct validity of the items of the ICF Core Set for OA, proposed as a scale for each ICF component, was tested by Rasch analysis.”

It is not easy to follow your argument that there are not enough items in the body structure component (=6 items) since there are many questionnaires containing subscales with 6 or less items.

5) RESULTS, “Activities and participation” component

“Although in the ICF, the domains in “Activities and Participation” are given as a
single list and the components of “Activities” and “Participation” are not
distinguished, it is also possible to designate some domains as activities and
others as participation (ref: ICF book)…”

This is absolutely right, it is correct to separate activities from participation as you
did. Nevertheless, according to the publications in this field it is unusual (for
example Whiteneck et al., 2009). Most frequently a separation along chapters is
used. Therefore, it is important to make your rationale for this separation clear,
otherwise the reader would not able to follow. In addition, the interpretation of the
results (validity) is also dependent on the rationale (or concept) behind your
separation from activities and participation.

6) DISCUSSION

I miss the discussion about the external construct validity.

Since the SF-36 and the WOMAC cover only parts of the ICF, it is obvious that at
least for some scales a low or moderate co-variation could be assumed. One
possibility to make – more or less precise – hypothesis about the validity is the
linkage of the SF-36 and WOMAC to the ICF scales. If you can assess for
example the numbers (or the ratio) of corresponding ICF core set items to the
respective scale form the questionnaire, you may use this for the expected
covariation. This approach could help to make the findings more meaningful.

- Discretionary Revisions

1) MATERIAL AND METHODS, internal construct validity, second to last
paragraph

“DIF can be detected both statistically and graphically. In the current analysis,
DIF was tested by age, gender, years of education, and disease duration.”

It is good to know, how DIF can be tested, but an explicit information about how
you detected DIF would be more interesting. (Obviously, you perform statistics,
but at this point it is not clear for the reader).

2) DISCUSSION, second paragraph

“The modifications included firstly the collapsing of the categories of some of the
items.”

If I count right, 35 out of 55 items had to be collapsed, therefore “some of the
items” seems very cautious written. In my view it shows clearly that the ICF
qualifier does not work (in the way WHO wants it). It may be helpful for the
reader to describe which consequences follow for example clinicians using the
qualifier but not perform Rasch analyses.

3) DISCUSSION, limitations

To my understanding RUMM 2030 calculates each scale fully independent. Since
we have to assume that the scales (or latent dimensions) were correlated, this
technique may provide a systematic bias. This bias is not considered as so
important to change the result significantly, but there are methods who were able
to handle more adequately with multidimensionality within IRT models.
4) INTRODUCTION/DISCUSSION

You are aiming to built the latent constructs/dimensions according to the ICF components. This is OK and not uncommon. As a consequence you assume that the constructs itself are unidimensional. At this time, there is no common agreement how many dimensions (or factors) ICF based data include. Other authors (Cieza et al., 2009, Disability and Rehabilitation) found that it is possible to perform Rasch analysis for body functions and structures and activities and participation item with one latent construct. As I have shown (Ewert et al., 2010; Disability and Rehabilitation) it is not surprising, that the number of dimensions have for example an impact on the numbers of fitting items. Therefore it may be helpful for the interpretation as well as the background of the manuscript to mention the (possible) dimensions of the data and the consequences for the results. Because your aim was not to explore or test different ways of latent constructs / dimensions, your decision to choose one option has in deed impact on the results and their interpretation.

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