

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

### ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Validity of increasing the number of motor control tests within a test battery for discrimination of Low Back Pain conditions in people attending a physiotherapy clinic – A case controlled study
<b>AUTHORS</b>	Biele, Christoph; Möller, Dirk; von Piekartz, Harry; Hall, Toby; Ballenberger, Nikolaus

### VERSION 1 – REVIEW

<b>REVIEWER</b>	Prof. Hannu LUomajoki, Head of masetr msk program ZHAW Winterthur Switzerland
<b>REVIEW RETURNED</b>	10-Jul-2019

<b>GENERAL COMMENTS</b>	<p>Thank you for this interesting paper, which has clear clinical implications. I think this work has good aspects and could add to body of knowledge for clinical physiotherapists. However, there are many small notices to make. It is all over quite messy written (Language, spelling, referencing, lots of slips of the pen... you should have checked it properly before submitting.</p> <p>Minor corrections / remarks:</p> <p>References have been done badly... please check all the way. Abstract; Start sentences with a capital letter. Put dots at the end of sentences. Strengths: I wonder whether the patients were really consecutive... (I'll come later back to this) Background Lane 32-33 Costs... 119£ ...is something missing like Billions? Or is that per capita... please clear. Lane 45...Ref Fritz... that is not her alone, but a bunch of others... bad referencing; also in the ref. list only Fritz 2015... wrong. Then lane 48; you name Jull et al... afterward you name a few of the authors... please check your references... and put in the style of the journal. Lane 52. You should not reference people but the content of the paper... Page 4. Lane 7-8 again... check referencing. Lane twelve position should be in plural. Lanes 27-28... now you name a few people instead of et al... be consistent. Lane 40... status of chronicity... is not true Luomajoki et al 2008 did this.</p> <p>Study sample page 5. Lane 15. What does consecutive mean here? Everybody who came in from the door was recruited, or asked or what? How do you manage this in a practice? Explain. Lane 38... spelling error Lane 57 ... no patients involved?? What are then patients with LBP?</p>
-------------------------	--

	<p>Page 6. Lane 53 don't start a sentence with a number (I will come back to tables afterwards)</p> <p>Analysis, page 8, lane 22. Explain youden index more understandably.</p> <p>Results page 9. Lane 30 Explain ODi scores is that out of 50 or is it %?</p> <p>Page 10 Table. I think you should use same tests names like authors earlier to avoid mismatch; Knee extension... should be sitting knee extension. Why only rocking forward but not also backwards?</p> <p>Page 12, lane 41 don't start a sentence with a number.</p> <p>Discussion</p> <p>Lane 13... I wonder whether the journal wants to have personalized language... "We" please check author guidelines.</p> <p>Lanes 40-45... check the naming od the testes... also knee flexion... should it not be prone knee flexion?</p> <p>Page 14, lane 59... you did not check the reliability...how about page 7 lane 24?</p> <p>References: check the whole list... like Fritz 2015 was not alone. Luomajoki 2012 is twice. O'sullivan and ... D.W: ?? Saragiotto 2016 was hardly alone making a cochrane review.</p> <p>Author contributions... All are declared the same... please specify better who did what.</p> <p>Figures are messy:</p> <ul style="list-style-type: none"> <li>- Check spellings, capital letters, omit the DPI</li> </ul> <p>Mc Descriptions:</p> <ul style="list-style-type: none"> <li>- Explain more clearly what was done and the criteria for correct and not correct performances.</li> <li>- Is symptom provocation also positive? Is that a typical sign of movement control problem?</li> <li>- Test 4. Do you really expect -10° Knee extension? If someone has tight hamstring, you call it movement control problems? (I would call it movement restriction)</li> <li>- One leg stance... how do you measure the 2 cm? DO standardise the position some how?</li> <li>- Test 7 side bending How do you check and measure the 2cm? When is this test positive?</li> <li>- Rocking test 8... why only forward?</li> <li>- Test 9 knee flexion... I would call it prone knee bend as in studies before. Why 10° difference makes it a movement control problem?</li> <li>- Test 11 This test has been shown to be unreliable (Luomajoki et al 2007)</li> <li>- Test 11 ... I don't understand the definition and the picture... on wall, where is the wall?</li> </ul>
--	--

<b>REVIEWER</b>	Buse Ozcan Kahraman Dokuz Eylül University, Turkey
<b>REVIEW RETURNED</b>	05-Aug-2019

<b>GENERAL COMMENTS</b>	<p>The manuscript is well written, has important clinical message. However, a few issues, need to be addressed:</p> <p>In abstract, Participants section 'y' was wrong 'ten' should be used as '10' in abstract</p> <p>You should define measurement of physical activity and ODI</p> <p>You should add p values to results section</p>
-------------------------	---

	What was the differences from other studies? Which test did they included? It can be more explain in discussion. References should be organized
--	--

### VERSION 1 – AUTHOR RESPONSE

<b>Reviewer: 1</b>	
Minor corrections / remarks:	
References have been done badly... please check all the way.	We corrected the reference style according to the BMJ Author Hub guidelines .
Abstract; Start sentences with a capital letter. Put dots at the end of sentences.	Thank you for pointing this out. We corrected this throughout the document.
Background Lane 32-33 Costs... 119£ ...is something missing like Billions? Or is that per capita... please clear.	Thank you for pointing this out. We corrected this (\$119-\$238 billion)
Lane 45...Ref Fritz... that is not her alone, but a bunch of others... bad referencing; also, in the ref. list only Fritz 2015... wrong.	We corrected the reference
Then lane 48; you name Jull et al... afterward you name a few of the authors... please check your references... and put in the style of the journal.	We corrected the reference style according to the BMJ Author Hub guidelines.
Lane 52. You should not reference people but the content of the paper...	We corrected the text according to the reviewer's recommendation
Study sample page 5. Lane 15. What does consecutive mean here? Everybody who came in from the door was recruited, or asked or what? How do you manage this in a practice? Explain.	Consecutive people referred with non-specific LBP were assessed for eligibility and enrolled. We explained this as follows: Subjects were recruited from four different private physiotherapy practises in Germany. Consecutive patients referred to the physiotherapy clinics with a doctor's diagnosis of non-specific LBP were assessed and consequently enrolled if they met eligibility criteria. Participants were subdivided into acute (<6 weeks), subacute (6 to 12 weeks) and chronic (> 12 weeks) pain.
Page 4. Lane 7-8 again... check referencing.	We corrected the reference
Lane twelve position should be in plural.	We corrected this mistake..
Lanes 27-28... now you name a few people instead of et al... be consistent.	We corrected the reference
Lane 40... status of chronicity... is not true Luomajoki et al 2008 did this.	We thank the reviewer for their helpful comment. As far as we understand Luomajoki's paper it was shown that patients with chronic LBP presented with more failed tests compared to acute and healthy subjects. However, no study has been conducted that investigated validity specifically including cut-off values for the number of failed tests. Hence we modified the text as follows: "Even though it has been shown in a three group comparison, that patients with chronic LBP demonstrated more failed tests, the accuracy and validity of tests in more than two groups (for example different levels of impairment, severity, and chronicity) has not been evaluated in detail (14)."
Lane 38... spelling error	We are not sure what is referred to here, but we

	have checked the manuscript for spelling error.
Lane 57 ... no patients involved?? What are then patients with LBP?	No patients were involved in the development or design of the study. This is a required statement from the journal.
Page 6. Lane 53 don't start a sentence with a number (I will come back to tables afterwards)	We corrected the text.
Analysis, page 8, lane 22. Explain youden index more understandably.	According to the reviewer's recommendation we added the sentence: "The Youden-index takes both sensitivity and specificity into account and ranges between 0 (correct classification only due to chance) and 1 (100 % correct classification)."
Results page 9. Lane 30 Explain ODi scores is that out of 50 or is it %?	We added further information to the document to explain the ODI score "absolute ODI score (out of 50)"
Page 10 Table. I think you should use same tests names like authors earlier to avoid mismatch; Knee extension... should be sitting knee extension. Why only rocking forward but not also backwards?	We changed knee extension in sitting knee extension and knee flexion into prone knee Flexion throughout the whole document.  Thank the reviewer for their question. As we did not want to exceed 11 test items we only used rocking forward. We did not include rocking backwards due to lower reliability values found by Luomajoki (2007) who found a higher ICC of rocking forward (0.68) compared to rocking back (0.57).
Page 12, lane 41 don't start a sentence with a number.	We corrected the text.
Discussion Lane 13... I wonder whether the journal wants to have personalized language... "We" please check author guidelines.	It is our understanding that personalized language is not an issue in the journal. However, we corrected this.
Lanes 40-45... check the naming od the testes... also knee flexion... should it not be prone knee flexion?	We corrected the text.
Page 14, lane 59... you did not check the reliability...how about page 7 lane 24?	We thank the reviewer for drawing attention to this . We added "Only rater reliability of the total test battery was assessed after the training programme in the pilot phase"
References: check the whole list... like Fritz 2015 was not alone. Luomajoki 2012 is twice. O'sullivan and ... D.W: ?? Saragiotto 2016 was hardly alone making a cochrane review.	We corrected the reference list and text.
Author contributions... All are declared the same... please specify better who did what.	We added more detail to the author contributions
Figures are messy: - Check spellings, capital letters, omit the DPI	Thank you. We changed this accordingly
Mc Descriptions:	
- Explain more clearly what was done and the criteria for correct and not correct performances.	The supposed ideal pattern of movement which Hodges calls "The Blueprint" defines the expected movement to pass each motor control test (8). Therefore we did not define criteria to pass a test. We described the single-

	<p>test-movement and defined the parameters to fail. These we divided in two groups: General reasons to fail any of the included tests and additional single-test-parameters to fail. If none of those occurred the test was rated passed. Generally the following parameters were determined to define a failed test: Only clear motor dysfunctions were rated as failed, following the instructions of Luomajoki (2008). It was also assumed that the participants performed the tests for the first time. This led to the consensus that participants should be given multiple testing options per test. If the participants were to improve from test to test, this would be interpreted as motor learning and the test would be assessed positively. Failed tests were assessed as negative as follows:</p> <ul style="list-style-type: none"> <li>- The subject was not able to perform the tests correctly within 3 attempts</li> <li>- Rater had to facilitate the subject to do the test correctly</li> <li>- Lumbar spine moves excessively during performance</li> <li>- Compensatory movements in other body areas</li> <li>- Expected movement does not achieve full range</li> <li>- Symptom provocation</li> </ul> <p>The test was "passed if:</p> <ul style="list-style-type: none"> <li>- The subject was able to perform the tests correctly within 3 attempts</li> <li>- Lumbar spine stayed stable while performing the test</li> <li>- No compensatory movements in other body areas</li> <li>- Expected movement with full range on the test</li> <li>- No symptom provocation</li> </ul>
<p>- Is symptom provocation also positive? Is that a typical sign of movement control problem?</p>	<p>Symptom provocation mandated that the test was stopped according to requirement of the Ethics committee, as repeat testing with pain might have irritated the participants problem. Symptom modification was also based on the inclusion criteria of the test battery of Luomajoki 2007.</p>
<p>- Test 4. Do you really expect -10° Knee extension? If someone has tight hamstring, you call it movement control problems? (I would call it movement restriction)</p>	<p>We apologise for this mistake.. We do not expect - 10° extension. If a patient reached that angle on both sides without any criteria to rate the test failed, the test was passed. If range was less, the side difference could not be more than 10° or it was deemed to have failed the test. We altered the description in the appendix of the manuscript to better explain this.</p>
<p>- Test 7 side bending How do you check and measure the 2cm? When is this test positive?</p>	<p>The patients longest fingertip was marked at the lateral side of the leg in the end-position of the test. The distance to the floor was measured and compared. More than 2 cm side-to-side difference</p>

	was rated as failed. We described this in the manuscript.
- Rocking test 8... why only forward?	We explained this in response to a previous reviewer's comment.
- Test 9 knee flexion... I would call it prone knee bend as in studies before. Why 10° difference makes it a movement control problem?	<p>We thank the reviewer for their comment. We changed the test names according to this proposal throughout the document.</p> <p>A 10° difference was used as we felt it was more visually assessable and might have had less effect on neural tissue of the lumbosacral plexus. We added a reference which supports the influence of knee flexion of neurodynamics.</p> <p>Herrington, L., Bendix, K., Cornwell, C., Fielden, N., &amp; Hankey, K. (2008). What is the normal response to structural differentiation within the slump and straight leg raise tests?. <i>Manual Therapy</i>, 13(4), 289-294.</p>
- Test 11 This test has been shown to be unreliable (Luomajoki et al 2007)	We thank the reviewer for pointing this out. As far as we know the reliability was 0.38 in Luomajoki. However, we found varying evidence about this test: v. Dillen 1998 0,6; White 2002 0,50; Luomajoki 2007 0,38; Enoch 2011 0,94. We included this as it involves rotation of the lower extremity which other tests in the MCT battery do not include.
- Test 11 ... I don't understand the definition and the picture... on wall, where is the wall?	We corrected the photo.
<b>Reviewer: 2</b>	
In abstract, Participants section 'j)' was wrong	We thank the reviewer for pointing this out. It has been corrected.
'ten' should be used as '10' in abstract	We thank the reviewer for pointing this out. It has been corrected.
You should define measurement of physical activity and ODI	According to the reviewer's recommendation we defined physical activity, adding to the paper: "physical activity (of at least moderate intensity with slightly elevated heart rate or breathing such as riding the bike or gardening)" and "ODI (Oswestry Disability Index) score (out of 50)"
You should add p values to results section	Thank you for the comment. In general we agree with the reviewer that the results section should provide p-values as parameters of inference. However, we decided not to present them due to two reasons. Firstly, we were primarily not interested whether our results deviate from the null and secondly, we would have to present a separate p-value for each parameter we calculated. As a consequence, we would have to

	correct the p-values for multiple testing (such as Bonferroni correction) giving us completely deflated p-values and indicating that our results are just due to chance which simply does not reflect reality. Besides which, it is quite usual to present psychometrics/clinimetrics without p-values
What was the differences from other studies? Which test did they included? It can be more explain in discussion.	<p>Thank you very much for the interesting question. We stated in the paper the different reasons for different authors to include different tests in the introduction. In order to emphacize this point we added further information to the discussion:</p> <p>“Furthermore, the findings of this study represent a clear enhancement to the existing body of knowledge as a conclusion can be drawn with respect to an optimal number of tests to included in the test battery. This holds true for a two and three-class categorisation which has not been subject to investigation yet in this field of research.“</p>
References should be organized	The reference style as well as the referencing have been modified throughout the complete document.
<p><b>FORMATTING AMENDMENTS (if any)</b>          Required amendments will be listed here; please include these changes in your revised version:          - Supplementary File          Please re-upload your supplementary files in PDF format.</p>	Many thank the reviewer for their help and we changed this accordingly.

### VERSION 2 – REVIEW

<b>REVIEWER</b>	hannu luomajoki ZHAW, Dep. Health, Technikumstr. 71, 8400 Winterthur, Switzerland
<b>REVIEW RETURNED</b>	21-Oct-2019
<b>GENERAL COMMENTS</b>	I think the paper is now good enough for publication.
<b>REVIEWER</b>	Buse Ozcan Kahraman Dokuz Eylül University, Turkey
<b>REVIEW RETURNED</b>	28-Oct-2019
<b>GENERAL COMMENTS</b>	I think the authors have tried hard to revise the paper. Revised verision of the study was good. In general, I believe this manuscript merits publication.