Reviewer's report

Title: Exploring integrative medicine for back and neck pain - a pragmatic randomised clinical pilot trial

Version: 1 Date: 31 March 2009

Reviewer: Stéphane Poitras

Reviewer's report:

Major Compulsory Revisions

1- The authors state that the study is justified because of a lack of knowledge regarding the efficacy of IM in chronic pain patients, since it has been demonstrated ineffective in acute patients. However, the inclusion criterion is two weeks pain duration. This inclusion criterion does not fit the study's purpose. This did not appear to have a major impact, since most patients suffered from chronic pain. However, the 2 weeks criterion should be justified.

2- It is not clear if the allocation was concealed to the research coordinator. This should be clarified

3- Who assessed patients for their eligibility?

4- I'm a bit puzzled by the power analysis. On page 9, it is stated : "The sample size for achieving a power of 80% in detecting a moderate to large difference between groups in favour for IM care was determined based on a pilot trial hypothesis of disability scores". However, on page 13 "Larger sample sizes may therefore be recommended to adequately power future clinical trials of the developed IM model targeting back/neck pain management."

Also, "Having no previously reported IM clinical trial data for estimating appropriate sample size our hypothesis underlying the estimation appears partly impaired." However, using the MCID of the SF36 and it's SD, the authors can easily calculate the power of the study, according to the number of subjects who completed the study.

5- What was the time interval between baseline data and allocation?

7- The types of conventional care received in the control group should be described. General information was given on page 6, but more detailed information should be provided.

8- On a note related to comment 7, most interventions included in IM are routinely used by PTs, at least here in NA. Since subjects in the control group were allowed to receive PT, it is difficult to assess how much they were different regarding therapy received.

9- The average number of sessions per week appears to be at 1. This seems to
be relatively small and should be justified. Also, the lack of difference between groups could be related to the lack of intensity of sessions per week. This should be discussed.

10- The psychometric proprieties and the justification of many of the outcomes selected have not been provided. Namely:
   - disability of activities in daily living due to back/neck pain, stress and wellbeing measured by 11 points numerical scales (0-10) where 0 indicated no and 10 indicated maximum levels respectively;
   - number of days with back/neck pain over the last two weeks (0-14);
   - self reported use (yes/no) over the last two weeks for prescription and non-prescription analgesics, conventional care and CTs

11- The outcome measures should be classified primary and secondary. I would suggest the SF36 as the primary outcome. This can then be used to calculate the power of the study (comment 4).

11- I do not understand the reason behind: "In order to standardise the administration of the questionnaires patients were matched as far as possible between the randomised groups"

13- The authors did not prohibit the use of CT during the study in the control group. To get a better understanding of possible contamination, the exposition of the control group to CT during the study period should be detailed.

14- The statistical analysis of change in a group is unnecessary. Amount of change in a group can be explained by several factors (placebo, natural recovery). We are more interested in the difference of change between groups. Thus, only the between group comparisons are necessary. (UC vs IM column (baseline and change)). This limits the number of analyses performed. The Bonferroni correction therefore becomes unnecessary, too harsh at the level of significance. A level of 0.05 can be used. This last column shows the lack of difference between the two groups.

15- Why were non-parametric analyses used?

16- Statistical analyses at a level of 0.05 should be performed to assess the differences in baseline characteristics between groups (table 1). At a minimum, a p-value should be provided. There appears to be differences.

Minor Essential Revisions

1- Data on the recruitment of patients (page 6) should be put at the beginning of the results section.

2- Data and analyses on drop-outs (page 6 and 14) should be put in the results section.

Discretionary Revisions
None

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests