Author's response to reviews

Title: The new COSMIN guidelines confront traditional concepts of responsiveness.

Authors:

Felix Angst (fangst@vtxmail.ch)

Version: 3 Date: 25 February 2011

Author's response to reviews:

Comments to the revisions

Preliminary remark: I have completely revised the whole commentary article (incl. abstract) and restructured it into 5 subchapters with subheadings. I think they help the reader to gain an overview and follow the arguments. If you consider the subheadings to be unnecessary they can be left out.

1. The response is raising questions… was developed using the Delphi method. Almost by definition this is likely to raise debate on conclusions.
   Comment: I completely agree with this comment and have included your remark in the first paragraph.

2.1. Response is not clearly written and unreferenced.
   Comment: I agree to this remark and have completely revised the following paragraphs in order to clarify the content and to improve the citations/referencing.

2.2. The original paper does not seem to suggest to use a gold standard…
   Comment: This is principally true. However, they do suggest the use of an external criterion, especially in the clarification study (to the guidelines) of Mokkink et al 2010, ref. (2). Further, it is not so clear what should be done in the absence of a gold-standard. I have clarified this by discussing items 8-14, p.40/41 of the guidelines, ref. (1) and (2), especially the issue of comparing correlations of score changes; see subchapters 1 and 2.

3. The original paper reports measures that were felt “inappropriate” ….but does not explain why they felt to be inappropriate. Perhaps the author can make more of an argument as to why he feels they are.
   Comment: I have tried to find explanations for this and found one more than those already stated in the text; see subchapter 4: “these measures are considered measures of the magnitude of an intervention or an event, rather than measures of the quality of the measurement instrument”, refs. (1,2). The additional one is: “Measures of treatment effects such as effect sizes of the t-test statistic are in itself only useful for interpretation of score changes, not for assessing the responsiveness of a measure, because it will not be possible to
infer if a corresponding change in the concept has taken place”, see p. 234 of (5). My arguments and criticism are in the same paragraph.