Reviewer’s report

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

**Version:** 2  **Date:** 7 December 2014

**Reviewer:** Jenny Keating

**Reviewer’s report:**

Thank you for the opportunity to review this report that describes the development of an instrument for assessing shoulder function following humeral fracture.

The assessment instrument has been developed using sensible principles and provides clinicians with a method for assessing shoulder function.

My suggested revisions are all minor.

**Major Compulsory Revisions**

I have no major revisions; the work has been well conducted and reported.

**Minor Essential Revisions**

Below are observations I recorded on reading the report that might be considered by the authors.

**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.

**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?

**Line 66**

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

reflecting activity limitations following proximal humeral fracture

**Background Line 101**

Clarify the time period referred to as ‘the initial stages after injury’

**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.

Line 103 104
Can the authors place some parameters around typical recovery times referred to here

Line 105
The ongoing disability after a proximal humeral fracture, is often experienced as..........
remove comma after fracture

Line 121
combines should read 'combine'

Line 123
assesses should read 'assess'

Line 125
explain sub-scores to the reader

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

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

Line 140
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

Line 140, 141
The point here has been made several times before and it is now feeling repetitive

Line 150
Development of the SFInX comprised two phases:
should read
Development of the SFInX was comprised of two phases:
same point in line 175

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

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
Table 3
add footnotes for the abbreviations

I found the break down of the AO classification difficult because I am not familiar with the classification system; can you help readers either with a breif summary of the relevance of the distribution of people across AO categories, or by removing these data and offering to 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
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..."

Can the authors add whether the instrument will be/is freely available (assumed from the provision of the manual but not stated)
Was time to complete collected?

Is it not possible to run any exploratory differential item functioning

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

Discretionary Revisions

These are recommendations for improvement which the author can choose to ignore. For example clarifications, data that would be useful but not essential.

Please note that both the comments entered here and answers to the questions below constitute the report, bearing your name, that will be forwarded to the authors and published on the site if the article is accepted.

**Level of interest:** An article of importance in its field

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests