Reviewer’s report

Title: Sleep hygiene intervention for youth aged 10 to 18 years with problematic sleep: a before-after pilot study of sleep and weight

Version: 2 Date: 29 July 2012

Reviewer: Dean W Beebe

Reviewer’s report:

This manuscript describes the findings from a pilot study of an uncontrolled behavioral intervention on the sleep quality and quantity, daytime activity, and body mass of 10-18-year-old participants with sleep complaints. The authors conclude that the intervention shows promise in improving both sleep quality and body mass. In light of recent epidemiological and experimental work that relates inadequate sleep quantity to increased body mass, the present findings are promising and intriguing. I appreciate the authors’ enthusiasm. Nonetheless, I had several areas of concern that dampened my enthusiasm as a reader/reviewer. To be clear, I don’t expect definitive answers from a pilot study. Pilot studies need only demonstrate feasibility, yield intriguing results, and suggest a need for follow-up work. However, I also expect the write-up of such a study to fully acknowledge its limitations, and the present paper in my estimation missed that mark, potentially misleading naïve readers and allowing savvy critics to dismiss promising work out as interpretively biased.

Major Compulsory Revisions:

1. Sampling children with self- (or more likely) parent-identified sleep problems in an uncontrolled study raises significant concerns about regression to the mean. Specifically, selection of a group based upon extreme levels of any parameter (e.g., poor sleep quality) almost invariably leads to mean scores on that parameter trending towards average upon re-testing. This should be addressed as a limitation that might have artificially inflated apparent effects.

2. The manuscript alludes on multiple occasions to developmental appropriateness. I agree that this is an important element of any pediatric intervention, but it was unclear in the current manuscript how this was achieved by the current intervention. Citing an unpublished dissertation is insufficient, as the vast majority of readers cannot be assumed to have access to unpublished dissertations. Also, the age range of the sample (10-18 years) is quite broad, covering an impressive developmental span; how was the developmental appropriateness ensured across such a wide developmental span?

3. Has the PSQI been validated in children as young as age 10? It was developed for adults, and I did not see any citations on its psychometric features in a pediatric sample. It seems incongruous for a paper that highlights the developmental appropriateness of the methods to cite adult validation work on
the PSQI, and prior infant work on the Actical, for a study of individuals aged 10-18 years.

4. The intervention was inadequately described. Indeed, it was not at all clear how some letters of the F.E.R.R.E.T. acronym were addressed during the intervention. For example, to what does the word “Restrict” relate? Does it refer to the behavioral intervention used for insomnia in which one limits time in bed? Does it relate to dietary intake? I strongly recommend a more detailed exposition of all intervention targets and process, perhaps even including a Table that summarizes each element.

5. Without more information on the intervention, it is extremely difficult for the reader to evaluate claims that this was a sleep-only intervention, rather than one that targeted other behaviors. For example, if waking activity levels and diet were targeted, the study could plausibly be reinterpreted as one that promotes a healthy weight, and the sleep quality changes could be secondary to improved weight, rather than the other direction.

6. While relatively sophisticated time-series analyses were described in the methods, the results and the associated tables all appeared to take into account only two time points. What happened to the sophisticated modeling and other measurement time points?

Minor Essential Revisions:

7. What was the timeframe of data collection on the calendar? For example, did recruitment and initial visits all occur within a particular part of the year, such that changes over time might be confounded with seasonal variations (e.g., poor sleep during winter months and/or the school year, followed by better sleep during the summer)?

8. What medications in particular were viewed as “medications for insomnia”? Were any other medications excluded? How about psychiatric conditions or medications?

9. While the 100% retention rate was impressive, it is highly unusual (almost unheard of) in clinical trials. It seems not only possible (as noted on page 21) that the sample was atypical, but almost certain. This additional concern about generalizability should be more explicitly stated, lest readers walk away with an overly optimistic reading.

10. I recommend moving much of the coverage of the validity of hip-mounted Actical measures sleep duration from the discussion to the methods section. Although the discussion should continue to reflect the limitations of that approach, most readers of scientific work expect measurement tool data to fall primarily in the methods.

11. I found arguments against placebo effects to be unconvincing, especially those that related to consistency of improvement across multiple sleep measures. If the recruitment and interventions were as transparent as they sound
to be, it is almost certain that both parents and children knew exactly what the authors were hoping to change, and consequently would show expectancy effects. I suggest that, rather than trying to argue away the potential for placebo effects, the authors use this potential to argue for the need for a more definitive study.

Discretionary Revisions

12. It would be helpful to include a Table or Figure that visually depicted the data collection timepoints and major measures.

13. Most of the full paragraph on page 18 (“From the results obtained…”) is redundant with other parts of the discussion, and could be substantially reduced or eliminated.

14. Given the limitations of the study, I strongly recommend against the detailed description of potential mechanisms of effect that runs from page 19 through most of page 20. It seems reasonable to very briefly allude to such mechanisms, but there are simply too many questions about whether the current effect is real to begin speculating on its mechanisms.

15. Table 3 seems unnecessary. I found it to be distracting, and to devote an inordinate amount of space to a single measure that lasts only 14 ½ minutes, especially since the take-home message was that none of 11 indexes reached statistical significance.

16. Finally, pilot studies often teach important lessons that can be addressed in follow-up work. Are there things the authors wish had been done differently? Do they have specific advice for researchers who wish to replicate and extend their intriguing work?

Level of interest: An article whose findings are important to those with closely related research interests

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.