Author's response to reviews

Title: Pregabalin in fibromyalgia - responder analysis from individual patient data

Authors:

Sebastian Straube (sebastian.straube@googlemail.com)
Sheena Derry (sheena.derry@pru.ox.ac.uk)
R Andrew Moore (andrew.moore@pru.ox.ac.uk)
Jocelyn Paine (jocelynpain@i-paine.org)
Henry J McQuay (henry.mcquay@pru.ox.ac.uk)

Version: 5 Date: 4 May 2010

Author's response to reviews: see over
To the
BioMed Central Editorial Team

MS: 214296498351434 - Pregabalin in fibromyalgia - responder analysis from individual patient data

Dear BioMed Central Editorial Team

In response to the reviewers’ comments we have undertaken a substantial revision of both the manuscript text and the figures and tables of our paper. We have paid particular attention to the wording and added explanatory phrases in several places. Re-designing the figures and tables was likewise done with the aim of improving readability.

New text in the manuscript is in red. A detailed response to the comments raised by the reviewers follows below (also in red):

Referee 1 report:

My initial review of this paper was an attempt to get the authors try to be clearer in their writing, since many things were not clear. Unfortunately, I still can not fully understand what is being done in this paper and why. Because the data set is unique and there is likely to be some interesting information here, I will try again,

The main problem is that I find this article is still hard to understand. Although it may be a result of being too close to the data and being telegraphic in their writing style, it seems that the authors expect the reader to intuit what they are doing, with little or no explanation. What is the primary purpose of the paper? The title suggests it is a responder analysis of multiple studies of pregabalin in fibromyalgia, but much of the paper talks about whether the responder analysis is appropriate for specific types of measurement scales or data. Which is it?

There is no reason why it cannot be both. After all, the IMMPACT team defined response in terms of pain, and importantly in other dimensions as well as pain intensity. It included physical and emotional functioning, as well as global rating of improvements.

In theory, any measurement on any scale could be used for a responder analysis, with a wide range of possibilities of what constitutes a responder. Using change from baseline, with several different levels of response, should allow an assessment of the utility of both the scale, and the level of response. Utility can be assessed by the occurrence of statistical or clinically significant differences between active therapy and placebo for a particular scale,
especially if there appears to be a dose response. The absence of a significant difference between an effective therapy and placebo at all levels of response is an indication that that particular scale lacks utility for measuring response in a particular circumstance.

The particular circumstance of fibromyalgia is an interesting test bed because many different measurements are made using different scales.

We have added a paragraph to the introduction to make this explicit.

In the methods section we are not told now the authors analyse the data to make a conclusion of the appropriateness of the responder analysis. As below, what is the justification for the criteria you choose to use to ?test? the appropriateness of the responder analysis and how do you know it is the analysis and not the measures which are at fault, assuming the criteria applied are reasonable. Certainly, what ever your findings, they will apply only to this data set, and generalizing to other data sets is at the level of a hypothesis only. For the most part the results section tells us that the data is presented in a figure or table, without pointing out the important components we need to focus on to understand what is being shown. In addition, the tables and figure do not have adequate legends or enough further explanation for the reader to understand what is being presented. The discussion section then concludes that the study has demonstrated something about a responder analysis but the reader is left wondering what the discussion is referring too.

The final paragraph of the methods defines this – we thought rather well. However, we do see your point, so have tried to be more explicit.

It is quite possible that there are some interesting findings in this work but at this point I can not tell. A significant re-organization and clarification of the writing is necessary before such a judgment can be made.

Some but not all of the issues include:

1) One issue is the addition of a small amount of interpretation of results and discussion in the results section. This adds to the readers confusion and should really be moved to the discussion section and better explained. For example on the top of page 10, the results are:

?The corresponding NNTs (Figure 2, Table 1) generally increased over
time for all response levels. At 12 weeks, 11 people need to be treated with pregabalin 450 mg daily rather than with placebo for one of them to achieve a moderate benefit of at least 30% pain relief.

The comment: “This may represent either discontinuation, perhaps because of adverse events, of patients previously achieving a response at a given level or a decrease in their magnitude of improvement to below the level in question.” is a discussion point. This interpretation of the results really belongs in the discussion section with a better explanation of what is meant.

In addition, it is not clear to me how either of these two explanations account for the decreased NNT, unless the changes occur more in the treated group than the placebo group. If this is what the authors mean, then it should be testable in their data. Did there a higher discontinuation rate in the treated group than placebo group? Did patients in the pregabalin treated group have a decrease in the magnitude of their improvement over time at a rate higher than in the placebo group?

Moving all the interpretation of the results, currently in the results, to the discussion would likely help to make the paper more easily understood.

We are willing to accommodate this by moving the two instances where we do this to the discussion section, with additional comments and references.

2) Legends are needed for all the figures and tables, so they can be self explanatory.

Legends for Figures are provided in the way the Journal suggests. Legends to Tables should not be necessary, as there are no terms that are not described in the text.

3) Legends or descriptions for the additional tables submitted as an Appendix are need and (again) it would be easier to understand if they were labeled (a, b, c?).
   a. Are the amounts cumulative? I assume so, but this is not clear.
   b. What does “not calculated” mean and why do they occur intermittently. Could it not be calculated?

Headings to the additional files have been provided.

We’re not completely sure about the comments about cumulative amounts. If this refers to percentage of patients with the outcomes, these data are cumulative (at least 15% improvement, at least 30% improvement, at least 50% improvement etc.). This is indicated in the table in the column “Levels”. 
We explain in the methods that NNTs were not calculated where there was no statistical difference.

4) Page 5 The last sentence of the background states that different information is given in this re-analysis, but does not clarify what that difference is. Were the authors striving to provide a more complete analysis of the fibromyalgia data, more clinically relevant analysis, more analytically appropriate, etc? Reading the rest of the paper would be easier if this was clarified up front.

We’re not quite sure what this comment refers to. This paper is a responder analysis based on individual patient data, something substantially different from reports of the individual trials or a meta-analysis of trial reports (that we have previously undertaken, new reference 31). We thought this was obvious from the Background section. Nonetheless, we have undertaken substantial re-wording of this section to make the purpose and scope of our analysis even clearer.

5) Page 8 ?Only data from trials that included a particular pregabalin dose were used for calculations for that dose; pooled placebo from all trials was not used.? I think the author means: Only data from trials that included a particular pregabalin dose were used for calculations for that dose. Only the placebo data from the specific trials which included that specific dose were used in each dosing comparison. This must be clarified

Thank you. The change is made and we have used your words.

6) I do not understand the statement: ?The intention was to analyse data only where there were at least 200 patients in at least two trials? We and many others have published on the importance of having sufficient numbers of events to avoid the random play of chance overpowering the data. We give one highly referenced paper (175 citations), but could give many more.

7) I am not clear what the authors mean by the word ?hypothesized? at the bottom of page 8. The last statement in the methods section seems, rather, to be a statement of the authors? criteria of what constitutes a useful responder analysis, since they go on to use these three
criteria to judge the usefulness of the specific responder analyses. While the criteria may be reasonable, they are not the only possible criteria. The authors should provide a basis for the selection of these criteria; why they are appropriate to judge the usefulness of a responder analysis; and how they differentiate the analysis from issues related to the use of a particular scale in this patient population.

We disagree with the above comment. The criteria are common sense criteria for judging the usefulness or otherwise of responder analyses of different outcomes used in fibromyalgia trials. They are the ones we chose, and, to be fair, we say that they were the ones we could think of. We’ve tried the arguments out on others, and no-one has come up with any better ones. The fun then, for those that can, will be in responding, or doing analyses of their own. After all, when you really look at many of the scales used with a cold and fishy eye, they are not exactly based on exceptionally large or detailed analyses.

8) Page 9: I am afraid I do not understand the statement in the results: ?Over time the number of patients reporting ?any improvement? fell and the number reporting the higher response levels of at least 50% or at least 70% improvement increased, demonstrating that change in recorded pain intensity was a sensitive indicator for a responder analysis.? Exactly how do these specific changes demonstrate that change in recorded pain intensity is a sensitive indicator? And what is meant by indicator? Please clarify.

According to our hypothesis, a good indicator of a sensitive scale in responder analysis will be a dose response. If there is a dose response, as here, our argument is that this scale is has utility. By our hypothesis, such a dose response “indicates” utility (in the same way that litmus turning red indicates the presence of acid).

9) Page 11 states: ?This demonstrates that Patient Global Impression of Change was a sensitive indicator for a responder analysis.? I am not sure what it means to be a ?sensitive indicator? for a ?responder analysis?? Do the author mean that the PGIC works well in a responder analysis? If this is so, what are the criteria used to make this judgement. Again this comment is in the results section and is really a discussion issue.

As above. The criteria are outlined in methods.

10) Page 12 ? The first paragraph starting with ?Most of these demonstrated sensitivity?? is again a discussion point not a result. I
do not understand what the authors mean by demonstrated sensitivity? I believe the authors are talking about the concept of sensitivity to change or responsiveness, which can be defined as the ability to demonstrate a difference when a real difference is present, based on their definition of sensitivity. Responsiveness is affected by both the type of analysis and the measure used. The fact that the difference between groups for the SF-36 is not a problem with the analysis, but rather a well known issue with the SF-36, which is that it is not very responsive in many pain trials. The authors need to clarify what they are trying to convey.

As above. The criteria are outlined in methods. It may well be that SF-36 is not responsive in pain trials, but then some would argue that neither is the difference between pain scores for treatment and placebo. Part of the problem is that this is based on average results, when the underlying distributions, especially for pain, are not normal (i.e. Gaussian) distributions, and treatment group average values are therefore not a particularly good way of assessing anything. We demonstrated this for pain about 15 years ago.

This analysis is different because it is not based on averages. SF-36 isn’t particularly useful in this type of analysis either, but to our knowledge it’s the first time anyone has tried analysing SF-36 in this way.

11) Page 13: An important issue in this paper is the interpretation of the results presented in the discussion of this analysis by the authors: With our approach (responder analysis based on percentage change from baseline) this was not generally the case with less simple outcomes, including fatigue, fibromyalgia impact questionnaire, anxiety, depression, and most domains of SF-36, apart perhaps from vitality. Therefore, responder analysis as performed here is probably not suitable for most of the outcome measures identified in fibromyalgia clinical trials [3].

a. The major problem is that the authors do not provide adequate justification for their definition of the criteria for a good responder analysis?, and

We think we do provide one. It may be somewhat unrefined, but since responder analyses have not been tried in this way as yet, one might expect that, and, indeed, welcome some lack of refinement in case over-refinement omits some important aspect someone else would pick up.

b. Even if we accept these criteria, they do not present any data in their results section to support this. In the tables submitted as an appendix, most of the cells state not calculated?. What does this mean or why was it not calculated. No discussion or explanation is offered anywhere in the paper that I can see.
NNTs were not calculated when statistical significance was not achieved (we now say this in the Methods section). If the analysis can’t pick up that there are differences between groups with different treatments in this type of analysis, it doesn’t argue well for it being useful. That’s the point.

c. If the data do exist and were better presented, how do the authors know that this is an issue with the responder analysis? as opposed to a problem with the measurement scale used to collect the data in this population. Pregabalin has been demonstrated to work for the pain of fibromyalgia, but we do not know if it works for the other symptoms measured by these other measures.

Exactly. But the measurement scales ARE used, while the responder analysis is a new way of looking at them.

12) I have not specific comments on the discussion at this time, since it will depend on what is being shown in the results. Clearly, as the rest of the paper is being revised the discussion section will need to be revised also.

Minor:
1) Page 8; ?Response data were pooled in an intention-to-treat analysis (number of patients randomised and receiving at least one dose of trial drug).?

I think the author means; ?The response data were pooled and used in an intention-to-treat analysis including all randomised patients who received at least one dose of trial drug.?

We are happy to make this change.

2) In the figures, the author?s state that the ?NNT compared with placebo ? Pregabalin xxxmg?

I think the author?s mean the ?NNT for the comparison of placebo with Pregabalin xxxmg treated groups?

These data are now tabulated and presented in an appendix to safe space in the paper.

3) There are lots of other such issues that should be clarified in order to make the paper understandable.
Referee 2 report:

This study is interesting, methodologically sound, and comes from a strong research group that has advanced the field. Responder analysis is a tool to enhance the interpretation of clinical trial findings. In this paper, the authors pool the results of 4 high quality trials of pregabalin vs. placebo for the treatment of pain, sleep, and PGIC among patients with fibromyalgia. The study has several strengths. In my opinion, a weakness of this paper is the way in which some of the data is presented. So much data is presented, that I'm concerned that the main message of the paper may be diluted. My suggestions are simply my opinion to improve the paper.

1. In the tables, several cells regarding the NNT say "not calculated." I think a brief explanation why they were not calculated would be helpful.

Thanks. The convention is not to calculate NNTs when there is no statistical difference. We have added this to Table 1 as a footnote, and added a sentence in the Methods section.

2. Figure 1-- could be condensed since the same pattern exists for all the pregabalin doses.

We have a new Figure 1, showing pregabalin 450 mg only. The present figure becomes an additional file, so that all data for all treatments are available. This could also be in the form of individual graphs if needed, and we are happy to discuss this with the editorial team.

3. Figure 2 is more confusing than useful. The NNTs presented in tabular form is clearer

The data are now in a tabular form and presented as an appendix.

4. Figure 3-- same as Figure 1--could likely be condensed.
We have a new Figure 3, showing pregabalin 450 mg only. The present figure becomes an additional file, so that all data for all treatments are available. Again, this could also be in the form of individual graphs if needed.

5. Same as Figure 2. The NNTs presented in tabular form is clearer

The data are now in a tabular form (and presented as an appendix).

Referee 3 report:

The article summarizes data from clinical trials with pregabalin in the treatment of FMS, published previously. The authors claim new informations from individual responder analyses which were not given by data published in the primary trial reports. Several modifications of the manuscript submitted first have been performed in response to reviewers' comments and questions.

Minor essential revision:

The last paragraph of the "discussion" section should be reviewed, again: The authors postulate, that FMS might be a heterogeneous group of clinical entities with "multifaceted patterns of pain". There is agreement on this. However, it should also be clarified, that different pathways and mechanisms are not clearly correlated with different pain patterns, so far.

We have added a phrase to that effect.

The article is about FMS, but I am not sure, to what extent references 34-39 include FMS, particularly, or just chronic pain. I agree: This can be the same, but not necessarily in all patients, and this should be evaluated a bit more sophisticated. Does ref. 34 really say, that pain pathways in FMS are complicated by increased age? Or do the authors think about premature aging, anticipating the statement about premature gray matter loss in the subsequent sentence?

Chronic pain leads to the loss of gray matter. Actually, things have moved on recently, and we have replaced the previous references 34-36 with a more recent wide ranging review by Irene Tracey and Cathy Bushnell, who argue that chronic pain is a disease. In that review they rehearse the evidence about the complexity of chronic pain, including the point that we were trying to make that chronic pain leads to a loss of gray matter in excess of that due simply to age.
We hope that we have answered the reviewers’ comments to your satisfaction.

With best wishes

Andrew Moore and Sebastian Straube