Author's response to reviews

Title: A randomised controlled trial of preventive spinal manipulation with and without a home exercise program for patients with chronic neck pain.

Authors:

Johanne Martel (johanne.martel@uqtr.ca)
Claude Dugas (claude.dugas@uqtr.ca)
Jean-Daniel Dubois (jean-daniel.dubois@uqtr.ca)
Martin Descarreaux (martin.descarreaux@uqtr.ca)

Version: 2 Date: 10 September 2010

Author's response to reviews:

Dear Editor in chief,

Please find included a revised version of the manuscript entitled: “A randomised controlled trial of preventive spinal manipulation with and without a home exercise program for patients with chronic neck pain” that we would like to submit to BMC Musculoskeletal Disorders. The manuscript was initially submitted to BMC Medicine but was deemed not to be in the highest priorities of the journal. However, the reviewers and editor of BMC medicine strongly suggested that the revised version of the manuscript be submitted BMC Musculoskeletal Disorders following their suggestions and recommendations.

Therefore we would like to submit the revised version to your journal and since all the reviewers’ comments were thoroughly addressed, we would greatly appreciate if the revision did no go trough a complete round of review with different reviewers.

As you will notice in the revised manuscript, clarifications regarding the study objectives and hypotheses were added. Randomization and concealment description have been improved and each of the study weaknesses and limitations have been described and considered in the discussion section of the manuscript. Tables and figures have also been updated according to the reviewers’ suggestions. Finally, we have also included a point by point response to all comments made by the reviewers.

This manuscript has not been published previously and will not be submitted elsewhere until a decision is made regarding its acceptability for publication in BMC musculoskeletal disorders. All authors acknowledge having read the paper and approved the content presented in this manuscript.

Yours Sincerely,
Martin Descarreaux

RESPONSE TO REVIEWERS

A randomised controlled trial of preventive spinal manipulation with and without a home exercise program for patients with chronic neck pain.

Reviewer 1

Major Compulsory Revisions

Comment 1
In the result section (abstract) it would be better to provide the results of the main objective of the study, namely to compare the effect of three secondary prevention methods than to concentrate on the pre-study period. The outcome before the actual RCT starts is irrelevant or near-irrelevant, in particularly in an abstract. By providing extensive information on these results, the reader might loose focus.

Response: Appropriate changes have been made to the revised version of the manuscript.

Comment 2

Here (abstract), and throughout the manuscript, statistical significance is noted but never was the clinical significance considered. Please relate also to this aspect.

Response: Please refer to the response to comment 7.

Comment 3

The conclusion (abstract) avoids dealing with the results in relation to the objective of the study. Please re-write.

Response: Appropriate changes have been made to the revised version of the manuscript.

Comment 4

The objectives of the study, p.5. second para, should include a clear mention of the control group also in the first part of the explanation. There were three groups, SMT only, SMT plus exercise and the control group. I had to read your objectives text several times to understand it. And what was your purpose exactly: to compare the two treatment groups with each other, two lump the two treatment groups vs. the control group, or to look at differences between each and every of the groups?

Response: The objective of the study was clarified to reflect the nature and design of the study. The study objective was to investigate if the SMT treatment was superior to the no treatment strategy and also if SMT combined to exercise
was superior to SMT alone as previously described in non specific neck pain studies.

Comment 5
I did not find a text on ethical issues, and there should be a register number for an RCT register, which the Journal probably requires in order to accept publication.

Response: Ethics issues are now addressed in the revised version of the manuscript. The Trial registration number [Clinical trials.gov: NCT00566930] was included in the first submission but may not have been submitted to the reviewers!

Comment 6
I have some major problems with the data reporting. First of all, you have used mean group data. However, it has become customary to (also) define a level of clinically acceptable improvement and to count the number of participants who fulfill this criterion. This will result in a number of people in each of the three groups who recovered at and above a certain level, which eventually allows for a numbers needed to treat-analysis.

And

Comment 7
The second problem I have is the "statistical" significance. You will have to relate to the "clinical" significance as well.

Response:
Clinical significance (clinically significant change or minimally important change) are now presented in the results section when available. They are now included the result section describing the Non-randomized symptomatic phases of the trial (first clinical changes observed during the trial). Based on the primary outcome of pain, subjects from each group who "maintained" a clinically low level of pain are now presented in the results section.

Comment 8
Your reporting is very detailed and clear. However, I think that if you want to present the course of events over time, which you did for NDI and the BQ in fig.2, you should do this for all the variables.

Further, this should be done for each of the three study groups, or it should be stated that because there were no significant differences between the groups, you have grouped their data together.

Response: Considering the reviewer’s comment, the figures now illustrates the three treatment arms of the study. The longitudinal data for all variables are presented in table 4 and 5 and adding a graph for each of the variable would probably exceed the journal figure per article limit. However we have included VAS score in the revised version.
Minor Essential Revisions

Comment 1

Introduction p.4, last para: When talking of the results of a previous pilot study on maintenance care, only the good outcome is mentioned (in relation to disability). It would be necessary also to mention that the results were different in relation to pain.

Response: Appropriate changes have been made to the revised version of the manuscript.

Comment 2

The outcome measures were mainly collected via questionnaire. It should be mentioned in the method section that this was done independently of the assessors, making this an objective measure, because - surely - this is how the data were collected? This means that your text on weaknesses of the study in the discussion section can be softened up, as the assessors - blind or not - could not influence this part of the data.

Response: According to reviewers’ 1 and 3 comments, appropriate changes regarding the questionnaire and study limits have been made to the revised version of the manuscript.

Comment 3

Is there any information on compliance in the two treatment groups (i.e. not drop-out data)?

Response: Data on compliance of group 2 and 3 have been added to the revised version of the manuscript.

Comment 4

Admittedly, I am a bit tired, but I do not understand the sentence at the end of para 1 on p.14: "The question remains regarding a possible beneficial effect..." Why should one of the co-interventions have been excluded from the protocol? And what is meant by "co-intervention"?

Response: We agree with the reviewer that this last sentence was confusing and not adding much to the discussion. Therefore it was deleted from to the revised version of the manuscript.

Comment 5

Table 2: I realize that this study was carried out in French-speaking parts of Canada, which probably means that everybody eats a lot of good food. A mean weight in kilos of approximately 150 seems a lot, but considering that the mean weight is about 66 meters, it looks like weight and height have been mixed up and that data are unrelated to over-eating.
Response: French-Canadians do eat a lot of “poutine” (Quebec traditional junkfood), however this could not explain the weight and height reported in the manuscript. These two values were mixed and changes have been made to the revised version of the manuscript.

Discretionary revisions

All discretionary revisions have been considered and appropriate changes have been made to the revised version of the manuscript.

Reviewer 2

Major Compulsory Revisions

Comment 1

Who was in the trial and who was excluded given numbers reported in the abstract are different from the numbers reported in the main text of the manuscript? They should match.

Response: The correct numbers were the ones included in figure 1 and appropriate changes have been made to the revised version of the manuscript.

Comment 2

Randomization: The description of randomization, concealment and blinding is not clear to me and needs to be clarified. From the description provided in the paper, it appears that the allocation was properly concealed from the subjects, scientists and care providers. But the method used to allocate patients to study groups does not appear random at all.

What I take from what is written is that the PI numbered the patients as they were enrolled and wrote each number on a card and then put all the number cards in a single opaque envelope. I take it only the numbers went on the cards and not the names of the patients. If it was only numbers, then what is the importance of the opaque envelope at this stage?

Then, an assistant not involved in the project and blinded to the process drew numbers from the original opaque envelope in groups of three and then placed them sequentially in three separate envelopes? Why was a random number table or a computer not used to apply a random treatment allocation to a study id to generate the randomization process? Using a person to do this manually leaves the process vulnerable to corruption. Sequential is not random – it is predictable. And if they were arranged sequentially in the 3 different envelopes corresponding to the three treatment arms, then that is a form of blocking. Also, if the process was done as I have described above, why were the groups so uneven in size with 36, 33 and 29 persons? Without proper randomization, the study should not be called randomized, but rather controlled.

Response:
We agree that the randomization process needs to be clarified and appropriate changes have been made in the manuscript. Even if a randomization table or program was not specifically used in this study, each participant had an equal probability of being assigned to a given group. Randomization was performed using numbers written on cards and at only one point in time. The simple random assignment used for this study ensures that the observed differences in size or composition are the results of chance and not bias. The group composition could not have been predicted by the assessors, researchers or the patients. Allocation was completed after all eligible patients (108) completed their initial assessment (1 point in time) and was concealed until the beginning of the preventive treatment phase of the trial. Unfortunately, 10 patients dropped out of the trial during the intensive care phase therefore leading to unequal group size.

Appropriate changes were made to the methods and results section regarding concealment and randomization.

Comment 3

Intention to treat analysis implies that the analysis will account for all persons eligible and randomized into groups – that means the 36+33+29 patients should be accounted for in this analysis. If you have no follow-up data for them, then you should be doing some kind of sensitivity analysis with suitable assumptions for them.

Response: As now described in the revised version of the manuscript was performed right after eligibility was confirmed for all patients. However, the research question in this study specifically addressed preventive care as it is widely used and controversial in chiropractic. The symptomatic phase was designed to reduce inter-subject variability in clinical outcomes because chronic neck pain by nature will vary a lot in time and between subjects. Ten subjects were not included in the intention to treat analysis because they dropped out before the preventive phase. Therefore, clinical scores from these patients (obtained at the beginning of the non-randomized phase) would not have been helpful to answer the research question.

Comment 4

The sample size calculation was devised to detect a different of two points on pain VAS between groups with 90% power and 0.05 probability of type 1 error. This gave a required sample size of 35 for each group which was not achieved, so the study did not have adequate sample size as was set out. Also, the specification of a two point difference in VAS seems high given the symptomatic phase of treatment was only able to reduce pain VAS on average by 1.3 points. These things cannot be rectified now, but at least should be acknowledged in the limitations – especially important given the negative findings.

Response: A 20% difference in VAS was chosen for sample size calculation even though a few studies have proposed minimum clinically significant difference of (9-10%). Differences of less than this amount, should therefore not be considered clinically relevant. Other studies showed that 13 mm on a VAS
represents, on average, the minimum change in acute pain that is clinically significant. Hagg et al. (2003) proposed that the minimum clinically significant change in Oswestry disability index was of 10 points. As suggested by the reviewer, these things cannot be rectified now but have been considered in the limitation section of the revised manuscript.

Comment 5
It appears the statistical approach was to analyse each follow-up time point separately with a separate ANCOVA model. Were the assumptions of ANCOVA met in terms of error structure, independence of observations etc.? And why not use some method appropriate for longitudinal data like a random effects model or generalized estimating equations or growth curves?

Response:
We agree that the statistical analyses description was incomplete and confusing. The revised version of the manuscript now includes a thorough description of all analyses.

The ANCOVA model used included a repeated measure analysis with treatment and time as the main factors (3 [groups] X 6 [time intervals]) and gender, age as well as pain improvement as covariates. The analyses were performed for all clinical primary and secondary variables. ANCOVA assumptions were met for all variables. Mauchley’s test for sphericity was used to test for compound symmetry assumption. The assumption was not violated for all variables therefore no correction to the degrees of freedom was applied.

Comment 6
There is considerable variation in baseline characteristics across the study arms as evidenced in table 2. I recommend that the modeling be repeated, adjusting for key variables shown to have some considerable differences - not necessarily just those that are “statistically significant” but also those that show substantial differences – e.g., age with more than 5 years difference in the mean, gender, duration of pain, baseline pain and FABQ-1.

Response: The ANCOVA analyses have now been performed using variables where clinically significant between group differences were noted at baseline. See also previous comment. Pain outcomes were not included because the differences did not reach the threshold of minimally important change.

Comment 7
The description of randomization is not completely clear as mentioned above.

Response: See response to previous comment

Comment 8
It would also help the reader to provide the scales for the different measures used – so for instance, pain VAS can range from 0 to 10 where 10 is worst pain, NDI can range from 0 to xx where xx is worst disability.
Response: Appropriate changes have been made to the revised version of the manuscript.

Comment 9
However, the 2nd paragraph under Study strengths and limitations doesn’t make any sense and does not address the fact that the pre-determined sample size was not achieved. Sample size needs are largely determined by the effect size one wants to find.

Response: Appropriate changes have been made to the revised version of the manuscript.

Comment 10
In the 3rd paragraph under Study strengths and limitations, dropouts are reported that do not match the information provided in figure 1.

Response: See response to previous comment about number of the drop outs.

Comment 11
The abstract mostly describes what was found although the numbers of subjects reported in the abstract differ from the number of subjects reported in the main paper.

Response: The abstract has been modified according to reviewer 1 and 2 suggestions.

Minor Compulsory Revisions

Comment 12
Abstract – Background Change the 2nd sentence to “The objective of this study was” – only one objective is given.

Response: Appropriate changes have been made to the revised version of the manuscript.

Comment 13
Background 1st sentence. It would be a more appropriate to reference the NPTF chapter on prevalence and incidence of neck pain in the general population rather than the assessment chapter you have referred to here. See the article on pages S39-S51 in the Spine supplement.

Response: Appropriate changes have been made to the revised version of the manuscript.

Comment 15
Methods Participants last sentence of the 1st paragraph. Is table 1 a method or a result? Were all these conditions listed a priori or were these the serious pathologies detected on screening?
Response: Appropriate changes have been made to the revised version of the manuscript. Table 1 was deleted and information is now written in the text.

Comment 16
Methods- Interventions 2nd paragraph, last sentence, change “it their personal diary” to “in their personal diary”
Response: Appropriate changes have been made to the revised version of the manuscript.

Comment 17
Methods- Interventions - Spinal manipulation combined with a home exercise program group 2nd paragraph: please change “Participants had to” to “Participants were advised to”. And then in the 3rd paragraph, please change “All participants followed the same” to “All participants were instructed to follow the same”.
Response: Appropriate changes have been made to the revised version of the manuscript.

Reviewer 3
Major Compulsory Revisions
Comment 1A
The authors need to clarify what is was that they were trying to prevent? These patients already had chronic neck pain. Was the goal to prevent more pain, more pain episodes, more disability, etc.
Response: The objectives and hypotheses have been clarified to address comments made the three reviewers. We also hope that clarification about maintenance care (preventive care) in chiropractic have clarified the manuscript.

Comment 1B
The primary outcome measure appears to be patient reported pain, as judged by the sample size calculations and the others (cROM, NDI and BD) may be key secondary, but there does not appear to be any statistical adjustment for multiplicity, so that only pain should be referred to as the primary outcome.
Response: As suggested by the reviewer, pain is now the primary outcome and all other variables are either secondary or explanatory.

Comment 1C
It is suggested that "all 3 groups with NCNP maintained improvment" yet there is no explicit definition given for either clinical definitions or statistical definitions of maintenance, nor is there an apparent examination for these. The use of the word "maintenance" is not appropriate in this context.
Response:

Maintenance care is, at least in a chiropractic a theoretical context, based on very little scientific evidence. It is defined by Axen et al (2008) as: In other words, maintenance care can be described as both an attempt at secondary prevention (preventing further events from occurring) and tertiary prevention (maintaining an incurable condition at an acceptable level). The current study was designed to tackle such concept and evaluate if there was any value or clinical benefit deriving from continuous care (even in asymptomatic phases of a chronic condition). A clear definition of maintenance care (at least from the common chiropractic standpoint) is now included in the revised version of the manuscript.

The term maintained or maintenance as they are confusing were removed from the manuscript and replaced by basic descriptions of the study clinical outcomes.

Comment 2A

What is remarkable about the patient population is that their neck pain at inception was already below 4.0 and in fact, well below 4, at a mean of 3.1 in the SMT alone group. Can the authors explain why patients with slight to mild pain were being treated in the first place? Recent literature on the phenomenon of the PASS - the Patient Acceptable Symptom State suggests that pain values of 3.5 or less are considered major therapeutic successes.

Response: The trial was designed to challenge a common chiropractic clinical assumption that regular care, even during asymptomatic phases, can prevent further pain and disability episodes in several musculoskeletal conditions. As now stated in the revised version of the manuscript, the patient population included in the study (low levels of neck pain) may impact the ability of detecting clinically or statistically significant differences. However, we still believe that our results (or the lack of significant between group differences) raise important issues regarding the clinical relevance of chiropractic preventive care.

Comment 2B

The randomization process is not clear. It appears that each patient got a unique ID, although that is not explicitly stated. Rather it is stated that the PI gave them a number, which could have been a number for the treatment group. It also not clear that the PI assigned a unique ID in the absence of knowledge of their clinical history. It also appears that an assistant potentially drew the unique IDs from one envelop and put them into three separate envelopes. Were these steps separated in time?

Response:

Please refer to comment 2 of reviewer 2.

Comment 2C

It seems as if the patients had variable lengths of time from the end of the open label treatment portion of the trial (symptomatic phase) until the entry into the
randomized portion. Is this correct? was there any drift in pain scores from the end of the first phase and into the second? was it different between groups?

Response:
All patients were recruited at the same time and were included in the open label treatment portion of the trial (symptomatic phase) in the same week or so. Therefore, the length of time from the end of the open label treatment portion of the trial until the entry into the randomized portion, although it varied a little from one patient to another, did not exceed 7 days. We believe it is unlikely that such a short period of time would have a significant impact on the course of treatment of a chronic condition.

Comment 2D
The principal outcomes were patient reported (Pain, NDI, BD) and would be hard to blind in experienced patients. The authors have appeared to dismiss the possibility of using a blinded evaluator to assess a Physician Global Response - why?

Response: All the questionnaires used in this study are not routinely used in clinical setting in the Trois-Rivières area and it is unlikely that French speaking patients would have been previously exposed to these tools. A suggestion to use Physician Global response is now included in the discussion.

Comment 2E
The description of the treatment interventions suggest that the patients randomized to the Control Group received half the physician visits but equal "face time" as those in the other two SMT groups. Why was the number of appointments not the same across all groups?

Response: The request that patient in the control group would have less encounter with the clinician was suggested by the funding body and the ethical review board.

Comment 2F
It appears that the treatment groups were each uniquely managed by one treater - is that correct? if so, then, the effect of the intervention is both the effect of the treater and the techniques applied - they cannot be separated. while not necessarily a limitation, it is not a straightforward comparison of techniques and that has to be acknowledged.

Response: Appropriate changes were made to the revised version of the manuscript. A sentence regarding the specificity of the effect has been added in the discussion of the placebo effect.

Comment 2G
What advice on co-interventions were the patients provided, in terms of what was allowed and what was not? was this advice the same across treatment groups?
Response:

Patients in all three groups were given similar written and verbal instructions about the use of ice or other co-intervention. This information is now included in the manuscript and precision about what was allowed were also added.

Comment 2H

Outcomes - as discussed above, it appears that there was one primary outcome, pain, some key secondary (cROM, NDI, BD) and some exploratory outcomes. Needs to be revised.

Response: see response to comment 1B

Comment 2I

Sample size. It appears that the authors targeted a 2/10 point reduction in the randomized portion. This does not seem realistic for a patient population that started at 3.1-3.8. Why do the authors think this is a reasonable assumption for sample size?

Response: Please refer to response to comment 4 of reviewer 2

Comment 2J

Fundamentally, as the authors state, this trial is about prevention. then it needs to be clear what it was that was being prevented. If it was an increase in pain, to what level? For instance, one might accept that a transition from mild to moderate pain, across a 4.0 threshold, would be one such option. Alternately, the authors might have specified a level of disability on the NDI that they were hoping would be prevented, such as a transition from mild to moderate disability. Or one might consider an increase in pain that was being prevented independently of threshold, which would require the specification of a change that was important. Here the work of Pool et al., Spine 2007 might be useful, since they define change in both a positive and negative sense (improvement and worsening). It is not clear how population mean changes help in the interpretation of the trial, and why then the trial was analyzed in this manner.

Response: À voir avec reviewer 1 ou 2

Comment 2K

Was time considered as a covariate in the analysis, or where there two time points for the trial (mid and end). If there were two time points, was the p-value for significance adjusted to reflect this?

Response: Please refer to response to comment 5 of reviewer 2

Comment 3A

The authors have to acknowledge the low pain scores coming into the trial and how this may have influenced the ability to detect differences between groups.
Response:
Appropriate changes have been made to the limitations section of the revised version of the manuscript.

Comment 4A
The authors state that the lack of uniform size of the treatment groups may have reduced the effect size. They further state that because no difference between groups existed, the difference in the size of the groups is not relevant. On the contrary, the smallish sample size may have contributed to the inability to detect differences, as may have the analytical approach and the low levels of pain at onset. This paragraph should be removed.

Response: Appropriate changes have been made to the revised version of the manuscript.

Comment 5A
The paper *does not* show that there is equivalence between the SMT and combined interventions groups. This is frankly misleading - the authors have not performed an equivalence trial, rather they have shown a lack of superiority of SMT + exercise in a superiority trial, and there are several methodological reasons that this might be true. At a maximum, they have not been able to demonstrate the superiority of SMT + exercise versus SMT alone. This entire paragraph needs to be removed as it is misleading. Based on statistical reasoning alone, they have not been able to demonstrate that SMT, with or without exercise, has greater value than just spending time with the patient, although this could be argued employing this data.

Response: We agree with the reviewer that this study was not designed as an “equivalence trial” and that we only showed that a lack of superiority of the combined approach (which was one of our hypotheses). Appropriate changes have been made to the revised version of the manuscript.

Minor Essential Revisions

Comment 1A
The authors should specify that this was tertiary prevention, to clearly separate this from the primary prevention of neck pain in patients who are at risk, but do not have neck pain, as well as from those patients who have had, but do not currently have neck pain (secondary prevention). That is, these are also areas of important interdisciplinary discussion, but the trial does not address primary or secondary prevention.

Response: please refer to comment 1C

Comment 2A
Not clear what the comparison of interest is in the stated goals. Was it really to compare SMT +/- exercise? Why then was there a control group of "watchful waiting" who only had visits and examinations? Later this group of "watchful
waiting" is characterised as "no treatment" which is not correct, as the authors themselves acknowledge, the provider-patient interaction also has therapeutic potential.

Response: The objectives of the study were clarified according to the reviewers’ comments.

Comment 2A
Why were age and gender in the models, yet not considered for stratification? are the differences in these two variables, post randomization clinically important? no comment has been made on these differences, from the clinical point of view.

Response: Age and gender are now considered in the analyses.

Comment 3A1
The Min- Max and SD should be given for the changes in the non-randomized phase.

Response: As suggested Min Max and SD (already included in the initial version) have been added to the manuscript.

Comment 3A2
It should be acknowledged that the non-randomized phase used before-after paired statistics, which have a higher power for the resulting p-values. These p-values should all be adjusted for the number of comparisons - perhaps not a Bonferroni correction, but the authors should decide what is appropriate and adjust for multiple comparisons.

Response: Data were pooled because all patients received the same treatment. Therefore there was no multiple comparison.

Comment 3B
How should clinicians interpret the changes in the NDI and the BD in the randomized portion? are these changes larger than the least detectible differences of these measures? are they clinically important ?

Response: Please refer to comment 7 by reviewer 1.

Comment 3C
The application of ice was clearly different in the watchful waiting control group. was this specific advice given by the treating clinician to this group?

Response: As for all other co-intervention, participants from the three groups received exactly the same explanation (they were written) regarding ice. See also response to comment 2G.

Comment 3D
Pragmatically, one might believe that the low rates of adherence with exercise reflect "real world" acceptable of these prescriptions. Do the authors believe that SMT alone is just as good as SMT + exercise, based on these results, when considering the exercise adherence levels?

Response: To answer reviewer’s comment, further discussion of adherence level and the role of exercise have been added to the discussion.

Discussion

Comment A
the authors acknowledge that the control group did not worsen as much as expected, then argue that enhanced patient expectation did not play a role. this does not seem to be a particularly credible argument.

Response: As suggested, the argument has been removed from the discussion.

Table and figures

Comment A
Figure 2 needs to show the beginning and end of the open label (treatment phase) and the beginning and end of the randomized trial (preventative phase) clearly.

Response: Figure 2 has been modified according to reviewers’ comments.

Comment B
It seems like this is a combination of all three groups, although this is not clear in the text and in the legend. If it is all three groups combined, it would seem that separating them might be much more interesting.

Response: Data are now presented separately.

Discretionary Revisions

Comment 1
Not clear to me that Table 1, detailing exclusion criteria needs its own table. This could be dealt with in 2 sentences in the methods.

Response: Table 1 has been removed.

Comment 2
Table 2 title is not clear. One presumes that this is at the baseline at the time of inception into the open label portion. if so, this should be stated, so as not to be confusing that is could represent that baseline of the randomized trial.

Response: Clarifications have been added regarding table 2 now table 1.