Author's response to reviews

Title: Reliability and validity of the Physical Activity Scale for the Elderly (PASE) in patients with hip osteoarthritis.

Authors:

Ida C Svege (ida.svege@hjelp24.no)
Elin Kolle (elin.kolle@nih.no)
May Arna Risberg (mayarna.risberg@hjelp24.no)

Version: 4 Date: 9 January 2012

Author's response to reviews: see over
Dear editor

We thank you and the reviewers for thoroughly reviewing our manuscript, entitled “Reliability and validity of the Physical Activity Scale for the Elderly (PASE) in patients with hip osteoarthritis”.

A point by point response to the proposed revisions given in the reviewer’s reports is given below. Changes made in the revised manuscript are indicated by ‘tracked changes’ (coloured and underlined text). We have to our best effort addressed the reviewers questions and comments. We hope you will find our responses satisfactory, and that you will consider our manuscript for publication in BMC Musculoskeletal Disorders.

Best regards
On behalf of all authors
Ida C. Svege
Corresponding author
Editorial request:
Please make the following changes during revision of your manuscript:
Tables: Please note that we are unable to display vertical lines or text within tables, no display merged cells: please re-layout your tables without these elements.

Table 1 and Table 3 have been changed to meet the request. Vertical text has been changed to horizontal text, and the tables no longer contain merged cells.

Reviewer #1:
Major Compulsory Revisions

Methods
In general the methods are very clearly written and easy to understand. The questionnaires are well described and the theory behind the research is well explained. Please clarify why a minimum of 3 days of wear of the accelerometer was considered sufficient for comparison with the PASE and IPAQ, both of which measure activity over the previous 7 days. If a patient had had a very active (or inactive) 3 days and these figures are being compared with the average figures from a 7 day period assessed by PASE then this may explain the relatively poor correlation. Please clarify how many days the patients wore the accelerometer for. How many people wore the acceleromter for the full 7 days? I am concerned that the authors have stated that the PASE is not valid based on a relatively small sample group and comparison with the accelerometer that captured data over a different period.

Discussion
The information for the test/re-test reliability of PASE and assessing validity with comparison with IPAQ and accelerometer is interesting and warrants publication but before this manuscript is accepted I would like further details as described above.

As commented by the reviewer the Actigraph GT1M registration period should reflect the recall period of the PASE questionnaire. Based on this we decided to conduct a new analysis for the correlation between the PASE and the Actigraph GT1M. In the new analysis we included patients with a period of registration reflecting the 7-day recall period of the PASE. We excluded the patients with less than 6 days of registration, but included those with 6, 7, 8 and 9 days of registration. Two patients with three days of registration (one of which was already exluded in the first analysis) and one patient with four days of registration were thus excluded in the new analysis. This left 33 patients to be included in the study. Five patients had 6 days of registration, 23 patients had 7 days of registration, 4 patients had 8 days of registration and 1 patient had 9 days of registration. The average number of days of registration for the patients included in the analysis was 7.0 (0.6).

The new analysis revealed very similar results as the first analysis, with a Spearman’s correlation coefficient between the total PASE score and the Actigraph GT1M total counts per minute of 0.30, compared to 0.28 in the first analysis. The correlation coefficients for the different intensity categories were 0.20 (light physical activity), 0.38 (moderate physical activity) and 0.29 (vigorous physical activity), similar to 0.10, 0.35 and 0.29, respectively, in the first analysis. If we were to include only those patients with exactly seven days of
registration in the analysis, the correlation between the total PASE score and the Actigraph GT1M total counts per minute would have been 0.27.

New calculations of the mean (SD) score for average counts per minute and minutes spent in the different intensity categories based on the 33 included patients were similar to the first calculations based on 35 patients.

Due to the new analysis, the following changes are made in the text:

Abstract:
- Page 2, third paragraph, second sentence: “0.28” was replaced by “0.30”, “(p=0.110)” was replaced by “(p=0.089)”, “0.10 to 0.35” was replaced by “0.20-0.38”

Results:
- Page 10, third paragraph, third sentence: “with three or more valid registration days” is removed
- Page 10, third paragraph, fifth sentence: “In addition, recordings from one patient were excluded because the patient had informed us that he had stopped wearing the Actigraph GT1M because he “was not able to be as active as usual”” was removed, and the sentences “Six or more days of registration were considered to be sufficient. Three patients had less than six days of registration and were thus excluded from the analysis” were added.
- Page 10, third paragraph, seventh sentence: “35” was removed and replaced by “33”
- Page 10, third paragraph, eighth sentence: The sentence “The average days of registration were 7.0 (0.6)” was added.
- Page 11, second paragraph, second sentence: “69 %” was replaced by “67 %”, “31 %” was replaced by “30 %”
- Page 11, second paragraph, third sentence: “44.4” was replaced by “45 (32)”
- Page 12, first paragraph, second sentence: “0.28” was replaced by “0.30”, “(p=0.110)” was replaced by “(p=0.089)”
- Table 3, page 29: mean (SD) scores for average counts per minute and minutes spent in the different intensity categories, and the correlation coefficients for correlation between PASE and Actigraph GT1M.

Discussion:
- Page 13, second paragraph, first sentence: “68 %” is replaced by “67 %”, “31 %” is replaced by “30 %”.
- Page 16, first paragraph, second sentence: “0.30” was replaced by “0.28”
- Page 17, first paragraph: the paragraph is divided in two between the seventh sentence, ending with “moderate to low reliability” and the eighth sentence, beginning with “Our a priori hypothesis”
- Page 17, second paragraph, third sentence: “0.41” is replaced by “0.46”
- Page 18, second paragraph, second sentence: “calculation of test-retest reliability was based on 33 patients who had completed the PASE at both test and retest, and calculation of construct validity of PASE compared to accelerometer were based on data from 35 patients” was removed and replaced by “Both analysis of test-retest reliability and construct validity by comparing PASE to the Actigraph GT1M were based on data obtained from 33 patients.”

The reviewer comments on the relatively small sample group in this study, which also is mentioned as a limitation of the study in the Discussion section, page 18, in the manuscript.
Sample size varies in similar studies. The studies by Washburn and Ficker (1999), Dinger et al (2004) and Hagiwara et al (2008) all evaluate the validity of the PASE by comparing it to an accelerometer, and have included 20, 56 and 325 subjects, respectively. According to Forsén et al (2010) sample size in studies evaluating realibility and validity of physical activity questionnaires in elderly varies from 17 to 325 subjects. Prior to study start we consulted a statistician regarding sample size in this reliability- and validity study. He assumed that inclusion of 30-40 patients would be sufficient, based on that this study includes outcome measurements using a continuous scale, allowing a somewhat smaller sample size. Based on the suggestions from the statistician, and that some other studies have used similar sample sizes, we decided to include 40 patients in this study.

We added the sentences “After referring a statistician, and based on that other studies have used similar sample sizes [19, 33], we decided to include 40 patients in this study. According to the statistician a sample size between 30 and 40 is usually sufficient when evaluating outcome measurements that uses a continuous scale” to page 18, first paragraph to provide supplementary information concerning the sample size.

Conclusion
I would currently remove the sentence “Overall, the test-retest reliability and the construct validity revealed that the PASE was not acceptable for recording PA intensity” from the Conclusion for the reasons described above. Otherwise the conclusion is accurate and a fair summary of the research.

We removed the sentence ”Overall, the test-retest reliability and the construct validity revealed that the PASE was not acceptable for recording PA intensity” from the Conclusion, page 19, as suggested by the reviewer.

Minor Essential Revisions

Outcomes
IPAQ. Please clarify what is meant by MET
The metabolic equivalent of task (MET) is a physiological measure expressing the energy cost of physical activities. It is defined as the ratio of metabolic rate during a specific physical activity, where 1 MET represents the resting metabolic rate.
The sentence: “The IPAQ is scored by using the Metabolic Equivalent of Task (MET) method, where different activities and levels of intensity are given different MET estimates” was added to the description of IPAQ on page 7, second paragraph.

Table 1
This data was collected at the baseline visit for the study, which was over 2 years ago. A note to this effect should be added either to the paper or to the table so this is clear. Has the data in the table (e.g. age) been altered to all to reflect this 2 year gap?
The patients were included to the original RCT between 2006 and 2008. We have tried to clarify this at page 5, first paragraph, fifth sentence by replacing “61 patients who had attended the two year follow up of the RCT” with “61 patients who had been included in the original RCT between 2006 and 2008”.

The data in table 2 was, as noted by the reviewer, collected at the baseline visit for the original RCT. Data on age have been altered to reflect the time of execution of this present
study (October 2010). To clarify this we added the sentence “Data on age has been altered to reflect the actual age at the time of data collection in this validation study” to page 5, second paragraph.

Discretionary Revisions

**Background**

*Second paragraph - please clarify what is meant by “doubly labelled water” to briefly explain the theory/process.*

Doubly labeled water (DLW) is water where both hydrogen and oxygen have been partly or completely replaced for tracing purposes, usually with the isotopes deuterium (²H₂O) and oxygen-18 (¹⁸O). DLW-method is a method for direct assessment of energy expenditure, where the average metabolic rate is measured over a period of time. This is done by administering an initial dose of doubly labeled water, and then measuring the elimination rates of deuterium and oxygen-18 over time by sampling isotope concentrations in urine. The method is highly advanced and is often considered to be the gold standard for measuring physical activity. However, it is expensive, time-consuming and reliant on expertise and equipment, and it also have some limitations regarding measuring of frequency and intensity. These factors are discussed on page 18, second paragraph.
Reviewer #2:
Minor Essential Revisions

1. Analysis, second paragraph, third sentence – reviewer recommends revising the wording. Wording is revised to “As recommended by Terwee [15], the most similar constructs of the PASE and the Actigraph GT1M were compared.”

2. Table 2 – Term “SDC (smallest detectable change)” appears in this table. Elsewhere in the manuscript the term “MDC (minimal detectable change)” is used. Reviewer recommends standardizing terminology throughout the manuscript, and suggests the term MDC be used. Table 2, page 28: “SDC (smallest detectable change)” is replaced by “MDC (minimal detectable change)”

3. Table 3 – Is currently labelled as Table 3. Table 3, page 29: labelled as Table 3.

Discretionary Revisions

1. Background, second paragraph, third sentence – consider changing “is” to “are” “Is” is replaced by “are”

2. Analysis, last paragraph – consider whether SPSS citation and referencing agrees with conventional reporting recommendations
“PASW 18 for Windows (SPSS Inc., Chicago, IL)” was replaced by “PASW Statistics 18 for Windows (IBM Corporation, Route, Somers, NY, USA)”

3. Discussion, third paragraph, fifth sentence – consider changing “is” to “are”. 4. Discussion, fourth paragraph, first sentence - consider changing “represent” to “represents”.

5. Discussion, fifth paragraph, fifth sentence – consider changing “distinguishing” to “distinguish”

6. Discussion, ninth paragraph, tenth sentence – consider inserting “the fact” after “This may be due to”.

These changes have been made