Author's response to reviews

Title: Efficiency of immediate postoperative inpatient physical therapy following total knee arthroplasty: an RCT

Authors:

Anton F Lenssen (alen@pmzl.azm.nl)
Yvonne H.F Crijns (YYcrij@pmzl.azm.nl)
Eddie M.H. Waltje (ewal@pmzl.azm.nl)
Mike J.A. van Steyn (Mst@sort.azm.nl)
Ruud G.T. Geesink (RGE@sort.azm.nl)
Piet A van den Brandt (pa.vandenBrandt@epid.unimaas.nl)
Rob A de Bie (ra.debie@epid.unimaas.nl)

Version: 2 Date: 22 August 2006

Author's response to reviews: see over
Reviewer’s report
Title: Efficiency of immediate postoperative inpatient physical therapy following total knee arthroplasty: a RCT ISRCTN49658674
Version: Date: 6 21 April 2006
Reviewer: John Licciardone
Reviewer’s report:
General
The authors have adequately addressed most of the previous reviewer comments. The remaining item, however, is of primary concern and involves the sample size computation and statistical power reported in the manuscript. I again recommend that a biostatistician review this aspect of the manuscript.
------------------------------------------------------------------------------------------------------------------------
Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

1. Further clarification is needed in the manuscript to indicate that the flexion (not extension) passive ROM is the outcome that served as the basis for sample size computation. Additionally, some rationale should be provided as to why this particular ROM measurement was selected as the basis for sample size rather than any of the other ROM (primary outcomes) measures.

We chose passive flexion range of motion as basis for sample size computation because this was our primary outcome measure and we had information on possible effects and variance out of other studies ([1, 2]). In research reports passive as well as active ROM is reported and thus we measured both. Most postoperative programs focus on flexion ROM, although we would agree with the reviewer if his opinion were that extension should be of major concern as well. In our opinion it is a common procedure to calculate and present sample size only on one primary outcome measure.

2. There are several concerns with regard to the sample size computation (page 8). First, it states that a 1-sided test was assumed. This is contrary to the conventional use of 2-sided hypothesis testing and would result in under-powering of the study. That said it appears that the sample size of 16 subjects per group actually corresponds to that which would obtain using a 2-sided test. This needs to be reviewed and clarified.

The statement of the 1-sided test was a mistake. In the second review round we consulted an epidemiologist not involved in the study and he questioned whether 1 sided testing would not be sufficient in our study. We have overseen this remark in our second submission; so “1 sided” was left in the text by mistake.

3. In estimating sample size, the authors appear to have overstated the potential benefit of twice daily PT (10° difference vs. daily PT) and understated the amount of variance in the measured outcome (passive ROM â€” flexion). The consequence of these assumptions is substantial under-powering of the study with respect to this selected outcome measure. More importantly, perhaps, the statistical power with respect to the many other outcome measures reported in this manuscript is not known.
When designing the study, the therapists and orthopaedic surgeons, involved discussed what affect would justify the effort of extra PT sessions daily. We had no information on possible effect sizes or minimally important clinical differences out of the literature, since we are unaware of a single similar study in this field. However we had some information on the effects of cpm in this same setting [2]. We felt that 10 degrees would be a result worth the extra PT effort. Of course one can question that decision and if one does one can state that the sample size was too small to detect smaller differences.

The variance that was chosen came out of a study on the clinical effects of CPM, conducted in 2000. Here we found a sd in rom of 10 degrees[2]. The variance we found in this study is slightly higher (11.9 in the experimental group and 12.8 in the control group, 12.0 for the total group). Of course the reviewer is right that when you target a substantially smaller effect and take into account that variance may be different in the new study this study is underpowered. To find an effect of 5 degrees with an SD of 12 would require a study sample of more than 90 patients per group!

So we fully agree with the reviewer that our sample size may be too small to find smaller effects.

The reviewer is again right in his statement that statistical power with respect to the secondary outcome measures is not known. One can estimate what sample sizes are needed to find significant results with the given point estimates of the different outcome measures.

4. The authors should explain and/or cite references for their assumed 10° difference between treatment groups. Given their assumed SD of 10°, they are expecting the twice daily PT intervention to yield an effect size of 1.0 (computed as Cohenâ€™s d). This is indeed a very large (and probably unrealistic) effect to expect and helps to explain why the study is underpowered.

For example, even accepting this unrealistically large effect for twice daily PT, the observed pooled SD of about 15° at six weeks (Table 4) would require about 36 subjects per group (44 subjects per group would be needed based on the authorsâ€™ pre-study maximal attrition scenario).

We believe this comment is similar to comment no. 3, and therefore our reaction is stated above.

-------------------------------------------------------------------------------

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

1. The units of measurement should be clearly indicated for all variables in Tables 2-4.

Adjusted

2. Figure 2: The number analyzed for the experimental intervention should be 21, not 12.

Adjusted
3. Miscellaneous minor formatting, spelling, style issues.

Discretionary Revisions (which the author can choose to ignore)
None.
Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions
An article whose findings are important to those with closely related research interests Level of interest:
Acceptable Quality of written English:
Yes Statistical review:
Declaration of competing interests:
I declare that I have no competing interests.