Reviewer's report

Title: The intra- and inter-rater reliability of five clinical muscle performance tests in patients with and without neck pain.

Version: 1 Date: 15 June 2013

Reviewer: Shaun SO O'Leary

Reviewer's report:

Minor essential revisions
1. Introduction: There is no description in the introduction regarding what each of the measure records ie. Time in seconds, pressure reached etc etc. Often these tests have multiple criteria for failure in clinical practice, especially when considering something like the CCFT. Therefore it would appropriate to briefly describe what each test is actually trying to establish.
2. Page 4 Paragraph 2 Line 2: “supine hook lying” – I have not seen this term before, is it meant to be crook lying?
3. Participants section could be written more efficiently. There is repetition regarding exclusion criteria.
4. Page 7, Paragraph 2 - “given cut-point for each test”. What is meant by this? Does the literature not already give cut off points for performance?
5. Page 8 Paragraph 1 – “The rate at which participants performed the movements was not formally controlled. However, all subjects were instructed to move at a comfortable pace” – What movements is being discussed here. If it is related to specific tests it should be described relative to the individual tests.
6. Page 9 Paragraph 1 – Regarding the line describing recording performance of the CCFT “The examiner recorded which level of pressure the participant successfully achieved.” What did “successfully achieved” mean. This is not clear. How did the therapist determine what was excessive superficial muscle use. Did they use some sort of biofeedback or just palpation which surely is very subjective?
7. Page 10 Line 2 – “laser exceeded the measuring” This is unclear what this refers to. Exceeds in what direction ie. do they loose CCF position or do they lower their head to a set point?
8. Page 10 Paragraph 2 – “The test was terminated when the laser exceeded the measuring tape due to head movement indicating fatigue (i.e., inability to maintain upper cervical flexion, increase in neck flexion or lowering of the head)”. This should be stated in the description of the main tests as per above comment 
9. It would be good to see if endurance measures were generally different between the two sessions undertaken on the same day (irrespective of the order of examiners). Can that data be shown as these were endurance measures and patients justifiably may have performed worse on the second session of the day
just due to fatigue? Was this the case? In this manner these reliability tests may not mimic clinical reality as endurance tests are usually repeated after at least several days to assess the impact of training, not on the same day with inadequate time for recovery.

10. The description of “exceeding the measure” used in all endurance tests needs clarification.

11. The discussion needs to acknowledge that the findings only relates to how the tests were performed in this study.

12. Not enough is made of the fact that these were recently graduated therapists. While I understand that the reliability needs to be generalizable to all therapists the level of subjectivity and clinical skill required for these tests (with the exception of the JPE) may be relevant for inexperienced therapists. It would be good to at least acknowledge this more specifically than is currently in the discussion.

13. I was not able to view any Figures?

**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'