Reviewer's report

Title: Six-week high-intensity exercise program for middle-aged patients with knee osteoarthritis - a prospective, randomized, and controlled study

Version: 1 Date: 26 November 2004

Reviewer: Marlene Fransen

Reviewer's report:

General
This manuscript needs substantial review as it has major methodological issues, is very confusing to read and presents unsubstantiated conclusions.

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

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

Outcome measures:
1. Drop the 12 month analysis completely, the study is no longer a RCT after 6 months.
2. Need to limit outcome measures, particular in such a small sample. Apart from 3 follow-up times, too many measures and too many tests (paired and unpaired) resulting in a high risk of chance findings (Type 1 error), which may be why some of your results are counter-intuitive.
3. Your main outcome is KOOS, so do not confuse with reporting firstly on the functional performance tests. Also, the functional performance measures should be viewed as secondary outcome measures, not measures of compliance.
4. As your baseline KOOS scores were dissimilar between the groups, probably best to stick to a paired analysis throughout (comparing change scores).
5. Tables are too congested and confusing. All tables would be easier to follow if the allocation groups (exercise, control) were the columns, and the variables were presented as rows (as per Table 1).
Table 1: Drop Total group column (redundant). Drop modified Edworthy scores - not explained. More enlightening to provide baseline KOOS scores to give indication of comparability in symptom severity between groups.
Table 2: Too many numbers - cannot 'see the forest'. Why not average results for left and right limb. This table should be Table 3 (secondary outcomes).
Table 3: Main outcome, so should be Table 2. Drop 12 month results. Stick to within group (paired analysis) (which looks like you did for the p values? between group = paired analysis and within group = independent analysis? Very confusing choice of words. Why are some of the 'between group' p values significant for the 6 week scores when they are not significant for larger differences at baseline e.g. QOL, ??)
Table 4: Stick to reporting one analysis (paired or independent). Drop 12 months. Why not choose the PCS and the MCS instead of the 8 separate domains.
Tables 1-4: Titles for tables should not include your conclusions or comments.

STatistical analysis:
Power analysis should be performed using the main study outcome measure in a similar population. The power analysis reported (RA patients, beta endorphins) is completely irrelevant to the current study. Taking KOOS QOL and even using a clinically significant change (probably unrealistic with this patients population and the short intervention) of 10 points, you will need a sample size of about 100. The likelihood that this study is underpowered for anything less than a very large effect size.
needs to be discussed. You cannot test new hypotheses emerging from a study within the same study! Drop this entire section or call it something else. However, the sample is too small for any meaningful subgroup investigation. If analyses were performed 'intention to treat', how were the missing values dealt with? Results: Second paragraph in 'Compliance with and efficacy of program' does not belong here. Report these secondary outcomes, after the main outcomes. Choose either between-group or within-group. There is complete confusion here - reporting paired and unpaired tests within the same headings etc. Drop section 'Hypothesis generating information'. Sample is too small for dichotomous analysis, possibly why results are uninformative and often contradictory. Discussion: Main message should indicate more clearly the limited possible generalisability of these results. Drop obvious explanations: intervention effective for some but not all (that applies to every intervention!); QOL should be used in the evaluation of OA interventions (already in international consensus guidelines!). Do not generalise results to all patients with OA: this was a select population with severe radiographic disease. Ref 34 did not measure mm joint space narrowing: only K&L grades, with only about 30% with severe radiographic disease. Explain what you mean by contradictory results for ref 33 and 34. Conclusions: Unsubstantiated. It is a very big leap from showing no effectiveness from this short intervention in a select OA population to concluding that therefore 'exercise should thus be individually prescribed'. Other obvious explanations for the lack of effect include: study underpowered; poor choice of content of exercise program (designed for RA patients, program not targetted to OA knee, spending more than 50% of time on abdominal/hip muscle strengthening); substantial proportion of participants had a complete loss of joint space and therefore less responsive to intensive, often high impact, exercise; short duration of program. Really need to emphasis these limitations, otherwise readers may conclude that exercise is not useful for most people with OA knee.

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

------------------------------------------------------------------------
Discretionary Revisions (which the author can choose to ignore)

What next?: Unable to decide on acceptance or rejection until the authors have responded to the major compulsory revisions

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

Quality of written English: Acceptable

Statistical review: No

Declaration of competing interests:
I declare that I have no competing interests.