Reviewer's report

Title: A Randomized Controlled Trial of the Effectiveness of Mechanical Traction for Sub-Groups of Patients with Low Back Pain: Study Methods and Rationale.

Version: 1 Date: 7 March 2010

Reviewer: Martin Underwood

Reviewer's report:

Thank you for asking me to comment on this protocol. I note that I have been specifically requested to comment on the statistical aspects of the trial. I will just focus on this aspect of the protocol. Please be aware that I am a trialist not a statistician. You may want to also consider a review from a statistician experienced in these types of studies. I also note that I am being asked to comment of the statistical aspects of the trial protocol for publication. Presumably, this trial protocol has already been through scientific review for funding and ethical review. Thus, the way the trial is to be done has already been decided. Thus, any comments I may have on the trial design should not actually affect if the trial is performed and the analyses proposed. This means that even if the study design is fundamentally flawed it is still appropriate that the protocol is published in order that the design is on record ahead of the study being completed.

I have number of concerns about the statistical aspects of this study

1. The origin of the statement that the clinically important difference in the OSW is not given. This clearly needs to be smaller than the clinically important change that many consider to be 10 points or a 30% improvement from baseline. My view is that quite small between group differences can be important at a population level. Thus, I would not reject a SMD of less than 0.47 as not being clinically important and would suggest that a larger study is needed. Considering that many consider Traction to be an ineffectual or possibly harmful treatment then a SMD of 0.47 seems rather optimistic

2. When I calculated the sample size for a mean difference of 6 points in a measure with a SD of 14 with 80% power with significance of 5% I found that 87 participants were required in each group - 174 in total. This is nearly three times the size estimated in the paper

3. The effects of clustering have been recognised but I ma not clear how the standard formulas for this has been applied. My experience is that ICC values for clustering are rarely >0.05 and that they may not be as important as we thought in individually randomised trials. Nevertheless, more information is needed on how various clustering factors have been considered (Site/Clinic/Therapist). The inflation factor proposed might be too generous.

4. I am surprised that a sub-group effect that is larger than the main effect has been postulated. If the sub-group effect is greater than the main effect the
consequence of this is that for the other subgroup that treatment will be 
detrimental. If one went for the more plausible, but still very optimistic effect size 
for the interaction, of 0.47 and sub-gotups were of equal size then a four-fold 
increase in sample size is needed. In any event for the subgroup effect of 0.71 
specified the authors have estimated the increased sample size to be 1.4 but the 
authors have only multiplied their sample size by 1.2. Pedantically 102*1.2=122.4 
not 120

5. Sample size estimate is assuming no loss to follow-up. It would be wise to 
allow for up to 25% loss to follow-up at one year.

6. I note that the analysis for sub-group effects is the three way interaction 
time*subgroup*randomised group. I think that the two-way interaction of 
subgroup*randomised group is the more appropriate analysis. The hazard with 
the three-way interaction is that a false positive result might be obtained 
because, in most studies, change with time irrespective of treatment has the 
largest effect on outcome.

Using the authors’ assumptions, I think that the minimum sample size should be 
as follows

1. Mean difference in change score of 6.0, SD =14 requires data on 87 per group 
   = 174
2. Do not include inflation factor for clustering N=174
3. To show interaction of 0.71 multiply sample size by 1.4 = N=244
4. Allow for 25% loss to follow up. N=325

Thus I think the study needs to be at least 2.7 times larger. This is probably too 
conservative to ensure that important effects are not overlooked since both main 
effect size and size of any interactions are probably too optimistic and no 
allowance for clustering has been included.

I would be delighted if the authors could persuade me that my analysis is 
correct, as like them, I would very much like to be able to work out how to 
define sub-groups of people with back pain who are more likely to benefit from 
different treatments.

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

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests