Reviewer's report

Title: Reliability of movement control tests in the cervical spine

Version: Date: 2 August 2014

Reviewer: Eva Rasmussen Barr

Reviewer's report:

'Reliability of movement control tests in the cervical spine'
Maja Patroncini, Susanne Hannig, Andre Meichtry and Hannu Luomajoki
BMC Musculoskeletal Disorders

The study is important as many tests used in the clinical situation are not studied for reliability and absolutely not in neck pain and movement impairment.

However there are several parts lacking in the manuscript. Most important 1) in the method section – we don’t know anything on how the rating was performed or done 2) the discussion part is also very short and lack important passages of both strength and limitations.

Major Compulsory Revisions

Is the question posed by the authors well defined?

I consider that the question of reliability is well defined. I question if you can base a recommendation of a test battery on only one study of reliability as the study is a reliability study performed on both healthy and those with neck pain. The external validity of the reliable tests will only be those with the same pain level as in the study and only those that are examined in the same context/system with ‘reference lines’ to compare and use as references with the patients movements/tests. This study is a reliability study primarily. The limitation of the context used during the ratings (ref lines) needs to be discussed for external validity.

In the background section there are some issues that need to be addressed:

Page 1, line 74. You state that in movement impairment the patient moves uncontrolled. This needs to be rewritten. I suppose that you mean that the patient has a problem to control the cervical or the lumbar spine during movement.

Page 1, line 75-77. You also needs to clarify this sentence in how patients with MSI differ from persons with movement impairment.

Page 1, Line 75-77. Persons with MSI have not only problem with specific positions but also very common with movements as you have already stated above. Pls change and clarify.

Page 1, line 79-81 I also find this passage strange and not relevant in this context.

2. Are the methods appropriate and well described?
The method section is partly well described. You need to complete the section on how the raters rated the video clips. Did they do it on the same day or did they have more time to do it. Pls be more detailed. In the discussion part you need also to discuss the system (with ref lines) you use as a both a strength and limitation.

Statistical

How did you choose this cut off of K=0.06. Sufficient for what - for clinical use, for research or? It is proposed a cut of off K>0.40 for the use in clinic. Do you have a ref for this?

Based on Sims et al the lower bound of CI should be 0.2 or more. What ref do you have of 0,4?

3. Are the data sound?
The data seem to be sound.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?
Yes

5. Are the discussion and conclusions well balanced and adequately supported by the data?
The discussion is very short and I find that there are parts in the discussion that are lacking.

Section two. You need to discuss why you included the 13 tests if you don’t find that they are movement control tests. I also question that you have included so many. This is something that may affect the raters; to watch and rate so many video clips (this is also not described how and when this was done). Some of the tests are not relevant for subjects with neck pain. This is a limitation to the study on which you should elaborate.

You also need to discuss the way these tests are rated as yes/no and how that may affect the kappa values - and discuss if a more complex scale score instead of a dichotomous will affect the kappa value. The information on yes and no for the clinician – what information do we get and how do we value it? You also need to discuss the system you used with ref lines. This is nothing you use in the clinic. It is a good help to rate a movement but not used by a regular physical therapist in the clinic. This is a real limitation of the study. Could you repeat your study and get the same Kappa values with no ref lines behind the moving person?

You should discuss the prevalence-bias or problem with Kappa statistics. Also what does it mean if you use a cut off of 0.6 instead of the cut offs proposed by Sim and other authors?

You should after adding information how the raters performed there ratings also discuss this in discussion section. Pros and cons with video recording instead of ‘live’ observation.
Limitations of ratings of several tests. Again how was that performed?
For what kind of patients are these test good for based on yr included subjects?
i.e. the external validity of the tests
Conclusion: I find that the last sentence in the conclusion is non-relevant. Of course we use all kind of examinations to target rehabilitation. However this is not part of yr conclusion.

6. Are limitations of the work clearly stated?
Pls see above
7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?
Yes previous studies on reliability.
8. Do the title and abstract accurately convey what has been found?
Yes
9. Is the writing acceptable?
I think that the manuscript needs a language revision.
Table 1
How come that the healthy subjects have rated on NDI and have a score even if low?
What kind of trauma for the healthy?
Table 2
Are the tests performed in sitting or standing. Should be specified
Figure 15 – I don’t approve of this diagram. The kappa values should be presented in the diagram not as staples but as ‘dots’ with their confidence intervals

Level of interest: An article of importance in its field

Quality of written English: Needs some language corrections before being published

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

Eva Rasmussen Barr