Author’s response to reviews

Title: The effectiveness of motorized lumbar traction in the management of lumbo sacral nerve root pain: A pragmatic randomised controlled trial

Authors:

Annette A Harte (aa.harte@ulster.ac.uk)
David G Baxter (physio.dean@otago.ac.uk)
Jacqueline H Gracey (jh.gracey@ulster.ac.uk)

Version: 2 Date: 16 July 2007

Author's response to reviews:

Response to reviewers

Julie Fritz

1. A major concern about this manuscript is the small sample size. The study is substantially under-powered to detect minimally important differences on the primary outcome measures. The authors portray the study as a "feasibility" study, however they proceed to report inferential statistics examining between-group differences. With the potential for Type II error so high, this could lead to conclusions about the effectiveness of the interventions that may be potentially misleading and erroneous. If the authors truly conceive of this project as a feasibility assessment, the results should be reported in more descriptive, less inferential terms.

Response: Comments accepted and I have now removed inferential statistics and reported findings as descriptive only. Medians and IQR are stated as data were skewed.

2. The authors perform a post-hoc power analysis and suggest that an appropriately-powered trial would not be possible. The key consideration, however, is not how many patients would make the difference estimated by this study reach statistical significance, but how many patients would be required to detect a minimally important difference if one did actually exist. There is a substantial body of literature documenting the MCID for the Roland Morris and VAS that would support appropriate powering of a study such as this. The paragraph on page 11 describing the post hoc power analysis should be removed. If an a priori power calculation was performed it should be described. If not, I recommend removing the between-group inferential statistical analyses as suggested above.

Response: I have removed the sample size calculation and replaced it with one
based on the MCID for the RMDQ which I had completed prior to the study. This has now been inserted into the discussion in planning for another trial.

3. Page 3 - The objectives of the study need to be consistent with the methods actually used. If the authors intend to examine feasibility, reporting inferential between-group statistics should be avoided. Also, appropriate sample size should be determined a priori based on meaningful differences, not post-hoc based on the difference estimated from a very small sample of patients.

Response: The aims have hopefully been more clearly addressed, sample size dealt with (as above) and inferential statistics removed.

4. Table 2, It appears that there may be baseline differences between the groups that are important, even though they did not reach statistical significance. This is not surprising given the small sample size. For example, 25% of the traction group was off work compared to 43% of manual therapy group. 81% of patients in the traction group do not appear to have had prior episodes of LBP compared to 36% in the manual therapy group. 63% of traction group participated in physical activity compared to 36% in manual therapy group. These differences have the potential to introduce bias into the results. The authors do not attempt to adjust for any of these differences in the analysis.

Response: Due to my use of inferential statistics I didn't think I needed to, however on removing these stats from my study I again have highlighted these points within the results and the discussion.

5. Drop-outs and compliance were an important consideration. Examining Fig 2 it appears that 77% of subjects originally randomized were included in the analysis. It does not appear that intention-to-treat principles were used in the analysis. This introduces another important potential source of bias in the results.

Response: Yes this is a fair comment and I have re analysed my results employing intention to treat (last available score method) and also presented 'completers' results for comparison.

Minor Essential Revisions

1. Do not use bulleted lists in the text (page 3-4). Either place this information into a table or describe it in sentence format.

Response: These have now been labelled a - d and I hope this is ok.

2. Page 4, what was the operational definition of a positive SLR test in terms of ROM? Did the symptom reproduction have to occur below 70 degrees? 45 degrees? These two thresholds are both described in the literature. The authors should clarify their definition of a positive test.
Response: SLR was recorded as positive if leg symptoms were reproduced during a SLR test there was no limit to where in the range this should occur. This is something I could consider for another trial.

3. Figure 2 - The authors should detail which of the study’s inclusion/exclusion criteria caused ineligibility among the 70 subjects who failed to qualify for the study.

Response: I don’t have the details to do this, but usually it was that they did not meet the 'nerve root' criteria or else the criteria of 4-12 weeks since onset.

Helen Frost

1. The title should make it clear that the paper describes a feasibility study not imply that it evaluates the effectiveness of lumbar traction.

Response: Title has been altered to address this.

2. The design of the randomised controlled trial is described as single blind but in fact the patients and the care providers are not blind to treatment allocation and therefore the study is investigator blind only.

Response: This has been changed with the term single blind removed from the description in the research design and replaced with 'investigator blind'.

3) Four primary outcome measures are included without references.

Response; References have now been added, this was an omission on my behalf.

4) There needs to be a clear explanation in the methodology of which outcome is used to calculate the sample size and what time point was measured.

Response: I have taken advice from Julie Fritz (the other reviewer), my statistician and co-authors and based my sample size calculation on the MCID score from the RMDQ and hopefully that is now clearly explained within text.

5) The treatment is described well and would be easily replicated but the description of the statistical analysis needs further clarification. Primarily there is no mention of an intention to treat analysis which would be appropriate for this design. The description of the patients that were lost to follow up is confused with those that failed to complete treatment (i.e. non compliers). It appears that seven patients were excluded from the analysis when actually only three failed to return the outcome measures. The decision to include or not include patients in the final analysis should be stated in the protocol and not decided post randomisation.

Response: 3 patients failed to complete treatment and 4 patients failed to
complete various follow up points. I have now reanalysed data including intention
to treat analysis (last available score forward method) and compared this with 'completers analyses'. I have decided following advice from the other reviewers
to use descriptive statistics to describe the findings (due to the small sample size
and the chance of Type 2 error) and I hope this meets with your approval.

6) The recruitment phase was slow and results suggest that the design of the trial
is not a feasible model for ethical and financial reasons. The outcome of the
study was predictable, firstly because the patients had acute and sub-acute pain
that was likely to improve spontaneously over time and secondly because there
was very little difference between the treatment protocols. The most notable
findings of recent larger trials have shown that only small treatment effects are
likely from most conservative treatments for back pain and the addition of traction
to a package of care was very unlikely to demonstrate any treatment gains. The
authors correctly point out that the small sample size, with only 77% follow up,
resulted in an underpowered study. Interpretation of the results is misleading as
the authors make a number of suggestions based on data that are clearly not
statistically robust.

Response: See above response re statistics. Your point is taken and in light of
this and the other reviewer's comments it was decided to use only descriptive
statistics and to more clearly discuss the results to reflect the findings.

6) In the discussion the interpretations of the results of the study are inconsistent
with the aims and design of the protocol. In the first paragraph of the discussion
the authors claim that patients benefited significantly from both treatment
protocols but there was no control group. The trial was designed to assess the
differences between treatments not whether both treatment protocols impacted
on disability or pain. It is likely that the patients simply recovered spontaneously.

Response: I accept your comments here and have now made it clear that the
research question addressed here is the difference between the two groups.
However I had made it clear pg 13 (implications for practice) that even though
both groups appeared to improve it could not be assumed that recovery was
spontaneous.

7) In the conclusion the authors identify sub grouping as the most important
design issue of the study although it was not stated as the main aim in the
methods.

Response: It was not intentional to suggest that this study had identified this
subgroup but that was done in the earlier survey but that one of its strengths was
that it had used an identified sub group and had used recommended parameters
which other studies had not addressed. This has been reworded to hopefully
clear up this misconception.
8) In the conclusion the authors state that the study demonstrated that the addition of traction to a standardised protocol did not give additional improvement. The results may suggest that traction is not effective but this study was underpowered and the conclusion should be that the trial wasn’t feasible. The very large number of patients that would be needed to demonstrate a significant difference (clinically or statistically) between groups would make a further trial expensive and a waste of patient’s time and resources.

Response: I accept these comments and have altered the discussion and conclusion to address these concerns. Due to the small sample size I have now based my sample size calculation on MCID and have found a smaller number required which would suggest this trial is now feasible.

Minor Essential Revisions

1) In the results section four of the figures don’t have a number.
Response: These 4 graphs are referred to under Fig 3, and referred to together within the text. However I have now numbered them separately as 1.3-1.6.

2) The SF-36 physical and mental component scores are baffling (range 0-350?) and do not relate to the scores in the table.
Response: This has been an error that had occurred on uplifting these tables as the original document show a decimal point i.e. it should have read 35.0 as these were Norm based scoring for SF36.

3) It would be informative to present confidence intervals of the differences between groups.
Response: As medians are now being quoted IQR have been used as measures of spread.

Christopher McCarthy

1.Pg 3: State that the patients have been referred to Physiotherapy
Response: This is already stated in line 4 of Methods pg3.

2.Pg 6 Insert references for the outcome measures
Response: Done

3 Pg 8 It is difficult to see why you excluded Mulligan and McKenzie. Approaches often used by Manual therapists.
Response: These were excluded as the earlier survey had not indicated that these were used in conjunction with traction in the management of nerve root
and as the study was looking at a pragmatic trial with the addition of traction it was important to exclude these so that the survey results were reflected in this study.

4 Pg8 Need to state assumptions for parametric testing were tested and met.
Response: Have Removed inferential stats due to small sample size and the high potential for type 2 error.

5. Pg 9 May be worth making a comment in the discussion that the patients received the same amount of attention from the physio regardless of treatment group.
Response: Both groups received the same number of treatments and each treatment was 30 min and this should be clear in the methodology.

6. Figures - Graphs RMDQ, VAS scales etc present the graphs using the full scale of the questionnaire ie 24 and 100 etc.
Response: The graphs appear clearer and would be difficult to read if completed with the full score so I would prefer to leave them as they are.

7. Tables. State the mean differences with 95% CIs.
Response: Not appropriate now as using medians and IQR