Author's response to reviews

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

Authors:

Rachel Chester Ms (r.chester@uea.ac.uk)
Lee Shepstone Professor (l.shepstone@uea.ac.uk)
Helena Daniell Ms (Helena.Daniell@nnuh.nhs.uk)
David Sweeting Mr (david.sweeting@nhs.net)
Jeremy Lewis Dr (jeremy.s.lewis@gmail.com)
Christina Jerosch-Herold Dr (c.jerosch-herold@uea.ac.uk)

Version: 2 Date: 17 May 2013

Author's response to reviews:

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

Authors: Rachel Chester, Lee Shepstone, Helena Daniell, David Sweeting,
Jeremy Lewis, Christina Jerosch Herold

Version: 2 Date: 17 May 2013

Author's response to reviews: see over

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 contribution of an economic justification for the review has been expanded
on pages 5 and 6

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.

This issue now features extensively within the discussion section and is
referenced by the supportive literature - see pages 21-23. It is also referred to in
the recommendations for future research on page 25.

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.

Page 7 states that “Studies which included anatomical regions in addition to the shoulder but did not report results for the shoulder as a distinct anatomical region were excluded”. The differentiation of shoulder/neck pain has been expanded upon on page 11. This issue is also added to the discussion section on page 22.

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.

Thank you for highlighting the need to clarify this

The authors would have preferred to use a tool previously tested and validated in earlier reviews of prognostic factors. However, to the authors’ knowledge there are at present no recognised validated appraisal tools for use in prognostic studies which 1) address a variety of study designs and 2) fitted our needs in terms of assessing the literature with respect to the objective of the review. We suspect that this is common to other reviews of complex interventions.

In addition to literature searches the lead author contacted a number of authors with expertise in prognostic reviews and a number of different appraisal tools were suggested as a basis for the review. These are referenced in the main text on page 9. None of the tools addressed all the clinical considerations which we felt were important.

We therefore utilised specific items from a number of the assessment tools to inform as many questions as possible to best address the objective of our review. Each question is listed in additional file 3 with respect to the tool from which it was sourced or adapted. Six items (K, L, M, N, O, V), were not covered by any of the tools.

The text on page 9, under the heading “Quality Assessment of external validity, risk of bias, and presentation of results” has been expanded to provide more information.

We have also sought to convey more clearly to the reader that the cut-off point for describing the results of each study was not selected based upon whether studies met/reported a certain number of criteria. Our discussion section has
been amended of any inconsistency in this respect (page 23). The section “Presentation of results” on page 13 describes the poor quality of reporting results by seven of the studies in the review and the in part pragmatic basis of our decision.

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

We have edited this paragraph on page 10 in the hope that it provides more clarity.

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

We do not feel that it would be appropriate to add any further detail as this would be supposition.

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.

Within the results section, under each subheading, the results are first presented for those studies which included the highest number of items in our quality assessment. This has now been highlighted within the introduction to the results section on page 15.

Results are not presented for any of the studies which did not include the basic criteria for reporting results (see page 13 of the main text and response to comments made elsewhere within this document).

The items within our quality assessment tool and listed under “transferability of findings”, “potential for bias” and “reporting quality” (see page 8) are discussed prior to the presentation of results for individual studies under the subheadings “study characteristics” on page 10 and “quality assessment” on page 12. It was our intention that by presenting factors related to quality assessment first the reader could view the results in this context.

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?

Thank you for asking for more clarification on this subject; additions have been made to the main text on page 13.

Seven studies omitted details fundamental to reporting of results; all seven
studies omitted random variability and measure of association or differences. The limited results of these seven studies were therefore not presented in our results section (see page 13 of main text).

Justification: There was an absence of adequate detail to report. We feel that providing point estimates in the absence of any indication of random variability or measure of association (or difference) does not add to the knowledge base and may deter from the more detailed and transparent results provided by the other nine 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.

“Pain relieving” has replaced “techniques to reduce pain” within the abstract

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

“Putative” has been removed or replaced with “potential”

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

Thank you, this has been removed

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

Thank you, this has been amended to 2013 in the main text.

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.

During the initial stages of drafting the manuscript significant attempts were made to summarise the results under subheadings for prognostic factors. However there were ~36 categories of prognostic factor identified within the selected studies, many using a range of tools and definitions. This made the manuscript even longer and more repetitious. We therefore presented results according to our pre-determined outcomes. This process led to identification of the prognostic factors summarised on page 19.
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.

This section was necessary in order to present and synthesise results. We agree that this section is long. However the review and topic area were both complex. The studies included within this review presented results in many different ways, and no two studies used the same set of outcome measures and prognostic factors. Although results are tabulated it was still necessary to talk the reader through the tables and place them in context. We felt that this transparency was important.

To assist the reader through this section we have provided an additional introductory paragraph on page 15. This also serves as a signpost for the structure/framework of the 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.

We agree that the summary is the most readable of the results section. However for the reasons outlined above we do not feel that it would be appropriate to change the framework of the results section.

Also, in tandem with one of the points above, did the authors consider undertaking any sub-group analysis, for example, by diagnosis/ classification?

Thank you for highlighting the need to discuss this. We have now added a paragraph to the discussion section (p21/22) to explain why this was limited. There were only two sub-classifications of shoulder pain which included more than one study; the potential for subgroup analysis was therefore limited.

We have added comments to the results section on page 20 to include the results of the limited subgroup analysis which could be undertaken.

It would also be useful to integrate quality appraisal into this section.

Under each outcome subheading results are first presented for those studies which included the highest number of items in our quality assessment. This has been highlighted within the introduction to the results section on page 15. Results are not presented for any of the studies which did not include the basic criteria for reporting results (see page 13 of the main text and response to comments previously made in this document).

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
 generalize​able prognostic factors might not be feasible beyond what is currently
 recognized.

The discussion section has been expanded, in part due to our response to
 reviewer’s comments elsewhere in this document.

Presentation of additional factors that may have contributed to the results are
discussed in the context of “study characteristics” and “quality assessment of
external validity, risk of bias, and presentation of results” prior to presenting the
results for each study. These are referred to again on page 21. We are unsure
what else to add without repetition.

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.

The paragraph on page 23 includes a range of factors which should be
considered for a large cohort study. We are unaware of any additional factors
that may be useful to add here without repetition.

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

The conclusion has been significantly reduced.