Author’s response to reviews

Title: Effects and clinical feasibility of a behavioral treatment for sleep problems in adult Attention Deficit Hyperactivity Disorder (ADHD): a pragmatic within-group pilot evaluation

Authors:

Susanna Jernelöv (susanna.jernelov@ki.se)
Ylva Larsson (ylvalarsson@live.se)
Milagros Llenas (milagros.llenas@gmail.com)
Berkeh Nasri (berkeh.nasri@ki.se)
Viktor Kaldo (viktor.kaldo@ki.se)

Version: 2 Date: 16 May 2019

Author’s response to reviews:

May 16, 2019

Dear Editorial Officer,

Please find enclosed our manuscript entitled A pragmatic within-group pilot evaluation of a behavioral treatment for sleep problems in adult ADHD by Susanna Jernelöv, Ylva Larsson, Milagros Llenas, Berkeh Nasri, and Viktor Kaldo, that we have now revised to the best of our abilities according to the reviewers’ comments and demands. We thank the reviewers and the editor for their time and commitment to giving thoughtful suggestions to improve the manuscript. Although we have tried our very best to accommodate the reviewers’ wishes, we are not unwilling to make additional changes, should this be deemed necessary.

Representing the authors of this paper, I hope you agree that the manuscript has improved substantially thanks to the reviewers’ suggestions, and hope you will find it suitable for publication in BMC psychiatry. I look forward to hearing from you.

Yours Sincerely

Susanna Jernelöv, corresponding author
Below are the reviewers’ comments, and our replies.

Reviewer reports:

Outi Mantere, M.D., Ph.D. (Reviewer 1):

Reviewer comment: This pilot study on a novel treatment for insomnia in ADHD provides an important addition to the non-pharmacological treatment of a common problem. The study is well conducted and the paper well written. The authors also report that a larger clinical trial to confirm the results is already ongoing. I have only a few suggestions.

1) Please detail how you measured the background information. It should be possible to repeat the methods based on your description, including the classification presented in the table. Part of that came probably form the Sleep-50? Did you have an open question for diagnosis or a list of diagnoses? How were the medications classified? The information you use to correlate with outcome variables or to control for confounder is the most relevant to be described here in detail.

Author reply: We thank the reviewer for pointing to the need for clarification on these points. For type of ADHD, patients were asked to choose which described their type best: Predominantly inattentive, Predominantly hyperactive/impulsive, Combined subtype, or NOS; for psychiatric diagnoses an open-ended question was used; for psychiatric problems during the last year, patients were given a list and asked to mark which applied to them; medication use was gauged with an open ended questions with examples, and medications were classified according to stated use and/or substance (e.g. zolpidem was classified as a z-drug; if e.g. quetiapine was described as a sleep aid it was classified as such under the antipsychotic heading, but if the patient had reported a psychotic disorder or psychotic symptoms quetiapine was classified as an antipsychotic). We have added a description of this to the measurements section on page 7, and made some clarifications in table 2.

2) Please describe the methods to measure chronotype and the definitions for the cut-offs used in the table. Did the chronotype change in follow-up?

Reply: Circadian type, like all measures in this study, was self-report, and it was gauged with the question on morning-eveningness from the Karolinska Sleep Questionnaire, on a five point scale (Extreme morning type, morning type, intermediate type, evening type, extreme evening type), see for instance (Ingre et al., 2014). This has now been clarified in the Measurements section on page 7 and in table 2.

Unfortunately, we did not ask for circadian type at post- or follow-up assessments.
3) Impression of the treatment was scored from 1 to 6; please define interpretation of the scores.

Reply: 1 represents “very poor”, 2 “poor”, 3 “quite poor”, 4 “quite good”, 5 “good” and 6 “very good”. This has been included in the manuscript on page 8.

4) You had missing data both due to attrition and as single items. Please specify use of imputation for missing items.

Reply: Imputation for missing items was performed using the Corrected Item Mean substitution (Huisman 2000, Quality and Quantity), for symptom outcome measures. Missing items from the evaluation questionnaires were not imputed, since responses were qualitative and only descriptively reported. This has been clarified in the manuscript under Data analysis on page 9.

5) Missing information was common to evaluate the feasibility. Please discuss reasons for that (methods used to ask or to collect the responses). The most important would be to know the reasons for not continuing - did you have any information on that?

Reply: Feasibility was measured in several ways, one of which was attendance to treatment sessions, and there we have no data loss. Thus the feasibility measures with missing information were data from the evaluation questionnaires. We did get evaluation forms from 13 out of 17 non-drop-outs (76%), which is not less than we usually get from other evaluations at the clinic. From the clinic’s perspective that was satisfactory. Treatment evaluation forms were handed out after the last treatment session and patients were asked to fill them out and return them before leaving. This has been added to page 6 in the manuscript. The type of questions used are similar to the questions used to evaluate other treatment groups at the ADHD clinic.

The two patients who did not continue treatment and were considered drop outs reported “personal reasons” for their discontinuation. The patients who were not considered drop-outs did not attend all sessions, but attended a substantial proportion of the sessions. Sessions that were missed were mostly missed due to illness or more pressing engagements, such as work or doctor’s appointments, and some patients had a hard time making it to the sessions, due to their difficulties with planning and organizing. A description has been added to the manuscript on page 12-13.

6) The finding about a correlation between the Hyperactivity subscale reflects arousal and has some neuroscientific implications for research. Since wrong timing and control of
arousal likely associates with wrong timing of the circadian rhythm this could be discussed in a few sentences.

Reply: This very important point and we are grateful for the opportunity to elaborate on it. Arousal is indeed an important aspect of sleep, and is included in several of the theoretical models of insomnia (e.g. the cognitive behavioral model of insomnia and the hyperarousal theory). Inability to control arousal or wrongly timed arousal can cause acute insomnia, and can be a contributing factor to chronic sleep problems. In ADHD, hyperactivity is a core symptom, and has been associated with insomnia (Weibel et al 2017). The decrease in hyperactivity symptoms seen in the patients in this study could conceivably be indicative of a decrease in arousal possibly associated with a shift in the circadian rhythm, following the treatment. This is an intriguing prospect that would warrant further study. The above is now discussed in the manuscript on page 15.

7) Of the main factors controlling circadian rhythms or chronotherapy, the content of the therapy lacks information about the role of timing/regularity of eating (see reviews of Potter).

Reply: Again, an interesting point, and as a matter of fact, the role of timing an regularity of eating was not entirely left out in this treatment; during the 3rd and 4th sessions, when circadian rhythms were covered, zeitgebers other than light and darkness were also discussed. We have changed the wording describing the session content in table 2. However, it is true that systematic work with timing and regularity of eating was not specifically included in this treatment. Given the state of the evidence, we shall consider including this aspect in the updated manual.

8) The authors list other similar therapies but it remains whether there are some major differences in the content or duration of the formalised therapies.

Reply: When we say we have adapted this treatment to patients with ADHD, our comparison is a "general" CBT-i, which is commonly delivered in 4-8 sessions and contains psycho-education, sleep hygiene, relaxation, sleep restriction, stimulus control, paradoxical intention and some component aimed at identifying and challenging dysfunctional beliefs about sleep, for instance as described in the meta-analysis by (van Straten et al., 2018). The alterations were made in several ways, for instance by adding an introductory session on strategies for organizing and planning, from the Safren manual on CBT for ADHD. We also added telephone calls from the therapist between each treatment session, to help support behavior change. Another adaptation was the addition of work with circadian zeitgebers such as systematic use of light and darkness, and adjustments of activity and meal times. Neither of these components is normally included in a CBT-i-treatment. In addition, we changed the number of session and the lenght of sessions.
Which of these adaptations were the important we do not know, perhaps none of them. However, our clinical impression is that most patients did find the treatment components useful, the number of sessions adequate, and the telephone calls between sessions very helpful, and this is not contradicted by the data in this study. The present study was not designed to test the usefulness of different treatment components, but that is an important area for future research studies. We have tried to clarify this in the discussion, on page 14.

9) The authors use mean; was the distribution of the dimensional variables tested? Please detail the methods and results and where needed use nonparametric tests in the results and tables.

Reply: This is an important point. We have indeed tested the distribution for dimensional variables, and none of the variables were statistically non-normally distributed at any of the time-points (Kolmogorov-Smirnov p’s .108 < .200). We therefore found it acceptable to use means and parametric statistical tests. This information is now included in the manuscript on page 9.

However, we did realize that the first version of the manuscript was missing the statistical analyses to investigate the change in insomnia symptom severity level. We are terribly sorry and apologize deeply for this mistake. We have now added the results from Wilcoxon signed ranks tests, as this is rank data. The results indicate that patients moved from a higher to a lower severity level, i.e. the median post-treatment ranks were statistically significantly lower than the pre-treatment ranks (Mdn Pre = 3, Mdn Post = 2, Z = .000, p = .003) and the lower level was maintained at three-month follow-up, i.e. median follow-up ranks were significantly lower than the pre-treatment ranks (Mdn Pre = 3, Mdn FU = 2, Z = 8.000, p = .000), but did not differ from the post-treatment ranks (Mdn Post = 2, Mdn FU = 2, Z = 15.000, p = .166). These statistics are now included in the manuscript on page 11.

10) The testing for a change does not take into account the fact that part of participants had an increase in insomnia; more notice in terms of secondary analysis could be taken.

Reply: We have reworded the paragraph and hope this is more to the point, and have also added a graph with individual ISI scores for all patients (without imputations). However, we were unsure which analyses the reviewer might be thinking of, but will be glad to make further changes if needed.

11) The gold standards for insomnia are Pittsburg and AIS. What is the reliability and psychometric properties of ISI as compared to these and within a population with ADHD?
Reply: We thank the reviewer for this comment, but do beg to differ on the statement that PSQI or AIS, and not the ISI, would be the gold standard measures. For instance, (D. J. Buysse, Ancoli-Israel, Edinger, Lichstein, & Morin, 2006), list both the PSQI, the AIS, and the ISI as suitable measures in insomnia research, and the ISI is a very much used tool for detecting changes in insomnia severity in intervention studies. (Out of curiosity we ran a quick search on Google Scholar, yielding 124 000 results for the Insomnia Severity Index, 81 800 for the Pittsburgh Sleep Quality Index, and 10 500 for the Athens insomnia scale.)

The reported internal consistency is similar across these measures. For PSQI, Chronbach’s alpha has been reported to be .83 (Daniel J. Buysse, Reynolds, Monk, Berman, & Kupfer, 1989), for the AIS .90 (Soldatos, Dikeos, & Paparrigopoulos, 2000), (Chung, Kan, & Yeung, 2011), for the ISI .74 < .78 (Bastien, Vallières, & Morin, 2001), and, as stated in Buysse 2006: each of these has demonstrated insomnia-control differences, changes with treatment or both differences and changes.

Within the current sample, Cronbach’s alpha for the ISI-a was .646 at pre, .831 at post, and .890 at follow-up, which can be compared to that of the ISI-a in pain patients, which has been reported to be .83 (Kanstrup, Holmström, Ringström, & Wicksell, 2014).

12) Is the content or efficacy of the therapy meant/expected to be specific to ADHD, or could be shared with other conditions with hyperarousal such as hypomania or borderline?

Reply: This is a very interesting question, and ultimately an empirical one. We do believe that other psychiatric populations may benefit from similar adaptations, but some patient groups may find the short sessions and the long treatment period demanding, and telephone calls in between sessions may be unnecessary for patients who do not suffer from executive difficulties. As a matter of fact, within the larger project (Better sleep in psychiatric care), we are aiming to develop and evaluate manuals for treating sleep problems in other patient groups as well. Many psychiatric patients have circadian problems, and could probably benefit from the circadian component in this manual. One component that was not included in this manual, but might be included in the manual for psychotic patients is a component for dealing with nightmares and psychotic experiences during sleep, since this can be very troubling for patients with psychotic disorders.
Landon Hurley (Reviewer 2):

Reviewer comment: Given the lack of precision and the admitted errancy in utilising scores from the outcome metric (ISI; i.e., the stated indeterminacy which follows from the reported Cronbach's alpha < 1) combined with the further split of the scale into an ordinal interpretation, it seems grossly inappropriate to base conclusions and validity assessments upon comparisons of the location parametre. Further, and on a similar note, the reliability of the imputed data would need to be assessed, since the techniques involved (presumably, although completely unstated would be what parametric distribution of errors was utilised to impute the missing values, and also whether further heterogeneity should be expected as a consequence of staying or dropping out of the study. Together, these concerns would be observed as artificially greater power than would be expected (more stability in the ISI scores, which are empirically and theoretically discrete rather than continuous) uncertainty with respect to the difference between true score and uncertainty due to measurement error (which, given the reported values statistics by the researchers are at best capable of placing scores to within two standard deviation of their true value upon a standard normal). This uncertainty, one should note, is twice the reported effect size, and thus within such a small sample and with no controls taken for uncertainty or a desirability response to the clinicians, deeply calls into question the validity of this sample to serve as an inductive platform for anything beyond the sample at hand; this therefore undermines the generalisability of the conclusions.

A trivial solution to the problem would be to conduct rank based assessments, and such that the order placement of the pre-post tests can be assessed. This would combine well with the interpretation of the score ranges for remittance, subthreshold, clinical insomnia moderate, and severe. However, as it stands, the modelled measurement errors, combined with the lack of strict correspondence between self-reports and clinician observations, would indicate that the results would require correction both pre and post impuation, before the lack of bias in the treatment effect in the sample could be interpreted with any hope of generalisability.

Author reply: We thank the reviewer for this elaborate comment and for the suggested solutions.

We have indeed tested the distribution for dimensional variables, and none of the variables were statistically non-normally distributed at any of the time-points (Kolmogorov-Smirnov p’s .108 < .200). We therefore found it acceptable to use means and parametric statistical tests. This information is now included in the manuscript on page 9.

We have also conducted a sensitivity analysis with rank based assessments for the ISI and the ASRS, and find the results to be quite similar to that of the parametric statistical analyses. The Wilcoxon signed ranks test for ISI scores at pre-, post- and follow-up assessments indicates that for the ISI, the median post-test ranks were statistically significantly lower than the pre-test ranks (Mdn Pre = 15.00, Mdn Post = 12.00, Z = 7.000, p= .001), and that the median follow-up ranks
were statistically significantly lower than the pre-test ranks (Mdn Pre = 15.00, Mdn FU = 8.00, Z = 3.000, p = .000), and also lower than the median post-test ranks (Mdn Post = 12.00, Mdn FU = 8.00, Z = 26.500, p = .031). For the ASRS, the median follow-up ranks were statistically significantly lower than the pre-treatment ranks (Mdn Pre = 40.00, Mdn FU= 39, Z = 16.000, p = .012), the median post-treatment ranks were not quite statistically significantly lower than the pre-treatment ranks (Mdn Pre = 40.00, Mdn Post = 36.00, Z = 28.000, p=.068), and the median follow-up ranks were not statistically significantly different from the post-treatment ranks (Mdn Post =36.00, Mdn FU = 39, Z = 33.500, p = .230). We have noted in the manuscript on page 10 that this sensitivity analysis was performed.

Unfortunately, in the first version of the manuscript, the statistical analysis was missing to investigate the change in insomnia symptom severity level. We apologize for this mistake and have now added the results from Wilcoxon signed ranks tests. The results indicate that patients moved from a higher to a lower severity level, i.e. the median post-treatment ranks were statistically significantly lower than the pre-treatment ranks (Mdn Pre = 3, Mdn Post = 2, Z = .000, p = .000) and the lower level was maintained at three-month follow-up, i.e. median follow-up ranks were significantly lower than the pre-treatment ranks (Mdn Pre = 3, Mdn FU = 2, Z = 8.000, p = .000), but did not differ from the post-treatment ranks (Mdn Post = 2, Mdn FU = 2, Z = 15.000, p = .166). These statistics are now included in the manuscript on page 11.

We have also looked at the language in the paper and tried to be more careful with the wording as to not imply causality or generalizability, for example “The improvement seen in this pilot is small in effect (d=0.34), and cannot be causally attributed to the treatment, but to our knowledge this is the first study showing possible improvements on ADHD symptoms after a behavioral sleep intervention.” and “We conducted a preliminary evaluation of possible effects of the intervention […],” and hope the reviewer agrees that these are reasonable ways to handle the questions at hand.

Reviewer comment: On a more pragmatic point, controlling for patient satisfaction with the treatment in the assessment of the difference (that is, turning the paired t-test into a repeated measures ANCOVA or the rank based equivalent thereof) would probably be warranted, since satisfaction is likely highly correlated with the desire to report and interpret results from CBT intervention, regardless of the actual affective treatment effect.

Author reply: As per the reviewer’s request we have looked into the effects of patient satisfaction on the reported insomnia severity, by conducting a repeated measures ANCOVA at post-treatment with the ISI scores at pre- and post-treatment, and at follow-up as a sensitivity analysis. The ISI scores correlated significantly over all time-points both with parametric and non-parametric correlations, but neither the parametric nor the non-parametric correlations
between treatment satisfaction and ISI scores at the three time points were statistically significant (p’s .165 < .859). The ANCOVA showed a significant effect of time (F(1,16) = 15.449, p = .001), but as expected, the interaction with treatment satisfaction was not statistically significant (F(1,16) = .184, p = .673). Mauchly’s Test, Chi2(2) = 1.34, p = .51, did not indicate any violation of sphericity. This sensitivity analysis too is summarized in the manuscript, on page 10-11.


