Reviewer's report

Title: Reliability of isometric subtalar pronator and supinator strength testing

Version: 1  Date: 28 May 2014

Reviewer: Kade Paterson

Reviewer's report:

Title: Reliability of isometric subtalar pronator and supinator strength testing

Summary

Overall, this is an interesting study investigating the intra- and inter-session reliability of maximal subtalar pronator and supinator strength testing, and I praise the authors for their attempt to improve measurement in this area. The study reports that although this measurement is reasonably reliably, particularly within-session, the secondary measures of plantar pressures and surface muscle activity are not. I feel the findings are interesting to researchers, although given the equipment needed to evaluate strength and the provide feedback, clinical application is probably limited at this stage (although perhaps an area for future research?). If the issues below are addressed, primarily concerning the inclusion of soleus, the normalization procedures used and discussion/explanation of their high EMG errors, I feel this manuscript would be a good contribution to the area.

1. Is the question posed original, important and well defined?

The research question is novel however I feel it could be defined better with some refinement of the introduction. Although JFAR does not have a word limit, the introduction could be much more succinct in reviewing the research problem and presenting the research question. For instance, a more readable background could address/be set out as per the following: why do we measure the strength of these muscles? How is it currently being done? What are the problems with current practices? What will this research add?

Major Compulsory Revision:

• Please review and re-structure the background to make it more succinct and logical in presenting the research problem being addressed by this study.

Minor Compulsory Revisions:

• There are too many grammatical errors to list – please have the manuscript reviewed by a proof reader.

• The term strength “diagnostics” should be replaced with “assessment” or “testing” (throughout the manuscript).

• Line 61: Tibialis posterior is primarily an invertor. FHL and FDL do not appear to be spelt in English.

• Lines 86-8: Statistical information should not be included here.
2. Are the data sound and well controlled?

The data is presented well, particularly the boxplots and the tables. The format within the text is easy to follow and the results coherent. There are some interesting findings here which the authors outline very well.

Minor Compulsory Revisions:
• Line 210: Are MDCs of 12% and 15% and LOAs of 18% and 26% considered “substantial reliability”? I think this needs to be tempered somewhat.
• Line 210: Change “subjects exhibit” to “participants exhibited”.
• Line 211: I’m not clear what the % values listed in parentheses are – percentage change in mean from day 1 to 2? This should be explained.
• Where are the RMSE values? These are mentioned in the methods but are not included in the results. Please clarify.
• Again, I think the term “variability” needs to be changed here.
• Line 217: change “not present when subjects perceived the real time force signal” to “not present when participants were provided with real time force FB”.
• Line 222: again, I’m not sure what this “variability” data is here.

3. Is the interpretation (discussion and conclusion) well balanced and supported by the data?

The discussion is generally good, in particular with respect to their main findings concerning the strength measures. I feel it falls down somewhat though when it comes to explaining their findings of poor reliability for the EMG and there is minimal synthesis with previous literature. For example, what have previous studies found using biofeedback for strength testing?

Major Compulsory Revision:
• Page 12: The authors mention their reliability/error values are quite high but give no discussion/justification as to why - this should be provided. Three obvious reasons are (1) the recruitment of other muscles, which would decrease the activation of the studied muscles and increase their variability and consequently reliability, (2) the use of an anatomical STJ axis that is not valid or anatomical (see discussions in Kirby [2001] for instance regarding variations in STJ axis position), and (3) the atypical normalisation technique performed (also see following point regarding this).

Minor Compulsory Revisions:
• Line 236; should be “The principal finding”
• Line 237: Again, are MDCs of 12% and 15% and LOAs of 18% and 26% considered “small measurement errors”?
• Line 244-55: This would be good if some past literature regarding the use of biofeedback was presented and your results synthesized with these.
• Line 256: Variability again.
• Line 260: Where did the 15% come from? My reading of the table is that the value should be 17.9%.

• Line 261-2: The authors state “Increasing variability is present in TA and PL when activated to a lesser extent during supination compared to the pronation task.” TA activity would be expected to increase during a supinatory task.

• Line 265: Are MDC values of 18-30%, and more importantly 32-39%, really acceptable? These are quite high. As per my previous points, the language used regarding the level of reliability and/or acceptability of these values needs to be tempered. And as mentioned in the methods, the use of descriptors/parameters would be helpful.

• Line 282: Again, tib post is primarily a supinator.

• Line 291-2: This line is not clear.

• Line 302-5: Whilst these concluding remarks may be true, it is not justified or qualified in any way. The authors also did not test these injured populations so I would omit them from this sentence – perhaps an area for future research though?

• Line 312 and 314-7: The authors rightly suggest that their normalization method may have contributed to the poor reliability of the EMG data but do not really discuss how to resolve the issue. Perhaps assessing the reliability of their normalization method would have been a better first study? The authors also mention there is no study on measurement error for normalizing shank muscles during isometric tasks. Ball and Scur (2010) have done such a study (and included soleus), and reported the optimal method was by using muscle activity during the squat jump form normalization. This should be addressed in some way.

• Line 328: “more or less stable” is not a suitable scientific or statistical expression. And I would also argue that your MDCs of >30% are not stable at all.

• Line 332-3: Sample size calculation was not an aim of the study and should not be the concluding remark. I would omit entirely.

4. Are the methods appropriate and well described, and are sufficient details provided to allow others to evaluate and/or replicate the work?

The authors have done a great job outlining how the machine works and have covered most of the procedures well. My main criticism is the inclusion of soleus in the data collection and analysis. This muscle is primarily a supinator and the authors do not justify why they have measured activation of this muscle. I think the normalization procedure also requires further explanation/justification.

Major Compulsory Revisions:
• Unless its inclusion can be justified, I think soleus should not be included.
• Line 176-8: I am unfamiliar with this normalization procedure. The authors should either provide a reference for this or justify why they have devised a new procedure and not used one of the many existing normalization processes (see discussion).
Minor Compulsory Revisions:
• What was the range and median time between sessions?
• Line 96: What is meant by “background information”?
• Line 97: “content” should be “consent”.
• The authors use 23° and 42° as the STJ axis position. This is by no means accepted as the actual position of the axis, but perhaps this point is better addressed in the discussion.
• Line 112: Change “m. gastrocnemius” to “the gastrocnemius muscle”.
• Line 127: “inspired” should be “instructed”.
• The section on strength testing does not list how many trials were performed.
• Line 156-8: This sentence is very unclear, please clarify.
• Line 177: Why were normalization procedures conducted before and after testing. And of the 4, was it 2 before and 2 after?
• Statistical analysis:
  o Systematic bias is a significant difference between days 1 and 2 rather than the simply mean change score.
  o It would also be good if the authors can include some descriptors/parameters for which they will decide what is good reliability or otherwise.
  o The authors use the term “variability”. They should be consistent and stick with “reliability” as these are different statistical concepts (here and elsewhere in the manuscript).
5. What are the strengths and weaknesses of the methods?
No additional comments or concerns other than those above.
6. Can the writing, organization, tables and figures be improved?
The writing was quite difficult to follow in places and there were far too many grammatical errors to correct in this review. The manuscript would benefit greatly from the assistance of a proof reader.
7. Do the title and abstract accurately convey what has been found?
The title is concise, informative and piques interest in the study. The abstract is mostly a good summary of the paper.
Minor Compulsory Revisions:
• “muscle strength signals” and “muscular activities”. These are referring to the same thing and a bit confusing when you haven’t read the paper. “Muscle strength biofeedback” may be clearer.
• The final line of the abstract is concerning sample size calculations. As per the discussion, this was not an aim of the study and so should not be a conclusion. You have very interesting results regarding your main research question – a line on this would be a much better conclusion. The poorer reliability of the other
measures is also not listed here. If words permit, a sentence on this would also help.

8. Are there any ethical or competing interests issues you would like to raise?
No ethical concerns or competing interests.

References


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

**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:**
No competing interests to declare