Reviewer's report

Title: Responsiveness of five condition-specific and generic outcome assessment instruments for chronic pain: A prospective cohort study showed superior properties of the Multidimensional Pain Inventory (MPI) and the Short Form 36 (SF-36).

Version: 1 Date: 27 December 2007

Reviewer: Harriet Wittink

Reviewer's report:

Overall, this is a well-written paper on an important topic in chronic pain measurement.

Underneath you will find more specific comments that are meant to help make this paper even better.

Background.

Minor Essential Revisions This section could be elaborated on a bit more, with eg a discussion on the difficulties on measuring responsiveness (which measure to use etc).

It seems there is a contradiction between the statement in the first paragraph “should be performed with generic measurement tools” and your hypothesis that a condition specific instrument is more responsive than a generic. It’s not quite clear as to how you arrived at this hypothesis.

Methods

I realize there are multiple opinions as to how to conduct statistical analyses in measurement and that these opinions diverge sometimes quite considerably. A handy “quality criteria..” paper by Terwee et al. 2007 could provide a guideline.

Minor Essential Revisions In my understanding the choice of calculating an ES versus a SRM depends on the correlation between pre- and post scores. With correlations higher than 0.5 the choice would be a SRM statistic (Norman and Streiner 3rd edition biostatistics 2008). I would suggest using just one effect size measure in this paper depending on your pre-post measurement correlations.

The Hedge’s g for one sample statistic you describe is equal to Cohen’s delta statistic, Cohen’s delta would be the more familiar term to use.

Minor Essential Revisions The use of the term minimally clinically important difference I find confusing as basically what you are doing is describing measurement error (noise) in 1 SEM, elsewhere this is referred to as minimal important difference. The term MCID tends to be used in anchor based approaches to change.

Minor Essential Revisions In Table 1 the ES and SRM of the MCID are
calculated, could you describe how you did this in the statistics section? I have not seen an effect size of a SEM calculated before.

Discretionary Revisions I suggest you calculate the smallest detectable change (SDC, or smallest detectable difference, SDD), which is the smallest difference that can be measured above measurement error. SDC can be measured for individuals and for groups.

$$\text{SDC_{ind}} = 1.96 \times \sqrt{2} \times \text{SEM}$$ or $$2.77 \times \text{SEM}$$. For groups the SDC_{ind} can be divided by $$\sqrt{n}$$, where $$n$$=sample size.

Minor Essential Revisions I am not sure why you would want to describe the overlapping constructs of your measurement tools as this was not introduced in your background. A paragraph in the background on overlap between measurement tools and choosing the right measurement tool for evaluation purposes would be helpful.

Discussion

Minor Essential Revisions Page 10. I would classify the HADS and the CSQ as generic instruments as they measure mood and coping in various disease conditions. Could you elaborate?

What next?: Accept after minor essential revisions

Level of interest: An article whose findings are important to those with closely related research interests

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’