PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Pharmacologic interventions to improve sleep in hospitalized adults: A systematic review
AUTHORS	Kanji, Salmaan; Mera, Alexandru; Hutton, Brian; Burry, Lisa; Rosenberg, Erin; MacDonald, Erika; Luks, Vanessa

VERSION 1 - REVIEW

REVIEWER	Dr. Wenjie Sun USA, Tulane university
REVIEW RETURNED	23-Apr-2016

GENERAL COMMENTS	The authors did not discuss some important reference in this research area. such as following: Sleep duration and quality among different occupations--China national study. Please added it

REVIEWER	Altay Lino de Souza Federal Univeristy of Sao Paulo - Brazil
REVIEW RETURNED	26-Apr-2016

GENERAL COMMENTS	The article presents some interesting review about the effect of pharmacological interventions on quality of sleep in hospitalized adults. This is an important subject because quality of sleep can also be associated with better prognostic and also other improve perceived quality of hospitalization by patients. However, some aspects of the paper could be improved and I hope that these suggestions can be interesting in order to help the readers to follow the main findings of the article and their implications: 1- According with PRISMA check-list (Appendix 2 - #20) the authors claim that this point was fulfilled - but the part b) - effect estimated and confidence intervals and measures of consistency, in my point of view, is not complete. In Figures 4 and 5, only the confidence intervals of the outcome measurement for each study were presented. There is no measurement of effect size presented (at least the way how they were calculated is not clear). Was just the confidence interval of absolute values (in %) presented on figures 4 and 5? If yes, how the differences in sample size among studies were addressed?

	In table 3, the "measure of effect" was presented only reporting the p values from each study. P values are not a measure of effect by any means; they are just (in a frequentist way) the probability associated with the chance of finding more extreme values of my statistics of interest (t, F, chi square values, for example) than the values achieved in a particular study. For a better understanding of this concept, I recommend a recent paper from ASA (American Statistical Association) who present some didactic explanations about p-values and their interpretation. I suggest that the authors present as a Measure Effect in Table 3 and also Figures 4 and 5, some kind of measure of effect that tries to address the sample size heterogeneity among studies. Most articles do not present effect sizes yet, but according with the column "Outcome by treatment group" in Table 3, the authors can easily calculate some effect sizes like r, Cohen d or Hegdes g depending of the sample size and level of measurement of outcomes. An article explaining the differences among the effect sizes is also attached for the authors. In case of not present the effect sizes, a justification must be made by the authors in Methods section, including a comment about how to deal with sample size heterogeneity among studies in order to present reliable measurements of outcome effect. 2- How the authors deal with the assessed quality of sleep related with the severity of the condition during hospitalization? According with Table 1, most of the patients were in ICU. Depending of the setting, the poor quality of sleep can be associated with the medication administered but the severity of condition is an important confounder. Some commentary about this could be done in Discussion section as a study limitation or a issue to tackle in future studies. 3 - Most of studies in RCT (Table 1) are compound only by male patients. How this counfounder is addressed by the authors in the data analysis? The data about that is not clear but, in general, we have more men than women in this kind of study (using some medication in different groups) Some comment about this gender effect could also be done, presenting some consequences for the generalization (external validity) of the systematic review findings. Thanks for the opportunity to review the article and hope that the authors could find this suggestions useful for the article improvement. The reviewer also provided a file in addition to this comments. Please contact the publisher for full details.

REVIEWER	Christian Chan (with the assistance of Fiona Ho, PhD Candidate) Assistant Professor Department of Psychology The University of Hong Kong, Hong Kong
REVIEW RETURNED	03-May-2016

GENERAL COMMENTS	This study is a systematic review of randomized controlled trials, comparative cohort studies, and prospective case series examining the efficacy and safety of pharmacologic intervention for improving

	sleep in hospitalized patients. A total of 15 studies were included in the review. Inconsistent results on sleep efficiency and sleep quality were reported. The manuscript was generally well-written. Below are suggestions for clarification and improvement. Background 1. The authors cited a review study that found that non-pharmacologic treatments for sleep were not efficacious. The non-pharmacologic treatments in question appear to only include relaxation techniques and regulation of light/noise exposure and were intended for patients in non-intensive care. The authors' justification for the current study might be strengthened by discussing the pros and cons of pharmacologic treatments and non-pharmacologic treatments and the unique context of hospitalized patients (especially those in acute/critical care units). Methods 2. The definition of sleep efficiency was not clear. Sleep efficiency is conventionally operationalized as the total sleep time divided by total time in bed. The authors might want to clarify what "observation" entailed in their calculation of sleep efficiency. 3. It is unclear why among comparative cohort studies and prospective case series, only those that evaluated sleep using polysomnography were included. 4. The authors are encouraged to perform an interrater reliability analysis whenever possible. In addition, did the authors also consider other possible biases (e.g., baseline characteristics and statistical power)? Results 5. Given the heterogeneity of the studies and the comparison groups, as underscored by the authors meta-analysis was not appropriate. However, it appears that meta-analysis was conducted for one outcome variable. For some of the variables, the authors included a forest plot but without conducting a meta-analysis. It would be helpful if the authors i) justified why meta-analysis was warranted for sleep efficiency, ii) kept consistent when to represent forest plots, and iii) indicated more consistently whether meta-analysis was conducted. 6. Whenever the data is available, please indicate the attrition rate measured at different periods/time points. 7. In addition to the study characteristics, please report demographic information of the participants (e.g., age and gender). 8. Whenever the data is available, please report the long-term efficacy of pharmacologic treatment for improving sleep. 9. It is unclear whether the results warrant the conclusion that benzodiazepines are superior to placebo in reducing sleep latency. Gallais et al., for example, did not find significant differences between oxazepam and placebo. Given that a meta-analysis was not performed, the authors might want to specify that the results concerning the efficacy of benzos was found in several but not all studies and qualify their conclusion accordingly. Discussion

	10. The discussion was clearly written. The authors explained the possible factors that interfered the sleep architecture in hospitalized patients. It would be helpful if the authors can more explicitly discuss the implications of the current systematic review, especially as both pharmacologic and non-pharmacologic treatments demonstrated conflicting and unpromising results. What is the alternative to pharmacologic treatment as first-line treatment in hospital settings, especially acute/critical ones? Minor comments 11. The authors may consider assigning study numbers in tables to make it easier to follow. 12. There are some grammatical errors in the manuscript. 13. Figure 3 (Risk of Bias for Prospective Cohort Studies) did not show properly (P.40). 14. There are 2 independent sections "selection of studies" and "study selection" in Methods. Please either combine the two sections or changing the subheading.

REVIEWER	Vineet Arora University of Chicago, United States
REVIEW RETURNED	05-May-2016

GENERAL COMMENTS	This systematic review studies the evidence on drug therapy to improve sleep for hospitalized adults. The group identified 15 studies involving 861 patients and investigating a variety of sedatives and hypnotics from different drug classes. Studies were limited by low quality, small size, and lack of objective sleep measures. The review concluded that there was insufficient evidence to suggest that pharmacotherapy improves sleep in hospitalized patients. The paper is well written and the conclusion is sound. The authors are to be commended on rigor of the systematic review. Major critiques focus on the presentation of bundling results from ICU and ward together, as well as the summary in the discussion. Some additional suggestions to improve the the review are also highlighted below. Background • I was surprised at how short the background was. The context is important. By hospital, do you mean ICU, general ward, surgical are all? These are very different populations and environments so teasing out some of the prior work may be helpful. I think some additional details of the P in PICO is warranted. Methods • Please incorporate a reference for PICO • It seems like sleep duration should be an outcome of interest in addition to sleep efficiency and other metrics included. • I would be more specific on why studies of self-reported sleep were excluded as opposed to just citing the reference and letting the reader look it up. Also consider highlighting how many studies were in this category and got excluded in the results as that is also important to drive the field forward. It comes later in the discussion but its probably too late for the reader as they are wondering when it first appears. • Likewise, some context or rationale for only including cohort or case series of including polysomnography is needed. Were any with actigraphy excluded? Polysomnography is often only done in the

	ICU so this seems like ICU studies would be overrepresented in this inclusion. • It seems like grey literature review missed several conferences that may discuss sleep like ATS or Sleep? Results • I personally think the tables are jarring when propofol infusion is compared with something like a po drug like melatonin. This speaks to the heterogeneity of clinical populations. I think subheaders in the table and corresponding organization in the results should discuss ICU studies separately from ward patietns. Being mechanically ventilated is just a very different thing and PSG is more likely in that population anyway. The current way the results are presented if you were an intensivist or hospitalist, you would have to read every paragraph in detail to find a study that applied to your clinical situation. Discussion • The first few lines of the discussion are out of place and should be in the introduction. I would replace this with an accessible summary of the review findings. • the Discussion should get into implications of these findings rather quickly. It doesn't come til much later. • Was there any difference in ICU vs ward studies? Any recommendations there would be helpful given the wide heterogeneity of the environment. • Another possibility to consider is the disruptions are due to the medical care environment, and therefore are iatrogenic, so medications would not necessarily help and staff-based protocols would be needed. This should be elaborated on in the discussion.

VERSION 1 – AUTHOR RESPONSE

Reviewer: 1
Reviewer Name: Dr. Wenjie Sun
Institution and Country: Tulane university, USA
Competing Interests: None

The authors did not discuss some important reference in this research area. such as following: Sleep duration and quality among different occupations--China national study. Please added it

AR: We reviewed the reference suggested and decided that although this is an important study evaluating sleep quality among different occupations in China it is not relevant to our systematic review which focuses on sleep metrics in response to a pharmacologic intervention among acutely ill hospitalized patients. The suggested study evaluates sleep quality in otherwise healthy individuals and does not incorporate a pharmacological intervention.

Reviewer: 2
Reviewer Name: Altay Lino de Souza
Institution and Country: Federal Univeristy of Sao Paulo - Brazil
Competing Interests: None Declared

The article presents some interesting review about the effect of pharmacological interventions on quality of sleep in hospitalized adults. This is an important subject because quality of sleep can also be associated with better prognostic and also other improve perceived quality of hospitalization by patients.

However, some aspects of the paper could be improved and I hope that these suggestions can be interesting in order to help the readers to follow the main findings of the article and their implications:

1- According with PRISMA check-list (Appendix 2 - #20) the authors claim that this point was fulfilled - but the part b) - effect estimated and confidence intervals and measures of consistency, in my point of view, is not complete.

AR: We agree that measures of effect are inconsistently described in our evidence tables. The reason for this is the great variability in outcome reporting from individual studies. Since most studies reported sleep related outcomes of interest as medians and interquartile ranges, while others used different measures of central tendency and variance. After much discussion it was decided to report the data as it was reported in the original studies. We have added text to explain this in the discussion section.

In Figures 4 and 5, only the confidence intervals of the outcome measurement for each study were presented. There is no measurement of effect size presented (at least the way how they were calculated is not clear).

AR: Figures 4 and 5 present mean differences and corresponding 95% confidence intervals for the mean difference for each study described. In cases where a confidence interval was not reported and no SD values were reported or could be imputed, only the point estimates for the mean difference are indicated (along with 'NE' to denote where confidence intervals could not be estimated). In cases where multiple studies were present for a comparison, pooling was not performed for reasons described in the manuscript.

Was just the confidence interval of absolute values (in %) presented on figures 4 and 5? If yes, how the differences in sample size among studies were addressed?

AR: Figures 4 and 5 summarize, for each study, the sample size, observed/imputed mean values and corresponding standard deviations in each group (where available), and also present the corresponding estimated 95% confidence intervals for each study estimate (whose calculation includes sample size as a component).

In table 3, the "measure of effect" was presented only reporting the p values from each study. P values are not a measure of effect by any means; they are just (in a frequentist way) the probability associated with the chance of finding more extreme values of my statistics of interest (t, F, chi square values, for example) than the values achieved in a particular study. For a better understanding of this concept, I recommend a recent paper from ASA (American Statistical Association) who present some didactic explanations about p-values and their interpretation.

I suggest that the authors present as a Measure Effect in Table 3 and also Figures 4 and 5, some kind of measure of effect that tries to address the sample size heterogeneity among studies. Most articles do not present effect sizes yet, but according with the column "Outcome by treatment group" in Table 3, the authors can easily calculate some effect sizes like r, Cohen d or Hegdes g depending of the sample size and level of measurement of outcomes. An article explaining the differences among the effect sizes is also attached for the authors.

In case of not present the effect sizes, a justification must be made by the authors in Methods section, including a comment about how to deal with sample size heterogeneity among studies in order to present reliable measurements of outcome effect.

AR: We agree with the reviewer's comment that p-values on their own can be greatly uninformative for readers as they do not provide insight as to the clinical relevance of observed differences in treatment. In the context of our table, we feel this is not a limitation in that we consistently provide to the reader the observed summary values for each endpoint specific to each intervention group for each study, as well as group sizes, and thus we feel this provides the reader with all of the information needed to be aware not only of the statistical significance of any between-group differences, but also information to assess their degree of clinical relevance by reviewing the group summary data provided in the same table.

2- How the authors deal with the assessed quality of sleep related with the severity of the condition during hospitalization? According with Table 1, most of the patients were in ICU. Depending of the setting, the poor quality of sleep can be associated with the medication administered but the severity of condition is an important confounder. Some commentary about this could be done in Discussion section as a study limitation or a issue to tackle in future studies.

AR: Very good point. As per your suggestion we have added a few sentences to this effect to the discussion section.

3 - Most of studies in RCT (Table 1) are compound only by male patients. How this counfounder is addressed by the authors in the data analysis? The data about that is not clear but, in general, we have more men than women in this kind of study (using some medication in different groups) Some comment about this gender effect could also be done, presenting some consequences for the generalization (external validity) of the systematic review findings.

AR: Given the small samples sizes of many included studies it does appear that, at least in some cases, males predominate. However, when calculated the proportion of males in RCTs is actually 46%. Therefore we have not added any comments to the discussion regarding a possible gender effect.

Thanks for the opportunity to review the article and hope that the authors could find this suggestions useful for the article improvement.

Reviewer: 3
Reviewer Name: Christian Chan (with the assistance of Fiona Ho, PhD Candidate)
Institution and Country: Assistant Professor, Department of Psychology, The University of Hong Kong, Hong Kong
Competing Interests: None declared

This study is a systematic review of randomized controlled trials, comparative cohort studies, and prospective case series examining the efficacy and safety of pharmacologic intervention for improving sleep in hospitalized patients. A total of 15 studies were included in the review. Inconsistent results on sleep efficiency and sleep quality were reported. The manuscript was generally well-written. Below are suggestions for clarification and improvement.

Background
1. The authors cited a review study that found that non-pharmacologic treatments for sleep were not efficacious. The non-pharmacologic treatments in question appear to only include relaxation techniques and regulation of light/noise exposure and were intended for patients in non-intensive care. The authors' justification for the current study might be strengthened by discussing the pros and cons of pharmacologic treatments and non-pharmacologic treatments and the unique context of hospitalized patients (especially those in acute/critical care units).

AR: Thank you for this comment. We have added clarification and context around the statement you identified as per your suggestion.

Methods

2. The definition of sleep efficiency was not clear. Sleep efficiency is conventionally operationalized as the total sleep time divided by total time in bed. The authors might want to clarify what “observation” entailed in their calculation of sleep efficiency.

AR: Thanks for point this out. It is imperative that readers understand these definitions. Typically the period of observation was defined by the study protocolsWe have clarified this statement to read : “Sleep efficiency is defined as hours spent asleep divided by period of observation (in hours) by study investigators s…”.

3. It is unclear why among comparative cohort studies and prospective case series, only those that evaluated sleep using polysomnography were included.

AR: When designing the protocol our original plan was to only include RCTs that evaluated sleep quality using objective measures. Our scoping exercises indicated that there were few studies that met both criteria. We felt that a review inadequately addressed sleep quality would have limited value so it was decided to include prospective case series and cohort studies provided they used the gold standard method for assessing sleep quality. We felt that this was a reasonable strategy to strike a balance between what we wanted to evaluate from high quality evidence and what was actually available in the literature.

4. The authors are encouraged to perform an interrater reliability analysis whenever possible. In addition, did the authors also consider other possible biases (e.g., baseline characteristics and statistical power)?

AR: Unfortunately we did not perform an inter-rater reliability analysis for the two abstracters and screeners. If you feel strongly that this should be mentioned as a limitation of our methods we are happy to consider. Regarding a discussion of possible biases, we have elaborated, in our discussion, the potential confounding from variability with respect to 1)drug dosing, 2)small sample sizes, 3)variable severity of illnesses, 4)variable methods of outcome assessment and 5) a general lack of safety outcome reporting.

Results

5. Given the heterogeneity of the studies and the comparison groups, as underscored by the authors meta-analysis was not appropriate. However, it appears that meta-analysis was conducted for one outcome variable. For some of the variables, the authors included a forest plot but without conducting a meta-analysis. It would be helpful if the authors i) justified why meta-analysis was warranted for sleep efficiency, ii) kept consistent when to represent forest plots, and iii) indicated more consistently whether meta-analysis was conducted.

AR: In the results we state: “Only the two placebo controlled melatonin trials (total of 56 patients) were amenable to pooling of this outcome that revealed a non-statistically significant trend in favor of melatonin as compared to placebo (mean difference of 8.0%, [95% confidence interval: -1.5, 17.5]).”. In the discussion we have modified our original statement to read “…clinical and methodological heterogeneity and variable outcome reporting were judged to preclude performance of reliable meta-analyses in all cases except for two trials comparing melatonin to placebo with similar methods for evaluating sleep efficiency.” Regarding the figure, I think we clearly identify the one pooled estimate (and lack of pooled estimates for the rest). Given that we set out to pool data if possible I think we are obliged to do so for this one case as it was warranted based on our assessment of heterogeneity. Hopefully the modifications in the text and figures clearly identify when and why meta-analysis was conducted.

6. Whenever the data is available, please indicate the attrition rate measured at different periods/time points.

AR: The attrition rate (as I understand it being patients loss to followup or withdrawing consent after randomization, etc) was not reported for any study as such. However, not every randomized/enrolled patients was included in the analysis but the timing of exclusion was typically not available. Therefore, in table 3 under treatment group we have reported the sample size as the number enrolled and the number analyzed.

7. In addition to the study characteristics, please report demographic information of the participants (e.g., age and gender).

AR: Age, gender, setting and patient characteristics are available in Table 1.

8. Whenever the data is available, please report the long-term efficacy of pharmacologic treatment for improving sleep.

AR: Longitudinal outcomes were not reported for any study.

9. It is unclear whether the results warrant the conclusion that benzodiazepines are superior to placebo in reducing sleep latency. Gallais et al., for example, did not find significant differences between oxazepam and placebo. Given that a meta-analysis was not performed, the authors might want to specify that the results concerning the efficacy of benzos was found in several but not all studies and qualify their conclusion accordingly.

AR: I think that is fair. We have tempered our conclusions as per your suggestion in both the abstract and the discussion/conclusion.

Discussion
10. The discussion was clearly written. The authors explained the possible factors that interfered the sleep architecture in hospitalized patients. It would be helpful if the authors can more explicitly discuss the implications of the current systematic review, especially as both pharmacologic and non-pharmacologic treatments demonstrated conflicting and unpromising results. What is the alternative to pharmacologic treatment as first-line treatment in hospital settings, especially acute/critical ones?

AR: We have attempted to modify the language in the conclusions/discussion to reflect that this review describes a lack of supporting evidence rather than evidence suggesting that drug therapy is truly ineffective. I think better designed future studies are imperative to answering this question and we make suggestions as to the optimal design of future studies. Recommendations for alternative strategies to improve sleep in hospitalized patients without drugs or non-pharmacologic interventions are beyond the scope of this review.

Minor comments
11. The authors may consider assigning study numbers in tables to make it easier to follow.

AR : We agree that flipping between text and tables is difficult. Hopefully that will be better when/if published. Rather than using study numbers we have attempted to refer to studies by first author wherever possible (while also trying to avoid cumbersome reading). Hopefully it is easier to follow.

12. There are some grammatical errors in the manuscript.

AR: We identified and corrected a few that Word found for us. Hopefully the editorial review will pick out the rest.

13. Figure 3 (Risk of Bias for Prospective Cohort Studies) did not show properly (P.40).

AR: Not sure why it didn't appear properly for you. We will pay special attention to that figure when uploading and converting to PDF.

14. There are 2 independent sections "selection of studies" and "study selection" in Methods. Please either combine the two sections or changing the subheading.

AR: Thankyou for noticing that. We have removed the second subheading and incorporated that paragraph in to "Literature Search Strategy".

Reviewer: 4
Reviewer Name: Vineet Arora
Institution and Country: University of Chicago, United States
Competing Interests: none

This systematic review studies the evidence on drug therapy to improve sleep for hospitalized adults. The group identified 15 studies involving 861 patients and investigating a variety of sedatives and hypnotics from different drug classes. Studies were limited by low quality, small size, and lack of objective sleep measures. The review concluded that there was insufficient evidence to suggest that pharmacotherapy improves sleep in hospitalized patients. The paper is well written and the conclusion is sound. The authors are to be commended on rigor of the systematic review. Major critiques focus on the presentation of bundling results from ICU and ward together, as well as the summary in the discussion. Some additional suggestions to improve the the review are also highlighted below.
Background
• I was surprised at how short the background was. The context is important. By hospital, do you mean ICU, general ward, surgical are all? These are very different populations and environments so teasing out some of the prior work may be helpful. I think some additional details of the P in PICO is warranted.

AR: Based on your comments and those of other reviewers we have added text and context to the background as you suggested. In the methods section we have further clarified what we meant by hospitalized patients and added more detail about the population of interest.

Methods
• Please incorporate a reference for PICO

AR: done

• It seems like sleep duration should be an outcome of interest in addition to sleep efficiency and other metrics included.

AR: Sleep duration is influenced by the period of observation. For this reason sleep efficiency is regarded as a better metric by sleep researchers. Sleep duration is rarely reported in addition to sleep efficiency.

• I would be more specific on why studies of self-reported sleep were excluded as opposed to just citing the reference and letting the reader look it up. Also consider highlighting how many studies were in this category and got excluded in the results as that is also important to drive the field forward. It comes later in the discussion but its probably too late for the reader as they are wondering when it first appears.

AR: As per your suggestion we have elaborated on the reason for excluding studies where the sole method of sleep assessment was via patients self report. We also found a second reference published in march which further corroborates our statement. We have also added the number of studies excluded for this reason early in the results.

• Likewise, some context or rationale for only including cohort or case series of including polysomnography is needed. Were any with actigraphy excluded? Polysomnography is often only done in the ICU so this seems like ICU studies would be overrepresented in this inclusion.

AR: Based on yours and other's comments about this issue we have elaborated on why we chose to include only cohorts or case series where polysomonography was used to assess quality of sleep. When designing the protocol our original plan was to only include RCTs that evaluated sleep quality using objective measures. Our scoping exercises indicated that there were few studies that met both criteria. We felt that a review that inadequately addressed sleep quality would have limited value so it was decided to include prospective case series and cohort studies provided they used the gold standard method for assessing sleep quality. We felt that this was a reasonable strategy to strike a balance between what we wanted to evaluate from high quality evidence and what was actually available in the literature.

• It seems like grey literature review missed several conferences that may discuss sleep like ATS or Sleep?

AR: Sleep was reviewed, as it is a conference endorsed by the American Academy of Sleep Medicine. Unfortunately we did not review conference proceedings from ATS.

Results

• I personally think the tables are jarring when propofol infusion is compared with something like a po drug like melatonin. This speaks to the heterogeneity of clinical populations. I think subheaders in the table and corresponding organization in the results should discuss ICU studies separately from ward patietns. Being mechanically ventilated is just a very different thing and PSG is more likely in that population anyway. The current way the results are presented if you were an intensivist or hospitalist, you would have to read every paragraph in detail to find a study that applied to your clinical situation.

AR: We agree that the tables are very cumbersome. We hope that if/when published the print version is easier to maneuver however we have re-organized table 3 according to ICU vs non-ICU. Hopefully this makes it easier to read.

Discussion

• The first few lines of the discussion are out of place and should be in the introduction. I would replace this with an accessible summary of the review findings.

AR: We agree with your suggestion. The first few sentences were moved to the introduction. The Discussion now opens with a summary of the review findings.

• the Discussion should get into implications of these findings rather quickly. It doesn't come til much later.

AR: The discussions section has now been modified in many ways. The implications discussion is now earlier in the discussion as per your suggestion.

• Was there any difference in ICU vs ward studies? Any recommendations there would be helpful given the wide heterogeneity of the environment.

AR: Unfortunately given the degree of heterogeneity within and between ICU and non-ICU studies no sensitivity analysis or recommendations can be made. Although that was our original intention it is our opinion that it doesn't make sense to pool these studies in order to compare groups divided by severity of illness or location within hospital.

• Another possibility to consider is the disruptions are due to the medical care environment, and therefore are iatrogenic, so medications would not necessarily help and staff-based protocols would be needed. This should be elaborated on in the discussion.

AR: Again, this is a fair point. We have added a few sentences highlighting the need to evaluate a multimodal approach to insomnia in hospital in future studies. We also highlight the fact that environmental confounders are poorly documented in included studies.

VERSION 2 – REVIEW

REVIEWER	Altay Lino de Souza Federal University of Sao Paulo - Brazil
REVIEW RETURNED	18-Jun-2016

GENERAL COMMENTS	The suggestions were in general followed and the paper seems suitable for publication in my opinion. Congratulations to the authors.

REVIEWER	Christian Chan (with Fiona Ho, PhD Candidate) The University of Hong Kong, Hong Kong
REVIEW RETURNED	03-Jul-2016

GENERAL COMMENTS	The authors have done an adequate job in addressing our concerns and revising the manuscript. The questions we had are now clearly explained.