Author's response to reviews

Title: Effectiveness of physiotherapy exercise following hip arthroplasty for osteoarthritis: a systematic review of clinical trials.

Authors:

Catherine J Minns Lowe (catherine.minnslowe@noc.anglox.nhs.uk)
Karen L Barker (karen.barker@noc.anglox.nhs.uk)
Michael E Dewey (med@aghmed.fsnet.co.uk)
Catherine M Sackley (c.m.sackley@bham.ac.uk)

Version: 2 Date: 26 March 2009

Author's response to reviews: see over
Thank you very much for your email on 5th March 2009 informing us that our paper has now been peer reviewed and requesting a revised copy of the manuscript by March 26th 2009.

We have carefully considered all the comments given by the five reviewers and we believe we have addressed/answered each point raised. Due to the wide ranging and disparate comments provided by the reviewers it has not been possible to implement all the separate views which were expressed.

We have detailed our point by point response below for your consideration.

**Reviewer 1. Cindy Veenhof.** We are glad this reviewer believes the manuscript to contain an extensive literature review on an interesting and important topic.

1. Selection of Interventions. The reviewer suggests the review should only include interventions commencing within the first three months of surgery. This point may be a reflection of differing practice in different countries. In the UK for example many post operative movement restrictions are not lifted until the three month stage is reached i.e. full rehabilitation is not permitted until after this time and therefore many interventions are deliberately not started until after this time period is completed. We believe therefore that our original inclusion criteria to search for all trials with a post operative physiotherapy exercise intervention, which clearly specifies the time at which the intervention started, is appropriate. No other reviewer has raised this as an issue.

2. Methodological quality of the trials. We are glad the reviewer believes the use, and discussion, of a component approach to be a good means by which to assess the quality of the trials. She believes that the poor quality of the trials should be more emphasised in the manuscript. We believe we have repeatedly stated this; on page 6 we have stated “the quality of the studies evaluated in this review was mixed and generally poor……only two papers were judged of sufficient quality to be included in meta-analyses” on page 7 we again mention the “generally poor quality of the trials” and we repeat this in the discussion (p.14 “the quality of the trials was mixed and generally poor”). Table 2 states which trials were of sufficient quality to be included in explanatory meta-analyses and shows the vast majority were not. No other reviewer has raised this as an issue.

3. The inclusion of meta-analytic summaries. Reviewer 1 believes that the manuscript should not include the meta-analytic summaries given the trials are of low quality but accepts that we have underlined these summaries are only summaries and not exploratory meta-analyses. No other reviewer suggests the removal of these summaries and we do not feel this reviewer has made a compelling case for the removal of these summaries which do follow an established approach (Gleser & Olkin, 1994). In addition the reviewer does not understand why we have stated that meta-analyses could be performed on several trials in Table 2 and then included all trial in the meta-analytic summaries. She assumes we “do not know the difference between explanatory meta-analyses and meta-analytic summaries”; we believe she has misunderstood what we have said. Table 2 states which trials would be of sufficient quality for us to have included in explanatory meta-analyses and shows these have been possible or appropriate. We have stated that differing outcomes prevented this but that is was possible to provide a formal presentation of findings via the use of meta-
analytic summaries. We have inserted an additional word “explanatory” in the text to make this clearer to readers and prevent further misunderstanding.

4. The reviewer mentions that table 2 does not include data for Kaee et al. This is correct. On page 6 it explains that this paper was written up in summary form rather than as a journal and so could not be included in this Table. Table 2 itself also states that Kaee et al is excluded and this is explained again in the key to that Table. No other reviewer has raised this as an issue.

5. In the introduction it does not clearly explain why the authors focus on function, quality of life, range of joint motion, and muscle strength rather than pain which is excluded. In response we have amended the selection section from “Effectiveness outcomes were measures of functional activities of daily living, walking, self report measures of quality of life, muscle strength and range of hip joint motion. As most trials use functional measures rather than specific pain outcomes, we did not include pain as an effectiveness outcome” to “Effectiveness outcomes included in trials were measures of functional activities of daily living, walking, self report measures of quality of life, muscle strength and range of hip joint motion. As most trials use functional measures, which include pain, rather than specific pain outcomes, it was not considered possible to include pain as a separate effectiveness outcome”.

6. Future directions. The reviewer asked us to rewrite the “the temptation to vote count trials” sentence. We have clarified this sentence from “The trials which included out-patient individual or group training following discharge tended to report between group differences with generally negative results [16,19,21] however the temptation to vote count trials should be avoided” to “The trials which included out-patient individual or group training following discharge tended to report between group differences with generally negative results [16,19,21] however the temptation to count up the number of these negative trials, rather than await future high quality trials, should be avoided”.

Reviewer 2. Michael Hurley. We are glad the reviewer believes the search strategy to be excellent and the review to be well designed and well structured.

1. The reviewer’s main comment is that the review “lacks sparkle” and that the authors have not given enough consideration to the presentation of their results. We believe that we have carefully considered the presentation of results, which is why we have included visual meta-analytic summaries of the data. We are reluctant to “jazz” up the findings to appear more important than they are since we believe the representation of the findings to be accurate, have tried to make the English as clear as possible for those readers for whom English is a second language and since no other reviewer has raised the tone, or style, of the review to be an issue in this way. Dr Hurley also states the quality of the written English to be acceptable at the end of his review. Dr Hurley also suggests we include the calculation and presentation of effect sizes. We have discussed this with our statistician Dr Dewey who, with respect to the non statistician reviewer, states that the use of weighted mean differences is methodologically superior to the use of effect sizes in this review.

2. The reviewer wonders whether it may be worth commenting more forcefully on the questionable relevance of some of the outcomes i.e. walking speed is of academic interest. We have considered this again but believe the lack of available research regarding hip replacement outcomes makes this point a premature one. We also disagree that walking speed is purely of academic interest. In our recent knee systematic review (the sister paper to this hip one) we were able to discuss this issue
since there is much more research available regarding the value and meaningfulness of certain outcomes, such as range of motion, following knee replacement [Minns Lowe et al., *BMJ*; **335**:812-815]. The same cannot be said for hip outcomes.

3. Reference citing. The reviewer makes the helpful comment that we could improve the consistency of our referencing style and we have addressed this throughout.

4. On page 9 we have amended “private communication” to “personal communication” as requested.

**Reviewer 3. Martin Steultjens.**

1. The reviewer felt that how study quality was defined was not entirely clear. No other reviewer raised this concern. We believe this comment reflects the wider issue raised in our paper regarding the lack of consensus that exists about how best to summarise the quality of physiotherapy trials. We believe that presenting information regarding key trial components (using the CONSORT statement which identifies key areas) is the most comprehensive approach to date. This approach is not a new one and is used within physiotherapy reviews (For example: Shamley et al.’s Delayed versus immediate exercises following surgery for breast cancer: a systematic review *Breast Cancer Research and Treatment* (2005) 90:263-271). It is obviously not possible to fully detail every decision and discussion in the space of this review manuscript. We disagree that we have given more weight to the non-random nature of treatment allocation than to non blinded outcome assessment since Table two shows that, due to the lack of blinded outcome assessment in the study by Trudelle & Smith, (2004) that only the self report Oxford Hip Scores were considered possible for inclusion into any explanatory meta-analyses since the objective measures were not collected by a non blinded assessor. But such judgements have to consider the type of outcome being measured and the likelihood of serious bias being created. The trial by Trudelle-Jackson for example contained measures such as manual muscle testing (Table 4). The risk of a non blinded assessor subconsciously providing different pressure when using a hand held muscle dynamometer is obviously much greater than someone undergoing a computed tomography scan to obtain muscle cross sectional area (as in Suetta’s trial) for example (Table 4). Suetta’s trial was a well randomised study where non blinding of some outcomes was unlikely to have introduced significant bias into the trial results whereas alternate assignment could have introduced significant bias into trial findings. We intend the final category in the table – is the trial sufficiently robust enough to be included within a meta-analysis – to summarise whether a trial is of high quality or not. We have therefore added the following sentence into the quality assessment section to try and make this clearer for readers “Consideration was paid to the likelihood of serious potential bias being created throughout the assessment of quality decision making processes and this was taken into account when assessing individual trials”.

2. The reviewer would like us to comment on the absence of pain outcomes included within the review studies. I believe we have addressed this by the changes already made in response for Reviewer 1 point 5 since pain is included in the functional outcomes.

3. The reviewer suggests we use the grades of recommendation provided by the Oxford Centre of Evidence-Based medicine. There are many approaches we could have included and which we considered when writing the protocol for our review. None are
widely accepted. We rejected the inclusion of a CEBM grade of recommendation since we believed this held the potential to be highly misleading to readers. The grading would appear artificially high since there are trials but generally poor ones. No other reviewer has raised this issue.

4. We have amended out patient to outpatient as the reviewer requests.

Reviewer 4. Birger Hagen. Again, we are pleased the reviewer believes this to be an original and interesting study.

1. No response necessary.
2. Methods. The reviewer suggests we separate the validity and reporting sections. We have looked at many other published reviews similar to ours and that use our approach. As an exercise we did try to do this; it was very difficult to separate those studies which are poor in quality and those which are poor in reporting – many are both – and the manuscript became truly repetitive and the subsequent reduced readability was significant. No other reviewer suggested this change.
3. Results. The reviewer finds the results extensive and we agree that we have tried to provide extensive results. Regarding the re-analysis suggestions. We do not accept that a re-analysed review would have the potential to determine the absolute effectiveness of exercise with no exercise as the reviewer states; the overall quality of the studies and, as Table 4 shows, the disparity between these studies, prevents this. We also disagree that within group analyses are not useful. The review was designed to identify and include all trials. In many small trials using functional and rehabilitation measures this approach is used and we believe it would be wrong to ignore this data when there is so little data available. Table 2 provides the randomisation data the reviewer requests.

We agree there are many different approaches to protocol design and the analysis strategy and presentation of results in trials. We are also aware of the lack of consensus regarding the optimum approaches to utilise. When we developed our protocol the analysis plans were extensively discussed with many independent experienced reviewers – including Prof Clarke of the Cochrane Collaboration in Oxford. No reviewer of the protocol or other reviewer of this paper make the suggestion to analyse by group comparison rather than by outcome. We agree that these re-analyses could be performed, and if you feel this is essential we would of course do them. From our point of view however, this re-analysis would be extensive and provoke two problems. Firstly we would need to re-apply for further funds to employ statistical time, and secondly, we are concerned that the time required to achieve funding and for the re-analysis to take place would lead to this review becoming significantly dated and less useful before it is published. We already have a list of people waiting for notification of the publication of this review and keen for its publication.

Minor Revisions.
We believe the selection criteria to be sufficiently clear since it follows a frequently used format and no other reviewer states that this requires alteration. We also feel that combining Tables 4 and 5 would create an unwieldy and confusing table and have therefore not done this. We believe we have adhered to the QUOROM statement throughout. Here is our QUOROM checklist:
The reviewer also believes the abstract findings need changing but does not provide any details regarding her problem with these findings and, again, no other reviewer raises these as an issue.
Reviewer 5. V Karatosun.

This reviewer recommended the paper for acceptance without revision and raised no issues or concerns.

Summary.
Thank you once again for considering this response letter and the revised version of this paper. We look forward to your reply,

Yours sincerely,

Catherine Minns Lowe, Karen Barker, Michael Dewey and Catherine Sackley