Author’s response to reviews

Title: Feasibility and cost-effectiveness of internet-based exposure treatment for irritable bowel syndrome in a clinical sample: a randomized controlled trial

Authors:

Brjánn Ljótsson (brjann.ljotsson@ki.se)
Gerhard Andersson (gerhard.andersson@liu.se)
Erik Andersson (erik.m.andersson@ki.se)
Erik Hedman (erik.hedman.2@ki.se)
Perjohan Lindfors (perjohan.lindfors@aleris.se)
Sergej Andréewitch (andreewitch@gmail.com)
Christian Rück (christian.ruck@ki.se)
Nils Lindefors (nils.lindefors@ki.se)

Version: 3 Date: 30 June 2011

Author’s response to reviews: see over
We would like to thank the reviewers for thorough reading of our manuscript and for valuable comments. As detailed below, they have resulted in large changes of our manuscript. We hope that the reviewers will find that we have considered all comments and followed almost all recommendations.

**Reviewer: Yanda van Rood**

**Major Compulsory Revisions**

1. The authors state: "Our aim was to investigate the feasibility and effectiveness of ICBT within regular clinical practice." However, only data on the effectiveness of ICBT are presented. Since, ICBT has already been shown to be more effective for IBS than waiting list, data on the feasibility (and cost effectiveness) of ICBT would be of more interest.

We have now changed the title to “Feasibility and cost-effectiveness of internet-based exposure treatment for irritable bowel syndrome in a clinical sample: a randomized controlled trial”. This means that we have changed the focus of the article from effectiveness to acceptability of the intervention to the patients that were consecutively recruited. We have also included cost-effectiveness data that was gathered as part of the study but not reported in the first submission. Both these changes have resulted in substantial rewriting of the manuscript.

2. Abstract. An abstract is missing.

An abstract has been included in the manuscript.

3. Background. Page 2 "Recent studies that have included extensive self-help material combined with qualified therapist contact through telephone [13], the internet [14], or a reduced number of sessions [15] have demonstrated effects at par with face-to-face treatments." Please include references demonstrating that these interventions are as effective as face-to-face treatments.

We now state how these treatment performed compared to their respective control groups. We have also added references to newly published studies by our group to the same paragraph.

4. "This study is reported in accordance with the CONSORT statement for nonpharmacological trials [20]." However, there is no mention of trial registration. Please include the registration number if the trial has been registered.

The trial registration ID is now reported at the end of the abstract

5. Exclusion criteria. The reasons for including exclusion criteria g and k need to be explicated. Specifically, the authors should explain what makes patients highly unsuitable for ICBT other than not having access to internet, and not reading or write Swedish. In total 7 patients were not suitable for ICBT, i.e. 5% this needs explanation.

Exclusion criteria g has been explained. Only 4 patients were excluded because they were judged to be highly unsuitable for ICBT, not 7. It has been clarified that these were psychological or somatic reasons as judged by the gastroenterologist.

6. By excluding all patients with severe diarrhea, the IBS patient group is not any more representative for IBS patients in general. Patients type of IBS, diarrhea, constipation or
mixed according to the ROME III criteria should be described in table 1. In the discussion this limitation should be described.

We have added the exclusion of patients with severe diarrhea as a limitation of the study in the discussion. Table 1 has been updated with IBS subtypes.

7. The authors violate their own inclusion criteria by including 2 patients who first visited the clinic outside the specified period. These patients should not have been included.

These 2 patients have been removed from all analyses and are not mentioned in the manuscript. N.B. that 1 of these patients dropped out of the study before randomization, thus the number of analyzed patients is only reduced by 1.

8. Please explain how 5 patients could be unreachable if they visited the clinic in the specified period. If patients were phoned after their first visit this should be described in the recruitment procedure.

We have now clarified that some patients were contacted by telephone in the Recruitment section.

9. The following sentence is not clear. "GSA is defined as “the cognitive, affective, and behavioral response stemming from fear of gastrointestinal sensations, symptoms, and the context in which these visceral sensations and symptoms occur” [21] and is proposed to be a maintaining and exacerbating factor in IBS.” GSA is already anxiety, i.e. an affect. The cognitive, affective and behavioral response are part of the fear response. They do not stem from it.

Interesting, we have not reflected upon this somewhat illogical definition of GSA by Labus et al. We have completely rewritten and substantially expanded the role of fear and avoidance in IBS without getting stuck in the exact definition by Labus et al.

11. Symptom reduction and increase of quality of life are presented as primary goals. The authors might want to describe the increase of quality of life in line with the outcome measures as a secondary goal, i.e. if symptoms are reduced the quality of life will increase. "In the long run, acceptance of symptoms instead of control or avoidance behaviors will increase quality of life. Together with the change in GSA this will lead to a decreased burden of IBS symptoms." The primary goal of treatment now seems to be an increase in quality of life, whereas before symptom reduction was the primary aim of treatment. The authors might want to check the text for consistency on this matter.

We have completely rewritten the description of the treatment.

12. "Exposure can be defined as exposing oneself to an aversive stimulus and simultaneously engaging in a behavior that is inconsistent with the emotion that the stimulus elicits. This serves to reduce the strength of the emotional reaction when exposed to the stimulus in the future [25]." The authors give an unusual definition of exposure. What they describe is usually known as systematic desensitization or counterconditioning. In both techniques an aversive stimulus is associated with a new neutral or positive stimulus. A stimulus which elicits an emotion which is not consistent with the emotion elicited by the aversive stimulus. Mindful acceptance is, as is the relaxation response, not compatible with anxiety and can be used to
change the acquired negative valence of the stimulus (i.e. gastrointestinal sensations and symptoms). Exposure with or without response prevention can be used to falsify catastrophic expectations or to extinguish the fear response, depending on the working mechanisms that are supposed to play a role. This form of exposure is indeed an integral part of most CBT treatments. Counterconditioning techniques are much less often explicitly mentioned.

We have clarified that we frame mindful exposure as behavioral flexibility in the presence of aversive stimuli rather than counterconditioning.

13. It is not clear how mindfulness exercises are used in the exposure exercises and what rationale is given to patients to motivate them for treatment. A table in which the content of the sessions is presented would be most welcome.

We have rewritten and expanded the section that explains the treatment components and their rationale. A table has also been added that presents the contents of each of the five treatment steps.

14. Exposure can be very effective, however in clinical practice many patients do not start treatment or drop out during treatment because they find the exercises to challenging. This might be brought forward in the discussion. Exposure exercises need to be tailored to the patient. How was this done? Where there individual goals set and if so, how many patients reached all their therapeutic goals.

We now mention that the challenge of independently working with exposure may have been too difficult for some patients in the discussion. In our implementation of ICBT we have no reliable way of determining whether individual treatment goals were reached or not. Although patients report individual exercises to their online therapist and ask for feedback, there is no actual monitoring of the setting and reaching of goals.

15. Patients were encouraged to send at least one message per week. The intensity of the contact, i.e. how active the patient was in treatment might explain the observed effect. In the results section the authors might want to describe how many contacts the patients on average did seek per week as well as the range. This is at least as interesting as the amount of therapist time spent per week. Here the range is of interest too.

We now provide this data.

16. Treatment consists of different steps and patients only could go to a next step if they finished the last one. This is an important aspect of the treatment and the authors might want to consider to present the statistics (average and range). The same holds for the number of steps in the treatment and the number of patients that finished all treatment steps.

We now present number of patients that reached each step in Table 2.

17. As the authors point out in their discussion a major limitation of the study is the use of a waiting list control condition. To control for the placebo effect which is known to be large in IBS patients the control condition needs to include an activity that is as credible to the patient as the activity in the experimental condition and the time spend on the activity must me comparable to control for attention. Not only do they not control for the placebo effect, but waiting for treatment when one expects to be treated (as is the case in normal clinical care)
may have a negative effect on symptom severity. Indeed patients in the waiting list condition deteriorate on the MADRS-S, GRS-IBS and IBS-QOL. It can not be ruled out that the group by time effect is significant for GRS-IBS and IBS-QOL as a result of a worsening of symptoms of patients in the waiting list condition. The authors might want to include this aspect in the discussion.

The waiting list did show a significant decrease on the IBS-QOL from pre- to post-treatment but dependent t-tests on the pre- to post-treatment change in the ICBT group are also significant on all measures. Although we appreciate the reviewers point we believe that reporting these tests and then discussing the possible role of the waiting list in producing significant interaction effects (a conclusion not supported by the tests) would very much look like post-hoc reasoning leading to a refuted hypothesis that is unlikely to begin with (that relatively small changes in the waiting list would produce the illusion of significance). We believe that it is best to rely on the single p-value of the interaction effect to determine superiority of the treatment as this was our design and the changes in the waiting list were relatively small. However, the reviewer is quite correct in pointing out using a waiting list is a major limitation of our study. We have expanded that discussion and also compare these results to our previous trial where an active control was used.

18. A psychiatric assessment was conducted before randomization. Please explain at what time, by whom and for what reason in the recruitment procedure the psychiatric assessment took place. If it was used to check on exclusion criteria this should be described in the paragraphs on exclusion and inclusion. If the psychiatric assessment was carried out with the aim to collect data, e.g. psychiatric diagnosis, than the results should also be presented.

The psychiatric assessment was conducted as part of our research team’s interest in exploring the relationship between IBS and psychiatric illness and the effect of IBS-directed CBT on that illness, an effort which also includes other studies. We do not believe that these issues are directly related to the aims of this paper and since we do not have a controlled post-treatment psychiatric evaluation (it was only performed at pre-treatment and 12-month follow-up) it would not really make sense to report these data in this paper. Since our decision to include a psychiatric assessment may have affected the willingness to participate in the trial we deemed it important to inform the reader about this procedure. We have now clarified that the psychiatric assessment was not used to exclude any patients.

19. 2 patients were included who were seen outside the time frame reserved for inclusion. The data have to be analyzed again with 131 patients.

We have now removed these patients from all analyses.

20. "Since IBS symptoms are known to vary considerably over time [36], the mean score of four weekly assessments of GRS-IBS was used to get reliable estimates of the participants’ symptom levels at each assessment point." If patients filled in the GRS-IBS at the first visit and every week for three weeks after that before they started ICBT this might be stated in the method section.

This was not the case, the pre-treatment assessment period did not start until a month after the inclusion, at the earliest.

21. The IBS-QOL and the SDS both assess symptom induced disability / interference. Why are
both included in the study? Why did the authors include an instrument for depressive symptoms. Generally IBS patients are not depressive. The scores indicate that most patients indeed were not depressed. Another point of consideration is that the correlation between MADRS-S and physicians’ MADRS is moderate ($r = 0.54$, $p < 0.001$) indicating that MADRS-S is complementary rather than redundant to the MADRS (ref: The self-reported Montgomery-Asberg Depression Rating Scale is a useful evaluative tool in Major Depressive Disorder. Fantino B, Moore N. BMC Psychiatry. 2009 May 27;9:26.). So it can not be used for the diagnosis of depressive disorder. Since no specific hypothesis was formulated concerning the outcome on the MADRS-S and taking into account the above point, the authors might want to consider to eliminate the data on the MADRS.

The reviewer has a good point and in accordance with our reasoning about not reporting the psychiatric assessment since the aim of this article is to assess the feasibility of ICBT for IBS, we have removed the reporting of MADRS-S from the study. However, we do not agree with the reviewer that IBS-QOL and SDS are mutually exclusive. The IBS-QOL is a multi-item IBS-specific measure that includes cognitive and affective aspects of IBS. SDS, on the other hand, is a general measure of symptom-induced disability with only three items. We believe that the SDS gives a valuable impression of the patients’ general disability.

22. Follow-up: the scores on the MADRS-S increased during follow-up. Please present the p-values and other relevant statistics for all follow-up measurements.

We have removed MADRS-S from the manuscript. We have added the test statistics for follow-up analysis of the other measures.

23. In the first sentence of the discussion an other aspect of the study not explicitly mentioned before is highlighted, i.e. that before it was an self-selected sample and now it is a consecutively recruited sample. If this is the most important difference with the study done before it should be mentioned in the background paragraph. How the clinical usefulness might be evaluated has not been defined. This should be done and the results need to be reported.

We now state in the background section that the study included a consecutive sample. We have also removed the phrase “clinical usefulness” from the manuscript and instead focus on the feasibility of the ICBT.

24. “Clinical guidelines recommend that IBS patients should be referred to psychological treatments primarily if they show depressive symptoms, anxiety, or poor coping strategies and wish to undergo such a treatment [7, 48]. However, in this study we deliberately did not let the gastroenterologists select the patients that they deemed would benefit from a psychological treatment.” These points are unrelated. It is unlikely that the selection criteria used by gastroenterologists are the same as the mentioned criteria (depressive symptoms, anxiety, or poor coping strategies and wish to undergo such a treatment).

We have removed the reasoning about the selection criteria applied by the gastroenterologist.

25. On page 16 the authors give as a possible explanation for the lower effect sizes than they observed in their earlier study the supposedly lower motivation of these patients. The authors might want to objectify this by reporting the number of patients that completed all steps of the treatment.
This is now reported in Table 2 and also referred to in the discussion.

26. *The discussion might be extended by including the negative effect of being on a waiting list, the non-representativeness of the IBS patients, and the unacceptability of exposure exercises as another possible explanation for the high number of drop outs.*

We have added the points about exposure and non-representativeness to the discussion.

27. *The title in itself is contradictory, suggesting that one can have a RCT in clinical practice (without changing clinical practice).*

We have changed the title of the manuscript.

28. Conclusions. *"These results might indicate that a larger proportion of IBS patients than is usually presumed might benefit from psychological treatment." On what grounds do the authors draw this conclusion? If this is indeed a conclusion data supporting this should be presented and discussed in discussion section.*

We believe that we went beyond our data when we drew this conclusion from our data. We have now completely removed this line of argument from the discussion and focus instead on discussing the feasibility of ICBT for clinical IBS patients.

**Minor Essential Revisions**

*Page 4. Power calculation.* The authors state: *"We aimed at having a power of at 85% to detect a standardized mean difference between the active treatment and waiting list of 0.8, which gave a sample of at least 60 participants." Words seem to be missing; i.e. the outcome measures used to calculate the necessary number of patients.*

We expanded the power calculation.

2. *Page 4. Inclusion criteria. Criterion c precedes b and the order might be changed accordingly.*

Changed

3. *Page 5 first sentences breaks between ‘in’ and ‘Stockholm’*

Fixed

*Page 5. "However, if the gastroenterologists judged ....... Etc" is one of the exclusion criteria and has been explained already. Sentence may be deleted.*

Deleted

*4. Page 5. "To ensure that the effects of this basic IBS management ..... , we applied at least a one-month “washout” period ...." The word “washout” is confusing because it refers to medication use. Please clearly explain what was done.*

We have changed the wording of this sentence.
5. Page 6. "However, for 16 eligible patients......" This sentence should follow after the following sentence "... period, and 131 of these were eligible according to their medical record."

Changed accordingly.

6. Page 7. "GSA is defined as "the cognitive, affective, and behavioral response stemming from fear of gastrointestinal sensations, [and] symptoms, and the context in which these visceral sensations and symptoms occur" [21] and is proposed to be a maintaining and exacerbating factor in IBS." Delete punctuation mark and insert and.

This quote has been removed

7. Page 8. "In ICBT the patients learn about the treatment interventions by reading self-help texts that contain both educational material and instructions on how to perform the exercises that constitute the treatment. The general principle is that the treatment should reflect face-to-face therapy in terms of content, but using an online therapist to guide the participants through the course of the treatment. The format allows for large patient volumes to be treated and an increasing number of controlled studies indicate that for common psychiatric disorders ICBT is as effective as face-to-face delivered treatment [30]." The authors might want to consider to place this paragraph in the introduction, page 2 last sentence.

Changed accordingly

8. Page 10. "The waiting list completed assessments immediately after and 12 months after having been crossed over to and finishing treatment and a psychiatric assessment of all patients was conducted before the randomization and at 12-month follow-up." This sentence is not clear. Suggestion: The waiting list completed assessments immediately after and 12 months after having been crossed over to and finishing treatment. A psychiatric assessment of all patients was conducted before the randomization and at 12-month follow-up.

Changed accordingly

9. Page 13. "Of the 30 patients in the ICBT group, 20 (67%) completed the 12-month follow-up assessment. One of these patients had not completed the post-treatment assessment and one patient completed all follow-up questionnaires but GSRS-IBS. Thus 24 patients were included in the analyses of change between post-treatment and follow-up, with 5 patients having missing data." Confusing sentence. Suggestion: One of the patients who completed the 12-month follow-up assessment did not complete the post-treatment assessment, but was included in the analysis of change between post-treatment and follow-up resulting in 24 patients in this analysis. Of these 24 patients 5 had missing data.

Changed accordingly

10. ".... and one patient completed all follow-up questionnaires but GSRS-IBS." No information is given on the other patients with missing data. The authors might want to consider to leave this information out here.

Changed accordingly
11. Page 16 "..... with a few exceptions relating primarily to age [include: of onset], presence of inflammatory bowel disease, severe diarrhea, poor language skills, and patient willingness." Patient willingness however was not defined as an inclusion criterion.

This sentence has been removed

12. Page 16 From ... "It has been difficult to empirically determine which patient characteristics ....[to the end of the paragraph]" might be placed after "Of 131 eligible patients 77 (59%) were included in the study. An inclusion rate of 59% stands in contrast to the estimated 25% of IBS patients that will benefit from a psychological treatment [48]."

The discussion based on percentages has been removed from the manuscript

13. According to the instructions to the authors, tables should not be submitted as figures but should be included in the main manuscript file.

The tables have now been pasted into the manuscript

**Reviewer: Rona Moss-Morris**

There are some key issues with the reporting of this trial and the data analysis which need to be revised. Most of these would fit under minor essential reasons, except for the issues raised about the analysis under point 3 which are major issues for revision.

1. The title of the paper should include that it a feasibility RCT. It is not presented as a definitive trial.

The title of the manuscript has been changed

2. The copy I downloaded from your site did not include an abstract. I am not sure if this is an error on the site but clearly it is difficult to provide a complete review without evaluating the summary of the paper.

An abstract has been added

3. The authors state that the trial is presented according to CONSORT guidelines but they deviate from this in a number of respects:
   • The question posed at the need of the introduction should be broken down so that it relates first to the primary outcome and then to the secondary outcomes. The time point for the assessment should also be included i.e. that differences are expected at the end of treatment and that changes in the intervention group will be maintained at 12 months follow-up.

   Changed accordingly

   • The flow chart does not correspond to a CONSORT diagram. It should include assessment periods (including basline), n’s for each assessment period, reasons for drop out at each time point, and the number of participants analysed. The outside arrow for the treatment group pointing to the final assessment box does not make sense.

   We have made a new flowchart.

   • The allocation concealment for the randomisation process is not clearly described
We have clarified the randomization procedure

• The statistical analysis section states that data are analysed by intention-to-treat, yet the effect sizes in the tables are presented as completers and only completers are included in the follow-up analysis. This is not correct and must be sorted before publication. All people randomised should be included in the analysis of the data and the data displayed should relate to these analyses.

We do believe that we have done an ITT analysis by doing full information maximum likelihood estimation and exploring the data missingness mechanisms. As we point out in the manuscript this is a superior method compared to last observation carried forward, even with substantial data loss. Regarding the follow-up analysis we are not sure what the reviewer is referring to. When looking at follow-up data we are interested in whether patients that participated in the follow-up showed any change from post-treatment to follow-up. Imputing data from the post-treatment would give the illusion of maintenance of improvement so we believe it to be appropriate to check if any significant change has occurred between post-treatment and follow-up, using observed data (as suggested by this reviewer we have added sensitivity analysis to the follow-up analysis). Effect sizes are indeed reported using only complete post-treatment data in Table 3. Since we have 22% missing data at post-treatment both means and standard deviations are enormously influenced by carrying forward the pre-treatment data, with effect sizes dropping to about .25 for GSRS-IBS, IBS-QOL, and the VSI. Reporting only these effect sizes would give the false impression of a low overall improvement, while we argue that for a proportion of patients that dropped out the treatment is clearly unsuitable while it does work well for those who stayed in treatment. Nevertheless, we now report the effect sizes based on last observation carried forward in the text for full transparency.

4. The power analysis is very sketchy and requires more detail. What study was the effect size it based on and was it 60 people per group? Was the effect size for end of treatment or follow-up?

We have clarified the power analysis

5. A one month wash-out period is alluded to on p. 5 but it is not clear what drugs this washout includes. There is also no data presented as to change of treatment during the treatment and follow-up periods and whether patients started new drug or psychology therapies during this time. This is crucial and these data should be included in a sensitivity analysis or the lack of data in this area needs to be discussed as a limitation.

We have added a sensitivity analysis of the follow-up data. We have also removed the phrase “washout period” to clarify that we intended to let the effect of the treatment as usual at the clinic have its effect before pre-treatment assessment. We have included change in medication as part of our cost-effectiveness data but have not used in sensitivity analyses.

6. The last paragraph on p. 5 is also confusing. Perhaps the word ‘inclusion’ in the first sentence should recruitment?

Changed accordingly
7. The quote on p.7 (reference 21) needs to have a page number.

The quote has been removed

8. Details of the psychiatric assessment referred to under data collection (p.10) should be provided i.e. what was this for, what did it include and who conducted it? It should also be explained under procedure or screening.

We have given more details about the psychiatric assessment in the Data Collection section. Since it was not part of the inclusion procedure we believe that it is appropriate to mention it only there in the manuscript.

9. There are no data on how many sessions were completed by the ICBT group and how long patients took to complete sessions? Adherence data to the programme is important and we need to know how many people dropped out of the treatment. Even patients who complete follow-up questionnaires do not necessarily complete treatment.

We have now added data about the number of patients completing each step in Table 2 and the number of completers is also discussed in the manuscript.

10. The discussion needs to be altered in accordance with a proper intention-to-treat analysis.

We now state that based on the whole sample, using last observation carried forward to replace missing data, the ICBT resulted in small effect sizes.

11. On p.16 – the authors raise the issue of predictors of treatment efficacy and who should treatment be offered to. In the penultimate sentence they refer to mediation in this regard but in fact they are referring to moderation. Analysis to work out who responds to treatment best looks at interactions and moderator effects. Mediators look at mechanisms of an intervention and process variables. Indeed, the authors could look at moderators or predictors of treatment outcome in their own data and this would strengthen their findings.

We have rewritten our discussion about who should be offered treatment for IBS. To aid the discussion we have followed the reviewer’s suggestion to perform an exploratory prediction analysis of outcome.

12. Finally, as a discretionary point for revision the authors argue in the discussion that psychological therapy may be best targeted to patients who are suited to it. However, this depends on the model of therapy. For instance, if the therapy is focusing on underlying psychopathology, or as in this case of this study anxiety, then this supposition may well be true. However, if the intervention is based on maintaining cognitions and behaviours and a more self-management approach, it could be argued that this is applicable for all IBS patients.

We have added a sentence concerning this point to the discussion.

**Reviewer: Melissa Hunt**

There is only one discretionary revision that I would like to see: the authors point out that their effect sizes were lower than in their own previous trial, but they do not make any effort
to calculate or report on the clinical significance of their findings. This is particularly important given the nature of the sample. They point out that the mean baseline scores on the VSI and the IBS-QOL were “markedly” less severe than in prior physician referred or self-referred samples. It would be quite interesting to know several things. First, what percentage of the entire sample achieved clinically significant symptom reduction.

For use in the cost-effectiveness analysis we have now calculated clinical significance using 50% reduction the GSRS-IBS. The proportion of patients demonstrating clinical significance, only 20%, was surprisingly low. However, in our previous cost-effectiveness trial we got 36% using the same method, while for instance the Adequate Relief measure based on the same data gave a responder proportion of about 59%. Given the arbitrariness of these measure of clinical significance measures we decided to focus on the differences in effect sizes while still reporting the proportion of responders according to the 50% criterion.

Second, if you divide the sample into more and less severe groups (based simplistically on a median split, or better on 1 SD below the mean of the symptom inventories from their prior study) would the percentage of individuals achieving clinically significant change look dramatically higher for the individuals who were more impaired to begin with? That is, simply giving us group means and effect sizes for treatment versus control doesn’t give us enough detail about actual symptom change across the range of patients enrolled.

We have also followed these suggestions to perform an exploratory prediction analysis of outcome. However, we chose the change score in symptoms as dependent variable since only 20% fulfilled the clinically significant change criteria.

The authors conclude, rightly, that we need more research on which to base recommendations about which IBS patients stand to benefit from receiving psychological treatment. Ironically, they may have data from this study which speaks to that point but which they have not fully exploited. In the US, at least, it is very difficult to get gastroenterologists to refer for psychological treatment at all! If we had clear cut clinical guidelines that could foster a better referral pipeline, all the stakeholders would benefit.

We hope that the additional calculations we have made will shed some light over the issue of determining which patients to offer psychological treatment. However, more data than we have available in this study is necessary to further investigate this important matter.

Finally, we would like to extend our gratitude to all three reviewers. Although the comments have resulted in major rework on the manuscript, we believe that it has greatly improved. We have added data that helps the reader to interpret our results better and the discussion is now better supported by empirical observations. We hope that the reviewers and editor are in agreement. But major rewriting like this always raises new questions, and we look forward to getting new comments on how the manuscript can be further improved.