Reviewer's report

Title: Movement control tests of the low back; evaluation of the difference between patients with back pain and healthy controls

Version: 2 Date: 21 October 2008

Reviewer: Andrew Briggs

Reviewer's report:

Major compulsory revisions

DISCUSSION

It is important to also comment on:

i) Not randomizing the order of tests and the implications for a learning effect (limitations of the study)

ii) The results of the subgroup analysis

iii) Interpretation of the results and offer an explanation as to why they are so (subgroup and overall). Moreover, the issue of face validity of each test has not been addressed – this is important for interpreting the results. Without some interpretation of the results – based on theory/literature/your results I feel that the Discussion lacks some depth (point 5v in my original review).

iv) Variability in reliability coefficients reported in the literature – why are they so different between studies even though they use the same test. I don’t think this has been addressed adequately in the revised manuscript.

v) Include in the limitations about non objective-based measurement – this has not been addressed adequately in the revision. Why should future research examine differences in joint angles?

vi) The future directions section needs some revision. My original comment referred to the use of imaging as being contentious in the diagnosis of LBP disorders, but this would have no bearing on the use of imaging as a reference for establishing validity of MC tests.

TABLE 1
The unit of measurement in the table should be referred to in the title. I assume these are Kappa coefficients?

TABLE 3
What is meant by ‘local back pain’… this is not referred to elsewhere.

FIGURE 3
Significant differences between the groups should be indicated on this Figure to enable it to act as ‘stand alone’.
Minor essential revisions

ABSTRACT
Data concerning the subgroup analysis I suggested should also be presented in the abstract.

INTRODUCTION
Page 4, para 2: by gold standard, do you mean a criterion measurement tool? This should be clarified.
Page 5, para 2: I am not convinced that the range of reliability coefficients you quote is acceptable to reliable. Would it not be more correct to infer that there is a very large range of coefficients suggesting low to high reliability.

DATA ANALYSIS
Remove “a one way ANOVA for non parametric data” as this sentence may also be interpreted that you used a 1-way ANOVA in addition to the KW test, which is not correct.

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

Quality of written English: Acceptable

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