Author's response to reviews

Title: The Shoulder Function Index (SFInX): a clinician-observed outcome measure for people with a proximal humeral fracture

Authors:

Alexander T.M. van de Water (A.Vandewater@latrobe.edu.au)
Megan Davidson (M.Davidson@latrobe.edu.au)
Nora Shields (N.Shields@latrobe.edu.au)
Matthew C. Evans (mevans@mog.com.au)
Nicholas F. Taylor (Nicholas.Taylor@easternhealth.org.au)

Version: 4 Date: 22 December 2014

Author's response to reviews: see over
**Title:** The Shoulder Function Index (SFInX): a clinician-observed outcome measure for people with a proximal humeral fracture

Dear Editor,

Thank you for your compliments and consideration to publish this manuscript in BMC Musculoskeletal Disorders. We have now made amendments based on the comments of the reviewers. Please find below a detailed point-by-point response to the reviewers comments (line numbering in our italic responses refers to tracked changes manuscript file).

We are looking forward to hearing from you regarding your decision on this revised manuscript.

Warm regards,
Alexander van de Water

**Reviewer:** Stephanie Muth

**Reviewer's report:**

Minor Essential Revisions:

Line 74 - remove extra period

• Thank you. We have changed this.

Line 97 - spelling error – Aging

• We have changed this.

Line 131-135 - It is important to stress the wide limits of agreement with regards to the psychometric properties of these tests rather than the lack of information about the psychometric properties as there are numerous studies assessing these.

• As the reviewer state, it is important to address the wide limits of agreement found in the studies. We have addressed this in lines 136-139 and added an example of the relatively wide confidence intervals associated with the DASH. It is however also important to stress that there is little and limited psychometric information of these measures available in people with proximal humeral fracture. There are only five studies that investigated psychometric properties to a limited extent in this population. We therefore think this is important to address as well.

Line 156-157 - Font change

• Thank you. We have changed this.
Major Compulsory Revision:
The authors should include the Penn Shoulder Score (PSS) in their discussion of the available measures of shoulder pain and function. The Penn Shoulder Score has a functional scale much like the one proposed in this study. The primary difference between the 2 tools is that the (PSS) is patient reported versus clinician rated. (I have uploaded a copy of this tool for your convenience)

- Thank you. We have recognised and now included the Penn Shoulder Score in the discussion on page 415-418.

Indeed, SFInX activities are similar to those as in other outcome measures including the Penn Shoulder Score. As the reviewer mentions a difference between the Penn Shoulder Score and the SFInX is the patient versus clinician-rated. Also other differences can be noted such as unidimensionality (Rasch) versus multidimensionality (Penn) and redundancy of Penn Shoulder Score items. These differences should be taken into consideration in choosing a preferred outcome measure for shoulder function for people recovering from a proximal humeral fracture.

Reviewer: Jenny Keating

Reviewer's report:
Minor Essential Revisions
Abstract Line 58
In the main the authors have referred to those with fractured shoulders as people ..in line 58 they are called patients ..I think that 'people' is a fairer representation of the target population

- Thank you. We have changed this.

Line 62 63
In addition, psychometric information of these measures is scarce in this population, and indicate measurement issues with reliability. can you clarify Is it the authors intention to say that more information in needed about measurement reliability

- Since this is an abstract we felt we should raise this issue, and elaborate on this in the main text. We have indeed elaborated on this issue in the Background of the main text. For this, we would like to refer to lines 134-139.

Line 66
reflecting the activity limitations a person has after a proximal humeral fracture...might be reworded reflecting activity limitations following proximal humeral fracture

- Thank you. We have used your suggested rewording.

Background Line 101
Clarify the time period referred to as 'the initial stages after injury'
• We have changed this to “The first weeks” (lines 101-102). We then introduced the term active phase of rehabilitation (line 103) as suggested in a later comment, and have elaborated on a typical timing of recovery (lines 101-106).

Line 102
by reduced or loss of arm function and severe pain. can this be reworded eg by reduced or lost arm function and severe pain.
• We have used your suggested rewording.

Line 103 104
Can the authors place some parameters around typical recovery times referred to here
• We have introduced the term active phase of rehabilitation (line 103) as suggested in a later comment, and have elaborated on a typical timing of recovery (lines 101-106).

Line 105
The ongoing disability after a proximal humeral fracture, is often experienced as............ remove comma after fracture
• Thank you. We have changed this.

Line 121
combines should read 'combine'
• We have changed this.

Line 123
assesses should read 'assess'
• Thank you. We have changed this.

Line 133
The term 'active phase of rehabilitation' is introduced can this be introduced earlier and defined when the typical timing of stages of recovery is described earlier
• We have introduced the term active phase of rehabilitation earlier as suggested (line 103), and have elaborated on a typical timing of recovery (lines 101-106).

Line 133
Also, the information that is available suggests that existing scales may have problems with relatively wide limits of agreement and structural validity [11,15-18]. Can the authors make an argument here for when the limits of agreement would be acceptably narrow and explain what would suffice as evidence of adequate structural validity
• In lines 138-139, we have now provided examples to strengthen the statements concerning measurement issues of existing scales used in people recovering from a proximal humeral fracture:

“Also, the information that is available suggests that existing scales may have problems with relatively wide limits of agreement (for example, ±15% of total scores for the DASH) and structural validity (for example, inclusion of multiple constructs and redundant items).”
At some stage in the text before this point, can you explain why shoulder function following proximal humeral fracture is likely to need its own assessment instrument; the rationale of interest is why you think that function might be assessed in a different way if the reason for the shoulder problem were something other than proximal humeral fracture. This bears on the potential generalisability of the utility of the instrument for assessing recovery from other shoulder conditions

- The rationale was not specifically that shoulder function requires its own outcome measure in this population. However, shoulder function as a single construct requires a well-developed outcome measure. This we aimed to clarify in the lines 117-119 “This requires a functional outcome measure that is unidimensional (measures the single construct of activity limitations), psychometrically sound, relevant to the patient and clinically feasible”. The population, people recovering from a proximal humeral fracture, has been a good starting point to develop the SFInX from.”

In the discussion (line 422-426) we have also elaborated on the point raised in this comment. (“...the SFInX has also potential for use in other populations...”)

The point here has been made several times before and it is now feeling Repetitive

- We have taken this sentence out to avoid repetition.

Development of the SFInX comprised two phases: should read Development of the SFInX was comprised of two phases: same point in line 175

- Thank you. We have changed this in both instances.

Table 2

- carry heavy object with 2 hand should read carry heavy object with 2 hands

- We have changed this.

Table 3

Time after fracture (weeks)
(¼ yearly distribution)
26.5 ± 15.1 (5-52)
20, 30, 19, 23

Is this Time after fracture in weeks mean, sd (range); mean for 1/4 year periods It took me a long time to unravel can you consider a more user friendly way to present

- We understand the reviewer’s comment and have changed this in the same style as other results in this table, reported as: time after fracture in months, n=x (x%)
make them available to anyone who would like to reconsider the outcomes in the light of
AO categories I had the same problem with the breakdown by Neer categories ..did not
know how to use the data other than to scan the distribution across categories.

We have made some changes to the presentation of these results: we have written
out \((n=xx)\) per category per classification. We have reduced the subcategories of the
AO classification. Also the addition of footnotes for the abbreviations will make the
Neer categories clearer.

We acknowledge that when a reader is not directly familiar with the categories of
these classifications it is harder to interpret these results. In that case, we refer to the
original publications of the classification systems in lines 224-227.

Table 5
I tried to work out why the items were ordered in the way they were, but failed Perhaps if
the table is introduced with 'ordered to show misfit averaged across infit and outfit or
whatever is the reason...)

- The Table was sorted by the bigger of Infit mean-square and Outfit mean-square, in
descending order. We understand this might not be directly visible or ideal, so we have
changed the order to INFIT descending, and have stated this at the heading of the Table.

Can the authors add whether the instrument will be/is freely available (assumed from the
provision of the manual but not stated)

- In lines 453-463 we have added that the outcome measure is freely available. Beside
publication of the scale in this Open Access BMC Journal, the SFInX will soon have a
website where instructions on clinical assessment and further information such as
measurement properties will be available.

Was time to complete collected?

- The time to complete the SFInX was collected. The estimation to complete all 13
items is 5-7 minutes. We see time-to-complete as part of feasibility of the instrument,
and therefore have not included this information in this SFInX development paper.
We aim to publish a second paper regarding Measurement properties and feasibility
in which we will include time-to-complete information.

Is it not possible to run any exploratory differential item functioning

- For two reasons we have not included differential item functioning (DIF) analysis in
this manuscript: size of subgroups, and relevance of DIF analysis outcomes from
small subgroups. Technically it is possible to run an exploratory analysis on DIF with small groups.
Relevant variables we would consider are gender, fracture of dominant side and age
groups. We have not included any DIF analysis since the subgroups would be rather
small. It has been suggested (Linacre, 1994) that \(n=30\) would be sufficient for
informal DIF analysis. For the above variables subgroup sizes would be: gender \((n=79
and n=13)\), fracture of dominant side \((n=44\) and \(n=48)\), and four relevant age groups
\((n=16\) to \(n=28)\). For two of these analyses the number of patients falling in the
subgroups is below \(n=30\). The group sizes for only one DIF analysis would suffice this
\(n=30\) criterion. This analysis would be informal, in other words, adequate conclusions
on differential item functioning cannot be drawn from this analysis \((n>200\) is
recommended for formal DIF analysis (Scott et al, 2009)). Therefore, we acknowledge the sample size as a limitation for DIF analysis in the current study (lines 435-440).  

Is there any reason to believe that the instrument might be suitable for assessing function in a range of shoulder conditions in addition to proximal humeral fracture

- Yes, we believe there is indeed potential for the SFInX to adequately measure shoulder function in other shoulder populations with a similar clinical pictures. We have now elaborated on this in the discussion in lines 422-426.
  "Through these methods the SFInX has also potential for use in other populations than people recovering from a proximal humeral fracture. People suffering from a shoulder condition with a similar clinical picture, experiencing limitations in their daily activities might also benefit from functional measurements of the SFInX, but this would need to be confirmed with further study."