Reviewer's report

Title: Reliability and Sensitivity of the 20-Meter Walk Test among Patients with Knee Osteoarthritis

Version: 1 Date: 2 February 2013

Reviewer: Mohammad Mansournia

Reviewer's report:

The current paper aims to assess the reliability and sensitivity of the 20-meter walk test among patients with knee osteoarthritis. The objectives of the study are clear and the data are well described. However, there are several concerns regarding the methodology of the study.

Major comments:

1) The sample size of the study is too small for precise estimation of 95% limits of agreement. The 95% confidence intervals for the limits of agreement can be calculated and these will be wide with a sample size of 15. Altman recommended a sample size of at least 50, but preferably rather larger, for a method comparison study (practical statistics for medical research, Page 402).

2) Page 3, line 54: "Changes in walk time between -2.59 seconds (walking slower) and 1.65 seconds (walking faster) should be considered within the range of normal variability of 20-meter walking speed."

This is an overstatement: the estimates of the limits of agreement (i.e., -2.59, 1.65) are imprecise due to small sample size of the study.

3) The SDD (usually defined as mean difference ± 1.96 times the SD of the differences) is computed based on 95% confidence interval (CI) for the median difference. This is a non-parametric alternative for 95% CI for the mean difference, but not for 95% limits of agreement. In fact, the 95% CI for the median difference converges to the median difference in the population as the sample size approaches infinity.

4) The Bland-Altman plots for differences between sessions and between trials within sessions 3 and 4 are not presented in the manuscript.

5) The ranges of agreement instead of 95% limits of agreement have been reported in Table 3. The range is not a satisfactory measure of variation, because it is determined by only two of the data points (which are the extreme values and perhaps outlier) and depends on sample size.

6) Page 10, line 208: "There was a lack of agreement between session 1 and sessions 3 or 4 with no discernible pattern. In contrast, session 2 had agreement with sessions 3 or 4."
Table 3 of the paper does not support the statement e.g., the range of agreement between sessions 1 and 3 and between sessions 2 and 3 are not markedly different.

Minor comments:

1) Page 8, lines 145 and 160: P value should be replaced with alpha level.

2) Page 9, line 184: The degree of freedom (df) for the t statistic is required.

3) Page 10, lines 193 and 194: "Table 2" is correct.

4) Page 11, line 212: Why did the authors report p value for the Spearman correlation?

5) Page 21, Table 2: Please report 25th and 75th percentiles of walk time instead of minimum and maximum values.

Level of interest: An article of importance in its field

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.