Reviewer’s report

Title: Efficacy of zoledronic acid for chronic low back pain associated with Modic changes in magnetic resonance imaging

Version: 1
Date: 5 December 2013

Reviewer: Charlotte Leboeuf-Yde

Reviewer’s report:

Review of Koivisto et al: Efficacy of zoledronic acid for chronic low back pain associated with Modic changes in magnetic resonance imaging.

I have followed your instructions by considering these 9 points and dividing my comments into discretionary (DISC), minor essential (MINESS) and major compulsory revisions (MJCOMP)

1. Is the question posed by the authors well defined?
   Yes

2. Are the methods appropriate and well described?
   Yes, overall. However, some comments:
   DISC: No need to say in Methods under Treatment intervention, 2nd sentence that the study was “blinded”. It is not completely explained at this point. As a reader, I immediately started to wonder how it was blinded. An excellent explanation of that comes later, so I suggest remove the word “blinded” the first come it appears and save it for later to avoid early confusion.
   MJCOM: Please describe what your definition of clinically acceptable improvement is, and the rationale for this. It is not acceptable to look at estimates and describe all such estimates that go in the “right” direction as “tendencies” or “trends”. You need a pre hoc “guideline” for this. To pronounce a “trend”, you need to do a test of trend!
   MJCOMP: You are performing at least 40 statistical tests. With p at 0.05, you will probably obtain at least two “falsely” significant values, so either you have to explain this very clearly in your discussion section, or you have to adjust your p-value with the Bonferroni method (0.05 divided by 40) or do something else to accommodate to this fact.
   MJCOMP: In the statistical analysis text, explain what you do with normally and not normally distributed data.

3. Are the data sound?
   Yes, as far as it is possible to judge that without having access to raw data. All looks reasonable though.

4. Does the manuscript adhere to the relevant standards for reporting and data deposition?
Yes, but the Result section is messy and your results are embarrassingly over interpreted. I have some comments.

MJCOMP: Report first on your primary outcome variable. Your research question was if there was a difference in outcome between treatment A and B. Then please report clearly on this, using a clinically relevant cut point as your point of demarcation. Show some courage; It is OK to report that it did not work the way you thought it would!

Thereafter, explain in words, if there were any clinically and statistically significant findings without trying to “fool” the reader with trends and tendencies. Either it is significant or it is not. If not, then just say that the other variables did not achieve significance. We do not want to read about non-significant information. You stated your level of significance in the Methods section, so please stick to it. Also, you have provided exact estimates in tables, so there is no need to reiterate these (uninteresting) findings in the text.

In fact, your only “real” results are the NSAIDs difference of 20% vs 60%.

DISC: You will have to choose if you want to report your results as tables or as graphs, you cannot use both methods for the same data.

MNESS: Your Figs 2 and 3 (was there a Fig 1?) relate to data collected three times. Do not combine these points with lines. These lines falsely make the reader believe that you have continuous data, which you have not. You have three cross-sectional data collections, so to speak. A bar graph is what you need here or you could indicate the mean values with confidence intervals around (if it is confidence intervals you are showing; you forgot to tell us that).

MJCOMP: Suddenly in the Result section you tell us that a 20% improvement in intensity of LBP tended to favour the ZA treatment…etc. Why twenty percent? Please explain (here or previously in text).

5. Are the discussion and conclusions well balanced and adequately supported by the data?

Fairly well balanced, but I have some suggestions. There is lack of discussion on whether the results can be explained by the natural course, regression towards the mean and suchlike.

MJCOMP

You seem fairly hooked on this treatment, which I can understand because it is plausible. However, the small improved estimates that you see sometimes (not always, as you claim), could well have occurred because most patients figured out that they were in the treatment group. Almost all of the real treatment patients had adverse effects but only few in the other group. You need to discuss this possibility, and in particular, you have to excuse yourselves for not having done a check of whether patients guessed which group they belonged to!

MJCOMP You did not do a power calculation. I accept that you did not. Now is the time to make up for this weakness and to calculate how many study subjects you would need in a “proper” study. However, you will need to look at the 30%
level, which is the usual improvement level in studies on LBP. Would you then get a difference between group estimates at all?

6. Are limitations of the work clearly stated?

Some are but the authors very clearly believe strongly in this method of treatment, their interpretation is really very optimistic.

MJCOMP: The major problem, in my view, is the possible lack of naïve study subjects. Because of the adverse effects, that most people in the “real” treatment group experienced, it would be easy for patients to assume they had been "treated", despite an excellent study design where everybody that could be blinded was blinded.

7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?

MINNEX: I dislike reviewers who tell the authors to quote work of the reviewer. But the fact is that you did forget a recent RCT that I co-authored. It dealt with chronic LBP in people with Modic changes (Jensen KR et al) which basically studies the same aspect of Modic changes as you do; namely the quality of the bone. Rest vs. physical activity were the two test groups and the results were that there was no difference between the groups.

8. Do the title and abstract accurately convey what has been found?

The title explains what the study is about, the abstract deals with the results. The result section does not, in my opinion, give a entirely fair picture of the results.

MJCOMP: Clean up the result section in abstract by being hard an honest with your own results.

9. Is the writing acceptable?

Yes, very good.

**Level of interest:** An article of importance in its field

**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.