Reviewer's report

Title: Predicting response to physiotherapy for musculoskeletal shoulder pain: A systematic review

Version: 1 Date: 16 April 2013

Reviewer: Chris Littlewood

Reviewer's report:

This is a relevant review but I think the following points warrant further consideration or clarification:

Major Compulsory Revisions

The argument for the review/ area of study could/ should be enhanced; considering the low relative cost of physiotherapy in relation to the overall cost, surely it is reasonable to expose all those who seek care to this aspect of the treatment pathway considering the probable uncertainty associated with any future predictive model.

The issue of why MSK shoulder pain was selected as a focus for the review as opposed to more specific musculoskeletal shoulder diagnoses/ classifications is not adequately considered and really is central to the potential value of this work. Hence I feel that robust justification should be provided particularly as heterogeneity is highlighted as an issue in the included studies.

Participants – it is unclear how shoulder pain of MSK origin was differentiated from, for example, shoulder/ neck pain. For example, was a physical examination required to determine the source or were other methods accepted, e.g. pain drawings. The reliability and validity of the selected approach should be considered in terms of how it might impact upon the results of the review.

Can the authors clarify why tools previously used in other reviews of prognostic factors were not used? The development of a previously unused and untested quality appraisal tool is justified with reference to study design only.

The choice of tool appears fairly central to the conduct of this review because a cut-off point was selected based upon whether studies met/ reported a certain number of criteria. Also, how was the cut-off point chosen? Currently this appears rather arbitrary.

Results/ summary measures – can the authors explain/ clarify the approach detailed in paragraph 2 in relation to points i and ii?

Presentation of results – can the authors clarify which studies might have suffered from type II error?

Individual quality scores are presented but it is unclear what these mean in terms of risk of bias in the individual studies and generally how the quality of the available data is perceived to affect the results of the review. Essentially quality doesn’t appear to be well integrated into the interpretation of the results.
Presentation of results – can the authors confirm that they omitted the results of studies that did not report the range of quality items required by the tool developed for this review? If this is the case then this should be justified because it seems that a lot of data is unnecessarily omitted. Why not just report this data in the context of it being from low quality studies?

Minor Essential Revisions

Pain neuromodulatory techniques – does this mean techniques that aim to relieve pain? If so I think such plain English should be used.

The use of the word ‘putative’ seems unnecessary and serves to confuse

Background/ 4th paragraph/ 2nd Line – unnecessary ‘and’

Search strategy in the main text reported to January 2012 but 2013 in the abstract. I presume the former is a typo?

Discretionary Revisions

Results – I understand that there is no word limit in the selected publication but generally the reporting of the results seems excessive to the point where readability is significantly compromised and the message lost. Will the authors consider condensing this section and focus on the most important messages? Also see recommendation below.

Results/ summary measures – can the authors clarify why the results weren’t simply summarised by prognostic factor? From my perspective this would improve the reporting of the review.

Results from individual studies – much of this section is already or could be tabulated and seems like unnecessary repetition. This section contributes significantly to what feels like an excessive results section.

The summary of results section is the most readable section of the review and I would recommend refocusing the results section and the review around this section. Also, in tandem with one of the points above, did the authors consider undertaking any sub-group analysis, for example, by diagnosis/ classification? It would also be useful to integrate quality appraisal into this section.

In contrast to the results section, the discussion appears rather short and abrupt. I was expecting to see some consideration of other factors that might contribute to the limited results and possibly some recognition that identification of generalizable prognostic factors might not be feasible beyond what is currently recognised.

Following on from this I feel that further justification of the need for a large cohort study is needed beyond the suggestion that some of the studies might have suffered from type II error.

Conclusion – needs to be more concise and reflective of the main messages rather than further repetition.

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

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.