Author's response to reviews

Title: The effectiveness of the McKenzie method in addition to first-line care for acute low back pain: A randomized controlled trial

Authors:

Luciana AC Machado (lmachado@george.org.au)
Chris G Maher (cmaher@george.org.au)
Rob D Herbert (rherbert@george.org.au)
Helen Clare (clare.ha@bigpond.com)
James H McAuley (jmcauley@george.org.au)

Version: 2 Date: 30 November 2009

Author's response to reviews: see over
Re: Revised manuscript MS 165696647295700 “The effectiveness of a specific exercise program in addition to first-line care for acute low back pain: A randomised controlled trial”

Dear Mr Aulakh,

The authors would like to thank the reviewers for their valuable comments on our randomised clinical trial. We believe that we have addressed their concerns in the revised manuscript. Changes to the previous version are highlighted in light grey in the revised manuscript. Responses to the reviewer’s comments are listed below.

Responses to Reviewer Thomas Ewert

Discretionary Revisions

P4&5 Background
Please provide more information about the mechanisms how the McKenzie exercise program is targeting factors relevant to LBP. Is it to strengthen the muscles, or to improve the blood flow, or to achieve a better posture, etc?
We have addressed the reviewer’s concerns by inserting information on the core components of the McKenzie method and on the main differences from other exercise programs (pages 4 and 5).

Minor Essential Revisions

P2 Abstract and Method / Result section
Methods: you have different primary outcome measures. Since you have multiple tests you should adjust the alpha risk.
This is not routine practice in clinical trials. When we planned the study we specified a p value of 0.05 and we are reluctant to change the analysis after the study is completed. If readers wish to adjust the critical p value they can and they would come to a similar conclusion to our own. Based upon our a priori p value of 0.05 we concluded that the treatment had no appreciable effect on pain. A reader who set a p value of 0.01 would conclude that the treatment had no effect on pain. In both cases the treatment would not be recommended.

P7 First line care
“Although there was no limit to the number of follow-up visits, physicians were instructed to restrict treatment to advice and simple analgesics.” Have the number of follow-up visits been documented? This would be interesting information about
attention patients got. Please, provide information about the number of follow up visits if available.
The number of follow-up visits to primary care physicians was not documented.

P10 Assessment Procedure
Please provide information about the data collection for health care utilization. Were all contacts documented /analyzed or just those in respect to LBP?
Only those contacts in respect to low back pain were documented. The text has been edited to clarify this issue (page 11).

P10 Statistical analysis
How exactly was the intent-to-treat analysis conducted?
When conducting the analyses we included subjects in the treatment group they were allocated to regardless of treatment compliance.

P 11
Information about the adjustment (Table 2) is lacking. For which parameters has been adjusted and why? This is not obvious since this you are reporting an RCT and both groups showed no difference at baseline.
In the footnote to this table we wrote: “Treatment effects are model-based adjusted differences in outcomes between groups. For example with the pain outcome we created a model to estimate pain scores at each time point based upon data for group membership and the pain scores at the other time points where pain was measured. Other covariates were not entered into the model.

P12
“Thirty-seven (53%) subjects in the Exercise Group and 32 (47%) in the First-line Care Group developed persistent low back pain;…” How is “persistent pain” defined? As ongoing, permanent pain since inclusion?
The definition used for persistent low back pain is listed on page 10, under “Secondary Outcome Measures”.

Discussion
A structured paragraph about limitations is missing
A paragraph on the strengths and limitations of the study has been inserted in the Discussion (page 17).

Figure 1
22 patients met the inclusion criteria, but were excluded. Why that? Please provide more information.
The exclusion criteria for the study are listed on pages 6 and 7: “Patients were excluded if they had any of the following: nerve root compromise, “red flags” for serious spinal pathology (e.g. infection, fracture), spinal surgery in the past 6 months, pregnancy, severe cardiovascular or metabolic disease, or inability to read and understand English.”

Responses to Reviewer Julie Fritz

Major Compulsory Revisions

The treatment program being examined in this study should be consistently labeled as “McKenzie Therapy” or something similar that clearly indicates the treatment
approach that is being examined. The term “Specific Exercise” is a broad term that could mean many different things. In addition, the group in this study received more than exercise, they received an approach consistent with McKenzie philosophy, which includes more than exercises. The clarity of the manuscript will be significantly enhanced by using a more precise and descriptive label for this group.

The manuscript has been edited in accordance to the reviewer’s suggestion.

The results of the study for the outcome of pain found a statistically significant difference between groups, however this mean between-group difference was below the threshold the authors’ established to define a minimally important difference in pain (1 point change). It is important to recognize, however, that a minimally important difference threshold is generally approached statistically and conceptually as referring to the change that is considered important at the level of the individual patient, not at the level of mean between-group differences. The reporting of the results would be improved if the authors reported these outcomes as the proportion of patients in each group achieving at least a minimum clinically important level of change, in addition to the mean between-group differences. This approach matches the individual-patient concept of an MCID with the reporting of the results, and is recommended to improve the interpretability of clinical trials (see for example - Guyatt GH, Schunemann HJ. How can quality of life researchers make their work more useful to health workers and their patients? Qual Life Res 2007;16:1097-105). The difference in the proportion of patients in each group attaining at least a minimally important level of change can be tested statistically, and will help the authors to interpret whether there is indeed an important difference between the groups.

We have followed best practice for clinical trials by pre-specifying our analyses in a trial registry and in a published protocol. We chose not to dichotomise continuous outcomes because this approach loses power and the choice of the cut-point can influence the results. As an illustration for pain improvement by day 7 with cut-offs of 1 or 2 for the MCID the exercise treatment had more participants achieving an MCID, when the cut-off for MCID was raised to 3 or 4 GP care was superior, for cut-offs of 5 and 6 exercise was superior. This pattern demonstrates the flaw in this approach.

We think our presentation of result is preferred. Our results are unambiguous; we have produced precise estimates of the effects of treatment and they are trivially small. This treatment should not be recommended for this group of patients.

It does not appear that the authors evaluated the adherence to protocol of either the physical therapists or primary care physicians. This issue should be dealt with more explicitly and acknowledged as a limitation unless some effort in this regard was in fact undertaken. A sensitivity analysis examining the results among patients in the exercise group who were adherent to the activities prescribed by their physical therapist would be of interest.

This approach is not considered good practice in clinical trials because it ensures that all the benefits of randomisation are lost. We would prefer not to undertake this post hoc analysis but would do so if directed by the editor.

Page 5, paragraph 1 – The contention that referral for McKenzie treatment for patients with acute LBP at the initial contact with primary care is “common clinical practice” requires justification and referencing.

The intention of the authors was to illustrate the usual pattern of care seeking of patients with low back pain in Australia, which is to first visit a primary care physician (Walker et al, J Manipulative Physiol Ther 2004). To avoid misinterpretation, the term “as common clinical practice” has been deleted from the Introduction.
While those who support McKenzie methods would endorse the ability of this approach to benefit patients with acute LBP, surveys of practice patterns in various countries do not indicate that McKenzie methods are the predominant treatment approach used by physical therapists, and I am not aware of any data supporting the contention that primary care providers are particularly likely to refer patients explicitly for McKenzie-based therapy.

All surveys cited on page 4 indicate the McKenzie method as a popular treatment for low back pain among physiotherapists. For example, in a survey conducted in the U.S., the McKenzie method was deemed the most useful approach for managing patients with back pain (Battié et al 1994); in a more recent survey conducted in Ireland, the McKenzie method appeared as the second most commonly used physiotherapy treatment: it was only less used than advice and more commonly used than Maitland mobilisations (Gracey et al 2002).

With regards to the last comment, published clinical practice guidelines for the management of low back pain in primary care provide assistance to primary care providers in deciding what treatment they should refer patients to and when. At present, the McKenzie method features in some of these guidelines, such as the guidelines from the American College of Occupation and Environmental Medicine (2005), the Danish Institute for Health Technology Assessment (1999), and the Clinic on Low-back pain in Interdisciplinary Practice (CLIP) guideline (2007).

Page 6, paragraph 1 – please clarify how the presence of nerve root compromise or “red flags” was determined when screening subjects for eligibility. Was this left to the judgment of the primary care provider or was an explicit process for screening used?

An explicit process was used for screening of nerve root compromise and red flags. Primary care practitioners used patient’s history and findings of the clinical examination to complete a standardized form in which the features and risk factors for each of these pathological conditions were clearly described. For example, the explicit criteria used to diagnose a nerve root compromise were the presence of two positive tests out of sensation, power and reflexes for the same spinal nerve root.

Page 10, paragraph 3 – In the Statistical Analysis section, please clarify how missing data points were handled.

The linear mixed models analysis estimates a value based upon available scores for that person, group membership and available scores for all participants. We would also point out that we had <6% missing data.

Page 11, paragraph 2 – The authors indicate that 260 consecutive patients were screened across 31 primary care providers over a 27-month period. This would equate to approximately 1 patient every 3 months per primary care provider. Unless these primary care providers had a rather atypical patient load, it would seem that many potentially eligible patients seen in primary care by these providers were simply not screened. It is understandable that the exact number of potentially eligible patients who were missed would be difficult to ascertain, but this needs to be acknowledged as a potential source of bias in the study. It may not be accurate to consider this screening of “consecutive patients”.

We understand the point that is being made but there is a misunderstanding here. The trial recruitment period was 27 months but not all 31 GPs recruited for 27 months. We began the trial when we had trained the first GP and progressively added more GPs as resources permitted and did not reach the full complement until the second year of the study. We would also note that not all practitioners worked full time.
Figure 1 – Please add the specific criterion on which the subjects were excluded based on not meeting the inclusion criteria (n=57) or meeting the exclusion criteria (n=22). Please add the follow-up at 3 months at which recurrent was determined. The figure has been edited in accordance to the reviewer’s suggestions.

Table 2 – It’s not clear why p-values are listed separately for the 1-week and 3-week time-points separately, while the outcomes of pain, disability and function have only an overall p-value that is reported. Please be consistent with the approach to presenting this information.

The P values for pain, disability and function are for the linear mixed models group x time interaction. As explained in the footnote to the Table we did not construct this model for the GPE outcome and so there are two P values; one for the main effect of treatment at one week and the other for the main effect of treatment at 3 weeks.

Minor Essential Revisions

Page 4, paragraph 2 – In the first sentence, please clarify to whom “these patients” refers.
The text has been edited accordingly.

Page 4, paragraph 2 – I believe it would be more accurate to state “However, the scientific evidence to support the use of these specific exercises is still scarce, particularly for patients with acute LBP”. It seems that the issue is the value in managing patients with acute symptom onset, not necessarily who the referral source may be.
The aim of our study was to evaluate the additional effectiveness of the McKenzie method in patients with acute low back pain first presenting to a primary care physician. It is possible that patients with acute low back pain treated with the McKenzie method under other circumstances (e.g. patients who first seek care from physiotherapist) have different outcomes than those reported here. Therefore, the referral source is relevant for the external validity of our study.

Page 9, paragraph 1 – please clarify the procedures for the daily pain rating – were subjects asked to rate their worst level of pain on each day? The average level of pain for that date?
This has now been clarified on page 9.

Page 9, paragraph 3 – Please clarify the procedures used by the physical therapists to assess participant compliance with the treatment program. It appears that at each physical therapy visit the therapist asked the participant if he or she was performing their prescribed exercises and activities – was the participant asked if they completed this every day (since their last visits), or just if in general they were doing these activities? How was it determined that a participant was compliant at a particular consultation – did it require that the participant did their exercises every day?
At each following visit to the physiotherapist, participants were asked if they were performing the prescribed exercises and if they were keeping the posture correction (i.e. using the lumbar roll) every day since their last visit. To be adherent with the McKenzie program in our trial, participants had to report be performing the exercises and keeping the posture correction in at least 50% of the times they were asked.

Page 15, paragraph 1 – The contention that “it is doubtful that effects superior to those reported here could be obtained by other therapists” is of dubious validity. There is no evidence in the literature that certification by the McKenzie Institute (or other organizations) results in superior outcomes (see for example Resnik and Hart. Phys
Ther 2003: 83:990-1002) It may be accurate to say that it is doubtful that other therapists would have been more likely to deliver the treatment with fidelity to McKenzie principles, but it is not defensible to state that other therapists would be unlikely to obtain superior effects.

The text has been edited according to the reviewer’s suggestion.

Responses to Reviewer Teresa Liu-Ambrose

Major compulsory revisions

In sample size justification, authors need clarify when they expected the 1 unit difference in pain score and the 1.2 unit difference in global perceived effect scale. In turn, those should be the primary outcome measures (e.g., pain at 3 weeks).

We designed the study to have sufficient power to detect the specified differences and it would not matter at what time point the test was conducted because time is not a parameter in sample size estimations.

The primary outcomes were specified a priori “primary outcomes were mean pain over the first week, pain at 1 week, pain at 3 weeks and global perceived effect at 3 weeks”

Who actually randomized the participants?

A research assistant not involved in recruitment, data collection (other than baseline data), or treatment was responsible for randomising patients into study groups.

Clarify in the introduction how this study is distinctively different than that conducted by Cherkin and coworkers. This was not clear when I read the manuscript.

This information is already provided in the introduction. The points of difference are in terms of the patients and the treatment contrast. Cherkin compare McKenzie vs. an educational booklet whereas we are evaluating McKenzie as a supplement to GP care. Our patients are recruited at the first visit to the GP whereas Cherkin recruited patients who had not responded to GP care by 1 week.

Minor Essential Revisions

Abstract: suggest in second last sentence reword “over the first week” to “over the first 7 days” as “over the first week” and “at 1 week” are potentially confusing.

The Abstract and the main text have been edited according to the reviewer’s suggestion.

Please clarify whether adverse events were monitored and if so, how, and the results.

Adverse events have not been monitored in the study.

Please clarify why adherence data were available for only a subset of the study participants.

Physiotherapists were trained prior to the start of the study to assess adherence at each visit using a standardized re-assessment form. However, in some re-assessment forms, the question on adherence was left in blank (missing data).

Discretionary Revisions

Suggest use of subtitles to improve clarity. For example, under outcomes, use subtitles such as: primary outcome measures, secondary outcomes measures, and adherence.

The text has been edited according to the reviewer’s suggestion.