Author's response to reviews

Title: Reliability and group differences in quantitative cervicothoracic measures among individuals with and without chronic neck pain

Authors:

Bahar Shahidi (bahar.shahidi@ucdenver.edu)
Cynthia L Johnson (cynthia.johnson@ucdenver.edu)
Douglas Curran-Everett (everettd@njhealth.org)
Katrina S Maluf (katrina.maluf@ucdenver.edu)

Version: 2 Date: 28 September 2012

Author's response to reviews: see over
Response to Reviewers RE: Biomedical Central Musculoskeletal Disorders MS #:
1554934425747208

We thank the reviewers for recognizing the contributions of this study to the musculoskeletal
assessment of neck pain, and for their thoughtful suggestions to improve the overall quality of
the manuscript. A point-by-point response to the reviewers’ critique is provided below.

Major Recommendations:

Reviewer #1:

Background:

1. Please outline whether hand held dynamometry is a robust tool

Further elaboration of hand held dynamometry as a robust and objective tool in clinical
practice is now clarified in the background. The references provided support the use of
hand held dynamometry for the assessment of isometric muscle strength across a
variety of test positions and muscle groups.

2. The authors outline that raters were not blinded to the clinical status of the participants.
However,

a. Were raters blinded to their previous findings?

Due to the sequential nature of the ordered testing, the raters were not blinded to their
own previous measurement as repeated measurements were performed consecutively.

b. Were raters blinded to the other rater’s findings?

Raters were blinded to the other rater’s measurements for all examination items.

c. Were raters blinded to other clinical information or cues that were not part of the test?

Raters were blinded to other clinical information/cues that were not part of the test.

The Procedures paragraph of the manuscript has been revised to clarify all procedures
related to blinding.

3. Paragraph 3-Procedures: Randomisation

The authors outline that the therapist order was randomised. Was the order of testing
randomised?

As acknowledged in the revised Procedures section, test order was not randomized.
This procedure was selected to ensure that any potential carry-over effects from
previous testing would be consistent across examiners and study groups.
4. Paragraph 5: Why was only one trial of cervical muscle strength taken when three trials were taken for scapulothoracic strength?

_Only one trial of cervical muscle strength was taken to limit the potential for increasing primary symptoms of cervical pain with repeated maximal effort testing. This was in an effort to avoid interference of pain with force production on subsequent tests. Maximum strength of accessory scapulothoracic muscles was measured 3 times based on pilot trials indicating that repeated strength testing of these muscles was well tolerated and did not increase primary cervical symptoms. This point is now clarified in Paragraphs 5 and 7._

5. Paragraph 5, Line 12: Is there a reference for the method of stabilisation utilised during side bending?

_We are unaware of any studies documenting this particular method of stabilization during isometric cervical side bending, as the only previous measurements of side bending strength have been performed using expensive instrumented equipment not typically available in the clinic. Our method of stabilization was chosen based on pilot studies designed to best stabilize the position of the thorax relative to head movement. This point is now clarified in Paragraph 5, and the novel test position is documented in Fig 1c._

6. Paragraph 6: Explain why only one test for endurance was taken.

_For similar reasons described for cervical strength testing, only one trial of cervical endurance was performed in each direction to minimize the potential for reproduction of cervical pain. This point has been added to Paragraph 6._

7. Paragraph 6, Line 5: How was the maintenance of upper cervical flexion assessed. I noted the authors refer to a change in chin position; please clarify how this was assessed and whether it was measured.

_Maintenance of upper cervical flexion was assessed visually. Specific criteria for test termination are now explained in more detail in Paragraph 6._

8. Paragraph 8: How many measures of muscle length were taken and why?

_To avoid influencing length measures by changes in the viscoelastic properties of muscle and connective tissue during repeated movements, only one measurement was performed for muscle length assessments. This point is now clarified in Paragraph 8._

Analysis and Results

10. Were data normally distributed?

_Analysis of normality of data using the Komologrov-Smirnoff test as well as Q-Q plots of distributions verified that the data for each outcome measure were normally distributed. This has been added to the Data Analysis section._
11. Results, Paragraph 1 and 2: In the data analysis, planned calculations ICC(95% CI) are described but later the ICC(SD) is presented in Results, Paragraph 1 and ICC(95%CI) in results, Paragraph 2. Please amend.

> **ICC(SD) values for intra-rater reliability presented in the Results have been revised to indicate the range of ICC(95%CI) values obtained from each Rater across all movement directions, rather than the average (SD) of ICC values for all measurements.**

12. Were MDC's calculated for intra-rater reliability? It would be helpful to present intra-rater reliability results in a table similar to the presentation of inter-rater reliability results.

> **Repeated assessments for cervical range of motion and scapular muscle strength were performed in consecutive order within the same session to help maximize test-retest reliability. As indicated previously, this method did not allow for blinding of the rater to their previous measurements. We feel that calculation of MDC’s for intra-rater reliability would not provide any meaningful information on the reliability of clinical measurements performed by the same therapist across different days, and have therefore chosen not to include this information in the revised manuscript.**

Discussion

13. Paragraph 2, Lines 1: I suggest deleting the first half of this sentence as it undersells your results and you have already acknowledged that some of the measures had poor agreement in paragraph 1.

> The first half of this sentence has been deleted as suggested.

14. Paragraph 2, 

a. Lines 4-6: Rephrase the sentence starting "Similar to our results..." so that it is clearer that the MDC values ranging from 10-19 degrees refers to results from previous studies.

> The sentence has been edited to reflect that the MDC values from 10-19 refer to previous literature.

b. Line 5: sentence starting "Together these findings": I think this could be mis-interpreted as meaning a change #10 degrees in any direction would be clinically significant when in fact your results show MDCs ranging from 9-21 degrees depending on the direction. Please address this sentence so that it more accurately reflects your data.

> This sentence was revised to clarify that 10 degrees does not indicate clinically significant change for all movement directions.

15. Paragraph 2, Line 12: What is different/better about your study versus previous studies in terms of methodology? Clarify what you mean about it having 'well-defined methodology', particularly observing queries relating to blinding and randomisation in the methods section above.
This sentence has been revised to clarify the methodological details which were missing from previous studies, and have been reported in the present study. We have also clarified that the clinical procedures used to assess cervical muscle strength are more clearly described in the present study (see Figure 1) compared to previous literature. Blinding and randomisation issues have been addressed in previous sections as recommended by the reviewer.

16. Paragraph 3, last sentence; why would reliability of neck endurance be poorer in healthy controls? Why would 'a ceiling effect of the endurance test' have affected this?

We agree with the reviewer’s assessment that a ceiling effect would not reduce the reliability of neck endurance measures. This sentence has been removed from the revised manuscript.

17. Paragraph 6, Lines 14-15: Explanation of the mechanism underlying impaired muscle performance in patients with chronic neck pain: While I don't think this is a key issue of this paper (and therefore could be omitted), I do think there are better explanations for impaired muscle performance in people with chronic neck pain than "exact mechanisms have not been determined" e.g. pain inhibition, morphological changes such as fatty infiltration and changes in muscle fibre type. If it is deemed important to include this explanation, the authors should expand on this briefly so that it is more reflective of recent research relating to changes in muscle function in chronic neck pain.

The section on postulated mechanisms underlying impaired muscle performance has been omitted from the revised manuscript as suggested.

18. Paragraph 8: Limitations:

a. Please acknowledge that the population tested were quite young and so data and particularly MDCs may not necessarily be applicable to other populations e.g. an older population.

The section now includes the narrow age range and disability level of the study population as a limitation to generalizability of the results to other populations.

b. Third paragraph: I would delete this as a limitation and maybe include as a possibility for future research, if you want to mention it. You answered your research question and this wasn’t part of it.

As both reviewers considered this concept to be noteworthy in that the present study was not designed to answer questions regarding how cervicothoracic impairments relate to recovery of pain, function, and disability, this paragraph was moved to a new section on future research directions. We believe this revision will highlight the importance of addressing recovery of impairments and their relationship to function and disability in future research.
Reviewer #2:

1. Most, but not all, tests performed showed reasonable inter and intra-observer reliability. As currently presented, the summary statistics ‘mask’ whether one observer consistently found higher or lower results or whether there was an ‘order’ effect so that measures tended to increase at the second examination or decrease. The Table 2 could be re-done so that each observer’s results are compared in order to make this transparent.

   Potential order or rater effects across groups were assessed with a 3-way RM ANOVA. Our results indicated that there was no significant effect of test order (P>0.05 for all outcomes), however, some outcomes did demonstrate a significant effect of rater. This additional analysis has been added to the Methods and Results sections. In order to make the rater effects transparent, group averages (SD) of measurements from each rater have been added to Table 2, with all significant effects indicated by an asterisk.

2. It appears that the main finding is that all the strength and endurance measures and the active range of motion are significantly poorer in the patients with neck pain than those without. Can the authors justify a 30 minute examination to discriminate people with and without neck pain when they can simply ask them?

   Although the data in this study demonstrate that these measures can discriminate between individuals with and without chronic neck pain, the primary goal of the study was not to establish diagnostic criteria for neck pain, but rather to identify measures that differ between individuals with and without neck pain that may serve as potential targets for the clinical assessment and management of neck related impairments. This purpose has been clarified in the revised Introduction and Discussion. Additionally, the term “discriminant validity” has been removed from the manuscript and replaced with terminology that more accurately reflects the study purpose.

3. If the aim is to create a reliable method of assessment of neck measurements that can be used to chart progress in response to treatment, would it not be better to correlate these measurements with PAIN levels and functional disability over time and demonstrate that they are performing reliably over time between two observers but also and perhaps more importantly that they are usefully correlating with reduced pain and/or increased functional capacity?

   The authors appreciate the vital importance of pain and functional performance in this clinical population, however, examination of the responsiveness to change and the association between recovery of impairments and function is beyond the scope of the current study design. The present findings help establish the measurement reliability of impairments that systematically differ between individuals with and without neck pain. With this foundation, the logical next step will be to establish the responsiveness to change and functional significance of relevant impairment measures. As recommended by the first reviewer, we have added a new section on Future Directions to emphasize this point.