Dear BioMed Central Editorial Team,

Please find below the detailed response to the reviewer’s comments regarding manuscript # 3565901415624315, entitled, “Preliminary Examination of the Efficacy and Safety of a Standardized Chamomile Extract for Chronic Primary Insomnia: A Randomized Placebo-Controlled Pilot Study.” Changes in the manuscript are highlighted in yellow.

Reviewer 1:

1. I am most concern with the study conclusion which states that there is a modest and mixed clinical benefit of chamomile for treating insomnia. It seems that the baseline severity of insomnia may be significantly higher in the chamomile group, resulting in a moderate within-group effect size after treatment. When the baseline severity of insomnia is controlled for, there is no significant group by time interaction. It is possible that the improvements in the chamomile group are due to events unrelated to treatment, such as natural course of illness and regression toward the mean.

-We agree and have added the suggestions of the reviewer to the last paragraph (p.16) and toned down the implications of our conclusions.

2. It may be worthwhile to provide more information of the allergic reactions and adverse effects of chamomile (references 14 and 15) and present in more details the two previous studies of chamomile that have included sleep-wake measures (references 16 and 17). For example, what is the sample size, is chamomile used as a single herb or in herbal mixture, what are the sleep-wake measures used, and what are the dosages of chamomile?

-Given limitations in manuscript length, we have summarized key points from these studies most germane to the study and we encourage readers to read more details in the references provided.
3. Please clarify whether individuals with depressive disorder are excluded and how the investigators determined the participants were in remission (Methods, 3rd paragraph).

-We removed “unless treated and in remission” (p. 6) to clarify that participants could not meet current criteria for a mood or anxiety disorder according to the Prime-MD.

4. The authors are encouraged to briefly describe the outcome measures used in the study and their range of scores (the paragraph before the Objectives and Outcomes section). Please also provide references of the scales. Explanation of the outcome measures can be grouped into a section. There are some repetitions in the paragraph before the Objectives and Outcomes section and the Intervention section.

-We have added references to the questionnaires in the Methods section. We have also added the range of scores and key cutoffs for each measure in a footnote in table 2.

5. Please explain the rationale for adjusting the baseline values in statistical analysis and whether there are significant differences between treatment groups at baseline. Why not using ANOVAs with repeated measures if there are no baseline differences? Please present more clearly whether there is time effect, group effect, and time by group interaction. The authors may need to explain whether the effect size refers to within-group or between-group effect size and provide a clearer definition of the term. When the authors used the term “effect size”, it is better to state whether it is within-group or between-group.

-Per our biostatistician we performed regression models that adjusted for any baseline differences. This is a very common approach to statistical analysis in RCTs with two time points. One advantage of analyzing data in this manner that it is more flexible than paired data since it allows for a non-unit slope for the baseline value. If you have many dropouts at time 2, then a longitudinal analysis (2 time-point repeated measures) is preferred since it is 'intent to treat'. We have clarified how we describe the results.

-We have clarified on p. 9 in the Statistical Methods and Sample Size section that we computed between groups effect sizes.

6. The authors are encouraged to present the RCT in the CONSORT format. It is necessary to revise Figure 1. Please present the demographics and clinical characteristics of the treatment groups in a Table and the percentage of missing data in Figure 1 and Table 1. The frequency of the adverse events in the treatment groups could also be presented in a Table, otherwise, it should be presented in greater details. Please explain more clearly the answers to the question on blinding in each treatment group (Results last paragraph).

-We have revised figure 1 and corrected the number of ineligible volunteers
We have added a demographics and adverse events table and changed the manuscript to reflect this.

We have added a sentence on p. 8 of the Methods section indicating what question was asked about blinding.

7. The authors are encouraged to omit the Cohen’s d values in the Discussion section (first paragraph) and rewrite the sentence “However, only differences in the FSS approached statistical significance” and “These measures of sleep quality …. 0.79” in the same paragraph. It is hard to say what level of p value can be regarded as being close to statistical significance. Please specify whether the pooled effect sizes ranging from 0.38 to 0.79 refers to sleep quality and fatigue. Sentences with unclear meanings are common in the manuscript. It is unwise to speculate that greater improvements could be attained if a sample with more severe insomnia were used (Discussion second paragraph). It is difficult to say that the study findings are similar to those of previous studies (Discussion third paragraph). In fact, the lack of significant difference between the treatment groups is contrary to the previous findings. It seems unclear only STAI-T, but not both STAI-T and STAI-S were presented. Since the authors discussed the treatment of anxiety disorder using chamomile, it is useful also to analyze the changes in STAI-S with treatment.

We agree about subjective evaluation of significance and have deleted the offending sentence.

We think that the moderate effect sizes for some of the outcome variables are worthy of highlighting, given their comparability to effect sizes in drug trials, and have therefore left those in the first paragraph of the discussion. We have removed references to small effect sizes.

We have clarified that the range of pooled effect sizes referred to sleep diary outcomes.

We have changed our wording to less definitely reflect that our less severely affected sample may have produced a floor effect on our outcome measures.

Our discussion of our findings in relation to the extant literature was balanced and accurately reflects the similarities and differences in the findings.

We had no expectation that chamomile would alter moment to moment anxiety levels, thus we did not include this subscale. Moreover, our use of the STAI-T only is consistent with recent recommendations for outcome measures in insomnia treatment trials (Buysse, DJ et al. Recommendations for a standard research assessment of insomnia. Sleep 2006;29(9):1155-1173).

8. The authors are encouraged to provide a definitive discussion on the dose and duration of chamomile for the treatment of insomnia and comment on whether a higher dose of chamomile and longer treatment duration should be used in future studies.
Unfortunately there is no information to directly address this question, e.g., a pharmacokinetics study in humans. We have adequately addressed the possibility that our dose and/or schedule of administration may have contributed to the mixed findings.

9. There are many grammatical mistakes and typo errors that require correction. Effect size is not a primary outcome measure (Abstract section). When the p value is provided in the Abstract, either all p values be presented or not to present the p values.

-We carefully reviewed this version of the manuscript for appropriate grammar and have corrected typographical errors.

-We have changed the Abstract for the designation of outcomes and removed all p-values.

Reviewer 2:

1. It would be very important for future researchers that a discussion of study numbers required to adequately power a study was included somewhere within the paper.

-We added two lines to the Discussion (p.17) that provide sample sizes based on calculated effect sizes for the largest and smallest effect sizes in our study.

2. While it is very interesting to discuss important findings such as the large change in sleep latency it would be important to make it clear that these findings are no more than passing observations. While it is clear on careful reading of the paper that the results are not statistically significant the conclusion in the abstract and the final paragraph of the paper both suggest that inference about the efficacy of chamomile can be drawn from these results.

--We agree and have added the suggestions of the reviewer to the last paragraph (p.16) and toned down the implications of our conclusions.

3. Insomnia is a problem which increases with increasing age. It might be worth a sentence or two in the discussion about why an age limit of 65 was imposed. The abstract merely states "over 18", this should be corrected.

-We have corrected the Abstract

-We have added a sentence in the Methods section (p.6) to address this issue.

4. I do not know the half-life of chamomile but was interested that it was given twice daily and think it might be worth some discussion. Assessment has been made about daytime effects and these have been attributed to insomnia and its treatment. Perhaps some of the changes noted were due to daytime dozing with chamomile.
-Given the absence of information on pharmacokinetic properties, we took a conservative approach to achieve steady state of key constituents of chamomile in the plasma - Discussion.

-We examined the sleep diary to look for increased napping between chamomile and placebo and found no difference.

5. I would prefer the supplementary table to be included in the main manuscript.

-Table 1 was never meant to be supplementary. We apologize for the error.