Author's response to reviews

Title: The relative timing of VMO and VL in the aetiology of anterior knee pain: a systematic review and meta-analysis.

Authors:

Rachel Chester Ms (r.chester@uea.ac.uk)
Toby O Smith Mr (toby.smith@nnuh.nhs.uk)
David Sweeting Mr (david.sweeting@nhs.net)
John Dixon Dr (John.Dixon@tees.ac.uk)
Sarah Wood Ms (sarah.wood@nnuh.nhs.uk)
Fujian Song Dr (fujian.song@uea.ac.uk)

Version: 2 Date: 3 October 2007

Author's response to reviews: see over
Response to reviewers comments (Re Manuscript: 1426035432151012 - The relative timing of VMO and VL in anterior knee pain: a systematic review and meta-analysis.)

We thank the reviewers for their constructive and helpful comments which have enabled us improve the quality of this resubmitted article. We have addressed the reviewers comments below.

Reviewer: Lee Herrington

“...the authors should consider their use of mean difference when comparing controls to symptomatic individuals muscle activation timing, though an appropriate tool, the significance of its use needs clearly explaining especially in light of the large between subject variability found in all the papers reported in this review, simplistically they have smoothed out these differences by only comparing the mean value.”

Large between subject variability is a significant factor in the studies reviewed and we agree that presenting mean values alone would be misleading. The standard deviation from the mean and 95% confidence intervals have therefore been included to demonstrate the extent of between subject variability in each study. The forest plots provide a visual representation of the means and corresponding 95% confidence intervals. The authors are unaware of a more suitable alternative which would accurately and explicitly present the extent of between subject variability in each study. We feel that readers should be provided with this information in order to visualise the presence and extent of between subject variability for themselves.

“I would question if it is appropriate to include studies which include insufficient data e.g. the absence of standard deviation which is then either estimated from graphs or expropriated from the data of other sources, this surely should exclude these studies in line with the rigor the authors have used elsewhere and the rationale for their inclusion may need consideration”.

We are concerned that by omitting data estimated from graphs or expropriated from the data of other sources, we may possibly be contributing to reporting bias. Some results (such as those with large standard deviations from the mean) may not present as favourably in their entirety and may therefore only be presented in part (Chan and Altman, 2005; Chan et al, 2004).


We dealt with missing data using methods suggested by Pigott (1994). We have included details in the legend of figures 2-4 indicating for which studies standard deviations have been expropriated from the data of other sources. In addition we have now also indicated in the legend which data has been estimated from graphs. We have also now drawn the readers’ attention to these factors in the text. Where data has been estimated from graphs this has been undertaken independently and agreed by three of the authors for each study.

We have now also completed a sensitivity analysis indicating the impact of missing/estimated data on the results and this has been placed in the results section and discussed in the discussion section. We hope that future studies will report all appropriate data so that such action is unnecessary.

“I feel that the discussion of the clinical implications of these very small and highly variable timing differences could have been expanded further. You pointed out that treatment appears to change activation timing, but failed to mention these improvements are within reported test-retest variability by the Cowan and Crossley group”.

The section on clinical implications has been expanded and the highly variable timing differences, both within and between groups and studies, has been made more explicit.

Crossley and Cowan [5] presented repeatability as standard error of the measurement (AKA typical error) of 6.22ms (95% CI 12.20ms) for the concentric and 5.90ms (95% CI 11.50ms) for the eccentric phase of stair stepping. According to Crossley and Cowan, the intraclass correlation coefficient of 0.91 for ascending and 0.96 for descending stairs demonstrates excellent reliability. We take the view proposed by Hopkins (2000) that it is the SEM or TE that is the best measure of absolute reliability, and feel that the changes reported are within reported test-retest variability by the Cowan and Crossley group.


“Likewise the discussion of the between subject and between mode of testing variability is mostly avoided in the discussion and again this is significant, at the very least the question of why are there timing differences between different modes of contraction for the same muscles requires some consideration. You have shown significant heterogeneity between groups yet are happy with pooled data, this needs some consideration within the discussion as well.”

a) The issue of between subject variability and heterogeneity has been dealt with above, in the clinical relevance section.
b) We thank the reviewer for raising the helpful point that we have not adequately discussed how mode of testing influences variability. We have inserted a section in the clinical relevance section on p 16-17

**Reviewer: Rudolf W Poolman**

2E) Were the criteria used for assessing the validity of the included studies reported?

“...the manuscript would benefit from a detailed description of this critical appraisal presented in the method section and its results in the results section. In the present manuscript the results of the appraisal are presented in the discussion. Please move from discussion [page 9] to results section. Limited info on primary study methodology is presented in the abstract’s results section.”

In the absence of a recognized methodological scoring system for comparative observational studies, a qualitative critical appraisal of each study was undertaken. As stated in the paper each study which met the inclusion criteria was independently assessed by two of four reviewers (DS, RC, SW and TS) and a data extraction form completed. This section has been expanded and the data extraction form has been included as an appendix. In addition to the table outlining population characteristics and procedural details, further details of study design and methodological issues are now presented in an additional table in the results section. These details have also been summarized in the results sections

Having expanded the results section we would prefer the critical appraisal to remain in the discussion section as in previous manuscripts published in this journal (Thomas et al, BMC Musc Dis 2006, 7;74; de Carvalho Leite et al, BMC Musc Dis 2006, 7;78).

More details on primary study methodology, have been provided in the abstract’s results section as suggested.

2F) Was the validity of all studies referred to in the text assessed using appropriate criteria (either in selecting studies for inclusion or in analysing the studies that are cited)?

“Can’t tell. Please provide more info on the validity assessment utilized in the methods section and present its results in the results section”

These factors have been addressed in conjunction and as described in 2E. We have expanded the results and discussion section to allow a fuller analysis of the validity of the studies included.

2H) Were the findings of relevant studies combined appropriately relative to the primary question the review addresses?

“Can’t tell. I am in doubt what is the best way to combine the results. Should the data be pooled? The pooled results suffer from strong heterogeneity. Therefore pooling of results may give flawed conclusions. Perhaps it is better to present the figures without totals or
We used a random effects model to incorporate heterogeneity across studies in meta-analysis. We have purposefully presented the results for individual studies so that readers are able to make their own judgements about the studies. In the presence of a number of small studies with similar patients, we feel a meta-analysis is appropriate. Whilst pooling of data is agreeably a grey area, we feel that readers should be presented with this data. Although there is considerable heterogeneity between subjects, as figures 2-4 illustrate, comparison between subject and control groups indicate at least a trend in delayed VMO recruitment for subjects with AKP. As stated in the abstract and conclusion we indicate that “not all AKP patients demonstrate a VMO-VL dysfunction, and that this is compounded by normal physiological variability in the healthy population. The clinical and therapeutic significance is therefore difficult to assess”.

2I) Were the conclusions made by the author(s) supported by the data and/or analysis reported in the overview?
“Partially. This point strongly correlates to the point discussed above. The authors conclude “a clear trend was demonstrated”. At least this should be rephrased to a trend [delete clear in abstract and discussion]. Furthermore, of the 18 included studies only a very small number had sufficient data facilitating data abstraction for subsequent pooling”

“Clear” has been deleted in both the abstract and discussion. In addition the heterogeneity within groups and consequently the caution which should be adopted when interpreting any trends has been emphasised.

6) Regarding the title it was unclear to me whether or not this was an intervention study. The title would benefit from adding “the aetiology of”: “The relative timing of VMO and VL in the aetiology of anterior knee pain:”

The title has been amended as suggested.

Minor Essential Revisions
It would be helpful to summarize the possible sources of heterogeneity in an additional table.

We have summarized the possible sources of heterogeneity, as helpfully suggested in an additional table.